University of Zagreb

Faculty of Science

Sven Ljubić

ALPHA SATELLITE DNA IN EVOLUTION OF GENE MODULATORY NETWORKS

DOCTORAL THESIS

Zagreb, 2026.

University of Zagreb

Faculty of Science

Sven Ljubić

ALPHA SATELLITE DNA IN EVOLUTION OF GENE MODULATORY NETWORKS

DOCTORAL THESIS

Supervisor:

Đurđica Ugarković, PhD, Professor Emerita

Zagreb, 2026.

Sveučilište u Zagrebu

Prirodoslovno-matematički fakultet

Sven Ljubić

ALFA-SATELITSKA DNA U EVOLUCIJI MREŽA GENSKE MODULACIJE

DOKTORSKA DISERTACIJA

Mentor:

prof. emer. Đurđica Ugarković

Zagreb, 2026.

This doctoral dissertation was carried out as a part of the Doctoral Programme in Biology at the University of Zagreb, Faculty of Science, Department of Biology. It was made at the Ruđer Bošković Institute, under the supervision of Đurđica Ugarković, PhD, Professor Emerita.

The research presented in this dissertation was supported by the Croatian Science Foundation project: "Alpha satellite DNA in evolution of gene modulatory networks" (AlphaSatGenNet, IP-2019-04-6915, PI: Đurđica Ugarković, PhD, Professor Emerita).

TABLE OF CONTENTS

Acknowledgements

First and foremost, I wish to express my heartfelt gratitude to my mentors, Prof. emer. Đurđica Ugarković and Prof. Isidoro Feliciello. Your guidance, understanding, and the wealth of knowledge you shared have been invaluable throughout my doctoral journey and the writing of this dissertation.

I would like to further express my sincere gratitude to Isidoro Feliciello, who from the very beginning believed in and actively shaped my doctoral research by involving me in his own original research projects and by continuously guiding their theoretical, methodological, and critical development. His contribution was fundamental to the definition of the scientific direction, the conceptual architecture, and the analytical framework of the thesis. Continuous discussion and rigorous intellectual exchange were an essential and constant reference throughout the entire doctoral journey.

I also want to thank my colleagues at the Laboratory of Evolutionary Genetics. Your collaboration and support in experimental work and joint projects over the years have made the whole endeavor truly rewarding.

My appreciation extends to the dedicated teaching and technical staff at the Ruđer Bošković Institute and the Faculty of Science in Zagreb. The professional environment and valuable work experience you provided have significantly contributed to my scientific growth.

Lastly, a big thank you to my family and friends for your unwavering support and encouragement during my studies. Your faith in me has been a tremendous source of motivation.

SUPERVISOR INFORMATION

Đurđica Ugarković is a molecular geneticist born in Gospić, where she completed elementary school and gymnasium. She enrolled in chemistry in 1978 at the Faculty of Science, University of Zagreb, and graduated in 1982. Since 1983, she has been working at the Ruđer Bošković Institute (IRB), where she received her doctorate in 1988. She was elected to the position of scientific advisor at IRB in permanent tenure in 2003, and since 2025, she has had the status of scientist emeritus at IRB. Since 2005, she has been an adjunct full professor at the Faculty of Science, University of Zagreb, where she taught biochemistry. Since 2016, she has been a full professor at the University of Osijek. At the Ruđer Bošković Institute, she held the following positions: from 2010-2012, Assistant Director for Science and Education; from 2001-2005, Head of the Department of Molecular Biology; from 1991 to 2024, Head of the Laboratory for Evolutionary Genetics. She is involved in scientific research in the field of molecular genetics of eukaryotes, especially non-coding DNA and RNA, and is internationally recognized for it. She has also worked on the study of gene expression regulation in various eukaryotic systems. She is the leader or principal investigator of fifteen projects, among which international projects funded by the European Commission stand out. She has given 30 invited lectures at international conferences, including those held in: Beijing (China), Dublin (Ireland), Leicester (England), Barcelona (Spain), and London (UK). Among her greatest achievements is the proof that non-coding satellite DNA has an important regulatory role in the adaptation of the organism to various stressful changes in the environment, from the effects of temperature and chemicals to stress caused by various infectious and tumor diseases. She also discovered the unexpected evolutionary conservation of certain satellite DNAs, as well as their possible role in the divergence of populations and the emergence of new species. She proved the existence of dispersed elements of satellite DNA in the genomes of the flour beetle and in the human genome, and clarified the mechanisms of their spread. Her election to the European Molecular Biology Organization (EMBO) in 2000, as the first scientist from Croatia, stands out among the recognitions and awards. In 2016, she was elected as an associate member of the Croatian Academy of Sciences and Arts (HAZU). From 2000-2007, she was the president of the Croatian Genetic Society, and since 2007, she has been the editor of the book series Progress in Molecular and Subcellular Biology, published by Springer. To date, she has published 87 scientific papers and chapters in books, which have been cited over 2600 times, with an h-index of 29 (WoS database as of 2025).

University of Zagreb
Faculty of Science
Department of Biology

Doctoral thesis

ALPHA SATELLITE DNA IN EVOLUTION OF GENE MODULATORY NETWORKS

SVEN LJUBIĆ
Laboratory of Evolutionary Genetics, Ruđer Bošković Institute
Bijenička cesta 54, 10 000 Zagreb

Satellite DNAs are highly abundant sequences that form functional centromeres and pericentromeric heterochromatin across various eukaryotic organisms. Beyond their structural significance, they have been implicated in gene expression modulation and reactions to stress, although the precise molecular mechanisms remain elusive. This study focused on the human alpha satellite as a model to investigate the roles of alpha satellite DNA transcripts in gene regulation and stress responses. Employing molecular biology techniques—including chromatin immunoprecipitation (ChIP), immunofluorescence assays, and a novel modified quantitative real-time PCR (qPCR) method designed for accurate quantification of repetitive DNA—the impact of alpha satellite transcripts on gene expression modulation among alpha-associated genes and their involvement in antibiotic stress responses was uncovered. Furthermore, the potential of these transcripts as cancer biomarkers was explored. The findings indicate that there is a positive correlation between the expression of exogenous alpha satellite RNA and the downregulation of alpha-associated genes, suggesting that alpha satellite RNA plays a significant role in regulating their transcription. Additionally, it was discovered that commonly used antibiotics in cell culture, which are also prescribed for bacterial infections, tend to enhance the transcription of a major human pericentromeric alpha satellite DNA in cell lines at standard concentrations. This response, however, varied among different cell lines. Also observed was a positive correlation between higher antibiotic concentrations and the higher levels of alpha satellite transcription, which was associated with either a decrease in H3K9me3 or an increase in H3K18ac histone modifications at the alpha satellite arrays. Notably, the results demonstrated a significant increase in intracellular alpha satellite RNA levels in the blood of patients with metastatic cancers, particularly in those with metastatic castration-resistant prostate cancer, compared to controls. This suggests a connection between prostate cancer pathogenesis and blood levels of intracellular alpha satellite RNA. These findings propose that alpha satellite RNA could be utilized as a novel diagnostic blood biomarker for prostate cancer, alongside widely accepted PSA (prostate-specific antigen). Overall, research presented here contributes valuable insights into the ever-expanding field of satellite DNA functional studies.

131 pages, 6 figures, 0 tables, 186 references, original in: English
Keywords: satellite DNA, transcription, gene expression, antibiotic, histone modification
Supervisor: Đurđica Ugarković, PhD, Professor Emerita
Reviewers: Petra Korać, PhD, Full Professor
Maja Matulić, PhD, Full Professor
Damir Đermić, PhD, Scientific Advisor
Barbara Milutinović, PhD, Research Associate

Sveučilište u Zagrebu
Prirodoslovno-matematički fakultet
Biološki odsjek

Doktorski rad

ALFA-SATELITSKA DNA U EVOLUCIJI MREŽA GENSKE MODULACIJE

SVEN LJUBIĆ
Laboratorij za evolucijsku genetiku, Institut Ruđer Bošković
Bijenička cesta 54, 10 000 Zagreb

Satelitske DNA su vrlo zastupljene sekvencije koje formiraju funkcionalne centromere i pericentromerni heterokromatin u različitim eukariotskim organizmima. Osim njihove strukturne važnosti, povezane su s modulacijom ekspresije gena i odgovorom na stres, iako precizni molekularni mehanizmi nisu u potpunosti objašnjeni. Ova studija je usredotočena na ljudsku alfa-satelitsku DNA kao model za istraživanje uloge njenih transkripata u regulaciji gena i odgovorima na stres. Koristeći tehnike molekularne biologije—uključujući kromatinsku imunoprecipitaciju (ChIP), imunofluorescencijska bojenja i modificiranu kvantitativnu lančanu reakciju polimerazom u stvarnom vremenu (qPCR), metodu dizajniranu za preciznu kvantifikaciju repetitivne DNA—otkriven je utjecaj transkripata alfa-satelita na modulaciju ekspresije gena povezanih s alfa-satelitskom DNA i njihova uključenost u odgovore na stres uzrokovan antibioticima. Nadalje, istražen je potencijal ovih transkripata kao biomarkera za maligne tumore. Dobiveni rezultati ukazuju na pozitivnu korelaciju između ekspresije egzogene alfa-satelitske RNA i smanjenja ekspresije gena povezanih s alfa-satelitima, što sugerira da alfa-satelitska RNA ima značajnu ulogu u regulaciji njihove transkripcije. Također, otkriveno je da uobičajeni antibiotici koji se koriste u staničnoj kulturi, a koji se također propisuju za bakterijske infekcije, mogu pojačati transkripciju glavne ljudske pericentromerne alfa-satelitske DNA u staničnim linijama pri standardnim koncentracijama. Ipak, ovaj odgovor je varirao kod različitih vrsta stanica. Primijećena je i pozitivna korelacija između rastućih koncentracija antibiotika i povećane razine transkripcije alfa-satelitske DNA, praćena smanjenjem količine histonskih modifikacija H3K9me3, odnosno povećanjem broja H3K18ac modifikacija na alfa-satelitskim sekvencijama. Također, rezultati demonstriraju značajno povećanje razine unutarstanične alfa-satelitske RNA u krvi pacijenata s metastatskim karcinomima, posebno kod metastatskog karcinoma prostate otpornog na kastraciju, u usporedbi s kontrolama, što sugerira vezu između patogeneze karcinoma prostate i razine unutarstanične alfa-satelitske RNA u krvi. Na temelju ovih nalaza, pretpostavka je da bi alfa-satelitska RNA mogla služiti kao novi dijagnostički krvni biomarker za karcinom prostate, uz opće prihvaćeni PSA (antigen specifičan za prostatu). Konačno, istraživanja predstavljena u ovom radu pružaju vrijedan uvid u sve kompleksnije polje funkcionalnih studija satelitskih DNA.

131 stranica, 6 slika, 0 tablica, 186 literaturnih navoda, jezik izvornika: engleski

Ključne riječi: satelitska DNA, transkripcija, genska ekspresija, antibiotik, promjena histona

Mentor: Prof. emer. Đurđica Ugarković

Ocjenjivači: Prof. dr. sc. Petra Korać
Prof. dr. sc. Maja Matulić
Dr. sc. Damir Đermić, znanstveni savjetnik
Dr. sc. Barbara Milutinović, znanstveni suradnik

1. INTRODUCTION

1.1. Satellite DNA

Repetitive DNA can be categorized into two main types based on their genome organization: tandem and interspersed repeats. Tandem repeats consist of clusters of identical sequence units that are situated next to each other, organized in patterns known as tandem repeats (in which the sequences are organized in a head-to-tail arrangement) or inverted repeats (where they appear head-to-head or tail-to-tail). Conversely, interspersed repeats are scattered across the genome without a defined arrangement and do not appear adjacent to each other. Additionally, repetitive DNA can be differentiated by the degree of repetition into two further categories: highly repetitive and moderately repetitive DNA. These two groups were originally identified through their different reassociation rates, known as C_0t values, following a high-temperature melting process. Highly repetitive sequences, like telomeres and satellite DNA, tend to reanneal faster compared to moderately repetitive DNA, which includes elements like retrotransposons and rDNA genes (Britten and Kohne, 1968). Satellite DNA, discovered in the 1970s, was noted for its unique low-density buoyancy, allowing it to separate from the bulk of genomic DNA during cesium chloride density gradient analyses (Yasmineh and Yunis, 1974) (Figure 1).

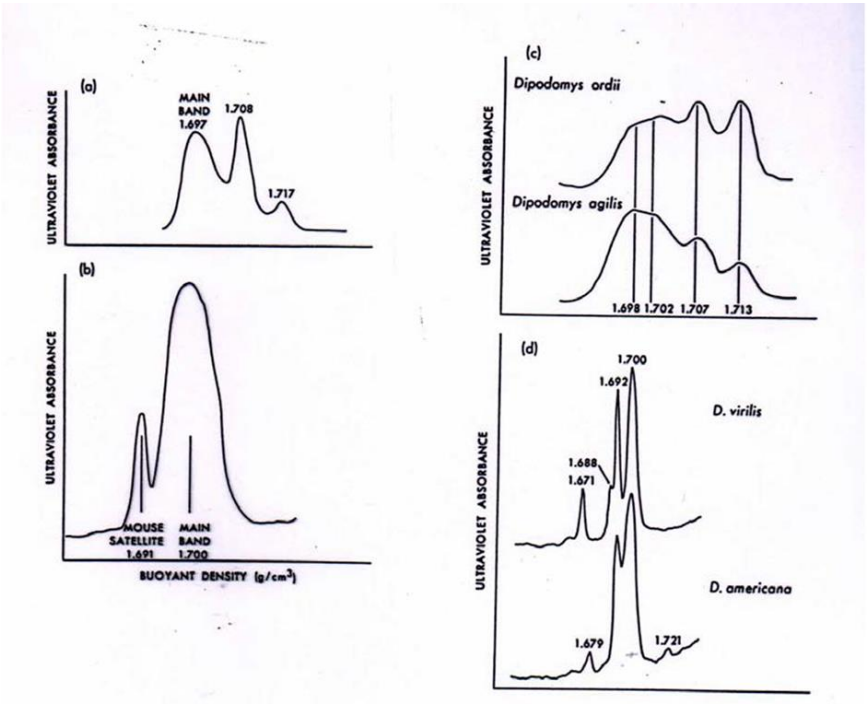


Figure 1. Buoyant density patterns of DNA preparations centrifuged to equilibrium in neutral cesium chloride. (a) *Ammospermophilus harrisi* (Mascarello and Mazrimas, 1977); (b) *Mus musculus* (Sutton and Walker, 1972); (c) *Dipodomys ordii* and *D. agilis* (Mazrimas and Hatch, 1972); (d) *Drosophila virilis* and *D. americana* (Gall and Atherton, 1974).

The repetitive tandem repeats or satellite DNA is typically described by three key features: the size of the repeat units, their sequence composition, and the overall length of the blocks or arrays it forms. The primary form of satellite DNA found in the human genome is known as alpha satellite (ASAT). This sequence is particularly concentrated around the centromeres of chromosomes and plays a crucial role in essential functions such as the assembly of centromeres and kinetochores, as well as in the formation of heterochromatin.

1.2. Alpha satellite DNA structure and organization

Alpha satellite DNA (ASAT) essentially consists of 171bp monomeric repeat units. It appears in two forms: higher-order repeat units (HORs), which are made up of organized and tandemly repeated 171bp monomers, and disordered stretches of monomers that lack any clear structural organization (Willard, 1985; Waye and Willard, 1987; Alexandrov et al., 1993b; Rudd et al., 2003) (Figure 2a). Typically, these two forms are found in close proximity, with unordered monomeric alpha satellites often positioned between large blocks of HOR alpha satellite DNA and the chromosome arms (Schueler et al., 2001; Rudd et al., 2003; Ross et al., 2005).

HOR alpha satellite arrays consist of a specific number of distinct 171bp monomers that are arranged in a head-to-tail fashion (Willard, 1985). The monomers within a higher-order repeat (HOR) unit exhibit 50–70% sequence identity, and the length of the HOR unit is defined by the point at which the next monomer closely matches the sequence of the first monomer in the HOR. Beyond these organized structures, monomers are scattered randomly, extending between the uniform array and the chromosome arm. Individual alpha satellite monomers are frequently mixed with other repetitive elements, including transposable elements and various forms of satellite DNA like satellite I and gamma satellite DNA (Trowell et al., 1993; Schueler et al., 2001; Kim et al., 2009) (Figure 2a). While HOR alpha satellite arrays are mostly consistent in structure, they are often interrupted by transposable elements, which can occur either between the HOR units or within them (Schueler et al., 2005; Miga, 2015; Jain et al., 2018). HOR units of alpha satellite DNA are defined based on

restriction enzyme sites, which typically make a single cut within each HOR. These cuts mark the end of one HOR monomer and the beginning of the next (Willard and Waye, 1987) (Figure 2b-e). Each chromosome contains these HOR units in extensive and largely uninterrupted repeats that can number in the hundreds or thousands. This arrangement results in a substantial, linear, and uniform sequence of nearly identical copies of tandem HOR units (Aldrup-MacDonald et al., 2016).

Alpha satellite DNA is commonly perceived as uniform across all centromeres in the human karyotype, but it actually displays a range of variations or polymorphisms that highlight its intricate nature and unique organization within the human genome, particularly its distinct chromosome-specific characteristics. Factors such as the sequence of a higher-order repeat, along with the quantity, type, and arrangement of the monomers that constitute the HOR unit, as well as the total number of HOR copies, contribute to this chromosome-specificity. While HORs within a specific chromosome can differ by only a few percent in sequence, those found on non-homologous chromosomes show only 50–70% similarity (Manuelidis, 1978; Willard, 1985).

Alpha satellite monomers vary in their sequences by about 10–40%, which depends on how closely they relate to the initially first identified human alpha satellite sequences (Wu and Manuelidis, 1980). While adjacent monomers can have noticeable sequence differences, there are shared similarities in their arrangement among different chromosomes, although the total count of monomers in a higher-order repeat (HOR) unit may differ. From analyses involving numerous individual monomers, twelve consensus alpha satellite monomers have been identified, labeled J1, J2, D1, D2, W1, W2, W3, W4, W5, M1, R1, and R2 (Alexandrov et al., 1988; Alexandrov et al., 1991; Alexandrov et al., 1993b; Rosandic et al., 2006; Shepelev et al., 2015). These monomers are categorized into five distinct suprachromosomal groups, or families (Figure 2b-e), defined by their sequence homology and the linear arrangement of monomers that form a HOR, which can be similar or even shared between chromosomes. The three primary suprachromosomal families (SF1–3) account for the majority of functional alpha satellite HORs located at the centromere's core (the kinetochore-forming region). These families form two dimeric and one pentameric HOR configurations. In contrast, SF4 and SF5 are monomer families that typically flank the functional HOR arrays, creating a boundary with the chromosome arms (Alexandrov et al., 1993b; Shepelev et al., 2009). SF4 consists solely of monomers that do not form HOR units, while SF5 can be organized into HOR units, but may also present an irregular structure that lacks HOR formation (Rosandic et al., 2006).

These classifications of suprachromosomal families highlight that there is considerable and intricate variation within alpha satellite DNA, stemming from differences in monomer composition and specific variations in the size and organization of HOR units for different chromosomes. Interestingly, within a single chromosome, the main HOR unit may show size variations, meaning that both variant HORs and canonical HORs can coexist in the same array (Durfy and Willard, 1987; Waye et al., 1987; Choo et al., 1990; Ge et al., 1992; Alexandrov et al., 1993a). These size variations in HORs are likely due to deletions that occur from unequal recombinations (Waye and Willard, 1986a, b; Warburton et al., 1993). Single nucleotide polymorphisms in specific monomers have been linked to particular HOR units. For instance, a SNP that introduces a *Hind*III site in monomer 13 of D17Z1 is found in a small portion of 16-mer HORs, as well as in a significant number of 13-mer HORs (Warburton and Willard, 1992, 1995).

These size and sequence variants, along with their spatial relationships within a specific HOR array, pose questions about how genomic variation may influence alpha satellite function. For instance, within D17Z1 on human chromosome 17 (HSA17), certain HOR variants—including SNPs and size variants—have been linked to issues with kinetochore architecture. This can lead to a diminished capacity to recruit or retain essential centromeric proteins (Maloney et al., 2012; Aldrup-MacDonald et al., 2016). It is possible that the organization or transcriptional patterns of HOR units (when comparing wild-type and variants) across the entire alpha satellite array could affect how well an alpha satellite array supports centromere assembly and kinetochore formation (Sullivan et al., 2017). The influence of such variations within regulatory and coding regions on gene expression is well-documented. Continued research aimed at identifying and characterizing structural and sequence polymorphisms within alpha satellite DNA, as well as their fundamental effects on basic chromosome function, is necessary to deepen our understanding of genomic variations and the roles of non-coding regions in the human genome.

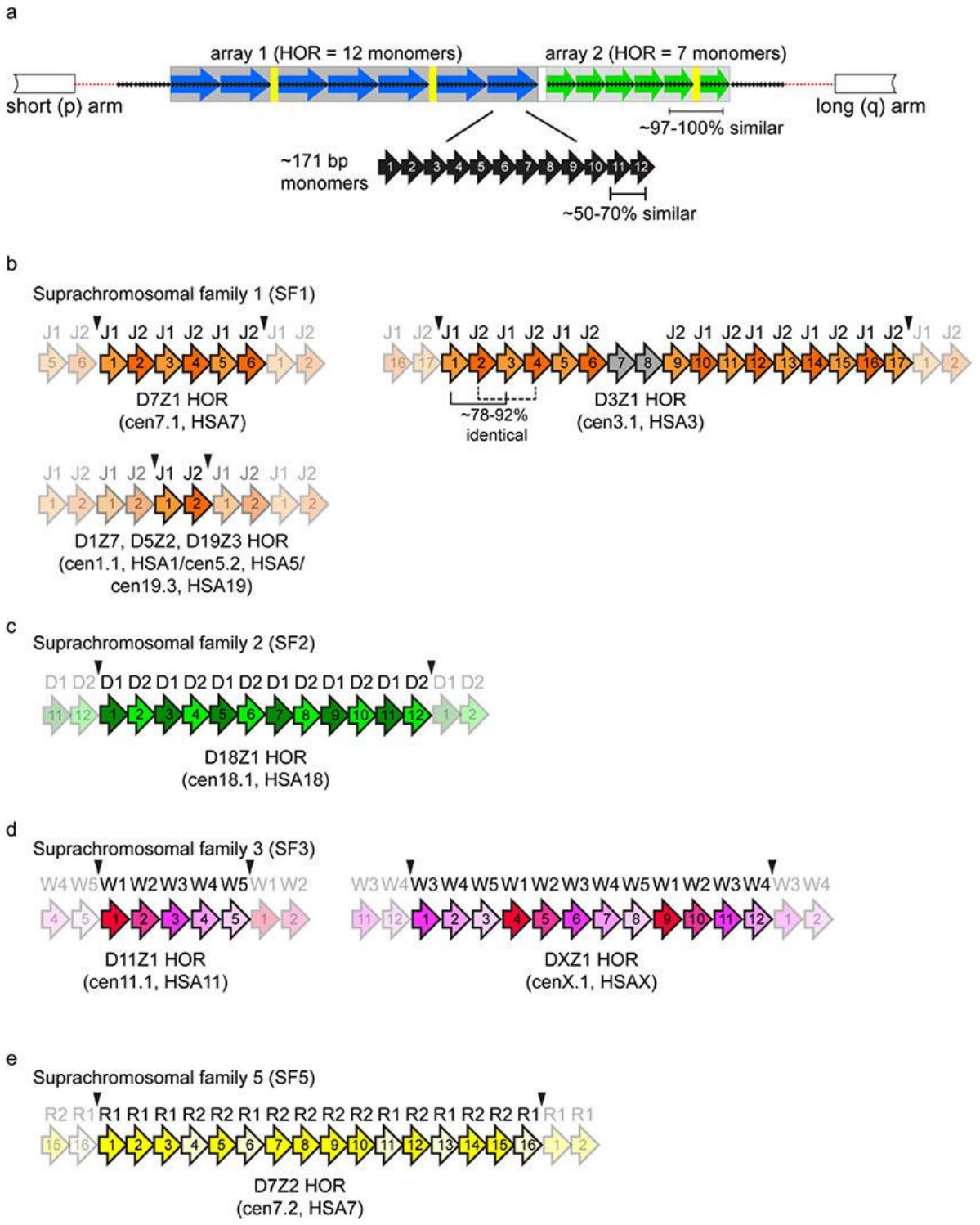


Figure 2. Array and chromosome-specific organization of alpha satellite DNA (McNulty and Sullivan, 2018). **(a)** Schematic of the general organization of alpha satellite DNA arrays at human centromere regions. Human chromosomes can have either one or more distinct higher-order repeat (HOR) arrays. HORs are array- and chromosome-specific. A defined number of individual monomers (black arrows) that are 50–70% identical in sequence are arranged tandemly to form a HOR unit; shown here as either a 12 monomer HOR (blue array)

or 7 monomer HOR (green array). Monomers are numbered by their position within the HOR and not based on their homology between two distinct HORs. The HORs are repeated hundreds to thousands of times to create homogenous arrays in which HOR within a given array are 97–100% identical. The HOR array is flanked by degenerate alpha satellite DNA monomers (small black arrays) that lack hierarchical structure and separate the HOR array from the chromosome arrays. HOR arrays are interrupted by other repetitive elements, such as transposable elements (TEs, yellow) but the extent of TE distribution across arrays is unclear due the lack of linear, contiguous assemblies of endogenous alpha satellite arrays. **(b)** Alpha satellite HOR arrays have been classified into suprachromosomal families (SF) that are related based on monomer type and organization. SF1 arrays are organized as alternating dimers of J1 and J2 monomers (D7Z1, cen7.1), although variation in the regular organization of monomers occurs on some chromosomes, like the D3Z1 (cen3.1) array of *Homo sapiens* chromosome 3 (HSA3). Additionally, the HORs can be shared among chromosomes, such as the D1Z7 (cen1.1) array that is also present as D5Z2 (cen5.2) on human chromosome 5 (HSA5) and D19Z3 (cen19.3) on HSA19. Each array-specific HOR unit is operationally defined by restriction enzyme sites (black arrowheads) that demarcate the last monomer of one HOR unit and the first monomers of the next HOR unit. Opaque shading illustrates the linear, reiterated nature of HOR units.
(c) SF2 is composed of a different dimeric structure based on D1 and D2 monomers. D18Z1 (cen18.1) on HSA18 has SF2 organization. **(d)** SF3 is based on a pentameric organization of monomers W1-W5. D11Z1 (cen11.1) is an example of a perfect pentameric HOR unit, while DXZ1 has an irregular organization of W1-W5 monomers. (**e**) SF5 arrays are defined by R1 and R2 monomers, although they largely lack the dimeric organization observed for SF1 and SF2 arrays. Some arrays have HOR unit structure, such as the D7Z2 (cen7.2) array of HSA7. D_chromosome_Z_number is the original Human Genome Project locus definition of alpha satellite arrays. The newer UCSC Genome Browser annotations of distinct HOR arrays (cen_chromosome number.array number) are also included.

1.3. Alpha satellite DNA role in centromere and kinetochore functional assembly and chromosome stability

The assembly of the human centromere and the formation of kinetochores involve the recruitment of approximately 100 proteins to regions of alpha satellite DNA (Musacchio and Desai, 2017). Essentially, the centromere is where a distinct type of chromatin is formed,

which serves as a crucial foundation for the attachment of architectural proteins. These proteins give structure to the kinetochore, a complex network that connects to microtubules and facilitates the movement of chromosomes along the spindle during cell division. Alpha satellite DNA is a major constituent of the centromeric region of the human chromosome. Functional centromere predominantly consists of ASAT higher-order repeat (HOR) units, flanked by pericentromeric regions characterized by disordered monomeric repeats, clearly establishing two different types of satellite organization (Jaggi et al. 2026; Figure 3).

CENP-A is a unique histone variant that plays a crucial role in defining centromere identity and serves as an important epigenetic marker for centromeres. The presence of CENP-A at alpha satellite DNA regions sets apart the centromere from the rest of the genome. Initially identified in sera from patients with CREST (Calcinosis, Raynauds phenomenon, Esophageal dysmotility, Sclerodactyly, Telangiectasia) syndrome, CENP-A was one of three antigens found biochemically that have since been confirmed as components of the centromere through immunostaining of mitotic cells (Earnshaw and Rothfield, 1985). This 17 kDa protein was labeled CENP-A, while the remaining two were classified as CENP-B (80 kDa) and CENP-C (140 kDa). Further research revealed that CENP-A co-purifies with nucleosome core particles and histones, underscoring its role as a centromere-specific histone involved in a crucial chromatin nucleoprotein complex (Palmer et al., 1987, 1991). It is found at all endogenous human centromeres and is particularly vital for the functional centromeres of dicentric chromosomes (Vafa and Sullivan, 1997, Warburton et al., 1997, Ando et al., 2002). Interestingly, in humans, CENP-A's unique connection with alpha satellite DNA goes beyond initial incorporation into chromatin during S phase—it is deposited instead during the late M and G1 phases (Shelby et al., 1997, Shelby et al., 2000, Jansen et al., 2007). This timing, where CENP-A synthesis occurs in G2 while its deposition happens in G1, is crucial for its association with the CENP-A specific chaperone protein, Holliday Junction Recognition Protein (HJURP) (Dunleavy et al., 2009, Bodor et al., 2013). The presence of CENP-A at the centromere is believed to be carefully regulated through its interactions with other centromere proteins, post-translational modifications of centromeric histones, and the transcription of alpha satellite DNA (Molina et al., 2016, Ohzeki et al., 2016, McNulty et al., 2017). Maintaining CENP-A effectively relies on its relationship with CENP-B and CENP-C, as well as its spatial positioning within alpha satellite DNA arrays (Fachinetti et al., 2013, Fachinetti et al., 2015, Ross et al., 2016).

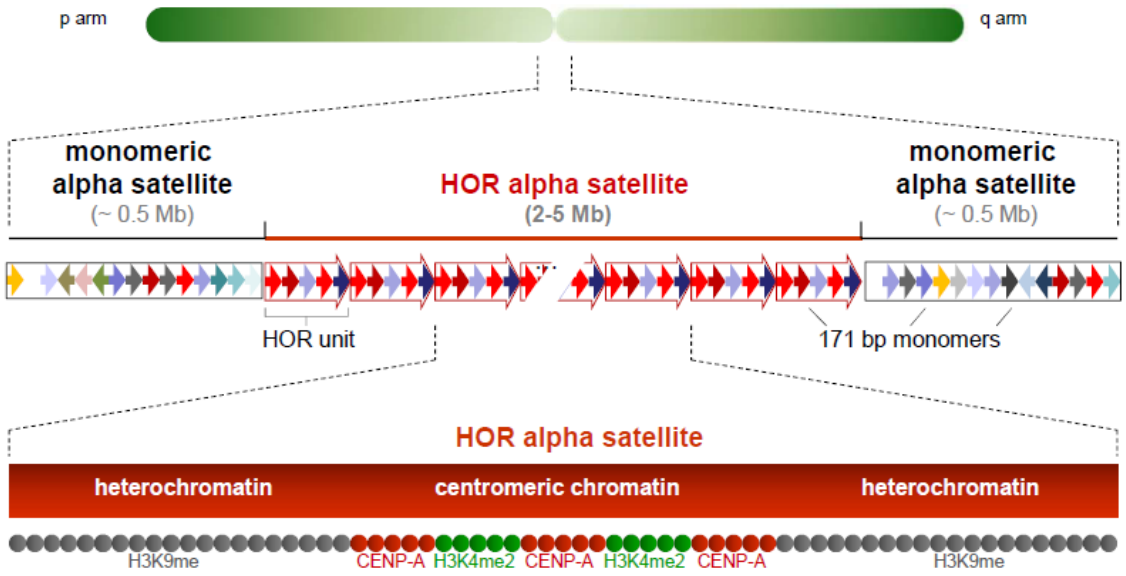


Figure 3. Alpha satellite organizational profile in the centromeric region of the human chromosome.

CENP-A is a part of a stable pre-kinetochore complex together with CENP-B and CENP-C (Ando et al., 2002). In mammals, CENP-B , an alpha satellite DNA binding protein, is an 80 kDa kinetochore protein that interacts with the CENP-B box, which is a specific 17-base pair sequence motif (5'-T/CTCGTTGGAAA/GCGGGA-3') (Masumoto et al., 1989). The CENP-B box is found in a subset of alpha satellite monomers (Muro et al., 1992; Ikeno et al., 1994) across all human chromosomes, with the exception of HSAY (Muro et al., 1992; Haaf and Ward, 1994). The positioning of CENP-B boxes can differ based on the chromosome-specific higher-order repeat (HOR) structures. Each suprachromosomal family contains alpha satellite monomers with particular sequences and higher-order characteristics, but they can be generally categorized into two main types: A-type and B-type monomers (Rosandic et al., 2006). A-type monomers include J1, D2, W4, W5, M1, and R2, whereas B-type monomers consist of J2, D1, W1-W3, and R1. The distinction between A and B monomers is evident in their sequences at positions 35–51, which relate to protein binding interactions. While B-type monomers have CENP-B boxes, A-type monomers feature a binding site for pJα (Rosandic et al., 2006), a protein whose properties and role remain largely unexplored. Notably, DYZ3 of HSAY entirely lacks CENP-B box-containing monomers but does include those with the pJα motif. Given that DYZ3 interacts with CENP-

A and other proteins associated with centromeres and kinetochores, this suggests that pJα might play a role in kinetochore assembly, though this mechanism is not yet completely understood.

Until recently, researchers largely overlooked the functional role of CENP-B in centromeric chromatin, particularly since it often appears at centromeres that are inactive (Earnshaw et al., 1989; Sullivan and Schwartz, 1995). Furthermore, CENP-B is found in the additional HOR alpha satellite arrays of multi-array chromosomes, such as HSA7 and HSA17. However, there has been a resurgence of interest in the importance of CENP-B in establishing, structuring, and maintaining centromere chromatin. The formation of new centromeres relies on CENP-B-associated alpha satellite DNA (Ohzeki et al., 2002; Okada et al., 2007). Additionally, it is believed that CENP-B plays a crucial role in positioning CENP-A nucleosomes and stabilizing both CENP-A and CENP-C within centromeric chromatin (Yoda et al., 1998; Okada et al., 2007; Hasson et al., 2013; Fachinetti et al., 2015).

CENP-C is a crucial component of the constitutive centromere-associated network (CCAN), which plays a vital role in connecting the inner and outer kinetochore. Its importance lies in facilitating the recruitment of CENP-A and aiding in the maturation of the kinetochore. The current understanding is that CENP-C stabilizes CENP-A nucleosomes by interacting with CENP-B and CENP-N, which is a part of the CENP-L-N complex (Carroll et al., 2009; Guo et al., 2017; Cao et al., 2018). Additionally, CENP-C, along with CENP-T from the CCAN, forms direct connections between the inner kinetochore and the NDC80/HEC1 complex in the outer kinetochore (Musacchio and Desai, 2017). Furthermore, CENP-C has the ability to bind to both alpha satellite DNA and RNA (Politi et al., 2002; Trazzi et al., 2002; Du et al., 2010; Shono et al., 2015; McNulty et al., 2017). Notably, while both CENP-B and CENP-C bind to the same type of alpha satellite DNA (specifically, HOR), they are found in spatially distinct locations, indicating that they likely engage with different HORs or various regions within the same HOR.

1.4. Chromatin characteristics of alpha satellite DNA-associated genomic regions

Genomic DNA is organized into chromatin by wrapping around two copies of core histones (H2A, H2B, H3, and H4) (Kornberg, 1974). This chromatin structure can be further compacted by chromatin remodeling proteins. Typically, gene-rich areas are organized as euchromatin, which features loosely packed nucleosomes that allow access for polymerase and transcription factors. In contrast, regions with fewer genes are structured as

heterochromatin, which tends to resist transcription and shows unique interaction patterns with transcription factors and other proteins. Post-translational modifications of the tails of histones (as well as certain non-coding RNAs) serve as signals, guiding the recruitment of chromatin remodeling proteins and transcription factors to specific genomic sites. Certain histone modifications clearly distinguish euchromatin from constitutive heterochromatin. For example, di- and tri-methylation of H3K4 and H3K36 (H3K4me2/3, H3K36me2/3), along with acetylation of H3 (K9, K14) and H4 (K5, K8, K12, K16), indicate transcriptionally active and open chromatin states (Peterson and Laniel, 2004). Conversely, modifications such as H3K9me2/3 and H3K27me3 are linked to repressive facultative or constitutive heterochromatin. Interestingly, studies using immunocytological techniques alongside chromatin immunoprecipitation (ChIP) have discovered that alpha satellite DNA can assemble into various types of chromatin, sometimes existing simultaneously on the same array (Lam et al., 2006; Mravinac et al., 2009; Ohzeki et al., 2012; Bailey et al., 2016).

Historically, repetitive DNA in mammals has been viewed as heterochromatic. However, the regions around centromeres show a unique pattern of histone modifications that sets them apart from both euchromatin and classic heterochromatin. Centromeric chromatin features a combination of nucleosomes that include the standard histone H3 alongside the specialized variant CENP-A, also known as CENH3 (Blower et al., 2002). This distinct organization of H3 and CENP-A nucleosomes is referred to as "centrochromatin" (Sullivan and Karpen, 2004). Within centrochromatin, H3 histones exhibit elevated levels of modifications such as H3K4me2 and H3K36me2, which are typically linked to transcription-friendly chromatin environments (Lam et al., 2006; Bergmann et al., 2011). Acetylated histones, usually found in euchromatin, are only briefly associated with centrochromatin. It is believed that these modifications play a key role in facilitating the loading of new CENP-A and help establish a boundary that prevents heterochromatin from invading centrochromatin (Molina et al., 2016; Ohzeki et al., 2016; Shang et al., 2016).

Centrochromatin is found next to pericentric heterochromatin, which is rich in modifications such as H3K9me2, H3K9me3, and H3K27me3 (Lam et al., 2006; Ohzeki et al., 2016) (Figure 4). In humans, about 35% of an alpha satellite array is incorporated into centrochromatin, while the rest is formed into heterochromatin (Lam et al., 2006; Mravinac et al., 2009; Sullivan et al., 2011; Bailey et al., 2016). The transition between centrochromatin and heterochromatin within a single alpha satellite array is not well defined. Due to variability in alpha satellite array sizes, the CENP-A domains differ according to the size of

the alpha satellite, and the surrounding heterochromatin also varies among homologous centromeres. Heterochromatin may serve as a significant boundary between the core centromere located on the alpha satellite and the chromosome arms, as its depletion or removal permits centrochromatin to expand and allows distinct chromatin domains to redistribute along the alpha satellite (Mravinac et al., 2009; Sullivan et al., 2011; Sullivan et al., 2016). The separation of heterochromatin and centrochromatin, as well as the formation of new CENP-A, seems to be regulated by the interaction between heterochromatin formation driven by SUV39H1/2 and the modification of nearby centrochromatin through the acetyltransferase KAT7/HBO1/MYST2 (Ohzeki et al., 2016). The action of these chromatin-modifying enzymes could be influenced by protein-protein interactions and/or by RNAs generated from alpha satellite regions (Johnson et al., 2017; McNulty et al., 2017).

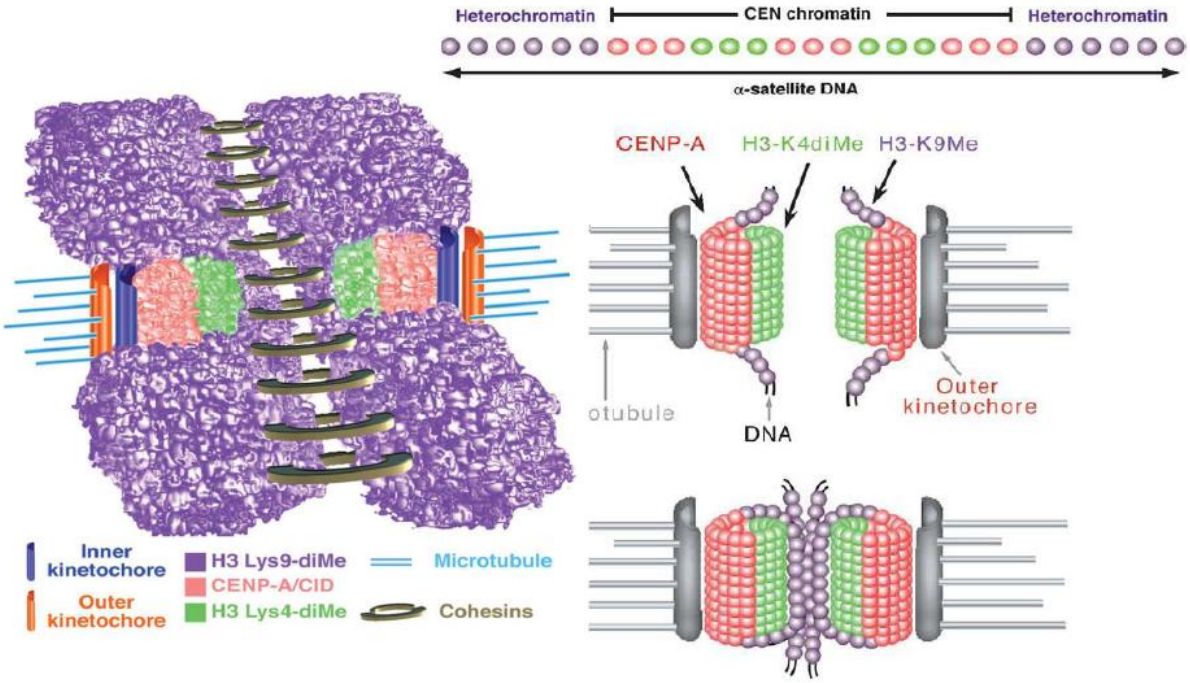


Figure 4. Chromatin signature of alpha satellite DNA regions during kinetochore assembly (modified from Gieni et al., 2008).

1.5. Transcription of alpha satellite DNA

The presence of repetitive regions in heterochromatin has led to the notion that these sequences are not actively engaged in transcription. However, recent evidence suggests that active transcription is actually quite common among various satellite DNAs, including alpha

satellite DNA. Notably, satellite RNAs are found in abundance in mammalian cells and are frequently stable associates of chromatin (Hall et al., 2014). Recent research indicates that the traits and roles of these transcripts are crucial for distinct chromosomal functions, as well as for development, cellular responses to stress, and cancer progression.

Studying the transcription of human centromeres poses a significant challenge due to the structural makeup of these regions. For chromosomes that feature single alpha satellite higher-order repeat (HOR) arrays, like HSAX, alpha satellite DNA becomes a part of centrochromatin, which is where the kinetochore develops, but also integrates into pericentric heterochromatin. This overlap means that research on bulk alpha satellite RNA cannot pinpoint whether the RNA comes from the centromere or the pericentromere. This emphasizes the importance of including protein-association data when examining alpha satellite DNA and RNA. Additionally, the potential presence of multiple unique arrays on a single chromosome adds another layer of complexity to understanding how alpha satellite DNA functions in both centromeric and pericentromeric roles.

Alpha satellite transcripts have been reported in various human cell types, but there is some inconsistency in these findings regarding their localization, length, binding partners, and functions (Wong et al., 2007; Chan et al., 2012; Ideue et al., 2014; Quenet and Dalai, 2014; Liu et al., 2015; McNulty et al., 2017). Early studies suggested that alpha satellite RNA was primarily located in the nucleolus before being relocated to the centromere at the start of mitosis, facilitated by CENP-C (Wong et al., 2007). However, further research has indicated that alpha satellite can also localize to centromeres during both interphase and metaphase (Ideue et al., 2014; Quenet and Dalal, 2014; McNulty et al., 2017), where it co-localizes with essential centromeric proteins like CENP-A. One of the more ambiguous aspects of alpha satellite RNA involves its binding partners and its overall role at the centromere. Two significant proteins, Aurora B and Sgo1, which are key players in cell division by coordinating spindle microtubule attachment and sister chromatid separation, respectively, appear to be influenced by alpha satellite transcription and alpha satellite RNA (Ideue et al., 2014; Liu et al., 2015). The process of RNAP II (RNA polymerase II) transcription is critical for Sgo1's relocation from the outer kinetochore to the inner centromere, an essential step for maintaining centromeric cohesion. Furthermore, alpha satellite RNA has been shown to directly interact with Aurora B, and when alpha satellite RNA is depleted, it results in abnormal cell shapes and cell division errors (Ideue et al., 2014). Similar effects were noted following the depletion of both minor and major satellites in mouse cells. These findings

highlight the importance of alpha satellite transcription and RNA transcripts for maintaining normal cell functions.

Research involving primary and transformed human cell lines has revealed that alpha satellite arrays generate sequence-specific non-coding transcripts. These transcripts interact with centromeric proteins CENP-A and CENP-C, in addition to the alpha satellite DNA binding protein CENP-B (Quenet and Dalal, 2014; McNulty et al., 2017). It is important to note that human centromeric regions often consist of multiple and distinct alpha satellite arrays. Interestingly, even inactive (non-kinetochore forming) alpha satellite arrays still produce alpha satellite RNA (Johnson et al., 2017; McNulty et al., 2017) (Figure 5a). The RNA from these different arrays seems to be integrated into functionally unique chromatin complexes and RNA from inactive arrays does not associate with CENP-A or CENP-C. In contrast, during the activation of kinetochore-forming arrays, alpha satellite RNA plays a crucial role in the loading of centromeric proteins (Figure 5b). While the specific regions of the RNAs that interact with CENPs, as well as the binding sites for RNA on these centromeric proteins, remain unidentified, it is known that CENP-C acts as an RNA-binding protein (Du et al., 2010).

Mammalian cells produce three types of RNA polymerases: RNAP I, II, and III, each responsible for transcribing different RNA variants. Specifically, RNAP I handles most ribosomal RNA genes except for 5S rRNA, while RNAP II focuses on protein-coding genes and microRNAs. RNAP III is in charge of transcribing tRNA genes, 5S rRNA, and some small nuclear RNAs. When it comes to alpha satellite RNA, there is no clear consensus on which polymerase takes the lead. All three have been proposed as candidates. However, research has shown that RNAP II is likely playing a significant role at human centromeres, especially given its presence and the influence of polymerase inhibitors on the associated transcripts. Studies suggest that while RNAP II may actively transcribe human alpha satellite DNA, RNAP I could be necessary for ensuring these transcripts are properly localized. Additionally, there is evidence that transcription from transposable elements dependent on RNAP III might help promote nearby alpha satellite transcription. To pin down which polymerase is truly responsible for transcribing repetitive DNA, higher quality assemblies of these repetitive regions could be crucial. They may reveal specific promoter elements and genetic signatures that offer more definitive insights into the polymerase involved in this complex transcription process.

Currently, there is not much data regarding the post-transcriptional processing of alpha satellite RNAs, including aspects like capping, splicing, and polyadenylation. Ideue et al. (2014) indicated that some alpha satellite RNAs may indeed lack poly-A tails. However, research has shown that repetitive RNAs, including alpha satellite RNA, tend to have a slow turnover rate, suggesting a strong inherent stability (McNulty et al., 2017). It remains unclear whether this stability is due to a poly-A tail or possibly results from another protective mechanism, such as RNA-DNA hybrid formation or various post-transcriptional modifications.

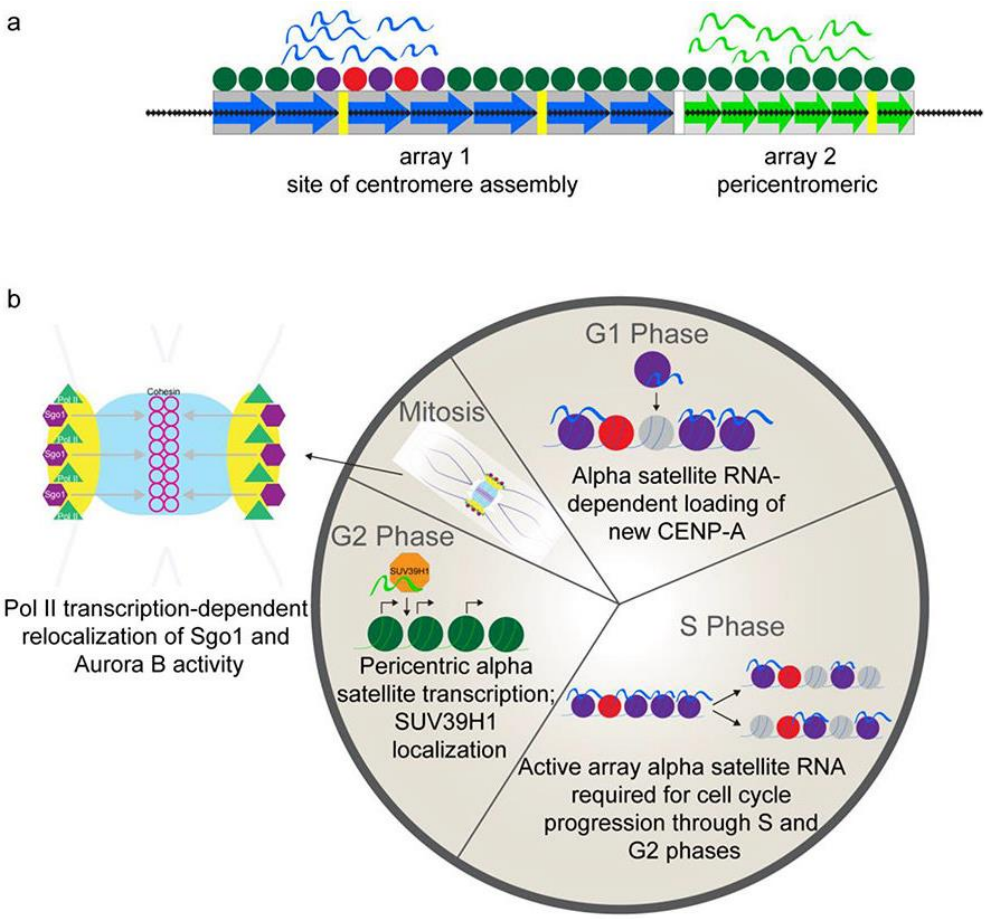


Figure 5. Alpha satellite transcription and non-coding RNAs play distinct roles at the centromere and pericentromere throughout the cell cycle (McNulty and Sullivan, 2018). (**a**) Schematic of the dual transcription observed at active and inactive alpha satellite DNA arrays at human centromere regions. The CENP-A domain (red and purple circles) forms on a portion of array 1 (blue arrows) and RNAs produced from this array (blue ribbons) remain associated with the centromere. Adjacent to array 1, array 2 (green arrows) is pericentromeric and associated with heterochromatic nucleosomes (green circles) but, like array 1, produces

alpha satellite RNAs (green ribbons) that localize *in cis*. (**b**) Summary diagram of the proposed roles of alpha satellite transcription and the resulting non-coding RNAs at each stage of the cell cycle. Alpha satellite RNAs produced from the active array help load new CENP-A at the centromere in early G1. In S phase, CENP-A is distributed semi-conservatively to each daughter strand. Although a precise role for alpha satellite transcription or RNA has not yet been elucidated, the presence of these transcripts is required for normal cell cycle progression through S and G2 phases. Alpha satellite transcription at inactive, pericentric arrays is thought to occur in G2 phase, shortly before the onset of mitosis. These RNAs are required for SUV39H1 (orange octagons) localization to the pericentromere. Sgo1 and Aurora B are both key players in mitosis and have been identified as alpha satellite RNA-binding partners. RNAP II-dependent transcription of alpha satellite is involved in relocalizmg Sgo1 (purple hexagons) from the kinetochore to cohesin (pink rings) in the inner centromere.

1.6. Satellite transcripts in heterochromatin formation and gene expression modulation

Satellite DNA repeats found in (peri)centromeric heterochromatin are not just silent parts of the genome; they are actively transcribed and play crucial roles in forming and maintaining heterochromatin and ensuring proper centromere function (Mihìc et al., 2021; Smurova and De Wulf, 2018; Leclerc and Kitagawa, 2021). In different organisms like insects, nematodes, and plants, the products of satellite DNA transcription include small interfering RNAs (siRNAs) and PIWI-interacting RNAs (piRNAs), which contribute to the epigenetic regulation of gene expression through RNA interference (RNAi) mechanisms (Grewal and Elgin, 2007; Fagegaltier et al., 2009; Holoch and Moazed, 2015). In the fruit fly *Drosophila melanogaster*, piRNAs derived from female germline satellite DNA help establish heterochromatin at their respective genomic locations, and their transcription depends on heterochromatin (Wei et al., 2021). For males of this species, siRNAs from satellite DNA play a role in modifying chromatin at specific X chromosome satellite repeats (Joshi and Meller, 2017).

Similarly, in another insect, the beetle *Tribolium castaneum*, the primary (peri)centromeric satellite DNA, TCAST1, is expressed into piRNAs in germline cells and siRNAs in somatic cells (Sermek et al., 2021). Both TCAST1 piRNAs and siRNAs are essential for the establishment and maintenance of heterochromatin, acting *in cis* at their genomic origins. It has been proposed that the different processing of TCAST1 transcripts is

facilitated by the existence of TCAST1 piRNA and siRNA-specific heterochromatic clusters, each regulated separately. This compartmentalization may enable the same satellite DNA to react to various signals and engage in multiple cellular activities (Sermek et al., 2021).

In mammals, RNA Polymerase II plays a crucial role by transcribing pericentromeric satellite DNA repeats into long non-coding RNA, utilizing a bidirectional approach. These double-stranded RNAs can be detected and potentially cleaved by Dicer1, which appears to regulate the levels of satellite transcripts during mitosis (Fukagawa et al., 2004; Huang et al., 2015). Conversely, during meiosis, the MIWI protein, guided by piRNAs and in collaboration with Dicer1, cuts back the availability of satellite RNA and manages its cellular concentration (Hsieh et al., 2020). Notably, the process by which Dicer1 influences the transcription of pericentromeric satellite DNAs, even when small RNAs are absent, seems to be a conserved mechanism from fission yeast to mammals (Gutbrod et al., 2022). In mice, transcripts of the major pericentromeric satellite DNA create RNA:DNA hybrids, which assist in the retention of heterochromatin proteins such as HP1 (Maison et al., 2011), along with methyltransferases SUV39h1 and SUV39h2 (Johnson et al., 2017; Velazquez Camacho et al., 2017). There is a suggestion that transcripts enriched with m6A RNA modification may enhance their interaction with heterochromatin (Duda et al., 2021). This heterochromatic condition at the pericentromere is significant for attracting and/or sustaining cohesin at the centromere, ensuring proper sister chromatid separation (Gutbrod and Martienssen, 2020). Additionally, certain microRNAs, including miR-30a-3p, miR-30d-3p, and miR-30e-3p, which have complementary sequences to major mouse satellite DNA, potentially regulated by the Argonaute protein 1 (AGO1), are implicated in modulating the expression of major satellite transcripts in mouse embryonic stem cells (Müller et al., 2022).

Satellite DNAs are primarily found in tandemly repeated sequences that are grouped together in vast arrays, mainly within gene-poor regions of constitutive heterochromatin located near centromeres and telomeres. While longer arrays of these tandem repeats in euchromatin are quite uncommon—likely due to the instability from intrastrand homologous recombination—there are instances of such blocks in the euchromatin of species like *D. melanogaster* (Kuhn et al., 2012) and *Triatoma infestans* (Pita et al., 2017). Interestingly, in the beetle *Tribolium castaneum*, some euchromatic satellite DNA arrays even exhibit higher-order repeats (Vlahović et al., 2017; Pavlek et al., 2015). Bioinformatic studies on sequenced genomes have uncovered various instances of single repeats or short arrays of satellite DNAs scattered throughout euchromatin, especially near gene regions in several insects, including *Tribolium castaneum*, *Drosophila melanogaster*, and *Locusta migratoria* (Ruiz-Ruano et al.,

2016; Brajković et al., 2012, 2018; Kuhn et al., 2012). In mammals, single repeats of major human alpha satellite DNA (Feliciello et al., 2020a, b) and major mouse satellite DNA (Bulut-Karslioglu et al., 2012) can also be found interspersed among genes or within introns. This suggests a widespread pattern where the majority of satellite DNA repeats are clustered in pericentromeric constitutive heterochromatin, while single repeats or short multimers are more widely distributed within euchromatin across various species.

Much like transposable elements, euchromatic satellite repeats are capable of undergoing cycles of proliferation. The observed variation in insertional polymorphism of these satellite repeats among different populations of the same species—or even among individuals within a single population—highlights their continued movement within euchromatin and show the mutational potential of satellite DNAs (Feliciello et al., 2015a, b). While a new insertion of a satellite repeat in euchromatin may not usually impact genes, under specific conditions, such as heat stress, it has the ability to influence the expression of adjacent genes through a specific epigenetic mechanism (Feliciello et al., 2015a). While processes of reverse transcription/integration and transposition may be associated with the dispersion of some alpha repeats, the primary mechanism of propagation appears to involve extrachromosomal circles of alpha satellite DNA. This process is aided by short sequences of homology between the alpha repeats and their target sequences (Feliciello et al., 2020b). In human cells, extrachromosomal circular DNAs (eccDNA) made up of alpha satellite repeats can be found, with sizes varying from less than 2 kb to more than 20 kb (Cohen et al., 2010), which highlights how tandemly arranged alpha repeats have a tendency to form eccDNA.

Heat stress in the beetle *T. castaneum* leads to increased transcription of TCAST1 satellite DNA. This surge in transcription aligns with a higher presence of repressive heterochromatin marks (H3K9me2/3) on satellite repeats found in constitutive heterochromatin, as well as on scattered TCAST1 satellite elements located within euchromatin and extending to regions up to 6 kb from the insertion site (Feliciello et al., 2015a). Consequently, the TCAST1 satellite DNA repeats dispersed in euchromatin appear to act as nucleation sites for the temporary formation of heterochromatin. This process results in partial suppression of nearby genes during heat stress, marking the first experimental evidence of satellite DNA's role in modulating gene expression (Feliciello et al., 2015a). Since this novel mode of gene expression regulation does not seem to be unique to a specific satellite DNA it is hypothesized that different satellites which are partially dispersed in the vicinity of genes and whose transcription is induced upon heat stress, could influence the

expression of associated genes by the same mechanism of temporary heterochromatinization (Figure 6).

Recent studies have shown that under heat stress, there is an increase in H3K9me3 levels associated with repeats of alpha satellite DNA found in both heterochromatin and euchromatin (Feliciello et al., 2020a). This increase of H3K9me3 at alpha repeats correlates with the activation of alpha satellite DNA transcription, while the spread of H3K9me3 up to 1–2 kb from their insertion sites indicates that euchromatic alpha repeats play a role in modulating local chromatin structure. All these results suggest that epigenetic effects, in particular H3K9me3 enrichment mediated by siRNAs and piRNAs, respectively, are common for some satellite DNAs and transposons, becoming pronounced upon stress and may affect neighboring gene expression. Additionally, some satellite DNA repeats are found within introns of particular genes, often associated with repetitive elements such as transposons, affecting their expression under specific conditions or developmental stages (Croisetière et al., 2010).

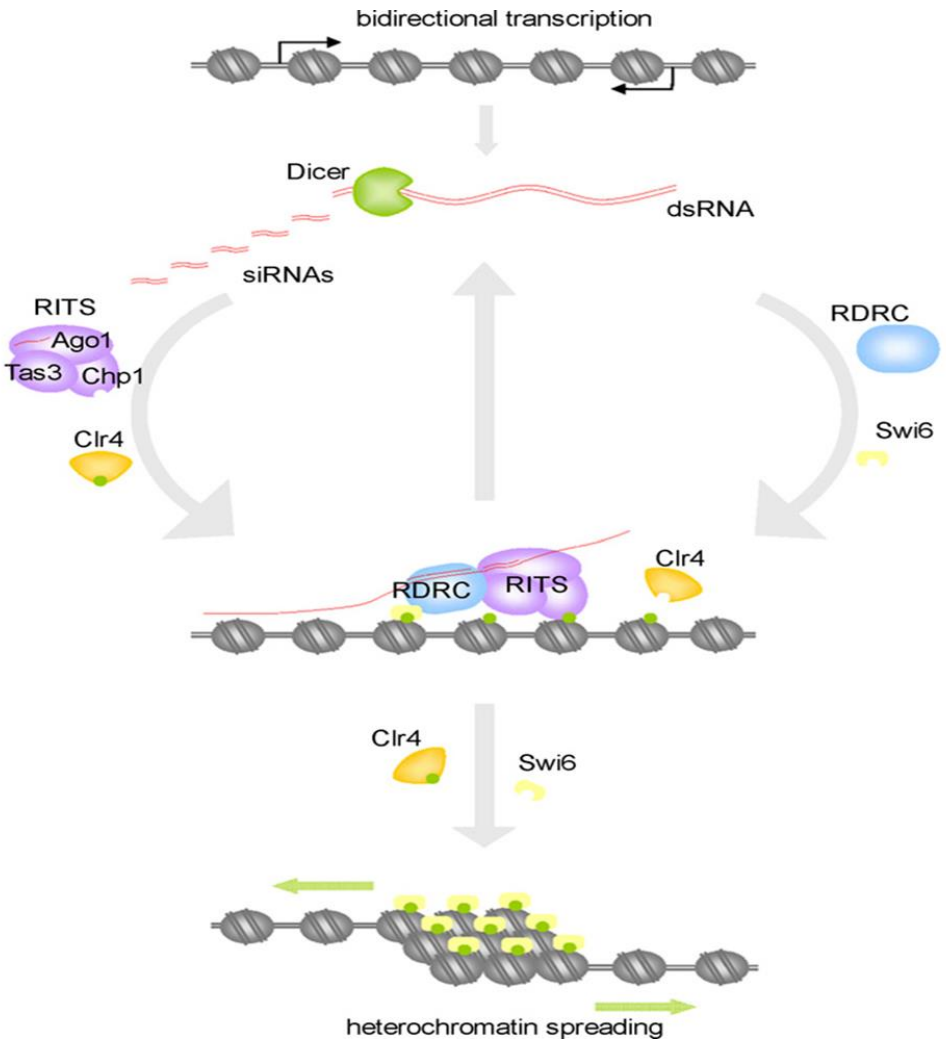


Figure 6. Heterochromatin assembly in fission yeast *Schizosaccharomyces pombe* (Pezer and Ugarković, 2008). Low-level bidirectional transcription from centromeric and pericentromeric regions followed by Dicer cleavage produces enough dsRNA to trigger RNAi-dependent heterochromatin formation. siRNAs guide heterochromatin assembly factors in the form of RITS (RNA-induced transcriptional silencing) complex to target nascent RNA sequence. Silencing signal is amplified by RDRC (RNA-directed RNA polymerase complex) which produces additional dsRNAs. Heterochromatin spreading across neighbouring sequences is induced by H3–Lys9 methylation (green dots) which results from activity of histone methyltransferase Clr4, followed by binding of chromodomain protein Swi6.

1.7. Satellite DNA in stress response and oncogenic transformation

It is clear that the expression of satellite DNAs is carefully controlled under normal physiological conditions. However, under certain circumstances and within various biological contexts, their expression can change significantly (Feliciello et al., 2021; Mihìc et al., 2021). Heat stress (HS) has a particular impact on heterochromatin across a range of organisms, including plants, insects, mice, and even human cells. This stress leads to the decondensation of heterochromatin and a reduction in nucleosome occupancy, which ultimately activates the transcription of heterochromatic satellite DNAs (Pezer and Ugarković, 2012; Jolly et al., 2004; Rizzi et al., 2004; Pecinka et al., 2010; Tittel-Elmer et al., 2010; Hédouin et al., 2017; Feliciello et al., 2020a). Notably, heat shock induces a dramatic increase in the expression of pericentromeric satellite III (HSATIII) due to factors such as heat shock, DNA-damaging agents, and hyperosmotic stress (Vourc'h and Biamonti, 2011). This process is largely mediated by heat shock transcription factor 1 (HSF1). HSF1 binds to satellite III DNA, recruits key acetyltransferases to interact with pericentromeric proteins, and facilitates the recruitment of Bromodomain and Extra-Terminal (BET) proteins—specifically BRD2, BRD3, and BRD4—that are essential for the transcription of satellite III (Col et al., 2017; Vourc'h et al., 2022). It seems that stress-induced activation of satellite III is a part of the general cellular response to stress, which provides protection against heat-shock-induced cell death (Goenka et al., 2016). Also, it was reported that satellite III and satellite II exhibit copy number variation (CNV) during the stress response, aging and pathology, and a close link between their transcription and CNV is postulated (Porokhovnik et al., 2021).

The activation of transcription for the major *T. castaneum* (peri)centromeric satellite DNAs, specifically TCAST1, and human alpha satellite, occurs in response to heat stress (Pezer and Ugarković, 2012; Feliciello et al., 2020a). For TCAST1, this is associated with demethylation of satellite DNA (Feliciello et al., 2013), highlighting the role of DNA methylation in regulating TCAST1 satellite expression. The transcripts of TCAST1 that are produced during heat stress, along with their derived siRNAs and alpha satellite transcripts, are believed to contribute to the maintenance and recovery of heterochromatin after heat stress. They achieve this by temporarily increasing the levels of silent histone modification H3K9me2/3 at satellite repeats within heterochromatin (Pezer and Ugarković, 2012; Sermek et al., 2021; Feliciello et al., 2020a). In contrast, the minor *T. castaneum* satellite DNAs do not exhibit induced transcription under heat stress, likely due to their positioning in euchromatin as part of the genome organization (Sermek et al., 2021; Brajković et al., 2018).

Different pathological conditions are known to trigger the transcription of satellite DNAs. In various epithelial cancers—including those of the pancreas, lung, kidney, colon, and prostate—there is a notable increase in satellite DNA transcripts (Ting et al., 2011). Interestingly, the transcription patterns of pericentromeric satellite DNAs are not solely altered in solid tumors linked to cancer cells but also in hematopoietic malignancies (Enukashvily et al., 2022). For instance, transcriptomes from multiple myeloma show a significant enrichment in pericentromeric tandem repeat transcripts from both hematopoietic and non-hematopoietic cells, like endothelial and mesenchymal stromal cells (de Jong et al., 2021). Moreover, co-culturing healthy donors' mesenchymal stromal cells with multiple myeloma cells can induce satellite transcription in those healthy cells (Enukashvily et al., 2022). The regions of (peri)centromeric satellite DNAs are under epigenetic control, and observations indicate that a reduced level of the repressive histone mark H3K9me3 at satellite repeats in cancer cell lines, compared to normal cells, along with the global hypomethylation typical of cancer cells, may contribute to the aberrant transcription of satellite DNA (Vojvoda Zeljko et al., 2021; Unoki et al., 2020).

Dysregulation of Polycomb repressive complexes, specifically PRC1 and PRC2, is prevalent in various cancers and has a notable impact on pericentromeric silencing (Blackledge et al., 2015). Besides these epigenetic alterations, an overabundance of satellite DNA is frequently linked to a deficiency in the tumor suppressor protein p53, which typically helps regulate the movement of repetitive elements (Wylie et al., 2016). Additionally, the absence of the tumor suppressor BRCA1 compromises the structural integrity of constitutive heterochromatin and influences the transcription of satellite DNA repeats (Zhu et al., 2011).

Finally, the heightened expression and activation of heat shock transcription factor 1 (HSF1), commonly observed in cancer cells, may contribute to increased levels of satellite DNA (Vourc'h et al., 2022; Dai et al., 2007).

Overexpressed heterochromatic satellite RNAs interact with BRCA1 and other crucial proteins, which play a significant role in maintaining the stability of the replication fork. This interaction can lead to DNA damage and genomic instability, thereby contributing to the development of breast cancer (Zhu et al., 2018). Furthermore, in mouse models with K-ras mutations, pericentromeric satellite DNA transcripts have been shown to hinder the DNA damage repair capability of the YBX1 protein, accelerating tumor growth by acting as intrinsic mutagens (Kishikawa et al., 2016, 2018). In human cancer cells, satellite II transcripts are recognized as immunogenic, stimulating the innate immune system to release cytokines (Tanne et al., 2015). Additionally, these same satellite II transcripts can lead to repeat expansions at pericentric heterochromatin through the formation of aberrant RNA:DNA hybrids (Bersani et al., 2015).

In general, overexpression of centromeric satellite DNAs promotes chromosome instability, which correlates with tumor metastasis (Zhu et al., 2011; Bakhoum et al., 2018). Satellite RNAs play a significant role in tumor progression through various mechanisms. They can induce mutations (Kishikawa et al., 2016), impact epigenetic regulators (Hall et al., 2017), promote tumor cell proliferation (Nogalski and Shenk, 2020), trigger inflammation (Tanne et al., 2015; Miyata et al., 2021), contribute to resistance against cancer therapies (Kanne et al., 2021), and even undermine genome integrity (Zhu et al., 2018; Zeller and Gasser 2017). Conversely, satellite transcripts may also be detected by the innate immune system, leading to an immune response that can facilitate the elimination of cancer cells and slow down tumor growth (Tanne et al., 2015; Rajshekar et al., 2018).

Satellite DNA overexpression is observed in cancer tissues, leading to the release of their transcripts into the bloodstream. This opens the door to using circulating satellite RNAs as potential biomarkers for various cancers (Ting et al., 2011). It is important to note, however, that the levels of satellite RNA present in the serum of cancer patients tend to be low and the RNA itself is unstable. To obtain reliable measurements of these RNA levels, new sensitive techniques have been developed. One notable method, tandem repeat amplification by nuclease protection (TRAP) combined with ddPCR, successfully quantified satellite II RNA in blood serum, allowing for the differentiation between healthy individuals and patients suffering from pancreatic ductal carcinoma (PDAC) (Kishikawa et al., 2016). Furthermore, elevated levels of human satellite II RNA found in the plasma of patients with

breast, gastric, lung, and bile cancers, along with sarcoma and Hodgkin's lymphoma, suggest its potential role as a diagnostic marker (Özgür et al., 2021). Additionally, research indicates that breast cancer patients with significantly high levels of alpha satellite RNA in their breast tissues are at a 10- to 20-fold increased risk of developing multiple cancers, even if they do not exhibit any BRCA-related clinical features (Kakizawa et al., 2019). Further studies are necessary to reveal satellite RNAs and DNAs as potential diagnostic, prognostic or therapeutic cancer biomarkers.

1.8. Objectives and hypothesis

Objectives:

1. To determine whether an elevated level of primary human alpha-satellite RNA can be induced by a specific external factor (for example, by applying various concentrations of antibiotics) or through exogenous expression by means of a vector.

2. To analyze if the elevated level of alpha-satellite RNA is connected with epigenetic changes within pericentromeric heterochromatic regions containing tandemly repeated alpha-satellite DNA, primarily histone modifications H3K9me3, H3K18ac and H3K4me2 and if these changes influence genome stability.

3. To test if the epigenetic changes in H3K9me3, H3K18ac and H3K4me2 histone modifications occur in euchromatin, specifically within regions containing dispersed alpha-satellite DNA repeats.

4. To assess the expression of genes with alpha-satellite DNA insertions within intronic regions, as well as of those closely flanked by alpha-satellite DNA elements.

Hypothesis:

Elevated level of human alpha-satellite RNA is tied to changes in epigenetic modifications of heterochromatic regions, as well as in euchromatic regions that contain alpha-satellite DNA elements, and can effect genome stability and differential gene expression.

2. DISCUSSION

2.1. Exogenous alpha satellite transcripts' effect on expression of alpha-associated genes and their epigenetic profiles after transfection

To delve deeper into the potential gene-modulatory role of alpha satellite transcripts and to eliminate the effects of heat stress or other stressors on gene expression, cell lines were developed with exogenous overexpression of alpha satellite RNA. In these modified cell lines, the expression levels of genes that contain alpha satellite repeats within their introns were closely monitored. While it is true that alpha repeats can also be found dispersed near genes (Feliciello et al., 2020b), the investigation focused on the impact of alpha satellite RNA on genes that contain these repeats within their gene bodies. The findings indicated a positive correlation between the exogenous expression of alpha satellite RNA and the downregulation of alpha-associated genes, providing strong evidence for the influence of alpha satellite transcripts on gene expression. Additionally, various histone marks on intronic alpha satellite repeats in cells with the overexpression of alpha satellite RNA were analyzed, proposing a potential molecular mechanism by which these transcripts modulate gene expression.

To explore the potential influence of alpha satellite RNA on gene expression, vectors that express alpha satellite monomers in both orientations (171Fw and 171Rev) were created. The findings demonstrated a notable level of exogenous alpha satellite expression from both the 171Fw and 171Rev vectors, peaking 24 hours after transfection before rapidly declining. This decrease was mainly due to plasmid loss associated with cell division and the absence of selective pressure. Moreover, both vector variants showed similar expression levels throughout the experiments. Endogenous alpha satellite expression remained stable at basal levels across all time points assessed, suggesting that the cells were in a standard physiological state without any stress or toxicity following the treatment.

The expression profiles of genes with dispersed alpha satellite repeats in their intronic regions (described in Feliciello et al., 2020b) were analyzed over four consecutive days following the transfection of the MJ90hTERT cell line with satellite-expressing vectors. These profiles were compared to controls that were transfected in the same manner with unmodified vectors. The genes examined included *SLC30A6, STAM, MYO1E, MAP7, ZNF675, VAV1, PRIM2*, and *DLG2*. Notably, there were no other genes with targetable intronic alpha satellite segments that met the design and expression criteria. This observation

aligns with the notion that the introduction of satellite DNA into euchromatin is typically harmful and thus is a rare occurrence, resulting in a limited number of genes that carry intronic alpha repeats.

After 24 hours post-transfection, a notable decrease in gene expression for six of the genes tested was observed when compared to the control samples. These genes included *SLC30A6*, *STAM*, *MYO1E*, *MAP7*, *ZNF675*, and *PRIM2*. Among them, the first five exhibited the most pronounced downregulation, while *PRIM2* also showed significant downregulation, albeit to a lesser degree. Importantly, there were no differences in the downregulation efficiency between the 171Fw and 171Rev transfected cells across all cases, indicating that both orientations of alpha satellite inserts were equally effective. When examining additional time points (48, 72, and 96 hours post-transfection), there were no significant changes in the expression of the candidate genes between the transfected samples and controls. This suggests that the expression dynamics of the vectors led to a peak effect at 24 hours post-transfection, followed by a rapid decline. Notably, the genes *VAV1* and *DLG2* did not show any detectable expression across all analyzed samples, indicating their tissue-specific nature and lack of expression in the MJ90hTERT cell line. The *GUSB* gene maintained stable expression throughout the experiments, with no significant differences identified between the samples and controls which reinforced its reliability as a normalization gene for relative RT-qPCR quantification. Additionally, we assessed the expression of five other housekeeping genes—*GAPDH, TOP3A, DEK, GPR68*, and *IFIT3*—after 24 hours of transfection with the alpha satellite-expressing vectors and an unmodified control vector. No significant differences in gene expression were found between samples transfected with the alpha satellite vectors and the control vector, suggesting that alpha satellite RNA did not affect the expression of these reference genes.

Additionally, it was investigated how the silent histone mark H3K9me3, the transcriptional activation mark H3K18ac, and the euchromatin-associated mark H3K4me2 were distributed among alpha repeats located within the introns of six genes previously tested for expression. It was also explored whether the downregulation of gene expression observed earlier was linked to these epigenetic changes. When the MJ90hTERT cells were transfected with the 171Fw alpha satellite expression vector, no significant changes were observed in the examined histone modifications at the alpha repeats across the six targeted genes compared to the control samples at the 24-hour mark. Likewise, the results were consistent in repeated experiments with the 171Rev vector featuring an inverted alpha satellite insert. Both vectors

presented similar outcomes to the controls, aligning with their comparable performance observed in gene expression analyses. On one hand, these results may stem from the limited transfection efficiency of MJ90hTERT cells or the reduced sensitivity of ChIP experiments. On the other hand, they might suggest the involvement of additional mechanisms that influence gene expression in the context of alpha satellite repeat-associated genes. Considering that the signal from individual alpha loci might have been too weak, ChIP-qPCR was conducted on tandem arrangements of alpha satellite arrays, which are indicative of heterochromatin. The findings highlighted a statistically significant rise in the silent histone mark H3K9me3 in MJ90hTERT cells 24 hours after they were transfected with the alpha satellite expression vector. However, the other two histone modifications that were tested, H3K18ac and H3K4me2, did not show any change.

To investigate whether alpha satellite transcripts actively generate RNA:DNA hybrids under standard physiological conditions, an RNase H digestion assay was conducted (RNase H specifically targets and degrades the RNA component of an RNA-DNA hybrid). The results showed a modest yet significant reduction (approximately 25%) in alpha transcripts in samples treated with RNase H when compared to untreated controls, indirectly confirming the existence of these hybrid formations.

This research introduced a new angle to understanding how satellite DNA transcripts contribute to the downregulation of gene expression. It was found that the temporary expression of exogenous alpha satellite DNAs, transcribed from both strands of DNA, consistently silenced alpha-associated genes, regardless of the strand they originated from. This observation, alongside the vulnerability of alpha satellite RNA to RNAse H treatment, suggests a direct interaction between alpha satellite RNA and homologous DNA within scattered intronic satellite regions, by forming hybrid structures such as triple helices (RNA:DNA:DNA) or R-loops. These interactions possibly influence the transcription of neighbouring genes. Prior research has demonstrated that gene expression modulation can occur through direct interactions between non-coding RNA and DNA (Statello et al., 2021). In contrast to the traditional Watson–Crick pairing in the DNA double helix, RNA-DNA hybrids are formed through Hoogsteen hydrogen bonding between the nucleic acid bases. This alternative binding allows for a weaker and more flexible connection between DNA and RNA, supporting the idea of transient interference based on the homology of alpha satellite sequences. This might clarify the observed variations in gene suppression levels among the alpha satellite repeat-associated genes. Considering the polymorphic nature of alpha satellite

DNA, which can exhibit up to 45% variability in monomer sequences, along with the fact that the expressed satellite RNA is a cloned sequence, it seems reasonable to infer that differences in gene repression can result from varying degrees of sequence homology.

Furthermore, the ability of alpha satellite RNA to modulate gene expression appears to be independent of gene polarity. Similar levels of gene downregulation were observed when using either vector, regardless of the transcription direction. This suggests that the construction of this hybrid genomic structure might be crucial, either directly or by directing RNA-associated regulatory proteins to specific genomic sites. The generation of triple helices and R-loops seems to be common and vital for the regulatory role of various non-coding RNAs (Li et al., 2016; Warwick et al., 2023). While numerous R-loop-forming long non-coding RNAs typically function *in cis*, R-loops can also occur *in trans*, impacting the expression of protein-coding genes (Ariel et al., 2020). Long non-coding RNAs have been shown to form triplex structures that regulate gene expression *in trans* as well (Warwick et al., 2023; Mondal et al., 2019). In mice, pericentromeric satellite DNA transcripts have been identified to form RNA:DNA hybrids that facilitate the retention of heterochromatin protein 1 (HP1) and the histone methyltransferases SUV39h1 and SUV39h2, which are essential for heterochromatin formation (Velazquez Camacho et al., 2017; Duda et al., 2021). The observed increase in H3K9me3 levels on the tandemly arranged alpha satellite arrays, typical of heterochromatin, following transfection with the alpha expression vector supports the notion of a potential mechanism in which alpha satellite RNA:DNA hybrids recruit chromatin modifiers.

In summary, the findings presented here indicate that alpha satellite RNA not only plays a crucial role in the assembly of centromeres and heterochromatin but also, for the first time, suggest its involvement in modulation of gene expression. Nonetheless, further research is essential to unravel the specific molecular mechanisms through which alpha satellite RNA influences gene expression.

2.2. Antibiotic-mediated alpha satellite DNA overexpression and epigenetic changes within pericentromeric heterochromatic regions as well as in euchromatin

This investigation examined the impact of various antibiotics, including geneticin and hygromycin B-commonly used in cell culture (Landers et al., 2021), alongside rifampicin, which is frequently utilized in treating various bacterial infections (Boeree et al., 2017). The focus was to determine whether these antibiotics influence the expression of a major human alpha satellite DNA, which is prominently located in (peri)centromeric regions across all human chromosomes (McNulty and Sullivan, 2018). Additionally, it was investigated whether any alterations in satellite DNA expression, prompted by antibiotic exposure, were associated with epigenetic modifications, such as specific histone marks at heterochromatic satellite arrays and at the satellite repeats found dispersed within euchromatin, as well as throughout the genome.

In order to determine the impact of antibiotics on alpha satellite DNA transcription, its transcription dynamics were examined in human cell lines under standard conditions, both with and without antibiotic treatment, using concentrations of antibiotics commonly employed for the routine treatment, selection, and maintenance of eukaryotic cells. The transcription of alpha satellite DNA was assessed immediately following antibiotic treatment and compared to an untreated control. This evaluation was conducted in immortalized fibroblasts (MJ90hTERT), the glioblastoma cell line A-1235, and HeLa cells derived from cervix carcinoma. The results showed an overall increase in alpha satellite DNA transcription across cell lines after application of different antibiotics at standard concentrations. Notably, the response varied among the cell lines: A-1235 cells experienced the highest increase with rifampicin at 82 µg/ml, while the other two antibiotics produced only minor changes. In contrast, HeLa cells exhibited their maximum transcription boost with geneticin at 400 µg/ml, and MJ90hTERT cells required a higher dose, specifically 600 µg/ml of geneticin, to achieve a notable change in transcription levels. Furthermore, the findings indicated a positive correlation between the concentration of antibiotics and the level of alpha satellite transcription.

The distribution of the silent histone mark H3K9me3, associated with heterochromatin, was examined alongside the H3K18ac mark, indicating transcriptional activation of heterochromatin (Tasselli et al., 2016), as well as of H3K4me2, which is characteristic of open euchromatin. This analysis focused on both tandemly arranged alpha satellite repeats and those dispersed throughout euchromatin, performed under standard

physiological conditions and following antibiotic treatment. Tandemly arranged satellite repeats and six alpha repeats located within gene introns (Feliciello et al., 2020b) were tested.

Results from experiments on HeLa and MJ90hTERT cells treated with geneticin, as well as A-1235 cells with rifampicin, showed a reduction of H3K9me3 at heterochromatic alpha repeats, which was linked to an increase in the transcription of alpha satellite DNA. In MJ90hTERT cells, the decrease in H3K9me3 following geneticin treatment was accompanied by a notable increase in H3K18ac, also correlating with heightened transcription of alpha satellite. For alpha repeats found within euchromatin, there were no observed changes in the histone modifications tested after any antibiotic treatment. Notably, at 41 μg/ml rifampicin, A-1235 cells exhibited an increase in alpha satellite transcription, despite only a slight alteration in H3K9me3 levels. Additionally, no statistically significant changes in H3K9me3 levels were detected at tandem alpha repeats in A-1235 cells, nor in H3K9me3 and H3K18ac levels in MJ90hTERT cells after treatment with 400 μg/ml geneticin, despite a modest increase in alpha transcription levels of 1.7x and 1.5x, respectively.

To investigate the impact of antibiotic treatment on epigenetic changes across the genome, immunofluorescence assays were conducted on above-mentioned cell lines, targeting the same histone modifications. The results demonstrated a genome-wide increase in H3K9me3 levels in HeLa and MJ90hTERT cells upon treatment with geneticin (300–600 μg/ml). In contrast, H3K4me2 levels were either downregulated or showed no significant change. Additionally, in HeLa cells treated with 300 μg/ml geneticin, H3K18ac levels were significantly downregulated. However, at higher concentrations, the changes in H3K18ac levels in both HeLa and MJ90hTERT cells were minimal, suggesting that the antibiotic's effects are concentration-dependent but not necessarily positively correlated. On the other hand, in A-1235 cells, a concentration of 400 μg/ml geneticin led to a genome-wide downregulation of H3K9me3 and an upregulation of H3K18ac, demonstrating that the response to the antibiotic varies across different cell lines. Furthermore, it is notable that different antibiotics can influence epigenetic marks differently within the same cell line; for instance, rifampicin triggered a decrease in H3K18ac modifications in A-1235 cells, while geneticin induced their upregulation.

Different antibiotics lead to the overexpression of pericentromeric alpha satellite DNA, but their impacts on heterochromatin vary across cell types. Some cells may show a reduction in H3K9me3 levels, while others may exhibit an increase in H3K18ac

modifications. These results suggest a regulation of alpha satellite transcription by epigenetic changes, in particular by histone marks H3K9me3 and H3K18ac. Furthermore, the broader effects of antibiotics on the genome are also cell-dependent. For instance, geneticin influences histone marks H3K9me3 and H3K4me2 similarly in HeLa and MJ90hTERT cells, but its impact diverges in A-1235 cells. Additionally, rifampicin affects these same histone marks differently in A-1235 cells. These findings indicate that both heterochromatin and other chromatin components respond to antibiotics in diverse ways, influenced by factors such as the specific cell line, the type of antibiotic used, and its concentration.

This research recognizes that gene expression and epigenetic responses differ across various cell types. During the investigation, unique regulatory patterns of alpha satellite expression were identified in the tested cell lines. These patterns are likely shaped by factors such as chromatin organization, baseline transcriptional activity, and cellular metabolism. These differences may indicate fundamental distinctions between cancerous and normal cells in how they respond to external factors such as antibiotics. Further research is necessary to uncover the molecular mechanisms that underline these cell-type-specific responses. It is possible that different signaling pathways, which become overactivated in various cell lines, provide resistance to certain antibiotics and influence how cells manage stress or damage. Additionally, it seems that the level of stress resistance varies between cell lines, and their strategies for coping with diverse stressors reflect their unique genetic backgrounds.

Finally, this research indicates that antibiotics could affect the transcription of satellite DNA by influencing specific histone marks. While the pathways for histone modification are well-established and conserved across eukaryotic organisms, the exact mechanisms through which antibiotics might engage with these pathways remain unclear. The roles of various effector enzymatic complexes are well understood; these include activating demethylases and acetyl-transferases, as well as inhibiting methyl-transferases and deacetylases, along with chromatin remodeling factors such as SWI/SNF. Additionally, the integrated stress response may play a significant role by interacting with these pathways. Different types of stress, antibiotics included, trigger this cellular machinery, leading to an increased production of specific transcription factors (such as ATF4). This surge stimulates downstream chromatin remodeling at targeted loci, consequently stimulating the activation of neighbouring promoters and thus clarifying the observed overexpression. Results of this investigation underscore broader implications for gene regulation, drug safety, and long-term antibiotic effects, highlighting the need for further research.

2.3. Alpha satellite RNA expression in prostate cancer, disease pathogenesis and as a potential diagnostic biomarker alongside prostate-specific antigen (PSA)

To explore the potential of alpha satellite RNA as a biomarker for prostate cancer, an analysis of intracellular alpha satellite RNA levels in the blood of patients diagnosed with prostate cancer was conducted, comparing these levels to those in healthy individuals. Intracellular RNA was collected from whole blood of prostate cancer patients categorized into four groups based on their disease stage. Group A consisted of patients with metastatic hormone-sensitive prostate cancer, Group B included those with metastatic castration-resistant cancer, Group C represented patients with localized hormone-sensitive prostate cancer, while Group D comprised newly diagnosed localized prostate cancer patients who had not yet received hormone therapy or other treatments. Unlike Group D, individuals in Groups A–C were all receiving hormone treatment with LHRH (Luteinizing Hormone-Releasing Hormone) agonists. Additionally, intracellular RNA from a control group of 27 healthy males was also collected.

To assess the levels of alpha satellite RNA within the total intracellular RNA extracted from whole blood samples of patients grouped into four categories (A–D), as well as a cohort of healthy controls, quantitative real-time PCR (qPCR) analysis was employed. The findings showed a marked increase in alpha satellite RNA levels in two cohorts of metastatic prostate cancer patients compared to the control group. In group B, comprising patients with metastatic castration-resistant prostate cancer, the increase was observed to be 2.8-fold, supported by strong statistical evidence ($p=2.7\times10^{-4}$). Conversely, group A, representing patients with metastatic hormone-sensitive prostate cancer, displayed a 1.4-fold increase, which was not statistically significant. For groups C and D, involving localized prostate cancer patients, there was no significant difference in alpha satellite RNA levels compared to the control group. The similarity in alpha satellite RNA levels in these two groups implies that drug treatment does not impact RNA levels. Notably, group B exhibited a significant elevation in alpha satellite RNA when compared to group A, with an increase of 2.0-fold ($p=4\times10^{-3}$), as well as relative to groups C and D, showing increases of 2.9-fold ($p=4\times10^{-6}$) and 1.7-fold ($p=0.017$), respectively.

These findings indicate that alpha satellite RNA levels can effectively differentiate between various disease stages, namely metastatic castration-resistant cancer versus metastatic hormone-sensitive cancer, as well as metastatic castration-resistant cancer compared to localized prostate cancer and healthy controls. This could position alpha satellite

RNA as a promising diagnostic biomarker for metastatic conditions, especially in the context of castration-resistant metastatic prostate cancer. Analysis using ROC curves and AUC values demonstrated that levels of alpha satellite RNA allow for a high degree of accuracy in distinguishing metastatic castration-resistant prostate cancer from primary localized tumors (AUC 0.85) and from healthy controls (AUC 0.85). Furthermore, the distinction between metastatic castration-resistant and metastatic hormone-sensitive prostate cancer showed a viable level of accuracy as well (AUC 0.74).

Prostate-specific antigen (PSA) is the most widely recognized biomarker for prostate cancer. This serine protease, related to kallikrein, is secreted by the epithelial cells of the prostate. In patients with prostate cancer, PSA levels tend to be elevated. The PSA levels in the blood of four groups of patients were analyzed alongside a control group and a significant difference was detected between the groups. Notably, PSA levels in groups A, B, and D were significantly higher compared to the controls and group C ($p<10^{-4}$). However, no significant differences in PSA levels were observed between group B and group A or between group B and group D. Also, there was no significant difference in PSA values between the control group and group C. The ROC curve analysis of PSA levels indicated a clear distinction between the control group and group D, which showed an AUC value of 0.912. Additionally, there was notable discrimination between the control group and two metastatic cancer groups, A and B, with AUC values of 0.8052 and 0.9256, respectively. However, the differentiation between metastatic hormone-sensitive (group A) and metastatic castration-resistant (group B) cases was minimal, with an AUC value of only 0.512. This underscored the superior effectiveness of alpha satellite RNA, which had an AUC of 0.744, compared to PSA in distinguishing between the two stages of metastatic prostate cancer. Furthermore, correlation analyses were performed to assess the relationship between alpha satellite RNA levels and PSA levels across all patient groups, but no statistically significant correlation was observed in any group.

Moreover, PSA is also used for identifying latent cases of prostate cancer, which may not progress to serious illness. However, it often increases in benign conditions, such as inflammation or hyperplasia, leading to concerns about over-diagnosis due to its lack of specificity (Prensner et al., 2012). While PSA is utilized to track disease progression, its levels in patients with metastatic prostate cancer frequently do not correlate well with disease stage or hormone sensitivity. This insufficiency in PSA specificity as a diagnostic and

prognostic marker has prompted efforts to identify alternative biomarkers for prostate cancer (Saini, 2016).

The aberrant overexpression of sequences in pericentromeric heterochromatin, where satellite DNAs like alpha satellite are prevalent, is a common feature in prostate cancer and characteristic of several other epithelial cancers, including those of the pancreas, lung, kidney, and colon (Ting et al., 2011). Moreover, the presence of satellite DNA in tumors can be attributed to a deficiency in other tumor suppressors like p53 or BRCA1. Such deficiencies compromise the integrity of constitutive heterochromatin and lead to heightened levels of satellite DNA expression (Wylie et al., 2016; Zhu et al., 2011). When satellite RNA levels rise, they destabilize the replication fork and compromise genome integrity, accelerating tumor transformation (Zhu et al., 2018).

One explanation for the increased alpha satellite RNA in the bloodstream of prostate cancer patients, particularly those in metastatic stages, may stem from the transfer of this RNA from cancer cells to blood cells, facilitated by exosomes. Exosomes are extracellular vesicles released by all types of cells; they are often found within tumor microenvironments and play a role in removing excess or unnecessary elements, including detrimental RNA and DNA (Kalluri, 2016; Takahashi et al., 2017). These vesicles are capable of transferring their RNA or DNA content to other cells and can also activate different signaling pathways in the cells they interact with (Takahashi et al., 2017; Valadi et al., 2007). This research proposes that the abundant satellite RNA from prostate cancer could be taken up and delivered by exosomes to blood cells, leading to an increased total RNA levels in the bloodstream. Furthermore, the interaction between exosomes and blood cells may trigger certain signaling pathways that could influence the structure of heterochromatin and the expression of satellite sequences found within. Circulating tumor cells (CTCs) found in the blood of patients with metastatic prostate cancer (Galletti et al., 2014) could interact with blood cells, potentially leading to elevated levels of alpha satellite RNA. However, since CTCs are relatively sparse compared to blood cells, their influence on this increase is likely minimal.

Following this investigation, it is proposed that alpha satellite RNA levels could serve as a complementary biomarker to PSA for tracking the progression of metastatic prostate cancer, as well as for diagnosing metastatic castration-resistant stages of the disease. Notably, another study demonstrated the potential of circulating satellite RNA levels in blood serum to act as a cancer biomarker. Using the sensitive technique of tandem repeat amplification by

nuclease protection (TRAP) combined with droplet digital PCR (ddPCR), researchers were able to distinguish patients with pancreatic ductal carcinoma (PDAC) from healthy individuals (Kishikawa et al., 2016). Moreover, elevated levels of circulating human satellite II have been identified in the plasma of patients with breast, gastric, lung, and bile cancers, as well as in sarcoma and Hodgkin's lymphoma (Özgür et al., 2021). This current study is the first to suggest that not just serum or plasma-circulating satellite RNA, but also alpha satellite RNA present in blood cells, could serve as an important indicator of specific cancer stages.

Satellite DNA is also emerging as a potential cancer biomarker, with variations in its copy number being linked to certain types of cancer (Bersani et al., 2015). However, detecting these variations can be quite complex, often necessitating the development of new assays and advanced technologies, including nanoplate-based digital PCR (de Lima et al., 2021). Further research is needed to clarify why the increased levels of intracellular alpha satellite RNA in the blood of prostate cancer patients were observed at a certain metastatic stage. It also needs to be determined whether this occurence is unique to this particular pathological condition alone.

2.4. A novel methodology for precise quantification of repetitive sequences irrespective of genomic DNA contamination

A significant issue with many amplification protocols currently in use is the frequent presence of DNA contamination in the samples. This contamination cannot be chemically distinguished from cDNA by the polymerase enzyme during PCR amplification, leading to false positive results (Verwilt et al., 2020; Bustin 2002; Kumar et al., 2006; Li et al., 2022; Padhi et al., 2016; Hashemipetroudi 2018).

To address this challenge, existing protocols incorporate a few DNA elimination steps both during RNA purification and the subsequent reverse transcription phase. In these cases, DNA is removed using specific mechanical filters, like silica-based columns, or through enzymatic digestion with a specific enzyme such as DNase I (Deoxyribonuclease I) (Green and Sambrook 2019). However, these treatments are not entirely successful in completely eliminating DNA, and traces often remain as contamination (Verwilt et al., 2020; Bustin 2002). This problem is particularly pronounced with the highly repetitive DNA sequences that make up a significant portion of the eukaryotic genome and are frequently transcribed at low levels. Additionally, it is important to note that any steps taken to reduce DNA

concentration in the sample will also inevitably lower the initial concentration of RNA, which is inherently unstable and prone to degradation. This research presents a new approach for transcriptome analyses using a PCR variant that effectively distinguishes between cDNA and genomic DNA. By eliminating DNA contamination, this method yields more accurate, dependable, and reproducible results.

The proposed method utilizes a modified primer (Modified Specific Primer, PSM) during the reverse transcription step of the protocol. This primer is constructed to be specific for the RNA molecules targeted for quantification, with its nucleotide sequence engineered to intentionally lack perfect homology to the retro-transcribed template DNA. Typically, introducing a few mismatches relative to the original sequence—preferably near the 3'-OH terminal region—is sufficient. These alterations render the primer only partially complementary to the target sequence, yet capable of hybridizing at the reverse transcription temperatures of 37–42 °C. During the PCR step, when temperatures rise to around 60 °C, the PSM will dissociate from the partially homologous genomic DNA sequence. The purpose of employing this specially modified primer is to facilitate amplification specifically from the cDNA template while effectively bypassing genomic DNA targets. It is essential to experimentally verify the optimal number of modifications, their effectiveness, and the proper discriminating temperatures for each transcript analyzed. This involves selecting parameters that exhibit differing amplification tendencies for DNA and cDNA targets, respectively. This optimization phase serves as a crucial preliminary step in this methodology, allowing for the establishment of negative and positive controls. Fortunately, this process only needs to be conducted once, as it remains applicable for a specific amplicon across various experimental conditions. In contrast, current protocols typically require the negative control (NC: –RT, without reverse transcriptase) to be prepared for each new sample, even if the target remains the same, due to the unpredictable effectiveness of DNase I treatment. By using a PSM, a cDNA can be generated that is subtly distinct from its genomic DNA counterpart, owing to the nucleotide mismatches incorporated into the sequence.

In the post-reverse transcription phase of the protocol, the PCR amplification of cDNA proceeds accordingly. This is done using the modified primers (PSM) from the earlier step along with unmodified specific primers (SP) oriented in the opposite direction. As a result, the generated amplicon is a cDNA copy, not a DNA copy, thanks to the specific annealing temperatures typically set between 55 °C and 62 °C. This method eliminates the need for removal of any co-purified DNA from the RNA sample, as it no longer competes as

a target and will not interfere with the assay's outcomes. In fact, under certain experimental conditions, having both DNA and RNA in the same sample can be beneficial, especially if there is a need to normalize the results with respect to gene copy number variation.

Satellite DNA stands out as an excellent candidate to showcase the effectiveness of this methodology. Its highly repetitive nature and abundance in non-coding genomic DNA regions make it a persistent presence in samples, rendering it challenging, if not impossible, to eliminate during RNA purification process. As a point of comparison the traditional method was used, which involved the removal of DNA during both RNA purification and reverse transcription. Despite this, alpha satellite DNA persisted in the negative control samples (–RT). Since satellite DNA is not structured into exons and introns, distinguishing it from satellite cDNA by length alone is impossible. Consequently, even minimal DNA contamination can lead to false-positive results. In contrast, the new method effectively eliminated contamination from alpha satellite DNA in the qPCR amplification results. This was further confirmed by the presence of the 126 bp ASAT amplicon exclusively in the +RT samples compared to the –RT samples (negative controls) when analyzed on an agarose gel.

Additional testing was performed on the satellite DNA known as TCAST1, a highly abundant primary satellite DNA accounting for about 30% of the genome of the beetle *Tribolium castaneum*. Using the new method, only cDNA was amplified, with the +RT samples showing clear amplification results, while the –RT samples revealed next to no contamination from genomic DNA. Importantly, these findings align closely with those previously reported for human alpha satellite DNA.

In summary, it can be concluded that this new method for quantifying various types of transcripts yields results that are more accurate, reproducible, and cost-effective than the protocols currently in use. This improvement is largely due to the method’s insensitivity to DNA contamination, which often leads to false positive signals, eliminating the need for removal of template DNA beforehand. By skipping the DNA elimination step, RNA is also better preserved from degradation, which in turn effectively mitigates the two main sources of inaccuracy found in transcriptome analyses.

3. CONCLUSION

1) A notable level of exogenous alpha satellite expression from both expression vectors was achieved and detected after transfection of MJ90hTERT cells, peaking 24 hours after transfection before rapidly declining, with both vector variants showing similar expression levels throughout the experiments.

2) A significant decrease in gene expression for six of the genes tested (*SLC30A6*, *STAM*, *MYO1E*, *MAP7*, *ZNF675*, and *PRIM2*) was observed 24 hours post-transfection. The genes *VAV1* and *DLG2* did not show any detectable expression across all analyzed samples.

3) A positive correlation was found between the exogenous expression of alpha satellite RNA and the downregulation of alpha-associated genes, providing strong evidence for the influence of alpha satellite transcripts on gene expression modulation.

4) No significant changes were observed in the examined histone modifications (H3K9me3, H3K18ac and H3K4me2) at the alpha satellite repeats across the six targeted genes compared to the control samples.

5) The results showed an overall increase in alpha satellite DNA transcription across all tested cell lines (MJ90hTERT, A-1235 and HeLa) after application of different antibiotics at standard concentrations with responses varying depending on cell type. Also, the findings indicate a positive correlation between the concentration of antibiotics and the level of alpha satellite DNA transcription.

6) HeLa and MJ90hTERT cells treated with geneticin, as well as A-1235 cells with rifampicin, showed a reduction of H3K9me3 modifications at pericentromeric heterochromatic alpha arrays. In MJ90hTERT cells, the decrease in H3K9me3 following geneticin treatment was accompanied by a notable increase in H3K18ac, both correlating with elevated alpha satellite

transcription in those genomic regions, impacting genome stability, and suggesting a regulation of alpha satellite expression by these particular epigenetic changes. For alpha repeats dispersed within euchromatin, there were no observed changes in those same histone modifications tested after any antibiotic treatment.

7) The impact of antibiotic treatment on total epigenetic changes across the genome varied significantly across different cell lines depending on the type of antibiotic used and its concentration.

8) The results demonstrated a significant increase in alpha satellite RNA levels in patients with metastatic castration-resistant prostate cancer compared to control group and other prostate cancer patients, allowing for precise discrimination of this particular type of cancer from other variants. Patients with metastatic hormone-sensitive variant, as well as those with localized prostate cancer did not significantly differ from control group.

9) Increasing levels of alpha satellite DNA expression in the blood of prostate cancer patients correlate with the higher severity of disease progression and significantly worse prognosis.

10) Prostate-specific antigen as a biomarker successfully discriminates between the control group and two metastatic cancer groups. However, the differentiation between metastatic hormone-sensitive and metastatic castration-resistant cases was minimal, underscoring the superior effectiveness of alpha satellite RNA compared to PSA in distinguishing between the two stages of metastatic prostate cancer.

11) Correlation analyses were performed to assess the relationship between alpha satellite RNA levels and PSA levels across all patient groups, but no statistically significant correlation was observed in any group.

12) These findings suggest that alpha satellite RNA may serve as a valuable diagnostic and prognostic biomarker for prostate cancer in blood tests, alongside already established prostate-specific antigen (PSA).

13) New method for quantification of various types of transcripts delivers results that are more accurate, reliable, and cost-effective compared to the protocols currently in use.

14) This method is unaffected by DNA contamination, which often leads to false positive results, circumventing the need for elimination of template DNA beforehand.

15) By excluding the DNA removal step, the RNA is effectively protected from degradation, which eliminates a significant source of inaccuracy in transcriptome analyses.

4. REFERENCES

Aldrup-MacDonald ME, Kuo ME, Sullivan LL, Chew K, Sullivan BA (2016) Genomic variation within alpha satellite DNA influences centromere location on human chromosomes with metastable epialleles. Genome Research. 26:1301–1311.

Alexandrov IA, Mitkevich SP, Yurov YB (1988) The phylogeny of human chromosome specific alpha satellites. Chromosoma. 96:443–453.

Alexandrov IA, Mashkova TD, Akopian TA, Medvedev LI, Kisselev LL, Mitkevich SP, Yurov YB (1991) Chromosome-specific alpha satellites: two distinct families on human chromosome 18. Genomics. 11:15–23.

Alexandrov IA, Mashkova TD, Romanova LY, Yurov YB, Kisselev LL (1993a) Segment substitutions in alpha satellite DNA. Unusual structure of human chromosome 3-specific alpha satellite repeat unit. Journal of Molecular Biology 231:516–520.

Alexandrov IA, Medvedev LI, Mashkova TD, Kisselev LL, Romanova LY, Yurov YB (1993b) Definition of a new alpha satellite suprachromosomal family characterized by monomeric organization. Nucleic Acids Research. 21:2209–2215.

Ando S, Yang H, Nozaki N, Okazaki T, Yoda K (2002) CENP-A, -B, and -C chromatin complex that contains the I-type alpha-satellite array constitutes the prekinetochore in HeLa cells. Molecular and Cellular Biology. 22:2229–2241.

Ariel F, Lucero L, Christ A, Mammarella MF, Jegu T, Veluchamy A, Mariappan K, Latrasse D, Blein T, Liu C, et al. (2020) R-Loop Mediated trans Action of the APOLO Long Noncoding RNA. Molecular Cell. 77:1055–1065.

Bailey AO, Panchenko T, Shabanowitz J, Lehman SM, Bai DL, Hunt DF, Black BE, Foltz DR (2016) Identification of the Post-translational Modifications Present in Centromeric Chromatin. Molecular & Cellular Proteomics. 15:918–931.

Bakhoum SF, Ngo B, Laughney AM, Cavallo JA, Murphy CJ, Ly P, Shah P, Sriram RK, Watkins TBK, Taunk NK, et al. (2018) Chromosomal instability drives metastasis through a cytosolic DNA response. Nature. 553:467–472.

Bergmann JH, Rodriguez MG, Martins NM, Kimura H, Kelly DA, Masumoto H,

Larionov V, Jansen LE, Earnshaw WC (2011) Epigenetic engineering shows H3K4me2 is required for HJURP targeting and CENP-A assembly on a synthetic human kinetochore. The EMBO Journal. 30:328–340.

Bersani F, Lee E, Kharchenko PV, Xu AW, Liu M, Xega K, MacKenzie OC, Brannigan BW, Wittner BS, Jung H, et al. (2015) Pericentromeric satellite repeat expansions through RNA-derived DNA intermediates in cancer. Proceedings of the National Academy of Sciences of the United States of America. 112:15148–15153.

Blackledge NP, Rose NR, Klose RJ (2015) Targeting polycomb systems to regulate gene expression: Modifications to a complex story. Nature Reviews Molecular Cell Biology. 16:643–649.

Blower MD, Sullivan BA, Karpen GH (2002) Conserved organization of centromeric chromatin in flies and humans. Developmental Cell. 2:319–330.

Bodor DL, Valente LP, Mata JF, Black BE, Jansen LE (2013) Assembly in G1 phase and long term stability are unique intrinsic features of CENP-A nucleosomes. Molecular Biology of the Cell. 24:923–932.

Boeree MJ, Heinrich N, Aarnoutse R, et al. (2017) High-dose rifampicin, moxifloxacin, and SQ109 for treating tuberculosis: a multi-arm, multi-stage randomised controlled trial. The Lancet Infectious Diseases. 17:39–49.

Brajković J, Feliciello I, Bruvo-Mađarić B, Ugarković Đ (2012) Satellite DNA-like elements associated with genes within euchromatin of the beetle *Tribolium castaneum*. G3: Genes, Genomes, Genetics. 2:931–941.

Brajković J, Pezer Ž, Bruvo-Mađarić B, Sermek A, Feliciello I, Ugarković Đ (2018) Dispersion profiles and gene associations of repetitive DNAs in the euchromatin of the beetle *Tribolium castaneum*. G3: Genes, Genomes, Genetics. 8:875–886.

Britten RJ, Kohne DE (1968) Repeated sequences in DNA. Hundreds of thousands of copies of DNA sequences have been incorporated into the genomes of higher organisms. Science. 161:529–540.

Bulut-Karslioglu A, Perrera V, Scaranaro M, de la Rosa-Velazquez IA, van de Nobelen S, Shukeir N, Popow J, Gerle B, Opravil S, Pagani M, Meidhof S, Brabletz T, Manke T, Lachner M, Jenuwein T (2012) A transcription factor-based mechanism for mouse heterochromatin formation. Nature Structural & Molecular Biology. 19:1023–1030.

Bustin SA (2002) Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): Trends and problems. Journal of Molecular Endocrinology. 29:23.

Cao S, Zhou K, Zhang Z, Luger K, Straight AF (2018) Constitutive centromere-associated network contacts confer differential stability on CENP-A nucleosomes in vitro and in the cell. Molecular Biology of the Cell. 29:751–762.

Carroll CW, Silva MC, Godek KM, Jansen LE, Straight AF (2009) Centromere assembly requires the direct recognition of CENP-A nucleosomes by CENP-N. Nature Cell Biology. 11:896–902.

Chan FL, Marshall OJ, Saffery R, Kim BW, Earle E, Choo KH, Wong LH (2012) Active transcription and essential role of RNA polymerase II at the centromere during mitosis. Proceedings of the National Academy of Sciences of the United States of America. 109:1979–1984.

Choo KH, Earle E, Vissel B, Filby RG (1990) Identification of two distinct subfamilies of alpha satellite DNA that are highly specific for human chromosome 15. Genomics. 7:143–151.

Cohen S, Agmon N, Sobol O, Segal D (2010) Extrachromosomal circles of satellite repeats and 5S ribosomal DNA in human cells. Mobile DNA. 1:11.

Col E, Hoghoughi N, Dufour S, Penin J, Koskas S, Faure V, Ouzounova M, Hernandez-Vargash H, Reynoird N, Daujat S, et al. (2017) Bromodomain factors of BET family are new essential actors of pericentric heterochromatin transcriptional activation in response to heat shock. Scientific Reports. 7:5418.

Croisetière S, Bernatchez L, Belhumeur P (2010) Temperature and length-dependent modulation of the MH class II beta gene expression in brook charr (*Salvelinus fontinalis*) by a cis-acting minisatellite. Molecular Immunology. 47:1817–1829.

Dai C, Whitesell L, Rogers AB, Lindquist S (2007) Heat Shock Factor 1 Is a Powerful Multifaceted Modifier of Carcinogenesis. Cell. 130:1005–1018.

de Jong MME, Kellermayer Z, Papazian N, Tahri S, Hofste op Bruinink D, Hoogenboezem R, Sanders MA, van de Woestijne PC, Bos PK, Khandanpour C, et al. (2021) The multiple myeloma microenvironment is defined by an inflammatory stromal cell landscape. Nature Immunology. 22:769–780.

de Lima LG, Howe E, Singh VP, Potapova T, Li H, Xu B, Castle J, Crozier S, Harrison CJ, Clifford SC, et al. (2021) PCR amplicons identify widespread copy number variation in human centromeric arrays and instability in cancer. Cell Genomics. 1:100064.

Du Y, Topp CN, Dawe RK (2010) DNA binding of centromere protein C (CENPC) is stabilized by single-stranded RNA. PLoS Genetics. 6:e1000835.

Duda KJ, Ching RW, Jerabek L, Shukeir N, Erikson G, Engist B, Onishi-Seebacher M, Perrera V, Richter F, Mittler G, et al. (2021) m6A RNA methylation of major satellite repeat transcripts facilitates chromatin association and RNA:DNA hybrid formation in mouse heterochromatin. Nucleic Acids Research. 49:5568–5587.

Dunleavy EM, Roche D, Tagami H, Lacoste N, Ray-Gallet D, Nakamura Y, Daigo Y, Nakatani Y, Almouzni-Pettinotti G (2009) HJURP is a cell-cycle-dependent maintenance and deposition factor of CENP-A at centromeres. Cell. 137:485–497.

Durfy SJ, Willard HF (1987) Molecular analysis of a polymorphic domain of alpha satellite from the human X chromosome. The American Journal of Human Genetics. 41:391–401.

Earnshaw WC, Rothfield N (1985) Identification of a family of human centromere proteins using autoimmune sera from patients with scleroderma. Chromosoma. 91:313–321.

Earnshaw WC, Ratrie H 3rd, Stetten G (1989) Visualization of centromere proteins CENP-B and CENP-C on a stable dicentric chromosome in cytological spreads. Chromosoma. 98:1–12.

Enukashvily NI, Semenova N, Chubar AV, Ostromyshenskii DI, Gushcha EA, Gritsaev S, Bessmeltsev SS, Rugal VI, Prikhodko EM, Kostroma I, et al. (2022) Pericentromeric Non-Coding DNA Transcription Is Associated with Niche Impairment in Patients with Ineffective or Partially Effective Multiple Myeloma Treatment. International Journal of Molecular Sciences. 23:3359.

Fachinetti D, Folco HD, Nechemia-Arbely Y, Valente LP, Nguyen K, Wong AJ, Zhu Q, Holland AJ, Desai A, Jansen LE, Cleveland DW (2013) A two-step mechanism for epigenetic specification of centromere identity and function. Nature Cell Biology. 15:1056–1066.

Fachinetti D, Han JS, McMahon MA, Ly P, Abdullah A, Wong AJ, Cleveland DW (2015) DNA Sequence-Specific Binding of CENP-B Enhances the Fidelity of Human Centromere Function. Developmental Cell. 33:314–327.

Fagegaltier D, Bougé AL, Berry B, Poisot E, Sismeiro O, Coppée JY, Théodore L, Voinnet O, Antoniewski C (2009) The endogenous siRNA pathway is involved in heterochromatin formation in Drosophila. Proceedings of the National Academy of Sciences of the United States of America. 106:21258–21263.

Feliciello I, Parazajder J, Akrap I, Ugarković Đ (2013) First evidence of DNA methylation in insect *Tribolium castaneum*—Environmental regulation of DNA methylation within heterochromatin. Epigenetics. 8:534–541.

Feliciello I, Akrap I, Ugarković Đ (2015a) Satellite DNA modulates gene expression in the beetle *Tribolium castaneum* after heat stress. PLoS Genetics. 11:e1005466.

Feliciello I, Akrap I, Brajković J, Zlatar I, Ugarković Đ (2015b) Satellite DNA as a driver of population divergence in the red flour beetle *Tribolium castaneum*. Genome Biology and Evolution. 7:228–239.

Feliciello I, Sermek A, Pezer Ž, Matulić M, Ugarković Đ (2020a) Heat stress affects H3K9me3 level at human alpha satellite DNA repeats. Genes. 11:663.

Feliciello I, Pezer Z, Kordiš D, Bruvo-Mađarić B, Ugarković Đ (2020b) Evolutionary history of alpha satellite DNA repeats dispersed within human genome euchromatin. Genome Biology and Evolution. 12:2125–2138.

Feliciello I, Pezer Ž, Sermek A, Bruvo Mađarić B, Ljubić S, Ugarković Đ (2021) Satellite DNA-mediated gene expression regulation: Physiological and evolutionary implication. Progress in Molecular and Subcellular Biology. 60:145–168.

Fukagawa T, Nogami M, Yoshikawa M, Ikeno M, Okazaki T, Takami Y, Nakayama T, Oshimura M (2004) Dicer is essential for formation of the heterochromatin structure in vertebrate cells. Nature Cell Biology. 6:784–791.

Gall, JG, Atherton, DD (1974) Satellite DNA sequences in *Drosophila virilis*. Journal of molecular biology. 85:633-664.

Galletti G, Portella L, Tagawa ST, Kirby BJ, Giannakakou P, Nanus DM (2014) Circulating tumor cells in prostate cancer diagnosis and monitoring: An appraisal of clinical potential. Molecular Diagnosis & Therapy. 18:389–402.

Ge Y, Wagner MJ, Siciliano M, Wells DE (1992) Sequence, higher order repeat structure, and long-range organization of alpha satellite DNA specific to human chromosome 8. Genomics. 13:585–593.

Gieni R, Chan G, Hendzel M (2008) Epigenetics regulate centromere formation and kinetochore function. Journal of Cellular Biochemistry. 104:2027–2039.

Goenka A, Sengupta S, Pandey R, Parihar R, Mohanta GC, Mukerji M, Ganesh S (2016) Human satellite-III non-coding RNAs modulate heat-shock-induced transcriptional repression. Journal of Cell Science. 129:3541–3552.

Green MR, Sambrook J (2019) Removing DNA contamination from RNA samples by treatment with RNase-free DNase I. Cold Spring Harbor Protocols. pdb.prot101725.

Grewal SI, Elgin SC (2007) Transcription and RNA interference in the formation of heterochromatin. Nature. 447:399–406.

Guo LY, Allu PK, Zandarashvili L, McKinley KL, Sekulic N, Dawicki-McKenna JM, Fachinetti D, Logsdon GA, Jamiolkowski RM, Cleveland DW, Cheeseman IM, Black BE (2017) Centromeres are maintained by fastening CENP-A to DNA and directing an arginine anchor-dependent nucleosome transition. Nature Communications. 8:15775.

Gutbrod MJ, Martienssen RA (2020) Conserved chromosomal functions of RNA interference. Nature Reviews Genetics. 21:311–331.

Gutbrod MJ, Roche B, Steinberg JI, Lakhani AA, Chang K, Schorn AJ, Martienssen RA (2022) Dicer promotes genome stability via the bromodomain transcriptional co-activator BRD4. Nature Communications. 13:1001.

Haaf T, Ward DC (1994) Structural analysis of alpha-satellite DNA and centromere proteins using extended chromatin and chromosomes. Human Molecular Genetics. 3:697–709.

Hall LL, Carone DM, Gomez AV, Kolpa HJ, Byron M, Mehta N, Fackelmayer FO, Lawrence JB (2014) Stable C0T-1 repeat RNA is abundant and is associated with euchromatic interphase chromosomes. Cell. 156:907–919.

Hall LL, Byron M, Carone DM, Whitfield TW, Pouliot GP, Fischer A, Jones P, Lawrence JB (2017) HSATII DNA and HSATII RNA Foci Sequester PRC1 and MeCP2 into Cancer-Specific Nuclear Bodies. Cell Reports. 18:2943–2956.

Hashemipetroudi SH, Nematzadeh G, Ahmadian G, Yamchi A, Kuhlmann M (2018) Assessment of DNA contamination in RNA samples based on ribosomal DNA. Journal of Visualized Experiments. 22:55451.

Hasson D, Panchenko T, Salimian KJ, Salman MU, Sekulic N, Alonso A, Warburton PE, Black BE (2013) The octamer is the major form of CENP-A nucleosomes at human centromeres. Nature Structural & Molecular Biology. 20:687–695.

Hédouin S, Grillo G, Ivkovic I, Velasco G, Francastel C (2017) CENP-A chromatin disassembly in stressed and senescent murine cells. Scientific Reports. 7:42520.

Holoch D, Moazed D (2015) RNA-mediated epigenetic regulation of gene expression. Nature Reviews Genetics. 16:71–84.

Hsieh CL, Xia J, Lin H (2020) MIWI prevents aneuploidy during meiosis by cleaving excess satellite RNA. The EMBO Journal. 17:e103614.

Huang C, Wang X, Liu X, Cao S, Shan G (2015) RNAi pathway participates in chromosome segregation in mammalian cells. Cell Discovery. 1:15029.

Ideue T, Cho Y, Nishimura K, Tani T (2014) Involvement of satellite I noncoding RNA in regulation of chromosome segregation. Genes to cells : devoted to molecular & cellular mechanisms 19:528–538.

Ikeno M, Masumoto H, Okazaki T (1994) Distribution of CENP-B boxes reflected in CREST centromere antigenic sites on long-range alpha-satellite DNA arrays of human chromosome 21. Human Molecular Genetics. 3:1245–1257.

Jaggi KE, Hoyt SJ, O'Neill RJ, Sullivan BA (2026) A genomic and epigenomic view of human centromeres. Nature Reviews Genetics. 10:1038–1052.

Jain M, Olsen HE, Turner DJ, Stoddart D, Bulazel KV, Paten B, Haussler D, Willard HF, Akeson M, Miga KH (2018) Linear assembly of a human centromere on the Y chromosome. Nature Biotechnology. 36:321–323.

Jansen LE, Black BE, Foltz DR, Cleveland DW (2007) Propagation of centromeric chromatin requires exit from mitosis. Journal of Cell Biology. 176:795–805.

Johnson WL, Yewdell WT, Bell JC, McNulty SM, Duda Z, O'Neill RJ, Sullivan BA, Straight AF (2017) RNA-dependent stabilization of SUV39H1 at constitutive heterochromatin. eLife. 6:e25299.

Jolly C, Metz A, Govin J, Vigneron M, Turner BM, Khochbin S, Vourc'h C (2004) Stress-induced transcription of satellite III repeats. Journal of Cell Biology. 164:25–33.

Joshi SS, Meller VH (2017) Satellite Repeats Identify X Chromatin for Dosage Compensation in *Drosophila melanogaster* Males. Current Biology. 27:1393–1402.

Kakizawa N, Suzuki K, Abe I, Endo Y, Tamaki S, Ishikawa H, Watanabe F, Ichida K, Saito M, Futsuhara K, et al. (2019) High relative levels of satellite alpha transcripts predict increased risk of bilateral breast cancer and multiple primary cancers in patients with breast cancer and lacking BRCA-related clinical features. Oncology Reports. 42:857–865.

Kalluri R (2016) The biology and function of exosomes in cancer. Journal of Clinical Investigation. 126:1208–1215.

Kanne J, Hussong M, Isensee J, Muñoz-López Á, Wolffgramm J, Heß F, Grimm C, Bessonov S, Meder L, Wang J, et al. (2021) Pericentromeric Satellite III transcripts induce etoposide resistance. Cell Death & Disease. 24:530.

Kim JH, Ebersole T, Kouprina N, Noskov VN, Ohzeki J, Masumoto H, Mravinac B, Sullivan BA, Pavlicek A, Dovat S, Pack SD, Kwon YW, Flanagan PT, Loukinov D, Lobanenkov V, Larionov V (2009) Human gamma-satellite DNA maintains open chromatin structure and protects a transgene from epigenetic silencing. Genome research. 19:533–544.

Kishikawa T, Otsuka M, Yoshikawa T, Ohno M, Ijichi H, Koike K (2016) Satellite RNAs promote pancreatic oncogenic processes via the dysfunction of YBX1. Nature Communications. 7:13006.

Kishikawa T, Otsuka M, Yoshikawa T, Ohno M, Yamamoto K, Yamamoto N, Kotani A, Koike K (2016) Quantitation of circulating satellite RNAs in pancreatic cancer patients. JCI Insight. 1:e86646.

Kishikawa T, Otsuka M, Suzuki T, Seimiya T, Sekiba K, Ishibashi R, Tanaka E, Ohno M, Yamagami M, Koike K (2018) Satellite RNA Increases DNA Damage and Accelerates Tumor Formation in Mouse Models of Pancreatic Cancer. Molecular Cancer Research. 16:1255–1262.

Kornberg RD (1974) Chromatin Structure: A Repeating Unit of Histones and DNA. Science. 184:868–871.

Kuhn GC, Küttler H, Moreira-Filho O, Heslop-Harrison JS (2012) The 1.688 repetitive DNA of Drosophila: concerted evolution at different genomic scales and association with genes. Molecular Biology and Evolution. 29:7–11.

Kumar SV, Hurteau GJ, Spivack SD (2006) Validity of messenger RNA expression analyses of human saliva. Clinical Cancer Research. 12:5033–5039.

Lam AL, Boivin CD, Bonney CF, Rudd MK, Sullivan BA (2006) Human centromeric chromatin is a dynamic chromosomal domain that can spread over noncentromeric DNA. Proceedings of the National Academy of Sciences of the United States of America. 103:4186–4191.

Landers CC, Rabeler CA, Ferrari EK, et al. (2021) Ectopic expression of pericentric HSATII RNA results in nuclear RNA accumulation, MeCP2 recruitment, and cell division defects. Chromosoma. 130:75–90.

Leclerc S, Kitagawa K (2021) The Role of Human Centromeric RNA in Chromosome Stability. Frontiers in Molecular Biosciences. 8:642732.

Li Y, Syed J, Sugiyama H (2016) RNA-DNA Triplex Formation by Long Noncoding RNAs. Cell Chemical Biology. 23:1325–1333.

Li X, Zhang P, Wang H, Yu Y (2022) Genes expressed at low levels raise false discovery rates in RNA samples contaminated with genomic DNA. BMC Genomics. 23:554.

Liu H, Qu Q, Warrington R, Rice A, Cheng N, Yu H (2015) Mitotic Transcription Installs Sgo1 at Centromeres to Coordinate Chromosome Segregation. Molecular Cell. 59:426–436.

Maison C, Bailly D, Roche D, Montes de Oca R, Probst AV, Vassias I, Dingli F, Lombard B, Loew D, Quivy JP, et al. (2011) SUMOylation promotes de novo targeting of HP1alpha to pericentric heterochromatin. Nature Genetics. 43:220–227.

Maloney KA, Sullivan LL, Matheny JE, Strome ED, Merrett SL, Ferris A, Sullivan BA (2012) Functional epialleles at an endogenous human centromere. Proceedings of the National Academy of Sciences of the United States of America. 109:13704–13709.

Manuelidis L (1978) Chromosomal localization of complex and simple repeated human DNAs. Chromosoma. 66:23–32.

Mascarello JT, Mazrimas JA (1977) Chromosomes of antelope squirrels (genus *Ammospermophilus*): A systematic banding analysis of four species with unusual constitutive heterochromatin. Chromosoma. 64:207–217.

Masumoto H, Masukata H, Muro Y, Nozaki N, Okazaki T (1989) A human centromere antigen (CENP-B) interacts with a short specific sequence in alphoid DNA, a human centromeric satellite. Journal of Cell Biology. 109:1963–1973.

Mazrimas JA, Hatch FT (1972) A Possible Relationship between Satellite DNA and the Evolution of Kangaroo Rat Species (Genus *Dipodomys*). Nature New Biology. 240:102-105.

McNulty SM, Sullivan LL, Sullivan BA (2017) Human Centromeres Produce Chromosome-Specific and Array-Specific Alpha Satellite Transcripts that Are Complexed with CENP-A and CENP-C. Developmental Cell. 42:226–240.e226.

McNulty SM, Sullivan BA (2018) Alpha satellite DNA biology: Finding function in the recesses of the genome. Chromosome Research. 26:115–138.

Miga KH (2015) Completing the human genome: the progress and challenge of satellite DNA assembly. Chromosome Research. 23:421–426.

Mihìc P, Hédouin S, Francastel C (2021) Centromeres Transcription and Transcripts for Better and for Worse. Progress in Molecular and Subcellular Biology. 60:169–201.

Miyata K, Imai Y, Hori S, Nishio M, Loo TM, Okada R, Yang L, Nakadai T, Maruyama R, Fujii R, et al. (2021) Pericentromeric noncoding RNA changes DNA binding of CTCF and inflammatory gene expression in senescence and cancer. Proceedings of the National Academy of Sciences of the United States of America. 118:e2025647118.

Molina O, Vargiu G, Abad MA, Zhiteneva A, Jeyaprakash AA, Masumoto H, Kouprina N, Larionov V, Earnshaw WC (2016) Epigenetic engineering reveals a balance between histone modifications and transcription in kinetochore maintenance. Nature Communications. 7:13334.

Mondal T, Subhash S, Vaid R, Enroth S, Uday S, Reinius B, Mitra S, Mohammed A, James AR, Hoberg E, et al. (2015) MEG3 long noncoding RNA regulates the TGF-β pathway genes through formation of RNA-DNA triplex structures. Nature Communications. 6:7743.

Mravinac B, Sullivan LL, Reeves JW, Yan CM, Kopf KS, Farr CJ, Schueler MG, Sulivan BA (2009) Histone modifications within the human X centromere region. PLoS Genetics. 4:e6602.

Müller M, Fäh T, Schaefer M, Hermes V, Luitz J, Stalder P, Arora R, Ngondo RP, Ciaudo C (2022) AGO1 regulates pericentromeric regions in mouse embryonic stem cells. Life Science Alliance. 5:e202101277.

Muro Y, Masumoto H, Yoda K, Nozaki N, Ohashi M, Okazaki T (1992) Centromere protein B assembles human centromeric alpha-satellite DNA at the 17-bp sequence, CENP-B box. Journal of Cell Biology. 116:585–596.

Musacchio A, Desai A (2017) A Molecular View of Kinetochore Assembly and Function. Biology. 6:5.

Nogalski MT, Shenk T (2020) HSATII RNA is induced via a noncanonical ATM-regulated DNA damage response pathway and promotes tumor cell proliferation and movement. Proceedings of the National Academy of Sciences of the United States of America. 117:31891–31901.

Ohzeki J, Nakano M, Okada T, Masumoto H (2002) CENP-B box is required for de novo centromere chromatin assembly on human alphoid DNA. Journal of Cell Biology. 159:765–775.

Ohzeki J, Bergmann JH, Kouprina N, Noskov VN, Nakano M, Kimura H, Earnshaw WC, Larionov V, Masumoto H (2012) Breaking the HAC Barrier: histone H3K9 acetyl/methyl balance regulates CENP-A assembly. The EMBO Journal. 31:2391–2402.

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

Okada T, Ohzeki J, Nakano M, Yoda K, Brinkley WR, Larionov V, Masumoto H (2007) CENP-B controls centromere formation depending on the chromatin context. Cell. 131:1287–1300.

Özgür E, Mayer Z, Keskin M, Yörüker EE, Holdenrieder S, Gezer U (2021) Satellite 2 repeat DNA in blood plasma as a candidate biomarker for the detection of cancer. Clinica Chimica Acta. 514:74–79.

Padhi BK, Singh M, Huang N, Pelletier GA (2016) PCR-based approach to assess genomic DNA contamination in RNA: Application to rat RNA samples. Analytical Biochemistry. 494:49–51.

Palmer DK, O'Day K, Wener MH, Andrews BS, Margolis RL (1987) A 17-kD centromere protein (CENP-A) copurifies with nucleosome core particles and with histones. Journal of Cell Biology. 104:805–815.

Palmer DK, O'Day K, Trong HL, Charbonneau H, Margolis RL (1991) Purification of the centromere-specific protein CENP-A and demonstration that it is a distinctive histone. Proceedings of the National Academy of Sciences of the United States of America. 88:3734–3738.

Pavlek M, Gelfand Y, Plohl M, Meštrović N (2015) Genome-wide analysis of tandem repeats in *Tribolium castaneum* genome reveals abundant and highly dynamic tandem repeat families with satellite DNA features in euchromatic chromosomal arms. DNA Research. 22:387–401.

Pecinka A, Dinh HQ, Baubec T, Rosa M, Lettner N, Mittelsten Scheid O (2010) Epigenetic regulation of repetitive elements is attenuated by prolonged heat stress in Arabidopsis. The Plant Cell. 22:3118–3129.

Peterson CL, Laniel M-A (2004) Histones and histone modifications. Current Biology. 14:546–551.

Pezer Ž, Ugarković Đ (2008) Role of non-coding RNA and heterochromatin in aneuploidy and cancer. Seminars in Cancer Biology. 18:123–130.

Pezer Ž, Ugarković Đ (2012) Satellite DNA-associated siRNAs as mediators of heat shock response in insects. RNA Biology. 9:587–595.

Pita S, Panzera F, Mora P, Vela J, Cuadrado Á, Sánchez A, Palomeque T, Lorite P (2017) Comparative repeatome analysis on *Triatoma infestans* Andean and non-Andean lineages, main vector of Chagas disease. PLoS One. 12:e0181635.

Politi V, Perini G, Trazzi S, Pliss A, Raska I, Earnshaw WC, Della Valle G (2002) CENP-C binds the alpha-satellite DNA in vivo at specific centromere domains. Journal of Cell Science. 115:2317–2327.

Porokhovnik LN, Veiko NN, Ershova ES, Kostyuk SV (2021) The Role of Human Satellite III (1q12) Copy Number Variation in the Adaptive Response during Aging, Stress, and Pathology: A Pendulum Model. Genes. 12:1524.

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

Rajshekar S, Yao J, Arnold PK, Payne SG, Zhang Y, Bowman TV, Schmitz RJ, Edwards JR, Goll M (2018) Pericentromeric hypomethylation elicits an interferon response in an animal model of ICF syndrome. eLife. 7:e39658.

Rizzi N, Denegri M, Chiodi I, Corioni M, Valgardsdottir R, Cobianchi F, Riva S, Biamonti G (2004) Transcriptional activation of a constitutive heterochromatic domain of the human genome in response to heat shock. Molecular Biology of the Cell. 15:543–551.

Rosandic M, Paar V, Basar I, Gluncic M, Pavin N, Pilas I (2006) CENP-B box and pJalpha sequence distribution in human alpha satellite higher-order repeats (HOR). Chromosome Research. 14:735–753.

Ross MT et al., (2005) The DNA sequence of the human X chromosome. Nature. 434:325–337.

Ross JE, Woodlief KS, Sullivan BA (2016) Inheritance of the CENP-A chromatin domain is spatially and temporally constrained at human centromeres. Epigenetics & Chromatin. 9:20.

Rudd MK, Schueler MG, Willard HF (2003) Sequence organization and functional annotation of human centromeres. Cold Spring Harbor Symposia on Quantitative Biology. 68:141–149.

Ruiz-Ruano FJ, López-León MD, Cabrero J, Camacho JP (2016) High-throughput analysis of the satellitome illuminates satellite DNA evolution. Scientific Reports. 6:28333.

Saini S (2016) PSA and beyond: Alternative prostate cancer biomarkers. Cellular Oncology. 39:97–106.

Schueler MG, Higgins AW, Rudd MK, Gustashaw K, Willard HF (2001) Genomic and genetic definition of a functional human centromere. Science. 294:109–115.

Schueler MG, Dunn JM, Bird CP, Ross MT, Viggiano L, Program NCS, Rocchi M, Willard HF, Green ED (2005) Progressive proximal expansion of the primate X chromosome centromere. Proceedings of the National Academy of Sciences of the United States of America. 102:10563–10568.

Sermek A, Feliciello I, Ugarković Đ (2021) Distinct Regulation of the Expression of Satellite DNAs in the Beetle *Tribolium castaneum*. International Journal of Molecular Sciences. 22:296.

Shang WH, Hori T, Westhorpe FG, Godek KM, Toyoda A, Misu S, Monma N, Ikeo K, Carroll CW, Takami Y, Fujiyama A, Kimura H, Straight AF, Fukagawa T (2016) Acetylation of histone H4 lysine 5 and 12 is required for CENP-A deposition into centromeres. Nature Communications. 7:13465.

Shelby RD, Vafa O, Sullivan KF (1997) Assembly of CENP-A into centromeric chromatin requires a cooperative array of nucleosomal DNA contact sites. Journal of Cell Biology. 136:501–513.

Shelby RD, Monier K, Sullivan KF (2000) Chromatin assembly at kinetochores is uncoupled from DNA replication. Journal of Cell Biology. 151:1113–1118.

Shepelev VA, Alexandrov AA, Yurov YB, Alexandrov IA (2009) The evolutionary origin of man can be traced in the layers of defunct ancestral alpha satellites flanking the active centromeres of human chromosomes. PLoS Genetics. 5:e1000641.

Shepelev VA, Uralsky LI, Alexandrov AA, Yurov YB, Rogaev EI, Alexandrov IA (2015) Annotation of suprachromosomal families reveals uncommon types of alpha satellite organization in pericentromeric regions of hg38 human genome assembly. Genomics Data. 5:139–146.

Shono N, Ohzeki J, Otake K, Martins NM, Nagase T, Kimura H, Larionov V, Earnshaw WC, Masumoto H (2015) CENP-C and CENP-I are key connecting factors for kinetochore and CENP-A assembly. Journal of Cell Science. 128:4572–4587.

Smurova K, De Wulf P (2018) Centromere and Pericentromere Transcription: Roles and Regulation … in Sickness and in Health. Frontiers in Genetics. 9:674.

Statello L, Guo CJ, Chen LL, Huarte M (2021) Gene regulation by long non-coding RNAs and its biological functions. Nature Reviews Molecular Cell Biology. 22:96–118.

Sullivan BA, Schwartz S (1995) Identification of centromeric antigens in dicentric Robertsonian translocations: CENP-C and CENP-E are necessary components of functional centromeres. Human Molecular Genetics. 4:2189–2197.

Sullivan BA, Karpen GH (2004) Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. Nature Structural & Molecular Biology. 11:1076–1083.

Sullivan LL, Boivin CD, Mravinac B, Song IY, Sullivan BA (2011) Genomic size of CENP-A domain is proportional to total alpha satellite array size at human centromeres and expands in cancer cells. Chromosome Research. 19:457–470.

Sullivan LL, Maloney KA, Towers AJ, Gregory SG, Sullivan BA (2016) Human centromere repositioning within euchromatin after partial chromosome deletion. Chromosome Research. 24:451–466.

Sullivan LL, Chew K, Sullivan BA (2017) Alpha satellite DNA variation and function of the human centromere. Nucleus. 8:331–339.

Sutton WD, Walker PM (1972) Self-renaturing fractions in the separated strands of mouse satellite deoxyribonucleic acid. The biochemical journal. 128:193-198.

Takahashi A, Okada R, Nagao K, Kawamata Y, Hanyu A, Yoshimoto S, Takasugi M, Watanabe S, Kanemaki MT, Obuse C, et al. (2017) Exosomes maintain cellular homeostasis by excreting harmful DNA from cells. Nature Communications. 8:15287.

Tanne A, Muniz LR, Puzio-Kuter A, Leonova KI, Gudkov AV, Ting DT, Monasson R, Cocco S, Levine AJ, Bhardwaj N, et al. (2015) Distinguishing the immunostimulatory properties of noncoding RNAs expressed in cancer cells. Proceedings of the National Academy of Sciences of the United States of America. 112:15154–15159.

Tasselli L, Xi Y, Zheng W, et al. (2016) SIRT6 deacetylates H3K18ac at pericentric chromatin to prevent mitotic errors and cellular senescence. Nature Structural & Molecular Biology. 23:434–440.

Ting DT, Lipson D, Paul S, Brannigan BW, Akhavanfard S, Coffman EJ, Contino G, Deshpande V, Iafrate AJ, Letovsky S, et al. (2011) Aberrant overexpression of satellite repeats in pancreatic and other epithelial cancers. Science. 331:593–596.

Tittel-Elmer M, Bucher E, Broger L, Mathieu O, Paszkowski J, Vaillant I (2010) Stress-induced activation of heterochromatic transcription. PLoS Genetics. 6:e1001175.

Trazzi S, Bernardoni R, Diolaiti D, Politi V, Earnshaw WC, Perini G, Della Valle G (2002) *In vivo* functional dissection of human inner kinetochore protein CENP-C. Journal of Structural Biology. 140:39–48.

Trowell HE, Nagy A, Vissel B, Choo KH (1993) Long-range analyses of the centromeric regions of human chromosomes 13, 14 and 21: identification of a narrow domain containing two key centromeric DNA elements. Human molecular genetics. 2:1639–1649.

Unoki M, Sharif J, Saito Y, Velasco G, Francastel C, Koseki H, Sasaki H (2020) CDCA7 and HELLS suppress DNA:RNA hybrid-associated DNA damage at pericentromeric repeats. Scientific Reports. 10:17865.

Vafa O, Sullivan KF (1997) Chromatin containing CENP-A and alpha-satellite DNA is a major component of the inner kinetochore plate. Current Biology. 7:897–900.

Valadi H, Ekström K, Bossios A, Sjöstrand M, Lee JJ, Lötvall JO (2007) Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. Nature Cell Biology. 9:654–659.

Velazquez Camacho O, Galan C, Swist-Rosowska K, Ching R, Gamalinda M, Karabiber F, De La Rosa-Velazquez I, Engist B, Koschorz B, Shukeir N, et al. (2017) Major satellite repeat RNA stabilize heterochromatin retention of Suv39h enzymes by RNA-nucleosome association and RNA:DNA hybrid formation. eLife. 6:817.

Verwilt J, Trypsteen W, Van Paemel R, De Preter K, Giraldez MD, Mestdagh P, Vandesompele J (2020) When DNA gets in the way: A cautionary note for DNA contamination in extracellular RNA-seq studies. Proceedings of the National Academy of Sciences of the United States of America. 117:18934.

Vlahović I, Glunčić M, Rosandić M, Ugarković Đ, Paar V (2017) Regular higher order repeat structures in beetle *Tribolium castaneum* genome. Genome Biology and Evolution. 9:2668–2680.

Vojvoda Zeljko T, Ugarković Đ, Pezer Ž (2021) Differential enrichment of H3K9me3 at annotated satellite DNA repeats in human cell lines and during fetal development in mouse. Epigenetics & Chromatin. 14:47.

Vourc'h C, Biamonti G (2011) Transcription of Satellite DNAs in Mammals. Progress in Molecular and Subcellular Biology. 51:95–118.

Vourc'h C, Dufour S, Timcheva K, Seigneurin-Berny D, Verdel A (2022) HSF1-Activated Non-Coding Stress Response: Satellite lncRNAs and Beyond, an Emerging Story with a Complex Scenario. Genes. 13:597.

Warburton PE, Willard HF (1992) PCR amplification of tandemly repeated DNA: analysis of intra- and interchromosomal sequence variation and homologous unequal crossing-over in human alpha satellite DNA. Nucleic Acids Research. 20:6033–6042.

Warburton PE, Waye JS, Willard HF (1993) Nonrandom localization of recombination events in human alpha satellite repeat unit variants: implications for higher-order structural characteristics within centromeric heterochromatin. Molecular and Cellular Biology. 13:6520–6529.

Warburton PE, Willard HF (1995) Interhomologue sequence variation of alpha satellite DNA from human chromosome 17: evidence for concerted evolution along haplotypic lineages. Journal of Molecular Evolution. 41:1006–1015.

Warburton PE, Cooke CA, Bourassa S, Vafa O, Sullivan BA, Stetten G, Gimelli G, Warburton D, Tyler-Smith C, Sullivan KF, Poirier GG, Earnshaw WC (1997) Immunolocalization of CENP-A suggests a distinct nucleosome structure at the inner kinetochore plate of active centromeres. Current Biology. 7:901–904.

Warwick T, Brandes RP, Leisegang MS (2023) Computational Methods to Study DNA:DNA:RNA Triplex Formation by lncRNAs. Non-Coding RNA. 9:10.

Waye JS, Willard HF (1986a) Molecular analysis of a deletion polymorphism in alpha satellite of human chromosome 17: evidence for homologous unequal crossing-over and subsequent fixation. Nucleic Acids Research. 14:6915–6927.

Waye JS, Willard HF (1986b) Structure, organization, and sequence of alpha satellite DNA from human chromosome 17: evidence for evolution by unequal crossing-over and an ancestral pentamer repeat shared with the human X chromosome. Molecular and Cellular Biology. 6:3156–3165.

Waye JS, Willard HF (1987) Nucleotide sequence heterogeneity of alpha satellite repetitive DNA: a survey of alphoid sequences from different human chromosomes. Nucleic Acids Research. 15:7549–7569.

Waye JS, Greig GM, Willard HF (1987) Detection of novel centromeric polymorphisms associated with alpha satellite DNA from human chromosome 11. Human Genetics. 77:151–156.

Wei X, Eickbush DG, Speece I, Larracuente AM (2021) Heterochromatin-dependent transcription of satellite DNAs in the *Drosophila melanogaster* female germline. eLife. 10:e62375.

Willard HF (1985) Chromosome-specific organization of human alpha satellite DNA. The American Journal of Human Genetics. 37:524–532.

Willard HF, Waye JS (1987) Hierarchical order in chromosome-specific alpha satellite DNA. Trends in Genetics. 3:192–198.

Wong LH, Brettingham-Moore KH, Chan L, Quach JM, Anderson MA, Northrop EL, Hannan R, Saffery R, Shaw ML, Williams E, Choo KH (2007) Centromere RNA is a key component for the assembly of nucleoproteins at the nucleolus and centromere. Genome Research. 17:1146–1160.

Wu JC, Manuelidis L (1980) Sequence definition and organization of a human repeated DNA. Journal of Molecular Biology. 142:363–386.

Wylie A, Jones AE, D'Brot A, Lu WJ, Kurtz P, Moran JV, Rakheja D, Chen KS, Hammer RE, Comerford SA, et al. (2016) p53 genes function to restrain mobile elements. Genes & Development. 30:64–77.

Yasmineh WG, Yunis JJ (1974) Localization of repeated DNA sequences in CsCl gradients by hybridization with complementary RNA. Experimental cell research. 88:340–344.

Yoda K, Ando S, Okuda A, Kikuchi A, Okazaki T (1998) In vitro assembly of the CENP-B/alpha-satellite DNA/core histone complex: CENP-B causes nucleosome positioning. Genes to cells. 3:533–548.

Zeller P, Gasser SM (2017) The Importance of Satellite Sequence Repression for Genome Stability. Cold Spring Harbor Symposia on Quantitative Biology. 82:15–24.

Zhu Q, Pao GM, Huynh AM, Suh H, Tonnu N, Nederlof PM, Gage FH, Verma IM (2011) BRCA1 tumour suppression occurs via heterochromatin-mediated silencing. Nature. 477:179–184.

Zhu Q, Hoong N, Aslanian A, Hara T, Benner C, Heinz S, Miga KH, Ke E, Verma S, Soroczynski J, et al. (2018) Heterochromatin-Encoded Satellite RNAs Induce Breast Cancer. Molecular Cell. 70:842–853.

5. PUBLICATIONS

I. Ljubić S, Matulić M, Đermić D, Feliciello MC, Procino A, Passaro F, Ugarković Đ, Feliciello I (2025) Downregulation of Gene Expression by Alpha Satellite Transcripts. International Journal of Molecular Sciences. 26:11204.

II. Ljubić S, Matulić M, Đermić D, Feliciello MC, Procino A, Ugarković Đ, Feliciello I (2025) Antibiotics induce overexpression of alpha satellite DNA accompanied with epigenetic changes at alpha satellite arrays as well as genome-wide. Epigenetics & Chromatin. 18:62.

III. Ljubić S, Sermek A, Prgomet Sečan A, Prpić M, Jakšić B, Murgić J, Fröbe A, Ugarković Đ, Feliciello I (2022) Alpha Satellite RNA Levels Are Upregulated in the Blood of Patients with Metastatic Castration-Resistant Prostate Cancer. Genes. 13:383.

IV. Đermić D, Ljubić S, Matulić M, Procino A, Feliciello MC, Ugarković Đ, Feliciello I (2023) Reverse transcription-quantitative PCR (RT-qPCR) without the need for prior removal of DNA. Scientific Reports. 13:11470.

PUBLICATION I.

Ljubić S, Matulić M, Đermić D, Feliciello MC, Procino A, Passaro F, Ugarković Đ, Feliciello I (2025) Downregulation of Gene Expression by Alpha Satellite Transcripts. International Journal of Molecular Sciences. 26:11204.

International Journal of *Molecular Sciences*

MDPI

Article

Downregulation of Gene Expression by Alpha Satellite Transcripts

Sven Ljubić [1], Maja Matulić [2], Damir Đermić [1], Maria Chiara Feliciello [3], Alfredo Procino [4], Francesco Passaro [5], Đurđica Ugarković [1,*] and Isidoro Feliciello [1,4,*]

1 Division of Molecular Biology, Ruđer Bošković Institute, Bijenička 54, HR-10000 Zagreb, Croatia
2 Department of Biology, Faculty of Science, University of Zagreb, HR-10000 Zagreb, Croatia
3 Department of Statistical Science, Alma Mater Studiorum, University of Bologna, 40126 Bologna, Italy
4 Department of Clinical Medicine and Surgery, University of Naples Federico II, 80131 Naples, Italy
5 Division of Urology, Instituto Nazionale Tumori, Fondazione G. Pascale, 80131 Naples, Italy
* Correspondence: djurdjica.ugarkovic@irb.hr (Đ.U.); isidoro.feliciello@unina.it (I.F.); Tel.: +385-014561197 (Đ.U.); +39-0817464317 (I.F.)

Abstract

Satellite DNAs are highly abundant sequences that build functional centromeres and pericentromeric heterochromatin in many eukaryotes. Apart from this structural role, their involvement in gene expression modulation has been demonstrated, although a detailed understanding of the molecular mechanisms is still lacking. Here, using the major human alpha satellite as a model system, we investigate the role of satellite transcripts in gene expression regulation. We generated cell lines with forced, exogenous overexpression of alpha satellite RNA and followed the expression levels of genes containing alpha satellite repeats within introns. Our results reveal a positive correlation between exogenous alpha satellite expression and the downregulation of alpha-associated genes, strongly suggesting that alpha satellite RNA affects their transcription. Notably, the elevated levels of exogenous alpha satellite RNA did not affect histone modifications characteristic of pericentromeric heterochromatin (e.g., H3K9me3 or H3K18Ac) or euchromatin (e.g., H3K4me2) at intronic alpha satellite loci. We propose that alpha satellite RNA directly interacts with homologous DNA at dispersed intronic satellite loci by forming RNA-DNA hybrid structures, which may affect chromatin structure and transcriptional activity. The results demonstrate that alpha satellite RNA is not only involved in centromere and heterochromatin assembly but, as shown here for the first time, also plays a role in modulating the expression of alpha-associated genes.

Keywords: satellite DNA; transcription; gene expression; RNA-DNA hybrid; heterochromatin; histone modifications; alpha satellite DNA

Academic Editors: Grzegorz Wegrzyn and Naoyuki Kataoka

Received: 9 September 2025
Revised: 17 November 2025
Accepted: 17 November 2025
Published: 20 November 2025

Citation: Ljubić, S.; Matulić, M.; Đermić, D.; Feliciello, M.C.; Procino, A.; Passaro, F.; Ugarković, Đ.; Feliciello, I. Downregulation of Gene Expression by Alpha Satellite Transcripts. *Int. J. Mol. Sci.* **2025**, *26*, 11204. https://doi.org/10.3390/ijms262211204

1. Introduction

Alpha satellite DNA represents a major human satellite that is composed of tandemly arranged, diverged 171 bp long monomers, often organized in complex higher order repeats [1]. The satellite comprises up to 10% of the human genome and is located in the centromeric and pericentromeric regions of all chromosomes in the form of long, Mb-size arrays [2]. In addition to the (peri)centromeric location, a bioinformatic search of the human genome revealed the presence of short arrays of alpha satellite repeats within the euchromatic regions of the genome, including introns and areas near genes [3]. As previously described, one possible origin of these euchromatic, dispersed forms

of satellite repeats is due to the molecular mechanism of satellite DNA evolution based on the rolling circle amplification [3–5]. This non-canonical occurrence of dispersed satellite repeats has been observed in various species to date and is proposed to be a characteristic feature of all satellite DNAs [6–9]. In the case of the beetle *Tribolium castaneum*, it was demonstrated that dispersed satellite repeats can influence the expression of neighboring genes upon specific conditions, such as heat stress [10]. In general, heat stress and other kinds of stress, like antibiotic treatment [11], induce transcription of satellite DNA. These transcripts are proposed to guide the deposition of repressive histone marks at homologous satellite sequences, affecting both heterochromatic and euchromatic regions [10,12].

In humans, RNA polymerase II (Pol II) transcribes alpha satellite DNA repeats into long non-coding RNA, and the predominant factors controlling transcription seem to be the presence of centromere–nucleolar contacts [13] and topoisomerase I (TopI) [14]. In addition, a role of N6-methyladenosine (m6A) modification in the regulation of alpha satellite transcription was proposed [15]. Alpha satellite RNA levels fluctuate throughout the cell cycle, peaking in the G2/M phase, and the transcripts are not exported to the cytoplasm [13]. Transcripts of alpha satellite DNA contribute to essential chromosomal functions such as centromere assembly and kinetochore formation [2] and are necessary for proper heterochromatin formation in humans [16,17]. Transcription of alpha satellite DNA is also induced upon heat stress, as shown by studies on different cell lines [18,19]. The increased levels of alpha satellite transcripts after heat stress correlate with the downregulation of genes containing alpha satellite repeats within introns or in the gene vicinity, indicating the possible influence of alpha satellite transcripts on gene expression modulation [19].

To analyze the possible gene-modulatory role of alpha satellite transcripts more deeply and to exclude the effect of heat stress itself or of other stressors on gene expression, we have now developed cell lines with forced, exogenous overexpression of alpha satellite RNA. In such modified cell lines, we monitored the level of expression of genes containing alpha satellite repeats within introns. Although alpha repeats can also be dispersed in the vicinity of genes [3], according to our opinion, the influence of alpha RNA can only be unambiguously established on genes having alpha repeats within the gene body. The results reveal a positive correlation between exogenous alpha satellite expression and downregulation of alpha-associated genes, strongly suggesting an effect of alpha satellite transcripts on gene expression. Furthermore, we performed an analysis of different histone marks on intronic alpha satellite repeats in cells with exogenous expression of alpha satellite RNA, and proposed a possible molecular mechanism by which alpha satellite transcripts modulate gene expression.

2. Results

2.1. Alpha Satellite Transcription After Transfection

To investigate the possible effect of alpha satellite RNA on gene expression, we constructed vectors expressing alpha satellite monomers in both orientations (171Fw and 171Rev) and monitored their transcription dynamics in the transfected human MJ90-hTERT cell line by RT-qPCR. The exogenous alpha satellite RNA was quantified 24, 48, 72 and 96 h after transfection and compared to its endogenous variant. Primers used for transcriptional analysis of endogenous alpha satellite RNA were able to amplify only tandemly arranged repeats (Figure S1), and since in human pericentromeric heterochromatin, alpha satellite DNA is organized in tandemly arranged monomers [2], it is expected that the primers preferentially recognize transcripts deriving from pericentromeric regions.

The results revealed a significant level of exogenous alpha satellite expression from both 171Fw and 171Rev vectors, the highest being achieved 24 h after transfection and

rapidly decreasing afterwards (Figure 1), primarily as a consequence of cell division-related plasmid loss and lack of selective pressure. Additionally, both vector variants demonstrated mutually similar levels of expression during the experiments. Endogenous alpha satellite expression retained its basal levels throughout all investigated time points, indicating standard physiological conditions and lack of stress or toxicity for the cells post-treatment.

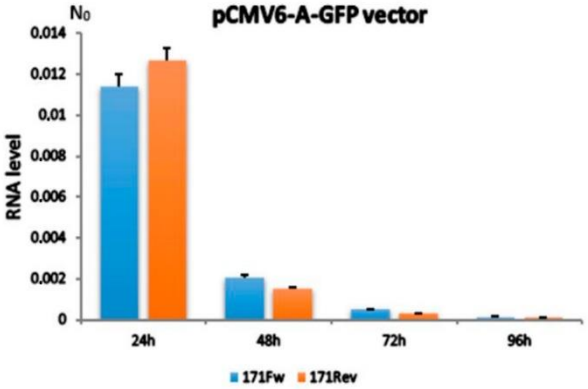


Figure 1. Levels of exogenous alpha satellite RNA expressed from pCMV6-A-GFP vectors 24, 48, 72 and 96 h after transfection. 171Fw indicates the 171 bp alpha satellite monomer insert in forward orientation and 171Rev its inverted form. N_0 value is expressed in arbitrary fluorescence units and is calculated by taking into account PCR efficiency and baseline fluorescence. Columns show averages of two independent experiments, and error bars indicate standard deviations. No significant difference in expression between vectors with 171Fw and 171Rev, respectively, was observed at any time point (Student's *t*-test, $p > 0.1$).

2.2. Expression Analysis of Alpha-Associated Genes After Transfection

The expression profiles of genes containing dispersed alpha satellite repeats within their intronic regions, described in [19] and listed in Table S1, were explored over the course of four consecutive days, after transfection of the MJ90-hTERT cell line with satellite-expressing vectors, and compared to controls transfected in the same way with unmodified vectors. These included the following: *SLC30A6* (solute carrier family 30 member 6, ID: 55676), *STAM* (signal transducing adapter molecule 1, ID: 8027), *MYO1E* (myosin IE, ID: 4643), *MAP7* (MAP7 domain containing 2, ID: 256714), *ZNF675* (zinc finger protein 675, ID: 171392), *VAV1* (vav guanine nucleotide exchange factor 1, ID: 7409), *PRIM2* (DNA primase subunit 2, ID: 5558) and *DLG2* (discs large homolog 2, ID: 1740). There were no additional genes with targetable intronic alpha satellite segments that met design and expression constraints. This is consistent with the idea that insertion of satellite DNA into euchromatin is generally deleterious and therefore rare, leaving few genes that naturally harbor intronic alpha repeats. The glucuronidase beta gene (*GUSB*, ID: 2990) was used as an endogenous control for expression normalization and a negative control against alpha repeat-associated genes.

Additionally, 24 h post-transfection, a significant downregulation of gene expression was detected for six tested genes relative to controls. These included *SLC30A6*, *STAM*, *MYO1E*, *MAP7*, *ZNF675* and *PRIM2*. Out of those, the first five genes demonstrated the highest level of downregulation between the transfected samples and controls ($p < 10^{-3}$); *PRIM2*, although to a lesser extent, was also significant ($p = 0.02$) (Figure 2). Also, there were no observed differences in vectors' downregulation efficiency between 171Fw and 171Rev transfected cells in all cases, indicating comparable levels of effectiveness

regardless of alpha satellite insert orientation. Other investigated time points (48, 72 and 96 h post-transfection) showed no significant differences in gene expression of candidate genes between transfected samples and controls (Figures S2–S4). This is likely due to vector expression dynamics, signifying the most impactful effect of satellite transcription 24 h post-transfection and subsequent rapid decline (Figure 1). The genes *VAV1* and *DLG2* showed no detectable expression in any of the samples at any of the analyzed time points, indicating they are tissue-specific and, therefore, not expressed in the MJ90-hTERT cell line. The *GUSB* gene was stably expressed at all times during the experiments, with no significant variability between samples and controls (Figure S5), confirming its role as a dependable normalization gene for relative RT-qPCR quantification. In addition to *GUSB*, we also tested expression of five more housekeeping genes, *GAPDH*, *TOP3A*, *DEK*, *GPR68* and *IFIT3*, in the MJ90hTERT cell line after 24 h transfection with alpha satellite-expressing vectors (171Fw, 171Rev) and an unmodified control vector. In all cases, no significant differences in gene expression were observed between samples transfected with alpha satellite-expressing vectors and control vector (Figure S6), indicating no influence of alpha satellite RNA on expression of reference genes. It is important to mention that all intronic alpha satellite repeats are flanked by other highly repetitive DNA (*Alu*, L1, SVA, etc.). This prevented the elimination of alpha repeats using CRISPR/Cas9 and enabled the creation of modified cell lines suitable for gene expression studies.

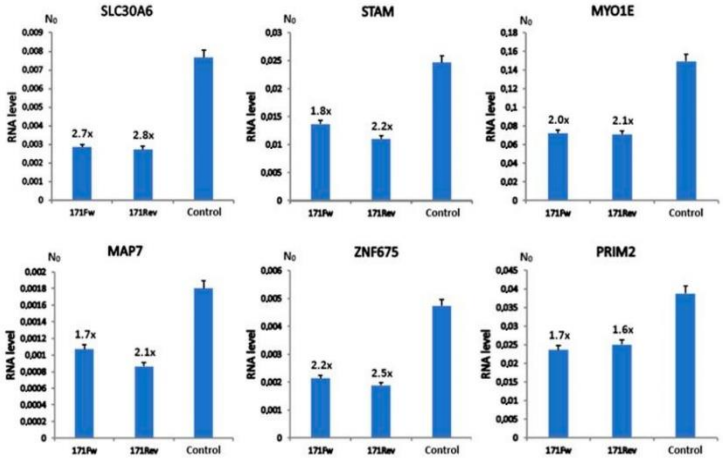


Figure 2. The expression profiles of genes containing alpha satellite repeats within intronic regions, in MJ90hTERT cell lines transfected with satellite-expressing vectors and controls, 24 h after treatment. 171Fw denotes the vector with satellite insert in forward orientation, and 171Rev denotes its inverted counterpart. Control refers to the unaltered pCMV6-A-GFP vector. Gene downregulation rates of transfected samples compared to controls are shown above the error bars, which represent standard deviations. Two independent RT-qPCR experiments were performed, and samples were analyzed in duplicates and averaged mean values are displayed. N_0 represents the normalized average N_0 value expressed in arbitrary fluorescence units. *SLC30A6*, *STAM*, *MYO1E*, *MAP7* and *ZNF675* demonstrated the highest level of downregulation between transfected samples and controls ($p < 10^{-3}$), and *PRIM2*, although to a lesser extent, was also significant ($p \approx 0.02$ for each vector, Student's *t*-test).

Additionally, it should be noted that our observed results are likely to be understated to a certain degree. The MJ90-hTERT human cell line is quite difficult to transfect effectively without using viral vectors (such as lentivirus), which is often accompanied by higher levels of toxicity and may induce adverse effects for the cells. To avoid these problems, using the lipofection method, we managed to achieve approximately 30–35% transfection

efficiency in our experiments by monitoring constitutively expressed GFP inside the cells, an inherent additional trait of our vectors. With higher transfection efficiency, it is likely that the observed downregulatory effects would be more pronounced or even become significant in later time points (48, 72 and 96 h after transfection).

2.3. H3K9me3, H3K18ac and H3K4me2 Levels at Alpha Repeats Dispersed Within Genes After Transfection

We analyzed the distribution of silent histone mark H3K9me3, H3K18ac mark characteristic for transcriptional activation of heterochromatin and H3K4me2, typical of open euchromatin, on alpha repeats dispersed within introns of six genes previously tested for expression. Histone marks were analyzed in MJ90hTERT cells 24 h after transfection with 171Fw, 171Rev and unmodified control expression vectors. In order to investigate whether the previously observed downregulation of gene expression could be related to the above-mentioned epigenetic changes, we performed chromatin immunoprecipitation (ChIP) coupled with quantitative real-time PCR, using specific primers for histone modification level analyses of alpha repeats associated with genes (Table S2). A ChIP assay was performed on chromatin isolated from MJ90hTERT cells. The levels of tested histone modifications were measured immediately after a 24 h transfection period and compared to the level of control transfected in the same way with an unmodified vector, using the unpaired *t*-test. In addition, we followed the IgG binding dynamics to investigate loci, and the amount of bound IgG was very low, resulting in a signal below the qPCR threshold.

The transfection of MJ90hTERT cells with 171Fw alpha satellite expression vector resulted in no significant differences with regard to tested histone modifications at alpha repeats in six genes of interest compared to control samples, 24 h after treatment (Figure 3). Similarly, no differences were observed after repeated experiments with the 171Rev vector containing an inverted alpha satellite insert (Figure S7). Both vectors showed similar results compared to controls, mirroring their close performance from gene expression analyses (Figure 2). On one hand, it is possible that this outcome is a consequence of the previously mentioned limited transfection potential of MJ90hTERT cells or the lower sensitivity of ChIP experiments. However, it might also be an indicator of other potential mechanism(s) modulating gene expression of alpha satellite repeat-associated genes. Suspecting that the signal originating from individual alpha loci was too weak, we performed ChIP-qPCR on tandemly arranged alpha satellite arrays characteristic of heterochromatin. The results revealed a statistically significant increase in silent histone mark H3K9me3 of 1.5× ($p < 0.05$) in MJ90hTERT cells 24 h after transfection with alpha satellite expression vector (Figure S8). Two other tested histone modifications, H3K18ac and H3K4me2, however, remained unchanged.

2.4. Alpha Satellite RNA Level Analysis After RNase H Digestion of RNA:DNA Hybrids

To determine whether alpha satellite transcripts actively form RNA:DNA hybrids under standard physiological conditions, RNase H digestion assay was performed. RNase H specifically degrades the RNA strand of an RNA-DNA hybrid. We detected a modest (around 25%) but significant (Student's *t*-test, $p < 0.05$) reduction in alpha transcripts in samples treated with RNase H compared to untreated controls (Figure 4a). Alpha satellite RNA was quantified by RT-qPCR using the same primer pair as in previous experiments (Fw 5′-CACTCTTTTTGTAGAATCTGC-3′; Rev 5′-AATGCACATATCACTATGTAC-3′). The *GUSB* gene was used as an endogenous housekeeper for normalization and a negative control, being stably and uniformly expressed, showing no significant variation between samples before and after RNAse H treatment (Figure 4b).

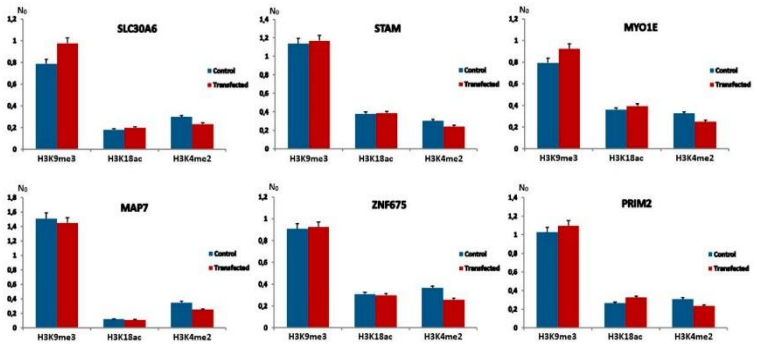


Figure 3. Levels of H3K9me3, H3K18ac and H3K4me2 histone modifications at alpha satellite repeats associated with six genes after 24 h transfection with 171Fw and unmodified control vectors. Levels of histone modifications were measured by ChIP coupled with quantitative real-time PCR on MJ90hTERT chromatin immediately after each treatment. N_0 values were normalized using N_0 values of input fractions and represent levels of histone modifications. Columns show averages of two independent experiments, and error bars indicate standard deviations. No significant differences with regard to tested histone modifications at alpha repeats in six genes of interest compared to control samples were observed ($p > 0.1$ in all cases, Student's *t*-test).

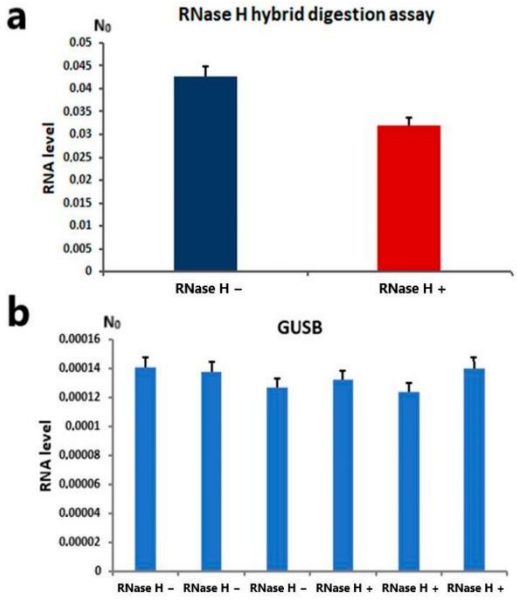


Figure 4. (**a**) Alpha satellite RNA levels in samples treated with RNase H and untreated controls. N_0 represents normalized average N_0 values expressed in arbitrary fluorescence units. Columns show averages of two different RT-qPCR experiments performed in triplicate, and error bars represent standard deviations. Statistical significance between controls and treated samples was calculated using Student's *t*-test ($p < 0.05$). (**b**) Expression profile of housekeeping gene glucuronidase beta (*GUSB*) in MJ90hTERT cell line after treatment with RNase H enzyme and in untreated controls. Three independent RT-qPCR experiments were performed. Error bars represent standard deviations, and averaged N_0 values are expressed in arbitrary fluorescence units. No significant differences in gene expression were observed between treated samples and controls in all cases (Student's *t*-test, $p > 0.1$).

3. Discussion

While transcription of satellite DNAs is tightly regulated under physiological conditions, upon specific conditions such as heat stress, it is significantly changed and satellite transcripts were proposed to be involved in gene expression modulation [20]. Human satellite III transcription is particularly strongly activated upon heat stress [21], and the transcripts mediate the recruitment of a number of RNA-binding proteins involved in pre-mRNA processing, participating in the control of gene expression at the level of splicing regulation [22,23]. In the case of the major TCAST1 satellite DNA of beetle *Tribolium castaneum*, increased levels of H3K9me2/3 are detected after heat stress at regions of (peri)centromeric heterochromatin and at dispersed satellite repeats and their flanking regions up to 2 kb from the insertion site, indicating that satellite transcripts can act in *trans*, targeting homologous regions in euchromatin. Increased levels of H3K9me2/3 at euchromatic satellite repeats correlate with transient suppression of neighboring genes and indicate the role of TCAST1 satellite siRNAs in the modulation of gene expression [10,12].

In the present study, we transiently transfected MJ90-hTERT fibroblasts with the plasmid expressing alpha satellite DNA monomer and followed the expression of alpha satellite and alpha-associated genes at 24, 48, 72 and 96 h post-transfection. Maximal exogenous expression of alpha satellite exceeding endogenous transcription for more than 10x was obtained 24 h after transfection, coinciding with the downregulation of all alpha-associated genes. The endogenous transcription of alpha satellite was not only significantly lower than the exogenous expression but also remained constant before and after transfection under all tested conditions, indicating that exogenous alpha satellite expression did not affect the endogenous expression regulation. Our present results showed that gene silencing by satellite RNA is not solely mediated by histone modifications but may also involve other forms of transcriptional interference. In humans, ChIP analyses did not reveal significant changes in histone marks typically associated with pericentromeric heterochromatin, such as H3K9me3 and H3K18Ac, nor in euchromatin-associated marks like H3K4me2 at the level of dispersed intronic alpha satellite repeats, although the H3K9me3 level was increased at tandemly arranged alpha repeats. This does not imply a fundamentally different mechanism between *Tribolium* and humans; rather, it suggests an additional silencing pathway alongside the one previously characterized in *Tribolium*. Although ChIP-seq is a powerful technique, conventional protocols often underperform when dealing with low-input material, as in the case of our experiments, where only about 30–35% of cells were efficiently transfected. To conclusively determine whether histone marks also contribute to gene silencing in humans, further experiments using higher-resolution methods are necessary. In particular, single-cell chromatin immunocleavage sequencing (scChIC-seq) may allow the detection of histone modification changes that remain undetectable using conventional bulk ChIP methods [24,25].

Our findings further add a layer of complexity to the mechanisms underlying gene expression downregulation by satellite DNA transcripts. Specifically, the transient expression of exogenous alpha satellite DNAs, produced by transcription from both DNA strands, exerts the same silencing effect on alpha-associated genes, independently of the strand of origin. Based on this observation and on the susceptibility of alpha satellite RNA to RNAse H treatment, we propose that alpha satellite RNA interacts directly with homologous DNA at dispersed intronic satellite loci by forming hybrid structures in the form of triple helices (RNA:DNA:DNA) or R-loops, thereby affecting the transcription of neighboring genes (Figure 5). A mechanism of gene expression modulation based on direct interaction between non-coding RNA and DNA was demonstrated [26]. Unlike classical Watson–Crick base pairing between the strands of the DNA double helix, RNA-DNA hybrids are characterized by Hoogsteen hydrogen bonding between nucleic acid bases. This type of interaction is

characterized by a weaker and more flexible association between DNA and RNA molecules, supporting the notion of a mechanism of transient interference based on alpha satellite sequence homology. This could also explain the observed differences in the level of gene suppression among the investigated alpha satellite repeat-associated genes (Figure 2). Given that alpha satellite DNA is polymorphic in nature, with potential monomer sequence variability of up to 45%, and the vector-expressed satellite RNA being a cloned sequence, it is plausible that the observed variability differences in gene repression can be attributed to different degrees of sequence homology. Additionally, the gene-modulatory capabilities of alpha satellite DNA do not depend on gene polarity since similar levels of gene down-regulation were observed using either vector, regardless of the direction of transcription (Figure 2). This would imply that the formation of this hybrid genomic structure plays a key role, either directly or by guiding RNA-associated regulatory proteins to specific genomic locations. The formation of triple helices and R-loops seems to be widespread and essential for the regulatory activity of many non-coding RNAs [27,28]. Although many R-loop-forming, long non-coding RNAs act in *cis*, R-loops can also form in *trans*, influencing the expression of protein-coding genes [29]. Triplex formation by long non-coding RNAs was also shown to be able to regulate gene expression in *trans* [28,30]. In mice, pericentromeric satellite DNA transcripts have been shown to form RNA:DNA hybrids that enable the retention of heterochromatin protein 1 (HP1) and the histone methyltransferases SUV39h1 and SUV39h2, which are necessary for heterochromatin formation [17,31]. The observed increase in H3K9me3 level on tandemly arranged alpha satellite arrays characteristic of heterochromatin, after transfection with alpha expression vector, speaks in favor of a potential similar mechanism of recruitment of chromatin modifiers by alpha satellite RNA:DNA hybrids (Figure 5). In order to test whether RNA Pol II accumulates or stalls at around intronic alpha loci, additional experiments using RNA Pol II ChIP should be performed.

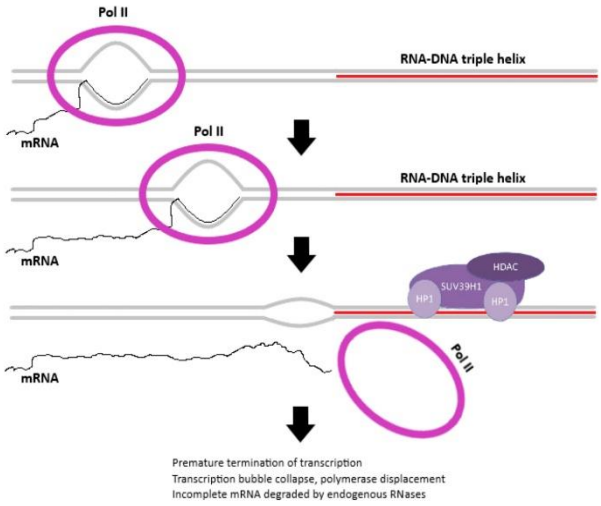


Figure 5. Proposed model of alpha satellite RNA gene-modulatory potential. Depicted is a part of an intronic region of an alpha satellite repeat-associated gene. Termination of transcriptional activity occurs at the boundary between transcriptional machinery and RNA-DNA triplex genomic structure. Pol II signifies eukaryotic RNA Polymerase II enzyme, mRNA is the nascent messenger RNA molecule, and the alpha satellite RNA sequence is highlighted in red. Potentially recruitable RNA-associated chromatin remodeling factors are also displayed (HP1—Heterochromatin Protein 1, SUV39H1—histone methyltransferase, HDAC—histone deacetylase).

In conclusion, our results demonstrate that alpha satellite RNA is not only involved in centromere and heterochromatin assembly, but, for the first time, reveal its role in the modulation of gene expression. Further studies, however, are necessary to reveal the detailed molecular mechanism of gene expression modulation mediated by alpha satellite RNA.

4. Materials and Methods

4.1. Human Cell Line

Human diploid fibroblast strain MJ90hTERT (HCA2hTERT) was a gift from Dr Olivia M. Pereira-Smith (University of Texas, San Antonio, TX, USA) [32–34]. Cells were cultured in appropriate medium (DMEM) supplemented with 10% FBS and 5% CO_2 at 37 °C.

4.2. Construction of Vectors

The alpha satellite 171 bp long monomer was amplified by PCR from MJ90-hTERT genomic DNA and cloned into the pCMV6-A-GFP plasmid vector (OriGene Technologies, Inc., Rockville, MD, USA). A map of the pCMV6-A-GFP plasmid vector is shown in Figure S9. Two separate constructs were created: the first containing the satellite in forward orientation (171Fw) and the second with the insert inverted (171Rev). Inserts were cloned into the vector's multiple cloning site using modified alpha satellite primers specific for the consensus sequence of 171 bp alpha satellite monomer [35] with inbuilt restriction site sequences. The forward-oriented insert was generated using forward primer (5′-TGCATTGGATCCCATTCTCAGAAACTTCTTTGTG-3′) and reverse primer (5′-TGCATTCTCGAGCTTCTGTCTAGTTTTTATGTGAAG-3′), while the inverted insert was obtained using forward primer (5′-GCATATCTCGAGCATTCTCAGAAACTTCTTTGTG-3′) and reverse primer (5′-TGGCTGGGATCCCTTCTGTCTAGTTTTTATGTGAAG-3′). The unmodified pCMV6-A-GFP vector was used as a negative control. Amplicon fidelity was tested by agarose gel electrophoresis before restriction and ligation into vectors. The JM109 bacterial strain (Promega Corporation, Madison, WI, USA) was subsequently transformed with the above-mentioned recombinant vectors; clones were selected on ampicillin-containing plates, screened by colony PCR, and plasmid vectors were isolated using GenElute HP Plasmid Midiprep Kit (Sigma-Aldrich, Burlington, MA, USA). The sequences of the cloned fragments were validated by Sanger DNA sequencing.

4.3. MJ90hTERT Cell Line Transfection

MJ90hTERT (immortalized human skin fibroblasts) cells were transfected with 171Fw, 171Rev and unmodified pCMV6-A-GFP control vector using Lipofectamine 3000 Transfection Reagent (Thermo Fisher Scientific, Waltham, MA, USA), according to the manufacturer's instructions. Afterwards, cells were incubated for 24, 48, 72 and 96 h at 37 °C in a complete medium.

4.4. RNA Isolation and Reverse Transcription

For RNA isolation from cell cultures, the RNeasy Plus Mini Kit (Qiagen, Venlo, Netherlands) was used at each specified time point, and the isolation and reverse transcription were performed using a previously published protocol [11,19]. The specifically modified primer for alpha satellite (Rev 5′-AATGCACATATCACTATGTAC-3′), designed to produce cDNA molecules that differ from genomic DNA in order to avoid DNA contamination, was used [36]. For all samples, negative controls without reverse transcriptase were used.

4.5. Quantitative Real-Time PCR (qPCR) Analyses

qPCR analyses were performed 24, 48, 72 and 96 h post-transfection, according to the previously published protocol [10]. Primers used for transcriptional analysis of endogenous alpha satellite RNA were constructed based on consensus sequence derived from cloned alpha satellite monomers of wide-ranging chromosomal origins [35], and the same modified primer used previously in reverse transcription (Rev 5′-AATGCACATATCACTATGTAC-3′) was used in qPCR amplification in order to avoid any potential DNA contamination along with the second primer (Fw 5′-CACTCTTTTTGTAGAATCTGC-3′). Exogenous 171 bp alpha satellite RNA derived from vectors in both orientations was quantified using forward (Fw 5′-CATTCTCAGAAACTTCTTTGTG-3′) or reverse (Rev 5′-CTTCTGTCTAGTTTTTATGTGAAG-3′) primers in combination with vector-specific T7 promoter forward primer (5′-TAATACGACTCACTATAGGG-3′). This strategy ensures the specific amplification of exogenous alpha RNA. Candidate genes of interest with intronic alpha satellite repeats were identified by bioinformatic analyses, as described in [3,19] and the primer combinations used for their expression analysis are listed in Table S3. Glucuronidase beta (*GUSB*) was used as an endogenous control for normalization in human samples as well as a negative control for the aforementioned genes. *GUSB* gene (Gene ID: 2990) was stably expressed at all time points without any variation among samples after transfection. Five additional housekeeping genes were used as negative controls: glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*, ID: 2597), DNA topoisomerase III alpha (*TOP3A*, ID: 7156), DEK proto-oncogene (*DEK*, ID: 7913), G protein-coupled receptor 176 (*GPR68*, ID: 8111) and interferon-induced protein with tetratricopeptide repeats 1 (IFIT3, ID: 3437). Sequences of primers of all reference genes are listed in Table S3. The thermal cycling conditions were described in [11,19]. Post-run data were analyzed using LinRegPCR software v.11.1. [37,38], which enables calculation of the starting concentration of amplicon ("N_0 value"). The N_0 value determined for each technical replicate was averaged, and the averaged N_0 values were divided by the N_0 values of the endogenous control. Statistical analysis of qPCR data was performed using GraphPad v.6.01, and the normalized N_0 values were compared using the unpaired *t*-test, which compares the means of two unmatched groups.

4.6. Chromatin Immunoprecipitation

Additionally, 24 h after transfection, MJ90hTERT cells were processed according to the published protocol [10,19], with the exception of the sonication step, which was performed 30 times for 30 s on ice, using a high-amplitude sonicator. The antibodies used were as follows: Anti-Histone H3 (tri methyl K9, ab8898, Abcam, Cambridge, UK), Anti-Histone H3 (acetyl K18, Abcam, ab1191), Anti-Histone H3 (di methyl K4, Abcam, ab7766) and IgG (sc2027, Santa Cruz Biotechnology, Inc., Dallas, TX, USA). Binding of the precipitated target was monitored by qPCR using the SYBR Green PCR Master mix (Bio-Rad Laboratories, Hercules, CA, USA). Primers used for H3K9me3, H3K18ac and H3K4me2 level analyses of genes with intronic alpha satellite repeats are listed in Table S2. The N_0 values were normalized using the N_0 values of the input fractions.

4.7. Alpha Satellite RNA:DNA Hybrid Detection Assay

Approximately 2.5×10^6 of MJ90hTERT cells per sample were washed in PBS, harvested in medium, transferred into tubes and centrifuged at room temperature for 5 min at 400 g. The pellets were resuspended in 600 µL of Lairds buffer (100 mM Tris pH8.5, 200 mM NaCl, 5 mM EDTA, 0.2% SDS). The samples were sonicated (45 s OFF, 15 s ON, 12 cycles) using a sonicator (high-amplitude), centrifuged for 5 min at 12,000 rpm (4 °C) and the supernatants transferred to new tubes. Total nucleic acids were purified with

phenol–chloroform–isoamyl alcohol (A156.3, Roth, Karlsruhe, Germany), precipitated with 2.5 volumes of cold ethanol (96%) and 0.1 volume of 3 M sodium acetate (pH 5.2), and washed with 70% ethanol. The air-dried pellets were resuspended in 50 µL of nuclease-free water. A total of 10 µg of chromatin-associated, phenol/chloroform-isolated nucleic acids per sample was incubated for 30 min at 37 °C with 13 U of RNase H (NEB) in 1× buffer (NEB) in a total volume of 50 µL. The untreated (without RNase H) controls (10 µg) were also incubated in the same way. RNase H-treated and untreated samples were then double-digested with DNase I (Rnase-Free Dnase Set, QIAGEN, Venlo, Netherlands), and total RNA was purified using RNeasy Mini Kit (QIAGEN), according to manufacturer instructions. All nucleic acid quantifications were carried out using a Qubit fluorometer (Invitrogen, Waltham, MA, USA).

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/ijms262211204/s1.

Author Contributions: Conceptualization, I.F. and Đ.U.; methodology, S.L. and M.M.; formal analysis, S.L., M.M., D.Đ., M.C.F., A.P. and I.F.; investigation, S.L., M.M., F.P. and I.F.; writing—original draft preparation, I.F. and Đ.U.; writing—review and editing, Đ.U. and I.F., D.Đ.; supervision, I.F.; All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Croatian Science Foundation grants: IP-2019-04-6915 to Đ. Ugarković and the Italian Ministry of Education, University and Research (MIUR), FFABR2017, fund for Investments on Basic Research (FIRB) by the International Staff Mobility Program of University of Naples Federico II to I. Feliciello.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article/supplementary material; further inquiries can be directed to the corresponding author(s).

Acknowledgments: We thank Marianna Feliciello for the graphical support.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Lee, C.; Wevrick, R.; Fisher, R.B.; Ferguson-Smith, M.A.; Lin, C.C. Human centromeric DNAs. *Hum. Genet.* **1997**, *100*, 291–304. [CrossRef]
2. McNulty, S.M.; Sullivan, B.A. Alpha satellite DNA biology: Finding function in the recesses of the genome. *Chromosom. Res.* **2018**, *26*, 115–138. [CrossRef] [PubMed]
3. Feliciello, I.; Pezer, Z.; Kordiš, D.; Bruvo-Mađarić, B.; Ugarković, Đ. Evolutionary history of α satellite DNA repeats dispersed within human genome euchromatin. *Genome Biol. Evol.* **2020**, *12*, 2125–2138. [CrossRef]
4. Feliciello, I.; Picariello, O.; Chinali, G. The first characterisation of the overall variability of repetitive units in a species reveals unexpected features of satellite DNA. *Gene* **2005**, *349*, 153–164. [CrossRef]
5. Feliciello, I.; Picariello, O.; Chinali, G. Intra-specific variability and unusual organization of the repetitive units in a satellite DNA from *Rana dalmatina*: Molecular evidence of a new mechanism of DNA repair acting on satellite DNA. *Gene* **2006**, *383*, 81–92. [CrossRef]
6. Brajković, J.; Feliciello, I.; Bruvo-Mađarić, B.; Ugarković, Đ. Satellite DNA-Like Elements Associated with Genes Within Euchromatin of the Beetle *Tribolium castaneum*. *G3 Genes/Genomes/Genet.* **2012**, *2*, 931–941. [CrossRef]
7. Feliciello, I.; Akrap, I.; Brajković, J.; Zlatar, I.; Ugarković, Đ. Satellite DNA as a driver of population divergence in the red flour beetle *Tribolium castaneum*. *Genome Biol. Evol.* **2015**, *7*, 228–239. [CrossRef]
8. Kuhn, G.C.; Küttler, H.; Moreira-Filho, O.; Heslop-Harrison, J.S. The 1.688 Repetitive DNA of Drosophila: Concerted Evolution at Different Genomic Scales and Association with Genes. *Mol. Biol. Evol.* **2011**, *29*, 7–11. [CrossRef]
9. Ruiz-Ruano, F.J.; López-León, M.D.; Cabrero, J.; Camacho, J.P.M. High-throughput analysis of the satellitome illuminates satellite DNA evolution. *Sci. Rep.* **2016**, *6*, 28333. [CrossRef]

10. Feliciello, I.; Akrap, I.; Ugarković, Đ. Satellite DNA Modulates Gene Expression in the Beetle *Tribolium castaneum* after Heat Stress. *PLoS Genet.* **2015**, *11*, e1005466. Erratum in *PLoS Genet.* **2015**, *11*, e1005547.
11. Ljubić, S.; Matulić, M.; Đermić, D.; Feliciello, M.C.; Procino, A.; Ugarković, Đ.; Feliciello, I. Antibiotics induce overexpression of alpha satellite DNA accompanied with epigenetic changes at alpha satellite arrays as well as genome-wide. *Epigenet. Chromatin* **2025**, *18*, 62. [CrossRef]
12. Sermek, A.; Feliciello, I.; Ugarković, Đ. Distinct Regulation of the Expression of Satellite DNAs in the Beetle *Tribolium castaneum*. *Int. J. Mol. Sci.* **2021**, *22*, 296. [CrossRef] [PubMed]
13. Bury, L.; Moodie, B.; Ly, J.; McKay, L.S.; Miga, K.H.; Cheeseman, I.M. α-satellite RNA transcripts are repressed by centromere-nucleolus associations. *eLife* **2020**, *9*, e59770. [CrossRef] [PubMed]
14. Teng, Z.; Yang, L.; Zhang, Q.; Chen, Y.; Wang, X.; Zheng, Y.; Tian, A.; Tian, D.; Lin, Z.; Deng, W.M.; et al. Topoisomerase I is an evolutionarily conserved key regulator for satellite DNA transcription. *Nat. Commun.* **2024**, *15*, 5151. [CrossRef] [PubMed]
15. Huang, C.; Shu, X.; Zhou, S.; Mi, Y.; Bian, H.; Li, T.; Li, T.; Ying, X.; Cheng, C.; Liu, D.; et al. Nuclear m6A modification regulates satellite transcription and chromosome segregation. *Nat. Chem. Biol.* **2025**. *online ahead of print*. [CrossRef]
16. Johnson, W.L.; Yewdell, W.T.; Bell, J.; McNulty, S.M.; Duda, Z.; O'Neill, R.J.; Sullivan, B.A.; Straight, A. RNA-dependent stabilization of SUV39H1 at constitutive heterochromatin. *eLife* **2017**, *6*, e25299. [CrossRef]
17. Velazquez Camacho, O.; Galan, C.; Swist-Rosowska, K.; Ching, R.; Gamalinda, M.; Karabiber, F.; De La Rosa-Velazquez, I.; Engist, B.; Koschorz, B.; Shukeir, N.; et al. Major satellite repeat RNA stabilize heterochromatin retention of Suv39h enzymes by RNA-nucleosome association and RNA:DNA hybrid formation. *eLife* **2017**, *6*, 817. [CrossRef]
18. Eymery, A.; Horard, B.; el Atifi-Borel, M.; Fourel, G.; Berger, F.; Vitte, A.-L.; Broeck, A.V.D.; Brambilla, E.; Fournier, A.; Callanan, M.; et al. A transcriptomic analysis of human centromeric and pericentric sequences in normal and tumor cells. *Nucleic Acids Res.* **2009**, *37*, 6340–6354. [CrossRef]
19. Feliciello, I.; Sermek, A.; Pezer, Ž.; Matulić, M.; Ugarković, Đ. Heat stress affects H3K9me3 level at human α satellite DNA repeats. *Genes* **2020**, *11*, 663. [CrossRef]
20. Ugarković, Đ.; Sermek, A.; Ljubić, S.; Feliciello, I. Satellite DNAs in Health and Disease. *Genes* **2022**, *13*, 1154. [CrossRef]
21. Jolly, C.; Metz, A.; Govin, J.; Vigneron, M.; Turner, B.M.; Khochbin, S.; Vourc, H.C. Stress-induced transcription of satellite III repeats. *J. Cell Biol.* **2003**, *164*, 25–33. [CrossRef] [PubMed]
22. Ninomiya, K.; Adachi, S.; Natsume, T.; Iwakiri, J.; Terai, G.; Asai, K.; Hirose, T. LncRNA-dependent nuclear stress bodies promote intron retention through SR protein phosphorylation. *EMBO J.* **2019**, *39*, e102729. [CrossRef]
23. Ninomiya, K.; Iwakiri, J.; Aly, M.K.; Sakaguchi, Y.; Adachi, S.; Natsume, T.; Terai, G.; Asai, K.; Suzuki, T.; Hirose, T. m^6A modification of HSATIII lncRNAs regulates temperature-dependent splicing. *EMBO J.* **2021**, *40*, e107976. [CrossRef]
24. Ku, W.L.; Nakamura, K.; Gao, W.; Cui, K.; Hu, G.; Tang, Q.; Ni, B.; Zhao, K. Single-cell chromatin immunocleavage sequencing (scChIC-seq) to profile histone modification. *Nat. Methods* **2019**, *16*, 323–325. [CrossRef]
25. Ku, W.L.; Pan, L.; Cao, Y.; Gao, W.; Zhao, K. Profiling single-cell histone modifications using indexing chromatin immunocleavage sequencing. *Genome Res.* **2021**, *31*, 1831–1842. [CrossRef] [PubMed]
26. Statello, L.; Guo, C.J.; Chen, L.L.; Huarte, M. Gene regulation by long non-coding RNAs and its biological functions. *Nat. Rev. Mol. Cell Biol.* **2021**, *22*, 96–118. Erratum in *Nat. Rev. Mol. Cell Biol.* **2021**, *22*, 159. [CrossRef] [PubMed]
27. Li, Y.; Syed, J.; Sugiyama, H. RNA-DNA Triplex Formation by Long Noncoding RNAs. *Cell Chem. Biol.* **2016**, *23*, 1325–1333. [CrossRef]
28. Warwick, T.; Brandes, R.P.; Leisegang, M.S. Computational Methods to Study DNA:DNA:RNA Triplex Formation by lncRNAs. *Non-Coding RNA* **2023**, *9*, 10. [CrossRef]
29. Ariel, F.; Lucero, L.; Christ, A.; Mammarella, M.F.; Jegu, T.; Veluchamy, A.; Mariappan, K.; Latrasse, D.; Blein, T.; Liu, C.; et al. R-Loop Mediated trans Action of the APOLO Long Noncoding RNA. *Mol. Cell* **2020**, *77*, 1055–1065.e4. [CrossRef]
30. Mondal, T.; Subhash, S.; Vaid, R.; Enroth, S.; Uday, S.; Reinius, B.; Mitra, S.; Mohammed, A.; James, A.R.; Hoberg, E.; et al. MEG3 long noncoding RNA regulates the TGF-β pathway genes through formation of RNA-DNA triplex structures. *Nat. Commun.* **2015**, *6*, 7743. Erratum in *Nat. Commun.* **2019**, *10*, 5290. [CrossRef]
31. Duda, K.J.; Ching, R.W.; Jerabek, L.; Shukeir, N.; Erikson, G.; Engist, B.; Onishi-Seebacher, M.; Perrera, V.; Richter, F.; Mittler, G.; et al. m6A RNA methylation of major satellite repeat transcripts facilitates chromatin association and RNA:DNA hybrid formation in mouse heterochromatin. *Nucleic Acids Res.* **2021**, *49*, 5568–5587. [CrossRef]
32. Cukusic Kalajzic, A.; Vidacek, N.S.; Huzak, M.; Ivankovic, M.; Rubelj, I. Telomere Q-PNA-FISH--reliable results from stochastic signals. *PLoS ONE* **2014**, *9*, e92559. [CrossRef] [PubMed]
33. Gorbunova, V.; Seluanov, A.; Pereira-Smith, O.M. Expression of human telomerase (hTERT) does not prevent stress-induced senescence in normal human fibroblasts but protects the cells from stress-induced apoptosis and necrosis. *J. Biol. Chem.* **2002**, *277*, 38540–38549. [CrossRef]
34. Young, J.I.; Smith, J.R. DNA Methyltransferase Inhibition in Normal Human Fibroblasts Induces a p21-dependent Cell Cycle Withdrawal. *J. Biol. Chem.* **2001**, *276*, 19610–19616. [CrossRef]

35. Choo, K.H.; Vissel, B.; Nagy, A.; Earle, E.; Kalitsis, P. A survey of the genomic distribution of alpha satellite DNA on all the human chromosomes, and derivation of a new consensus sequence. *Nucleic Acids Res.* **1991**, *19*, 1179–1182. [CrossRef] [PubMed]
36. Đermić, D.; Ljubić, S.; Matulić, M.; Procino, A.; Feliciello, M.C.; Ugarković, Đ.; Feliciello, I. Reverse transcription-quantitative PCR (RT-qPCR) without the need for prior removal of DNA. *Sci. Rep.* **2023**, *13*, 11470. [CrossRef] [PubMed]
37. Ruijter, J.M.; Ramakers, C.; Hoogaars, W.M.H.; Karlen, Y.; Bakker, O.; Hoff, M.J.B.V.D.; Moorman, A.F.M. Amplification efficiency: Linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Res.* **2009**, *37*, e45. [CrossRef]
38. Ruijter, J.M.; Pfaffl, M.W.; Zhao, S.; Spiess, A.N.; Boggy, G.; Blom, J.; Rutledge, R.G.; Sisti, D.; Lievens, A.; de Preter, K.; et al. Evaluation of qPCR curve analysis methods for reliable biomarker discovery: Bias, resolution, precision, and implications. *Methods* **2013**, *59*, 32–46. [CrossRef]

Table S1. List of alpha satellite repeats dispersed on human chromosomes. Genomic positions of repeats (genome assembly GRCh38/hg38), their composition and similarity to consensus alpha satellite sequence is indicated as well as genes associated with each repeat.

chr No	alpha repeat No	Genome position	Similarity to alpha satellite consensus/size	Associated gene
2	1	32213935-32214051	82%, 0.7 monomer	solute carrier family 30 member 6, (***SLC30A6*** ID: 55676), **intron**
6	14	57377750-57378048	70%, 1.7 monomer	DNA primase subunit 2 (***PRIM2*** ID: 5558) **intron**
10	18	17653288-17653378	76%, 0.5 monomer	signal transducing adapter molecule 1; (***STAM*** ID: 8027), **intron**
11	21	85431074-85431184	85%, 0.7 monomer	discs large homolog 2, (***DLG2*** ID: 1740), **intron**
15	25	59219764-59219973	88%, 1.2 monomers	myosin IE (***MYO1E***, ID: 4643), **intron**
19	28	6819160-6819369	70%, 1.2 monomers	vav guanine nucleotide exchange factor 1, (***VAV1***, ID: 7409), **intron**
19	29	23661142-23661242	86%, 0.6 monomer	zinc finger protein 675, (***ZNF675***, ID: 171392), **intron**
X	31	20101282-20101513	71%, 1.3 monomers	MAP7 domain containing 2 (***MAP7***, ID: 256714), **intron**

Table S2. List of primers used for histone modification levels analyses of alpha repeats located within introns of genes, in ChIP-qPCR experiments.

Gene	Primers Fw	Primers Rev
SLC30A6	GCCTCCTGAGTTCAAGCAAC	GCATGGTGCCTCATTCCTAT
STAM	TCCCAGTCCATCGAAACCTA	GAAGCTTCATCACCCTCCAA
MYO1E	CGACATGGGTCCAGTCTGAT	AGGAATCTGGATATGTCTTCCA
MAP7	CGCTAATGCTGAAGACATGC	GAAGTGGAATGGGATCTGAAA
ZNF675	GCTTACCTTGGCTTCTCAAAGT	TAGTAGACACCGGGTTTCACC
VAV1	AGACTCCATCCCCCTCAAAA	TCAAGGTGTCAGTAGGGTTGG
PRIM2	GTTCCTGTAAGGGCTCAACG	AGGCTGCAGTAAGCCATGAT
DLG2	GCACCACACTGGTCTTCC	CTGTATTTCAGCATGAGTGACAG

Table S3. List of primers used for expression analysis of alpha repeat-associated genes and reference genes.

Gene	Primers Fw	Primers Rev
SLC30A6	TGATCTTGCTGGAGCATTTG	AAACATGGGGTGGTGTTGTCT
STAM	AACAAAGGCAGCAGTCAACC	TTGATGGGTTTCACCTTTCC
MYO1E	CTGGGAGGAAAGCAGGGTAA	ACACTTTACTCCTCCCCAGC
MAP7	TTCCTGTTGTGAACTTCGGG	CCTTCCCTTTCCTTGTTGCT
ZNF675	ACACTGCACAGCGGAATTTA	GGGGTTCATTCACCATCTCA
VAV1	TGCTTCAAGTCTCTGGACACCAC	TCTCGGGCGCAGAAGTCATA
PRIM2	TGGACTTAAGTTGGGGTTCG	CAAAGCCTTGGACAGTTTGG
DLG2	TTGCATGTTACTGTGCACTCC	CAGAGGAGAAATATGAGACTGCAA
GUSB	GAAAATACGTGGTTGGAGAGCTCATT	CCGAGTGAAGATCCCCTTTTTA
GAPDH	CCACTCCTCCACCTTTGAC	ACCCTGTTGCTGTAGCCA
TOP3A	TCGACTCTTTAACCACACGG	AGATCTGACCTCTACCACAG
DEK	ATGTGGGTCAGTTCAGTGGC	CCAGAAGGCTTTGGATGCAT
GPR68	CTTCCTCTTCCCCATCTGCC	GCAGGAGGGAGAAGTGGTAG
IFIT3	AGGTTCTCTTGGGCCTGAAA	CCTTGTAGCAGCACTCAATC

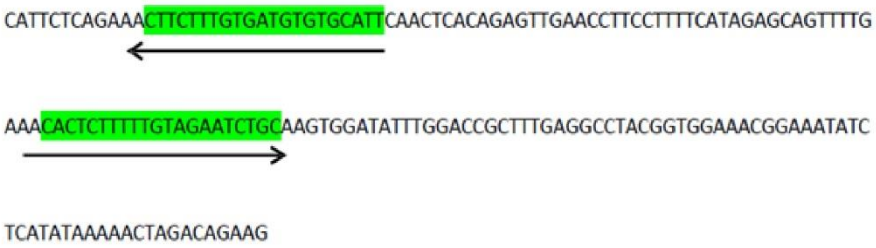


Figure S1. Consensus sequence of 171 bp alpha satellite monomer (Choo et al. 1991) and the annealing positions of primers used in qPCR.

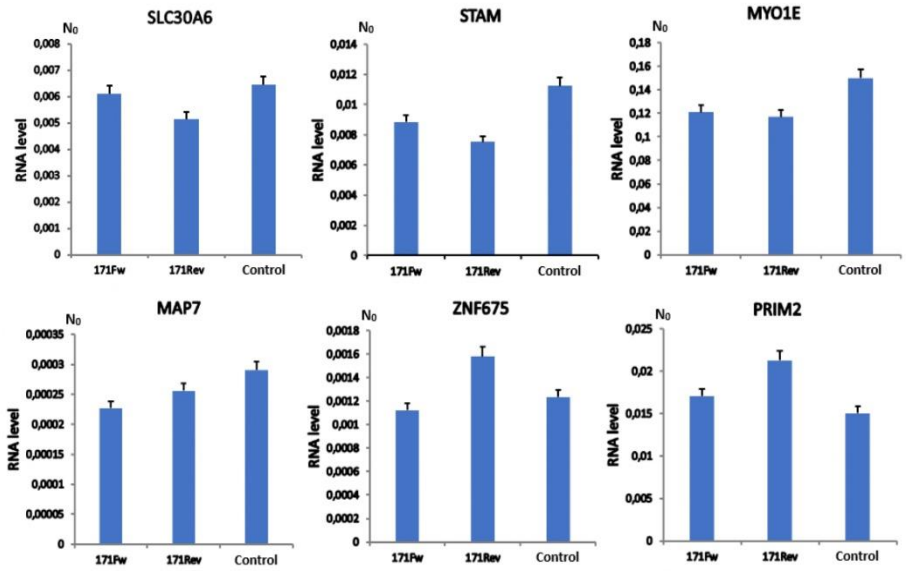


Figure S2. Expression profiles of genes containing alpha satellite repeats within intronic regions, in MJ90hTERT cell lines transfected with satellite expressing vectors and controls, 48 hours after treatment. 171Fw denotes the vector with satellite insert in forward orientation and 171Rev its inverted counterpart. Control refers to unaltered pCMV6-A-GFP vector. Error bars represent standard deviations. N_0 represents normalized average N_0 value expressed in arbitrary fluorescence units. No significant differences in gene expression of candidate genes between transfected samples and controls were observed (Student's t-test, P>0.05 in all cases).

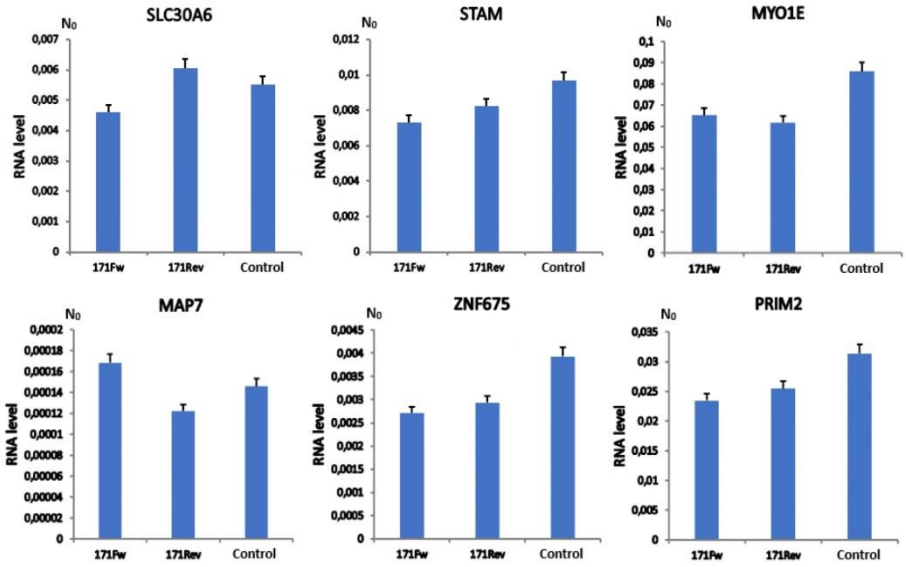


Figure S3. Expression profiles of genes containing alpha satellite repeats within intronic regions, in MJ90hTERT cell lines transfected with satellite expressing vectors and controls, 72 hours after treatment. 171Fw denotes the vector with satellite insert in forward orientation and 171Rev its inverted counterpart. Control refers to unaltered pCMV6-A-GFP vector. Error bars represent standard deviations. N_0 represents normalized average N_0 value expressed in arbitrary fluorescence units. No significant differences in gene expression of candidate genes between transfected samples and controls were observed (Student's t-test, P>0.05 in all cases).

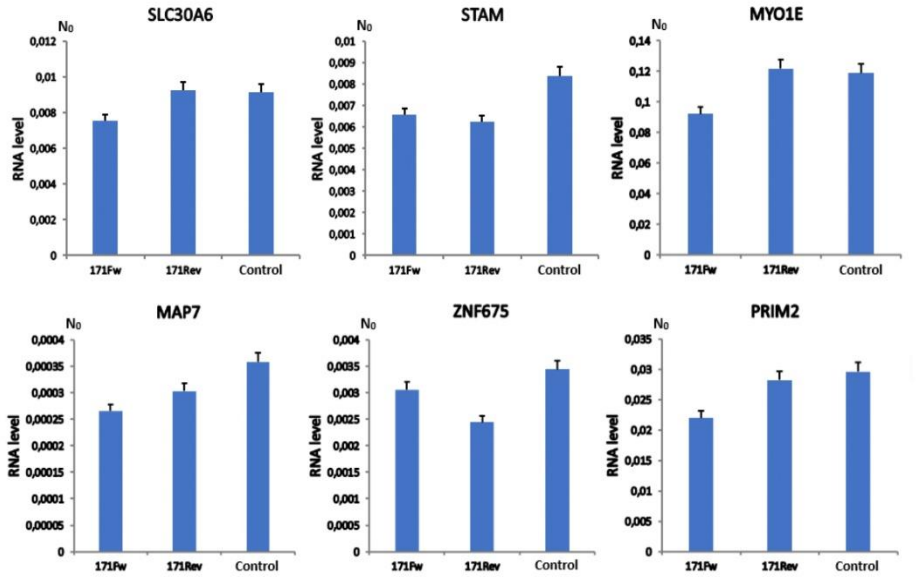


Figure S4. Expression profiles of genes containing alpha satellite repeats within intronic regions, in MJ90hTERT cell lines transfected with satellite expressing vectors and controls, 96 hours after treatment. 171Fw denotes the vector with satellite insert in forward orientation and 171Rev its inverted counterpart. Control refers to unaltered pCMV6-A-GFP vector. Error bars represent standard deviations. N_0 represents normalized average N_0 value expressed in arbitrary fluorescence units. No significant differences in gene expression of candidate genes between transfected samples and controls were observed (Student's t-test, P>0.05 in all cases).

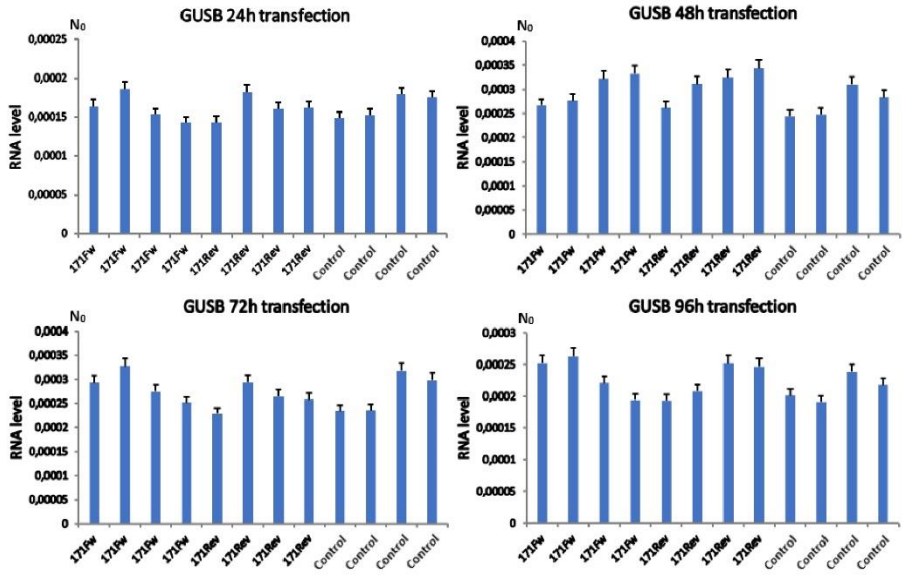


Figure S5. Glucuronidase beta gene expression profile in MJ90hTERT cell line 24, 48, 72 and 96 hours after transfection. Each experiments has been repeated four time. 171Fw denotes the vector with satellite insert in forward orientation and 171Rev its inverted counterpart. Control refers to unaltered pCMV6-A-GFP vector. Error bars represent standard deviations. N_0 values are expressed in arbitrary fluorescence units.

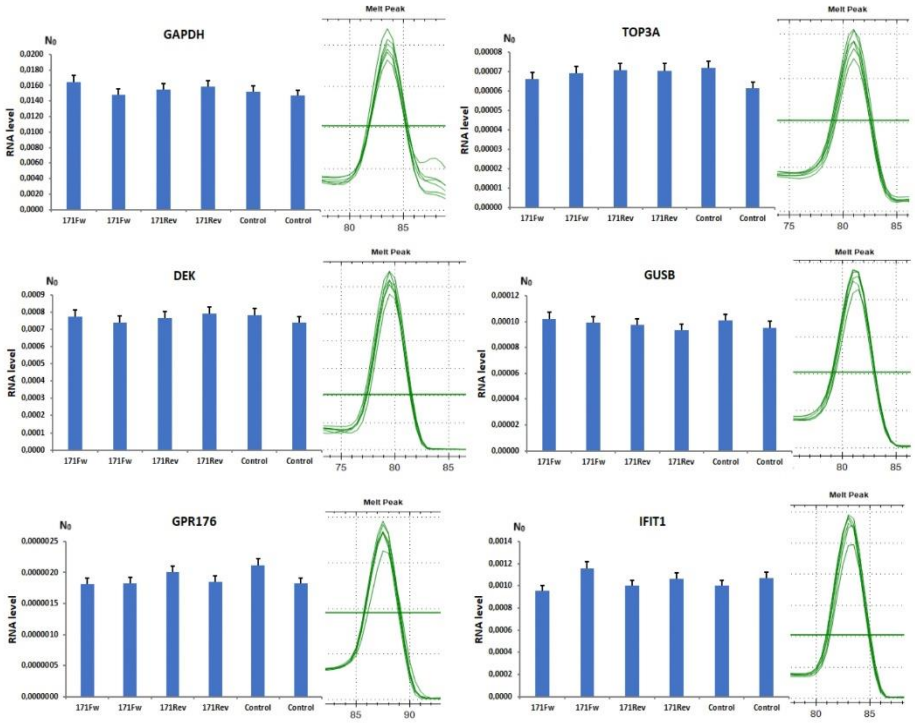


Figure S6. Expression profiles of six housekeeping genes in MJ90hTERT cell line after 24 hr transfection with alpha satellite-expressing vectors (171Fw, 171Rev) and unmodified control vector with corresponding melting curves. Two independent RT-qPCR experiments were performed and averaged values are displayed. Error bars represent standard deviations and averaged N_0 values are expressed in arbitrary fluorescence units. No significant differences in gene expression were observed between transfected samples in all cases (Student's t-test, P>0.1).

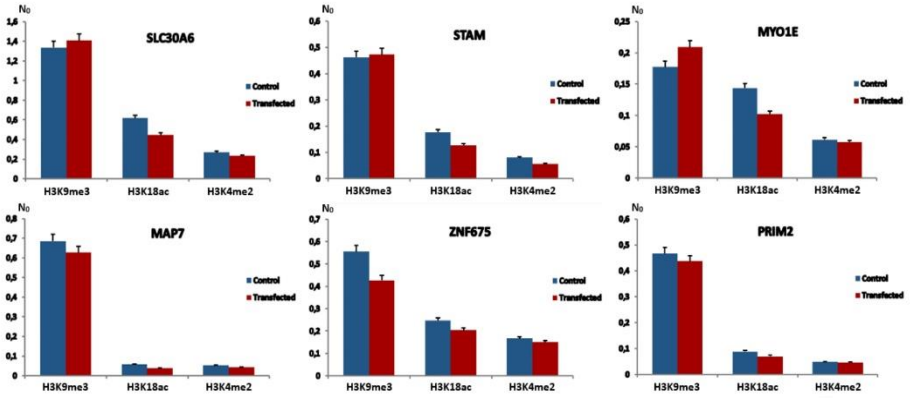


Figure S7. Levels of H3K9me3, H3K18ac and H3K4me2 histone modifications at alpha satellite repeats associated with six genes after 24 hour transfection with 171Rev and unmodified control vectors. Levels of histone modifications were measured by ChIP coupled with quantitative real-time PCR on MJ90hTERT chromatin immediately after each treatment. N_0 values were normalized using N_0 values of input fractions and represent the levels of histone modifications. Columns show averages of two independent experiments and error bars indicate standard deviations. No significant differences with regard to tested histone modifications at alpha repeats in six genes of interest compared to control samples were observed (P>0.1 in all cases, Student's t-test).

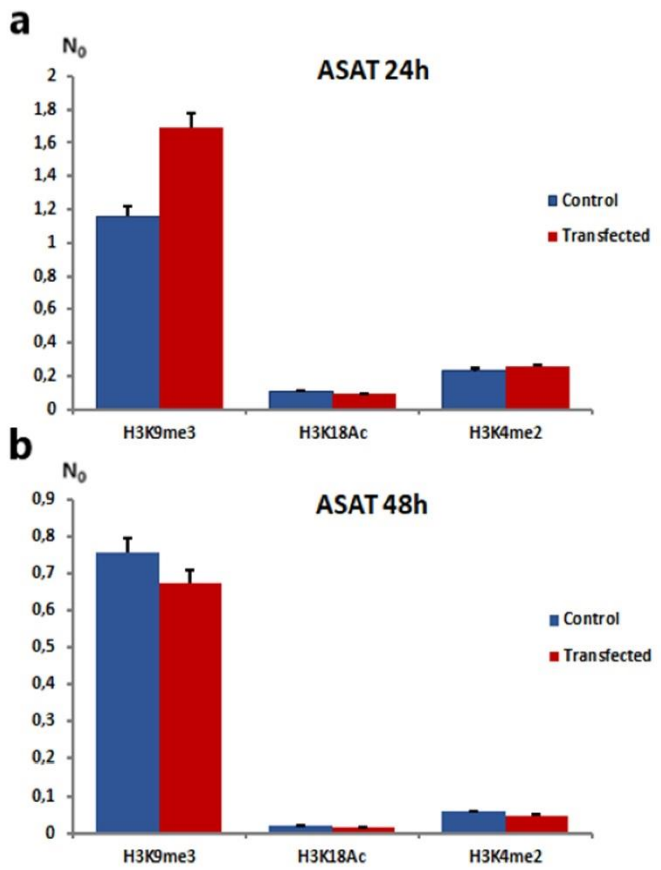


Figure S8. Levels of H3K9me3, H3K18ac and H3K4me2 histone modifications at tandemly arranged alpha satellite repeats after **(a)** 24 hour and **(b)** 48 hour transfection with 171Fw and unmodified control vectors. Levels of histone modifications were measured by ChIP coupled with quantitative real-time PCR on MJ90hTERT chromatin immediately after each treatment. N_0 values were normalized using N_0 values of input fractions and represent the levels of histone modifications. Columns show averages of two independent experiments and error bars indicate standard deviations. H3K9me3 histone modification was upregulated by ≈ 1.5x fold after 24 hour transfection with 171Fw vector compared to control samples ($P<0.05$, Student's t-test). After 48 hour transfection no statistically significant change in H3K9me3 is detected.

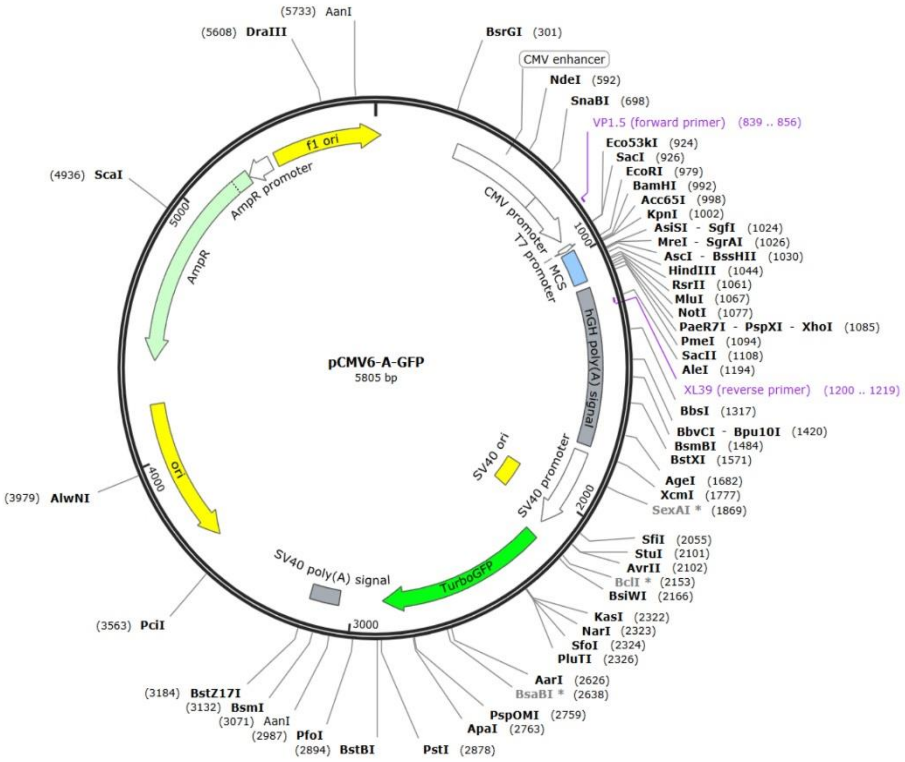


Figure S9. Map of the pCMV6-A-GFP plasmid vector (OriGene) used for transfection. Restriction sites *BamH*I and *Xho*I were selected for ASAT insertion, maintaining the approximately same size of modified and control vectors.

PUBLICATION II.

Ljubić S, Matulić M, Đermić D, Feliciello MC, Procino A, Ugarković Đ, Feliciello I (2025) Antibiotics induce overexpression of alpha satellite DNA accompanied with epigenetic changes at alpha satellite arrays as well as genome-wide. Epigenetics & Chromatin. 18:62.

Ljubić *et al. Epigenetics & Chromatin* (2025) 18:62
https://doi.org/10.1186/s13072-025-00628-z

Epigenetics & Chromatin

RESEARCH Open Access

Antibiotics induce overexpression of alpha satellite DNA accompanied with epigenetic changes at alpha satellite arrays as well as genome-wide

Sven Ljubić[1], Maja Matulić[2], Damir Đermić[1], Maria Chiara Feliciello[3], Alfredo Procino[4], Đurđica Ugarković[1] and Isidoro Feliciello[1,4*]

Abstract

The transcription of satellite DNA is highly sensitive to environmental factors and represents a source of genomic instability. Therefore, tight regulation of (peri)centromeric transcription is essential for genome maintenance. Antibiotics are routinely used for in vitro studies and for medical treatment, however, their effect on pericentromeric satellite DNA transcription was not investigated. Here we show that antibiotics geneticin and hygromycin B, conveniently used in cell culture, as well as rifampicin (along with five other antibiotics), used to treat bacterial infections, increase transcription of a major human pericentromeric alpha satellite DNA in cell lines at standard concentrations. However, response differs among cell lines - maximal increase in A-1235 cells is obtained by rifampicin while in HeLa cells and fibroblasts by geneticin. There is also a positive correlation between antibiotic concentration and the level of alpha satellite transcription. The increase of transcription is accompanied with either H3K9me3 decrease or H3K18ac increase at tandemly arranged alpha satellite arrays while H3K4me2 remains unchanged. Our results suggest that induced alpha satellite DNA transcription upon antibiotic stress could be linked to epigenetic changes - histone modifications H3K9me3 and H3K18ac, which are associated with transcription of heterochromatin.

Keywords Satellite DNA, Heterochromatin, Transcription, Antibiotic, Histone modifications

*Correspondence:
Isidoro Feliciello
isidoro.feliciello@unina.it
[1]Division of Molecular Biology, Ruder Bošković Institute, Bijenička 54, Zagreb HR-10000, Croatia
[2]Department of Biology, Faculty of Science, University of Zagreb, Zagreb HR-10000, Croatia
[3]Department of Statistical Science, Alma Mater Studiorum, University of Bologna, Bologna 40126, Italy
[4]Department of Clinical Medicine and Surgery, University of Naples Federico II, Naples I- 80131, Italy

Introduction

Antibiotics are routinely used for in vitro studies while culturing cells in order to avoid bacterial contamination and for selection purposes. However, different studies have shown that use of antibiotics can affect gene expression and could modify the results of studies focused on drug response, cell cycle regulation and cell differentiation [1–3]. Antibiotic rifampicin (10 μM for 24 h) was shown to induce genome-wide drug dependent changes in gene regulation and expression in human hepatocytes, some of them linked to changes in histone marks H3K4me1 and H3K27ac [1]. Cells cultured with standard 1% Penicillin-Streptomycin (PenStrep) supplemented media showed significantly altered gene expression and regulation, as observed in a common liver cell line such as HepG2. Drug-associated genes were differentially expressed following PenStrep treatment and differential enrichment of active promoter and enhancer regions marked by H3K27ac was reported [2]. The human peripheral blood mononuclear cells (PBMCs) expressed DNA damage features such as activation of a serine/threonine kinase ATM and p53, as well as epigenetic changes - phosphorylation of H2AX and H3K4me2/3 modifications at some promoter sites after the in vitro exposure to antibiotic oxytetracycline (OTC, 2 μg/ml or 4,3 μM) [4]. Since OTC is largely employed in veterinary practices, this reveals a potential influence of OTC on animal and human health. While the effect of antibiotic treatment on gene expression was previously at least partially characterized, the influence of antibiotics on non-coding regions of genome, in particular on (peri)centromeric satellite DNAs which are related to genome stability, is poorly investigated. Therefore, the molecular consequences of growing human cell lines with antibiotics at standard cell culture concentrations as well as of antibiotics use in veterinary and medical practice have yet to be thoroughly investigated.

Satellite DNAs are tandemly repeated sequences preferentially clustered in (peri)centromeric regions of eukaryotic chromosomes [5]. Recent studies reveal that their expression is highly sensitive to environmental factors such as heat stress, DNA damaging agents, genotoxic and hyperosmotic stress [6–11]. Satellite DNA expression is also significantly increased under different pathological conditions, such as in diverse types of cancer [5, 12, 13]. Transcription of satellite DNA may represent a source of genomic instability through collision between replication and transcription forks, formation of secondary structures and cytotoxic DNA–RNA hybrids known as R-loops [14]. Therefore, tight regulation of centromeric and pericentromeric transcription is essential for the maintenance of genome stability [15] and cells with aberrant satellite DNA expression can feature substantial mitotic defects and large-scale genetic aberrations, including chromosomal instability and aneuploidy [16]. Satellite DNAs located within heterochromatin seem to be at least partially under epigenetic control and their arrays in cancer cells are characterized by lower level of repressive heterochromatic histone modification H3K9me3 as well as by global DNA hypomethylation relative to normal cells [17, 18]. On the other hand, heat stress induces the increase of silent histone mark H3K9me3 at (peri)centromeric satellite repeats as well as at the satellite repeats dispersed within euchromatin [9, 19], resulting in downregulation of expression of nearby genes [19].

Here, we analyzed whether antibiotics such as geneticin and hygromycin B, which are conveniently used in cell culture [20], as well as rifampicin which is used to treat several types of bacterial infections [21], affect expression of a major, most abundant human alpha satellite DNA clustered in (peri)centromeric regions of all human chromosomes [22]. We also studied if potential changes in satellite DNA expression under antibiotic stress were accompanied by epigenetic changes such as histone marks at heterochromatic satellite arrays and at the satellite repeats dispersed within euchromatin, as well as genome-wide. Using different cell types, we show that all three antibiotics induce overexpression of alpha satellite DNA at concentrations routinely used for in vitro studies and for medical treatment. In addition, overexpression is accompanied by changes in epigenetic modifications on histone marks at alpha satellite arrays located in heterochromatin as well as genome-wide. We proposed that epigenetic changes in heterochromatin, such as decrease of silent histone modification H3K9me3 whose loss affects satellite DNA expression [23], and increase of H3K18ac, which is characteristic for transcriptional activation of heterochromatin [16], could be linked to induced alpha satellite DNA expression upon antibiotic stress.

Results

Alpha satellite DNA transcription after antibiotic treatment

To investigate whether antibiotics affect the transcription of alpha satellite DNA we followed its transcription dynamics in human cell lines by RT-qPCR under standard conditions and after antibiotic treatment. Primers used for transcriptional analysis were able to amplify only tandemly arranged repeats (Figure S1) and since in human pericentromeric heterochromatin alpha satellite DNA is organized in tandemly arranged monomers [22], it is expected that the primers preferentially recognize transcripts deriving from pericentromeric regions. Cells were incubated for 48 h at 37 °C in complete medium with concentrations of antibiotics used for routine treatment, selection and maintenance of eukaryotic cells: geneticin 300–600 μg/ml, hygromycin B 50–100 μg/ml as

well as with rifampicin 8.2–82 µg/ml (10–100 µM). The transcription of alpha satellite DNA was checked immediately after antibiotic treatment and compared with a control. The transcription of alpha satellite DNA was monitored in immortalized fibroblasts (MJ90hTERT), glioblastoma cell line A-1235 and cervix carcinoma HeLa cells.

Treatment of A1235 cells with hygromycin B (50 µg/ml), geneticin (400 µg/ml) and rifampicin (82 µg/ml) revealed the increase of transcription by 1.6x ($P = 0.01$), 1.7x ($P = 0.008$), and 3.0x ($P = 0.02$), respectively (Fig. 1b). To test if transcription responds to antibiotic concentration, A-1235 cells were treated with three different concentrations of rifampicin (8.2, 41, and 82 µg/ml). The results showed no effect after 8.2 µg/ml treatment while 41 µg/ml and 82 µg/ml induced increase of 1.8x ($P = 0.009$) and 3.0x ($P = 0.02$), respectively (Fig. 1c), suggesting positive correlation between transcription and antibiotic concentrations.

Hela cells treated with hygromycin B (50 µg/ml), geneticin (400 µg/ml) and rifampicin (82 µg/ml) showed the increase of transcription by 3.1x ($P = 0.02$), 4.9x ($P = 0.01$), and 1.5x ($P = 0.01$), respectively (Fig. 1a), while decreased concentration of geneticin (300 µg/ml) showed no effect on alpha satellite transcription (Fig. 1a).

Immortalized fibroblast MJ90hTERT cell line showed a modest increase of transcription of 1.5x ($P = 0.01$) after geneticin (400 µg/ml) while treatment with hygromycin

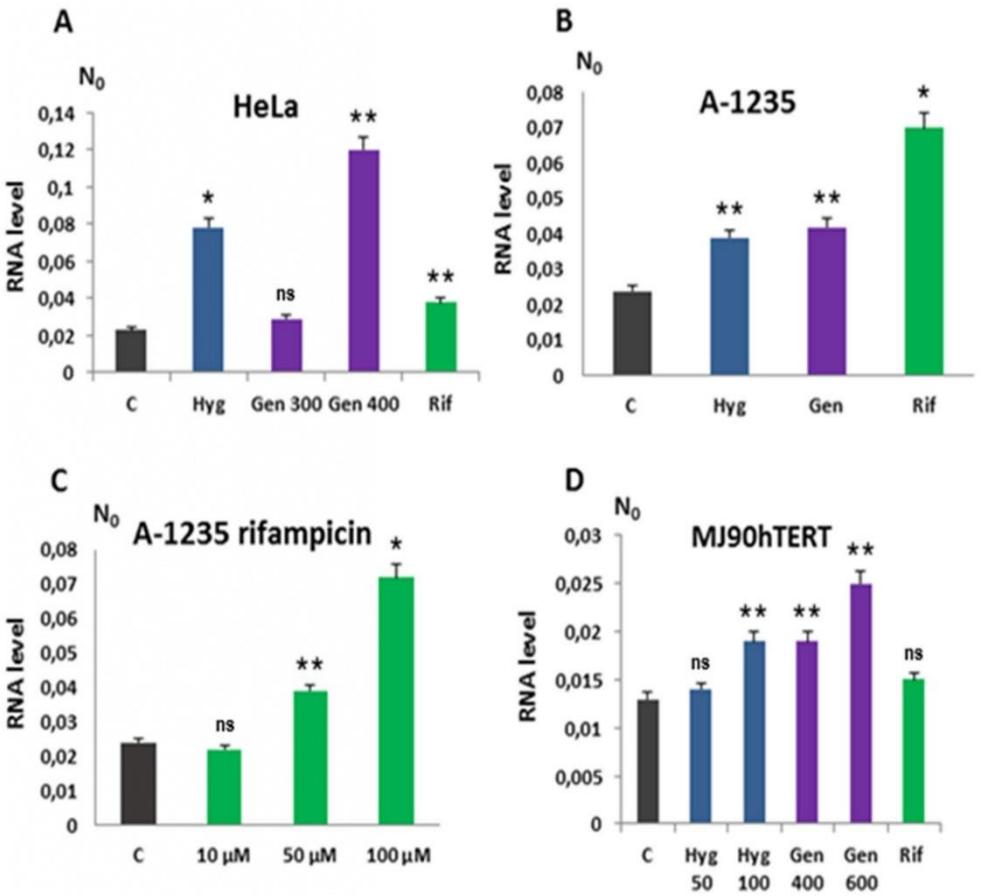


Fig. 1 Transcription of alpha satellite DNA in different cell lines after antibiotic treatment for 48 h: (A) HeLa cells treated with hygromycin B (50 µg/ml), geneticin (300 µg/ml and 400 µg/ml), rifampicin (82 µg/ml); (B) A-1235 cells treated with hygromycin B (50 µg/ml), geneticin (400 µg/ml), rifampicin (82 µg/ml); (C) A-1235 cells treated with rifampicin 8.2 µg/ml, 41 µg/ml, and 82 µg/ml; (D) MJ90hTERT cells treated with hygromycin B (50 µg/ml and 100 µg/ml), geneticin (400 µg/ml and 600 µg/ml), rifampicin (82 µg/ml). Two independent experiments were performed on each cell line. N_0 represents normalized average N_0 value and C denotes control. Columns show averages of two different RT-qPCR experiments performed in triplicates and error bars represent standard deviations. Statistical significance between controls and treated samples was calculated using the t-test and is indicated by stars (** $p < 10^{-2}$, * $p < 0.05$, ns - no significant difference)

B (50 µg/ml) and rifampicin (82 µg/ml) did not show a significant change of transcription. Higher concentration of hygromycin B (100 µg/ml) showed 1.5x ($P = 0.01$) increase of alpha satellite, however, with high percentage of dead cells. On the other hand, higher concentration of geneticin (600 µg/ml) induced 1.9x ($P = 0.008$) increase of alpha satellite transcription (Fig. 1d), while preserving the number of cells and their morphology.

The results revealed a general increase of alpha satellite DNA transcription in cell lines after treatment with different antibiotics at standard concentrations. However, the response differs among cell lines - maximal increase in A-1235 cells was obtained by rifampicin (82 µg/ml) while other two antibiotics showed a modest change (Fig. 1b). On the contrary, in HeLa cells the maximal effect on alpha satellite DNA transcription was obtained by geneticin (400 µg/ml; Fig. 1a), while in MJ90hTERT only higher concentration of geneticin (600 µg/ml) induced a significant change in transcription (Fig. 1d). The results also revealed a positive correlation between antibiotic concentration and the level of alpha satellite transcription.

Additionally, we tested influence on alpha satellite transcription of generally used prokaryotic antibiotics at standard concentrations routinely used in cell culture: amoxicillin (80 µg/ml), ampicillin (50 µg/ml), cefepime (131 µg/ml), cefuroxime (50 µg/ml) and streptomycin (50 µg/ml). The results, after 48 h treatment of MJ90hTERT cells, showed consistent overexpression of alpha satellite DNA in all samples relative to controls, aligning with our previous findings for geneticin, hygromycin B, and rifampicin (Figure S2).

H3K9me3, H3K18ac and H3K4me2 levels at alpha satellite repeats after antibiotic treatment

We analysed the distribution of silent histone mark H3K9me3 characteristic for heterochromatin, H3K18ac mark which is characteristic for transcriptional activation of heterochromatin [16] and H3K4me2, typical for open euchromatin, at tandemly arranged alpha satellite repeats as well as at those dispersed within euchromatin, under standard conditions and after antibiotic treatment. We performed chromatin immunoprecipitation (ChIP) coupled with quantitative real-time PCR, using specific primers for tandemly arranged satellite repeats as well as for six alpha repeats dispersed within introns of genes: AR 1, 10, 21, 25, 29 and 31 [9], (Table S1). Sequences flanking dispersed alpha repeats were used to construct single locus-specific primers. ChIP assay was performed on chromatin isolated from A-1235, HeLa and MJ90hTERT cells subjected to antibiotic treatment of 48 h at 37 °C. The level of tested histone modifications was measured immediately after antibiotic treatment and was compared to the level of control using the unpaired t-test. In addition, we followed the dynamics of IgG binding to dispersed alpha satellite repeats and tandemly repeated satellite arrays and the amount of bound IgG was very low, resulting in a signal below the qPCR threshold.

HeLa cells were treated with geneticin 400 µg/ml which exhibited the strongest effect on alpha satellite transcription (Fig. 1a) and decrease of H3K9me3 level of 2.1x ($P = 0.011$) at tandemly arranged heterochromatic alpha repeats was observed, while no significant change at six euchromatic repeats located within introns was found (Fig. 2a).

Decrease of geneticin concentration to 300 µg/ml resulted in no significant change of H3K9me3 level at tandemly arranged satellite arrays (Figure S3a) corresponding to no significant change of alpha satellite transcription at this concentration (Fig. 1a). The levels of H3K18ac and H3K4me2 were not significantly changed at tandemly arranged alpha satellite repeats as well as at alpha repeats dispersed within euchromatin after treatment with geneticin 400 µg/ml (Figure S5).

The treatment of A-1235 cells with 82 µg/ml rifampicin revealed 2.0x ($P = 0.02$) decrease of H3K9me3 level at tandemly arranged alpha satellite DNA repeats and no significant change at dispersed alpha satellite repeats (Fig. 2b). No significant change in H3K18ac or H3K4me2 level was detected, either at tandem or dispersed alpha repeats (Figure S4a). Lower rifampicin concentration of 41 µg/ml resulted in a slight but not statistically significant decrease of H3K9me3 level at tandem alpha arrays (Figure S3a), while an increase of alpha transcription of 1.8x was detected at this concentration (Fig. 1c).

The treatment of MJ90hTERT cells with geneticin 600 µg/ml, which showed the strongest effect on alpha satellite transcription (Fig. 1d), revealed a significant increase of 2.3x ($P = 0.01$) of H3K18ac level at tandemly arranged alpha satellite arrays but not on dispersed alpha repeats (Fig. 2d). Additionally, histone mark H3K9me3 was significantly decreased by 2.1x ($P = 0.02$) (Fig. 2c). On the other hand, no significant change in H3K4me2 level was found at tandemly arranged or dispersed alpha repeats (Figure S4b).

Results from HeLa and MJ90hTERT cells treated with geneticin and A-1235 cells with rifampicin revealed a decrease of H3K9me3 at heterochromatic alpha repeats which corresponds to increased transcription of alpha satellite DNA. In MJ90hTERT cells however, alongside H3K9me3 decrease after geneticin treatment, H3K18ac was significantly increased, which also corresponds to increased transcription of alpha satellite. On alpha repeats dispersed within euchromatin we did not detect changes in tested histone modifications after any antibiotic treatment. Although at 41 µg/ml rifampicin treatment an increase of alpha satellite transcription

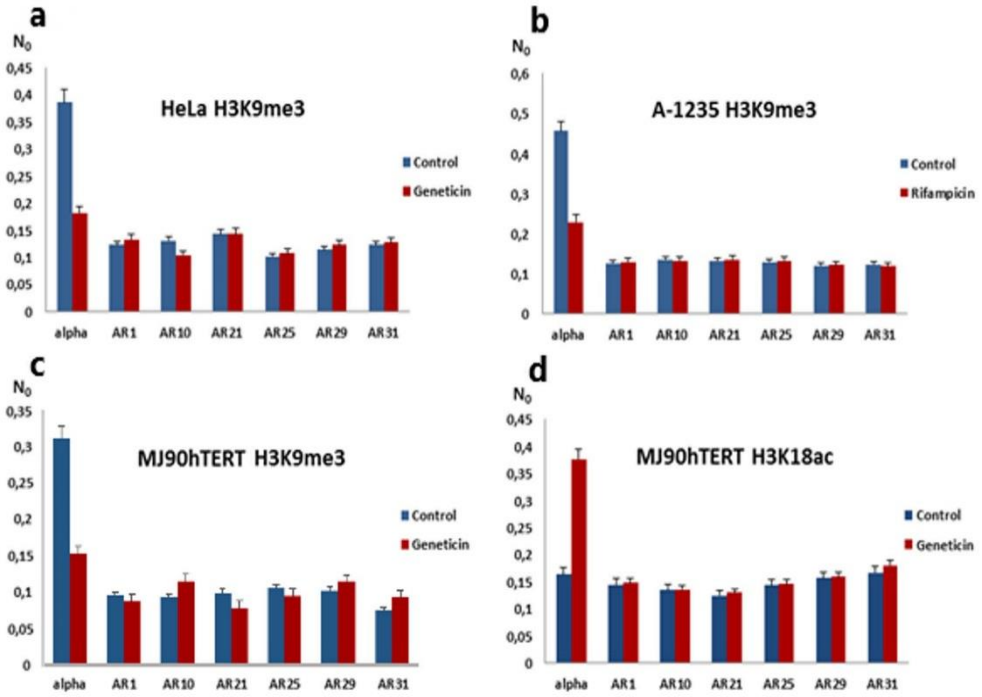


Fig. 2 Levels of histone modifications at tandemly arranged alpha satellite repeats characteristic for heterochromatin and at alpha repeats (ARs) dispersed within euchromatin after antibiotic treatment for 48 h: (A) H3K9me3 level in HeLa cells after treatment with geneticin (400 µg/ml). Significant decrease of H3K9me3 level of 2.1x ($P=0.011$) at tandemly arranged heterochromatic alpha repeats was observed; (B) H3K9me3 level in A-1235 cells after treatment with rifampicin (100 µM). 2.0x ($P=0.02$) decrease of H3K9me3 level at tandemly arranged alpha satellite DNA repeats was detected; (C) H3K9me3 level in MJ90hTERT cells after treatment with geneticin (600 µg/ml). Histone mark H3K9me3 was significantly decreased by 2.1x ($P=0.02$) at tandem alpha repeats. (D) H3K18ac level in MJ90hTERT cells after treatment with geneticin (600 µg/ml). A significant increase of 2.3x ($P=0.01$) was detected at tandemly arranged alpha satellite arrays. Levels of histone modifications were measured by ChIP coupled with quantitative real-time PCR at standard conditions (control) and after antibiotic treatment. N_0 values were normalized using N_0 values of input fractions and represent the levels of histone modifications. Columns show averages of two independent experiments and error bars indicate standard deviations

was observed in A-1235 cells despite a slight change in H3K9me3 level, this could be explained by low sensitivity of ChIP experiments. Similar to that, we observed no statistically significant change in H3K9me3 level at tandem alpha repeats in A-1235 cells as well as in H3K9me3 and H3K18ac levels in MJ90hTERT cells after geneticin 400 µg/ml treatment (Figure S3a, b) despite the slight increase in alpha transcription level of 1.7x and 1.5x respectively (Fig. 1b, d).

H3K9me3, H3K18ac and H3K4me2 levels genome-wide after antibiotic treatment

To see if antibiotic treatment affects epigenetic changes genome-wide we performed immunofluorescence on HeLa, A-1235 and MJ90hTERT cells using primary antibodies against histone marks H3K9me3, H3K18ac and H3K4me2, followed by secondary antibody marked with Alexa Fluor® 488.

After treatment of HeLa cells with geneticin 400 µg/ml for 48 h H3K9me3 level was increased genome-wide by 2.03x ($P<10^{-3}$), H3K18ac showed slight increase of 1.19x ($P<10^{-3}$), while H3K4me2 was decreased by 1.45x ($P<10^{-3}$) (Fig. 3; Table 1). Decrease of geneticin concentration to 300 µg/ml resulted in a slight increase of H3K9me3 of 1.32x ($P<10^{-3}$) while H3K18ac and H3K4me2 levels were downregulated 1.81x and 1.5x ($P<10^{-3}$), respectively (Table 1).

In MJ90hTERT cells after treatment with 600 µg/ml geneticin H3K9me3 and H3K18ac levels were increased 1.86x and 1.27x ($P<10^{-3}$), respectively, while H3K4me2 was slightly increased, 1.1x, but not statistically significant ($P>0.05$) (Fig. 4; Table 1). Decrease of geneticin concentration to 400 µg/ml induced the increase of H3K9me3 level of 2.28x ($P<10^{-3}$) while H3K18ac and H3K4me2 were not significantly changed (Table 1).

Different profile of histone changes was detected after treatment of A-1235 cells with 82 µg/ml rifampicin.

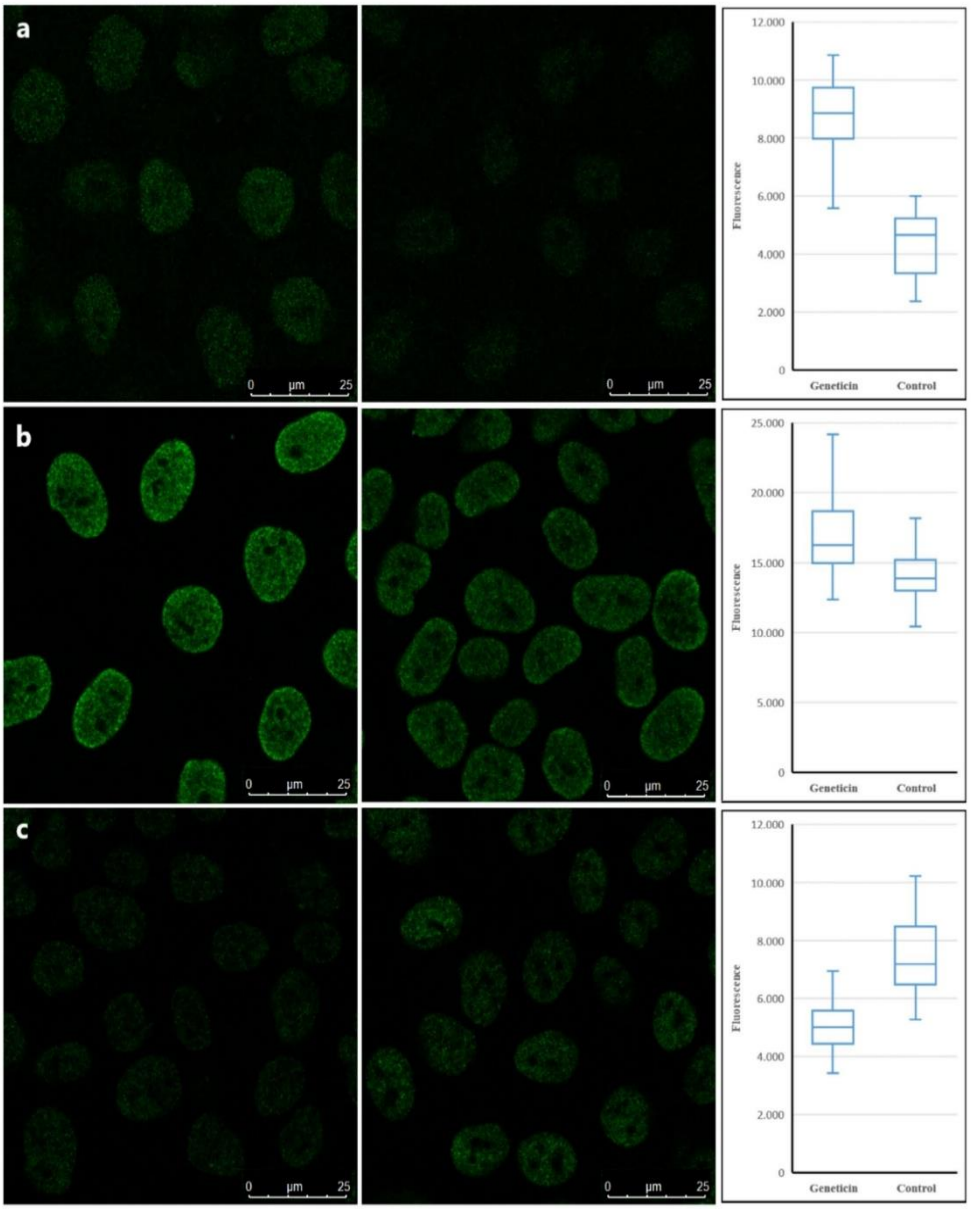


Fig. 3 Genome-wide analysis of H3K9me3 (a), H3K18ac (b) and H3K4me2 (c) levels in HeLa cells after geneticin treatment (400 µg/ml, left panels) and in controls (right panels). H3K9me3 level was increased genome-wide by 2.03x ($P<10^{-3}$), H3K18ac showed slight increase of 1.19x ($P<10^{-3}$), while H3K4me2 was decreased by 1.45x ($P<10^{-3}$). The fluorescence intensity is shown by box plots. Median values are indicated and error bars represent standard deviations

Namely, the level of H3K9me3 was slightly changed (1.08x, $P<10^{-3}$), while H3K18ac level was decreased by 2.38x and H3K4me2 increased 1.33x ($P<10^{-3}$) (Fig. 5; Table 1). Decrease of rifampicin concentration to 41 µg/ml resulted in H3K9me3 and H3K18ac decrease of 2.27x and 2.28x ($P<10^{-3}$), respectively, while H3K4me2 was upregulated by 1.4x ($P<10^{-3}$) (Table 1). In the same cell line, geneticin 400 µg/ml induced the decrease of

Table 1 The genome-wide fold changes of epigenetic modifications H3K9me3, H3K18ac and H3K4me2 in different cell lines after treatment with varying concentrations of antibiotics geneticin (Gen) and rifampicin (Rif) (up-upregulation; dw-downregulation; n.s.c.-no significant change)

Antibiotic	Cell line	H3K9me3 (X)	H3K18ac (X)	H3K4me2 (X)
Gen 400 µg/ml	HeLa	2.03 up	1.19 up	1.45 dw
Gen 300 µg/ml	**HeLa**	**1.32 up**	**1.81 dw**	**1.5 dw**
Gen 600 µg/ml	MJ90hTERT	1.86 up	1.27 up	n.s.c.
Gen 400 µg/ml	**MJ90hTERT**	**2.28 up**	**n.s.c.**	**n.s.c.**
Gen 400 µg/ml	A-1235	2.86 dw	2.58 up	1.35 up
Rif 82 µg/ml	**A-1235**	**1.08 up**	**2.38 dw**	**1.33 up**
Rif 41 µg/ml	A-1235	2.27 dw	2.78 dw	1.40 up

H3K9me3 level by 2.86x and the increase of H3K18ac and H3K4me2 levels by 2.58x and 1.35x ($P<10^{-3}$), respectively (Table 1). Changes of histone modifications genome-wide in three cell lines are summarized in Table 1.

The results reveal genome-wide increase of H3K9me3 levels in HeLa and MJ90hTERT cells induced by geneticin (300–600 µg/ml), while H3K4me2 levels were either downregulated or not significantly changed (Table 1). On the other hand, the level of H3K18ac was significantly downregulated in HeLa cells treated with 300 µg/ml geneticin, while higher concentrations of geneticin only slightly changed H3K18ac levels in HeLa and MJ90hTERT cells, indicating that the effect depends on the concentration of antibiotic but is not positively correlated with it. In A-1235 cells however, geneticin in the concentration of 400 µg/ml induced genome-wide downregulation of H3K9me3 and upregulation of H3K18ac, indicating that response to the antibiotic differs among cell lines. Also, different antibiotics affect differently epigenetic marks in the same cell line as shown by rifampicin which in A-1235 cells stimulates H3K18ac decrease while geneticin induces H3K18ac upregulation (Table 1).

Although different antibiotics induce overexpression of pericentromeric alpha satellite DNA, their effect on heterochromatin differs among cells, characterized either by decrease of H3K9me3 or increase of H3K18ac (Fig. 2). In a similar way, the effect of antibiotics genome-wide also differs among cells. Namely, while the effect of geneticin on histone marks H3K9me3 and H3K4me2 in HeLa and MJ90hTERT cells is similar, it differs from the one observed in A-1235 cells, as well as from the effect of rifampicin on the same marks in A-1235 cells. The results suggest that the heterochromatin, as well as the rest of chromatin, respond to antibiotics in diverse ways, depending on the cell line, type of antibiotic and antibiotic concentration.

Discussion

It is well known that antibiotics influence human microbiome and change its composition which can have a negative impact on host health including reduced microbial diversity and selection of antibiotic-resistant strains, making hosts more susceptible to infection [24]. However, besides targeting bacterial cells, antibiotics affect metabolism of eukaryotic cells as revealed by studies in vitro, on human cell lines [3, 4]. Some in vivo studies, such as those performed on male pseudoscorpions treated with the antibiotic tetracycline, showed significantly reduced sperm viability, which was passed to the next generation and suggests that a similar effect could occur in other species [25]. It was shown that clinically relevant doses of bactericidal antibiotics quinolones, aminoglycosides and β-lactams cause mitochondrial dysfunction and ROS overproduction in mammalian cells, and mice treated with these antibiotics exhibited elevated oxidative stress markers in the blood as well as oxidative tissue damage [26]. It is also known that treatment of some diseases requires high doses of antibiotic, e.g. for tuberculosis 35 mg/kg rifampicin per day is used (which corresponds to approx. 43 µM conc.) [21], and it is therefore interesting to know how similar doses affect metabolism of mammalian cells. Rifampicin is usually well-tolerated and rarely causes serious toxicity in eukaryotic cells. In extreme doses, however, rifampicin is known to produce hepatic, renal and hematological

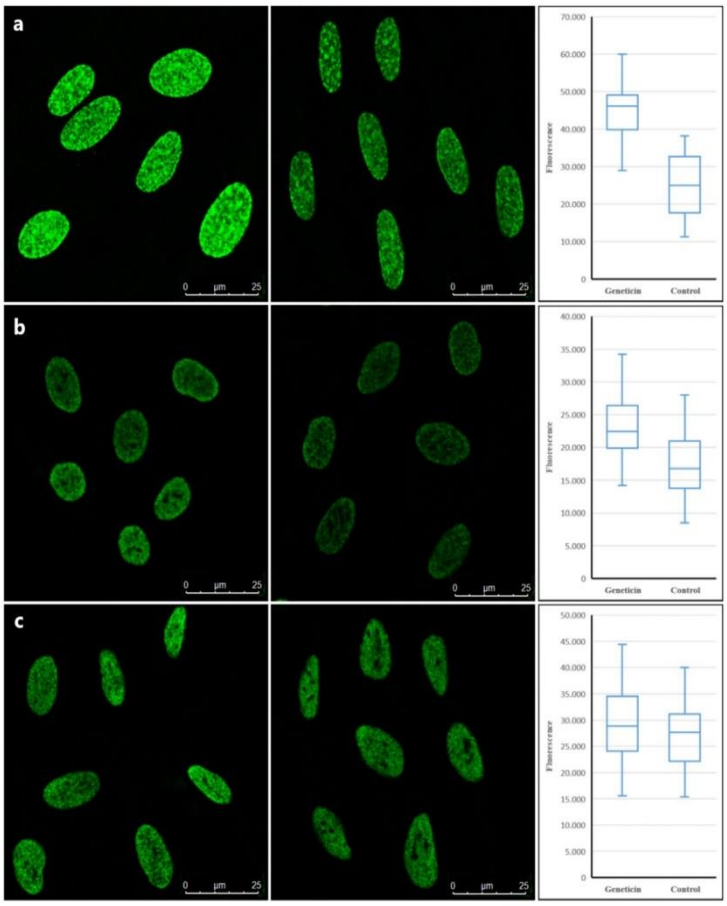


Fig. 4 Genome-wide analysis of H3K9me3 (a), H3K18ac (b) and H3K4me2 (c) levels in MJ90hTERT cells after treatment with geneticin 600 µg/ml (left panels) and in controls (right panels). H3K9me3 and H3K18ac levels were increased 1.86x and 1.27x ($P < 10^{-3}$), respectively, while H3K4me2 was slightly increased, 1.1x, but not statistically significant ($P > 0.05$). The fluorescence intensity is shown by box plots. Median values are indicated and error bars represent standard deviations

disorders and metabolic acidosis [27]. Its toxicity is predominantly hepatic and immuno-allergic in character. Rifampicin induces a dose-dependent hepatotoxicity in HHL-17 cells (IC50; 600 µM), and increases the serum levels of liver injury markers, e.g., alanine transaminase (ALT) and aspartate transaminase (AST) [28]. Also, it was found that rifampicin at exorbitant concentration exerts adverse effects on the proliferation of MSCs in human bone marrow and the differentiation of osteoblasts [29].

Tandemly arranged satellite DNA repeats represent a challenge for the maintenance of genomic stability, during normal cellular functions such as replication and transcription [15]. The increased pericentric satellite DNA transcription has negative effects on cellular physiology, leading to defects typically associated with tumorigenesis and ageing. Overexpressed transcripts of pericentromeric major satellite DNA in mice sequester BRCA1-associated network, cause accumulation of RNA loops, DNA damage and induce breast cancer [30]. Satellite DNAs are sensitive to different exogenous and endogenous stress conditions and here we investigated if the antibiotics commonly used in cell culture studies affect the expression of non-coding major human alpha satellite DNA. We used the aminoglycoside antibiotics geneticin G418 and hygromycin B which are effective against both eukaryotic and prokaryotic cells and are used to select for cells that express antibiotic resistance. Both antibiotics affect protein synthesis. We also tested rifampicin, an ansamycin antibiotic used to treat some bacterial infections, including tuberculosis. Rifampicin blocks bacterial DNA transcription by inhibiting bacterial RNA

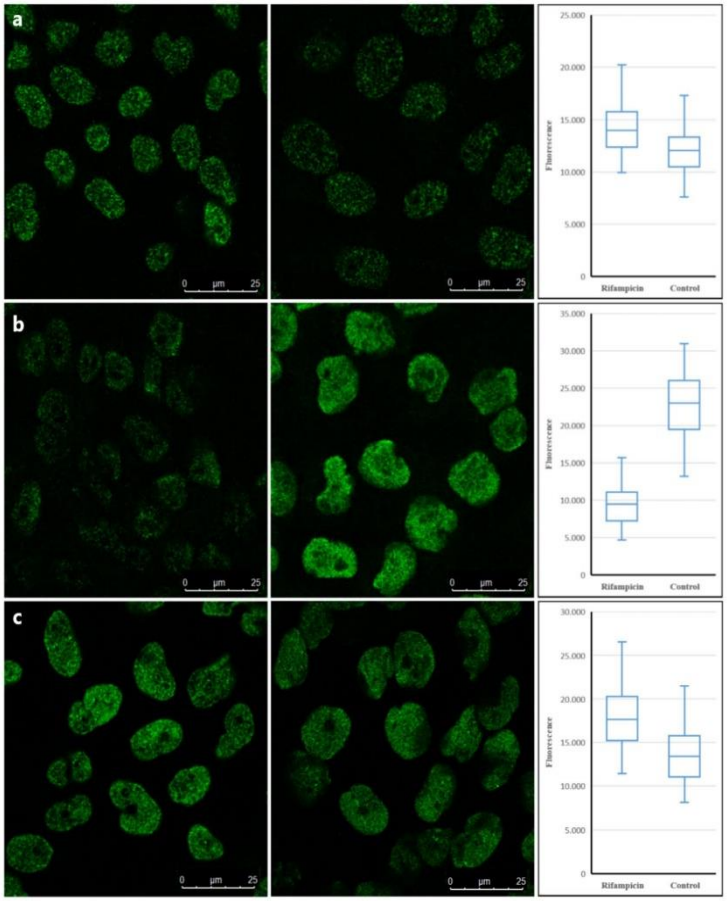


Fig. 5 Genome-wide analysis of H3K9me3 (a) H3K18ac (b) and H3K4me2 (c) levels in A-1235 cells after treatment with rifampicin 100 µM (left panels) and in controls (right panels). The level of H3K9me3 was slightly changed (1.08x, $P < 10^{-3}$), while H3K18ac level was decreased by 2.38x and H3K4me2 increased 1.33x ($P < 10^{-3}$).The fluorescence intensity is shown by box plots. Median values are indicated and error bars represent standard deviations

polymerase, whereas in our study it had an opposite effect on human heterochromatic DNA transcription, i.e. it stimulated it. The present study revealed that alpha satellite DNA is highly susceptible to antibiotic stress. Namely, using three cell lines: glioblastoma A-1235, HeLa and MJ90hTERT, we observed increased transcription of alpha satellite DNA using all three antibiotics at standard concentrations, although the response differed among cell lines. Maximal increase in A-1235 cells was obtained by rifampicin while in MJ90hTERT and HeLa cells geneticin induced the most significant increase of transcription under standard concentrations. The results also reveal a positive correlation between antibiotic concentration and the level of alpha satellite transcription.

Several other antibiotics that specifically target prokaryotes, such as amoxicillin and ampicillin (both penicillin derivatives that inhibit bacterial cell wall synthesis), cefepime and cefuroxime (both β-lactam antibiotics that also inhibit cell wall synthesis), and streptomycin (which binds to bacterial ribosomes and inhibits protein synthesis), also induced increased transcription of alpha satellite DNA. This suggests that their effects may reflect the activation of a general cellular stress response. This raises the question of how these antibiotics affect human cells, independent of their known actions on prokaryotic physiology. Further research is clearly needed to better understand the physiological consequences of applying these commonly used antibiotics.

The cell lines used as models in our research were selected for two specific reasons. First, in our previous experiments they demonstrated robustness and high viability (>80%; using sublethal, physiological doses) during treatments and manipulations, providing results that were both highly reproducible and significant. Second,

while the effect of antibiotics on satellite expression was observed in all cases, as is clearly shown in Fig. 1., the most significant combinations were chosen for further epigenetic testing. It should also be noted that MJ90 are fibroblasts, immortalized with hTERT, therefore considered non-transformed cells, both morphologically and physiologically.

The increase of alpha satellite expression upon treatment with antibiotics is associated with significant decrease of silent histone mark H3K9me3 at heterochromatic alpha satellite repeats in HeLa and MJ90hTERT cells treated with geneticin and A-1235 cells with rifampicin, respectively, suggesting possible influence of this epigenetic change on alpha satellite DNA transcription. At concentrations of antibiotics which did not significantly affect satellite transcription, no change of H3K9me3 level in HeLa and A-1235 cells was observed, also supporting a possible relation between satellite transcription and H3K9me3 levels at heterochromatin. In MJ90hTERT cells, however, H3K18ac level was increased upon treatment with geneticin alongside the decrease of H3K9me3 level. Since it is known that H3K18 hyperacetylation leads to aberrant accumulation of pericentric transcripts [16], we propose that increased level of H3K18ac might also be responsible for overexpression of alpha satellite DNA in MJ90hTERT cells after geneticin treatment. Previous studies have shown that alpha satellite transcription is not simply a byproduct of deregulation of other genetic elements under normal or stress conditions. While both transposons and satellite DNAs are expressed during the cell cycle, there is no evidence that transposon activity influences alpha satellite transcription. Instead, the alpha satellite transcription seems to be controlled by the presence of centromere–nucleolar contacts [31] and by CENP-B protein which promotes the binding of the zinc-finger transcriptional regulator (ZFAT) responsible for activation of RNA Pol II transcription [32]. It was shown that Topoisomerase I (TopI) promotes the transcription of α-satellite DNAs which is also stimulated in response to DNA double-stranded breaks (DSBs) [33]. However, the results presented here, as well as the increased transcription of alpha satellite DNA in cancer which is associated with decreased H3K9me3 level at satellite repeats [17], suggest a regulation of alpha satellite transcription by epigenetic changes, in particular by histone marks H3K9me3 and H3K18ac. Loss of epigenetic heterochromatic marks was shown to be responsible for overexpression of satellite DNA in ageing [34] and neurodegenerative diseases [23]. It is also important to mention that although present at lower levels than transposon transcripts, alpha satellite RNAs are essential for centromere assembly, kinetochore formation, heterochromatin organization, and may modulate expression of associated genes. These roles highlight the importance of studying their activation and function under stress conditions, alongside other repetitive elements.

Apart from H3K9me3 and H3K18ac we did not observe a change in histone mark H3K4me2, characteristic for open but inactive euchromatin [35], at tandemly arranged heterochromatic alpha satellite repeats. Also, no change in all three epigenetic modifications was detected on alpha satellite repeats dispersed within euchromatin upon antibiotics treatment. More diverse changes, including not only H3K9me3 and H3K18ac but also H3K4me2, were detected genome-wide using immunofluorescence and these changes depended on cell type, antibiotic and antibiotic concentration.

We acknowledge that gene expression and epigenetic responses vary depending on cell type. In our study, we observed distinct regulatory patterns of alpha satellite expression across the tested cell lines, likely influenced by differences in chromatin organization, baseline transcriptional activity, and cellular metabolism. These variations may reflect intrinsic differences between cancerous and normal cells in their response to external stimuli, including antibiotics. Further studies are needed to dissect the molecular mechanisms driving these cell-type-specific effects. Different signaling pathways, overactivated in different cell lines, possibly confer resistance to certain antibiotics as well as types of stress or damage that they induce. It is likely that the level of certain type of stress the cells can resist varies between different cell lines and that their coping mechanisms against different types of stress are also diverse as a general consequence of their specific genetic background.

Our findings suggest that antibiotics may influence satellite DNA transcription by modulating specific histone marks. While the histone modification pathways are well known and highly conserved among all eukaryotes, the underlying mechanisms by which antibiotics potentially interact with said pathways are currently unknown. Ubiquitous effector enzymatic complexes and their functions; such as activating demethylases and acetyl-transferases; suppressing methyl-transferases and deacetylases, as well as chromatin remodeling factors such as SWI/SNF are well understood. Integrated stress response could also play a role by interacting with above-mentioned pathways. Many types of stress, including antibiotic, activate integrated stress response cellular machinery resulting in overproduction of specific transcription factors (such as ATF4), stimulating downstream chromatin remodeling of specific loci and neighboring promoter activation, consequently explaining their overexpression.

It should be noted that transcription of alpha satellite DNA upon antibiotic treatment in some cases exceeded normal transcription rate almost 5-fold (e.g. HeLa cells treated with geneticin 400 μg/ml), possibly affecting cell physiology as well as genome stability in the process.

This paper shows that commonly used antibiotics not only affect bacterial cells but also induce significant transcriptional and epigenetic changes in eukaryotic cells, particularly at alpha satellite DNA regions. The observed epigenetic changes, such as the reduction in H3K9me3 and increase in H3K18ac, provide a potential mechanistic explanation for the transcriptional upregulation of alpha satellite DNA. These findings suggest that researchers using antibiotics in cell culture studies should consider the results described in this paper when such experiments are performed.

Our research highlights the following key points:

1. (1) Routine antibiotic use in cell cultures affects metabolism and genomic stability, warranting caution.
2. (2) Antibiotic treatment in humans may have systemic effects beyond bacterial targeting, varying across tissues.
3. (3) Differential antibiotic effects on cancer vs. healthy tissues could inform potential anticancer strategies.
4. (4) Antibiotics may influence gene expression beyond alpha satellite DNA, potentially affecting other genomic regions and regulatory elements.
5. (5) Understanding individual variability in antibiotic response could improve drug safety and personalized treatment strategies.
6. (6) Chronic antibiotic exposure may have cumulative effects on genomic stability and epigenetic regulation.

These findings underscore broader implications for gene regulation, drug safety, and long-term antibiotic effects, highlighting the need for further research.

Materials and methods

Human cell lines

The following human cell lines were used in experiments: MJ90hTERT (immortalized human skin fibroblasts), HeLa (human cervical cancer) and A-1235 (human glioblastoma). Human diploid fibroblast strain MJ90hTERT (HCA2hTERT) was kindly provided by Dr Olivia M. Pereira-Smith (University of Texas, Health Science Center, San Antonio, TX, USA) [36–38]. HeLa and A-1235 cells were were provided by Thermo Fisher Scientific. Cells were cultured in appropriate medium (DMEM) supplemented with 10% FBS and 5% CO2 at 37 °C. Cells were incubated for 48 h at 37 °C with antibiotics geneticin, rifampicin and hygromycin B (Carl Roth) in complete medium. The concentration range of the antibiotics were: geneticin 300–600 µg/ml, hygromycin B 50–100 µg/ml, as well as with rifampicin 8.2–82 µg/ml (10–100 µM). Data regarding cell viability testing are shown in Figures S7 and S8. Trypan blue assay demonstrated approx. 5–10% of cell death across all antibiotic treatment combinations as well as in untreated controls.

RNA isolation and reverse transcription

For RNA isolation from cell cultures lysis buffer was added directly after the PBS washing step, avoiding trypsin treatment. RNA was quantified with the Quant-IT RNA assay kit using a Qubit fluorometer (Invitrogen). Integrity of RNA was checked by gel electrophoresis. Approximately 1 µg of RNA was reverse transcribed using the PrimeScript RT reagent Kit with gDNA Eraser (perfect Real Time, Takara) in 20 µl reaction using specifically modified primer for alpha satellite rev AATGCAC ATATCACTATGTAC, designed to produce cDNA molecules that differ from genomic DNA in order to avoid DNA contamination [39]. For all samples, negative controls without reverse transcriptase were used.

Quantitative real-time PCR (qPCR) analysis

qPCR analysis was performed according to the previously published protocol (Feliciello et al. 2015). Primers used for transcriptional analysis of alpha satellite DNA were constructed based on consensus sequence derived from cloned alpha satellite monomers of wide-ranging chromosomal origins [40] and the same modified primer used previously in reverse transcription was used in qPCR amplification along with the second primer fw C ACTCTTTTTGTAGAATCTGC. In this way, amplification was unaffected by any potential DNA contamination [34]. Glucuronidase beta (GUSB) [41] was used as an endogenous control for normalization in human samples. GUSB gene (Gene ID: 2990) was stably expressed without any variation among samples after antibiotic treatment. Three additional reference genes (GAPDH, TOP3A and DEK) were also tested and they showed stable expression in all samples, not affected by geneticin treatment (Figure S6). The primers for these genes are listed in Table S2. The following thermal cycling conditions were used: 50 °C 2 min, 95 °C 7 min, 95 °C 15 s, 60 °C 1 min for 40 cycles followed by dissociation stage: 95 °C for 15 s, 60 °C for 1 min, 95 °C for 15 s and 60 °C for 15 s. Amplification specificity was confirmed by dissociation curve analysis and specificity of amplified products was tested on agarose gel. Control without template (NTC) was included in each run. Post-run data were analysed using LinRegPCR software v.11.1 [42, 43]. which enables calculation of the starting concentration of amplicon ("N_0 value"). N_0 value is expressed in arbitrary fluorescence units and is calculated by taking into account PCR efficiency and baseline fluorescence. N_0 value determined for each technical replicate was averaged and the averaged N_0 values were divided by the N_0 values of the endogenous control. Statistical analysis of qPCR data was done using GraphPad v.6.01 and normalized N_0 values were compared

using the unpaired t-test which compares the means of two unmatched groups.

Chromatin immunoprecipitation

MJ90hTERT, A-1235 and HeLa cells were grown to subconfluence, washed in PBS, scraped in Nuclear Isolation buffer (10 mM MOPS; 5 mM KCl; 10 mM EDTA; 0.6% Triton X-100) with protease inhibitor cocktail CompleteMini (Roche) and chromatin immunoprecipitation was performed according to the published protocol (Feliciello et al. 2015, 2020), with the exception of sonication step which was performed 30 times for 30 s on ice, high sonicator amplitude. The antibodies used were: Anti-Histone H3 (tri methyl K9, Abcam, ab8898), Anti-Histone H3 (di methyl K4, tri methyl K4, Abcam, ab6000), Anti-Histone H3 (acetyl K18, Abcam, ab1191), Anti-Histone H3 (di methyl K4, Abcam, ab7766) and IgG (Santa Cruz Biotechnology, sc2027). Binding of precipitated target was monitored by qPCR using the SYBR Green PCR Master mix (Bio-Rad). Primers used for H3K9me3, H3K18ac and H3K4me2 level analyses at heterochromatic alpha regions as well as at dispersed alpha repeats are listed in Table S1. The N_0 values were normalized using N_0 values of input fractions.

Immunofluorescence

Cells were grown on cover slips up to 70% confluence, washed with PBS and fixed for 5 min in cold methanol. Permeabilization was done by 0.5% triton X-100 for 5 min and blocking with DAKO Protein Block Serum-free ready to use reagent for 1 h at RT. Primary antibodies anti-H3K9me3 (Abcam, ab8898), anti-H3K18ac (Abcam, ab1191) and anti-H3K4me2 (Abcam, ab7766) were diluted in DAKO Antibody Diluent according to the instructions of Abcam, and incubation was performed overnight at 4 °C. After washing in PBS, goat polyclonal secondary antibody to rabbit IgG (ab150081) was diluted 1/1000 in DAKO Antibody Diluent and incubation was performed for 1 h at RT in the dark. Cells were stained with 1 µg/ml DAPI and a drop of DAKO Anti-Fade Fluorescence Mounting Medium was added. Cell slides were sealed with nail polish and analysed by confocal microscopy (Laser Scanning Confocal Microscope Leica SP8 X FLIM). For each sample slide (control and treated), five images were taken and the mean fluorescence values of all structurally and morphologically intact nuclei were quantified using „ImageJ“ software [44]. The Shapiro–Wilk test was used to test data normality. Mean fluorescence values of treated samples and controls were tested for statistical significance using the parametric 2-tailed Welch's t-test if the data had normal distribution and non-parametric Mann–Whitney test when it did not. P-values less than 0.05 were considered statistically significant.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13072-025-00628-z.

Supplementary Material 1 :The following supporting information can be downloaded at: https://www.mdpi.com/xxx/s1. Figure S1: Consensus sequence of 171 bp alpha satellite monomer; Figure S2: Transcription of alpha satellite DNA in MJ90hTERT cell line after treatment with prokaryote-specific antibiotics. Figure S3a: H3K9me3 level at tandemly arranged alpha satellite arrays in HeLa cells; Figure S3b: H3K18ac level at tandemly arranged alpha satellite arrays in MJ90hTERT cells; Figure S4a: H3K18ac and H3K4me2 levels at tandemly arranged alpha satellite repeats and dispersed alpha repeats; Figure S4b: H3K4me2 levels of MJ90hTERT cells treated with geneticin; Figure S5: H3K18ac and H3K4me2 levels in HeLa cells after treatment with geneticin; Figure S6: Expression profiles of four housekeeping genes in MJ90hTERT cell line after 48 h treatment with geneticin. Figure S7: Images of tested cell lines treated with most significant antibiotic concentrations. Figure S8: Cell counts after viability assays. Table S1: List of primers used in ChIP-qPCR experiments; Table S2: List of housekeeping genes used as endogenous controls and their primers.

Acknowledgements
We thank Lucija Horvat for the technical support.

Author contributions
Conceptualization, I.F.; methodology, M.M., S.L.; formal analysis, S.L., M.M., D.Đ., M.C.F., A.P., I.F.; investigation, S.L., M.M., I.F.; writing original draft preparation, I.F. and Đ.U.; writing—review and editing, Đ.U., I.F., D.Đ.; supervision, I.F; All authors have read and agreed to the published version of the manuscript.

Funding
This work was funded by the Italian Ministry of Education, University and Research (MIUR), FFABR2017, fund for Investments on Basic Research (FIRB) and by the International Staff Mobility Program of University of Naples Federico II to I. Feliciello, as well as by Croatian Science Foundation grant: IP-2019-04-6915 to Đ. Ugarković.

Data availability
No datasets were generated or analysed during the current study.

Declarations

Competing interests
The authors declare no competing interests.

Received: 14 May 2025 / Accepted: 2 September 2025
Published online: 26 September 2025

References

1. Smith RP, Eckalbar WL, Morrissey KM, et al. Genome-wide discovery of drug-dependent human liver regulatory elements. PLoS Genet. 2014;10:e1004648.
2. Ryu AH, Eckalbar WL, Kreimer A, Yosef N, Ahituv N. Use antibiotics in cell culture with caution: genome-wide identification of antibiotic-induced changes in gene expression and regulation. Sci Rep. 2017;7:7533.
3. Elliott RL, Jiang XP. The adverse effect of gentamicin on cell metabolism in three cultured mammary cell lines. Are Cell Cult Data Skewed?? PLoS One. 2019;14:e0214586.
4. Gallo A, Landi R, Rubino V, et al. Oxytetracycline induces DNA damage and epigenetic changes: a possible risk for human and animal health? PeerJ. 2017;5:e3236.
5. Ugarković Đ, Sermek A, Ljubić S, Feliciello I. Satellite DNAs in health and disease. Genes. 2022;13:1154.
6. Vourc'h C, Biamonti G. Transcription of satellite DNAs in mammals. Prog Mol Subcell Biol. 2011;51:95–118.
7. Pezer Ž, Ugarković Đ. Satellite DNA-associated SiRNAs as mediators of heat shock response in insects. RNA Biol. 2012;9:587–95.

8. Hédouin S, Grillo G, Ivkovic I, et al. CENP-A chromatin disassembly in stressed and senescent murine cells. Sci Rep. 2017;7:42520.
9. Feliciello I, Sermek A, Pezer Ž, et al. Heat stress affects H3K9me3 level at human alpha satellite DNA repeats. Genes. 2020;11:663.
10. Sermek A, Feliciello I, Ugarković Đ. Distinct regulation of the expression of satellite DNAs in the beetle tribolium castaneum. Int J Mol Sci. 2021;22:296.
11. Fonseca-Carvalho M, Veríssimo G, Lopes M, et al. Answering the cell stress call: satellite non-coding transcription as a response mechanism. Biomolecules. 2024;14:124.
12. Ting DT, Lipson D, Paul S, et al. Aberrant overexpression of satellite repeats in pancreatic and other epithelial cancers. Science. 2011;331:593–6.
13. Ljubić S, Sermek A, Prgomet Sečan A, et al. Alpha satellite RNA levels are upregulated in the blood of patients with metastatic castration-resistant prostate cancer. Genes. 2022;13:383.
14. Kim N, Jinks-Robertson S. Transcription as a source of genome instability. Nat Rev Genet. 2012;13:204–14.
15. Black EM, Giunta S. Repetitive fragile sites: centromere satellite DNA as a source of genome instability in human diseases. Genes (Basel). 2018;9:615.
16. Tasselli L, Xi Y, Zheng W, et al. SIRT6 deacetylates H3K18ac at pericentric chromatin to prevent mitotic errors and cellular senescence. Nat Struct Mol Biol. 2016;23:434–40.
17. Vojvoda Zeljko T, Ugarković Đ, Pezer Ž. Differential enrichment of H3K9me3 at annotated satellite DNA repeats in human cell lines and during fetal development in mouse. Epigenet Chromatin. 2021;14:47.
18. Unoki M, Sharif J, Saito Y, et al. CDCA7 and HELLS suppress DNA:RNA hybrid-associated DNA damage at pericentromeric repeats. Sci Rep. 2020;10:17865.
19. Feliciello I, Akrap I, Ugarković Đ. Satellite DNA modulates gene expression in the beetle tribolium castaneum after heat stress. PLoS Genet. 2015;11:e1005466.
20. Landers CC, Rabeler CA, Ferrari EK, et al. Ectopic expression of pericentric HSATII RNA results in nuclear RNA accumulation, MeCP2 recruitment, and cell division defects. Chromosoma. 2021;130:75–90.
21. Boeree MJ, Heinrich N, Aarnoutse R, et al. High-dose rifampicin, moxifloxacin, and SQ109 for treating tuberculosis: a multi-arm, multi-stage randomised controlled trial. Lancet Infect Dis. 2017;17:39–49.
22. McNulty SM, Sullivan BA. Alpha satellite DNA biology: finding function in the recesses of the genome. Chromosom Res. 2018;26:115–38.
23. Smurova K, De Wulf P. Centromere and Pericentromere transcription: roles and regulation in sickness and in health. Front Genet. 2018;9:674.
24. Davies J, Davies D. Origins and evolution of antibiotic resistance. Microbiol Mol Biol Rev. 2010;74:417–33.
25. Zeh JA, Bonilla MM, Adrian AJ, et al. From father to son: transgenerational effect of Tetracycline on sperm viability. Sci Rep. 2012;2:375.
26. Kalghatgi S, Spina CS, Costello JC, et al. Bactericidal antibiotics induce mitochondrial dysfunction and oxidative damage in mammalian cells. Sci Transl Med. 2013;5:192ra85.
27. Sridhar A, Sandeep Y, Krishnakishore C, Sriramnaveen P, Manjusha Y, Sivakumar V. Fatal poisoning by Isoniazid and rifampicin. Indian J Nephrol. Sep; 2012;22(5):385–7.
28. Kainat KM, Ansari MI, Bano N, Jagdale PR, Ayanur A, Kumar M, Sharma PK. Rifampicin-induced ER stress and excessive cytoplasmic vacuolization instigate hepatotoxicity via alternate programmed cell death paraptosis in vitro and in vivo. Life Sci. 2023;15:333:122164.
29. Zhang Z, Wang X, Luo F, Yang H, Hou T, Zhou Q, Dai F, He Q, Xu J. Effects of rifampicin on osteogenic differentiation and proliferation of human mesenchymal stem cells in the bone marrow. Genet Mol Res. 2014;25(3):6398–410.
30. Zhu Q, Hoong N, Aslanian A, et al. Heterochromatin-encoded satellite RNAs induce breast cancer. Mol Cell. 2018;70:842–e8537.
31. Bury L, Moodie B, Ly J, et al. Alpha-satellite RNA transcripts are repressed by centromere-nucleolus associations. Elife. 2020;9:e59770.
32. Ishikura S, Yoshida K, Hashimoto S, et al. CENP-B promotes the centromeric localization of ZFAT to control transcription of noncoding RNA. J Biol Chem. 2021;297:101213.
33. Teng Z, Yang L, Zhang Q, et al. Topoisomerase I is an evolutionarily conserved key regulator for satellite DNA transcription. Nat Commun. 2024;15:5151.
34. Larson K, Yan SJ, Tsurumi A, et al. Heterochromatin formation promotes longevity and represses ribosomal RNA synthesis. PLoS Genet. 2012;8:e1002473.
35. Soares LM, He PC, Chun Y, et al. Determinants of histone H3K4 methylation patterns. Mol Cell. 2017;68:773–e7856.
36. Cukusic Kalajzic A, Vidacek NS, Huzak M, Ivankovic M, Rubelj I. Telomere Q-PNA-FISH–reliable results from stochastic signals. PLoS ONE. 2014;9(3):e92559.
37. Gorbunova V, Seluanov A, Pereira-Smith OM. Expression of human telomerase (hTERT) does not prevent stress-induced senescence in normal human fibroblasts but protects the cells from stress-induced apoptosis and necrosis. J Biol Chem. 2002;277:38540–9.
38. Young JI, Smith JR. DNA methyltransferase Inhibition in normal human fibroblasts induces a p21-dependent cell cycle withdrawal. J Biol Chem. 2001;276:19610–6.
39. Đermić D, Ljubić S, Matulić M, et al. Reverse transcription-quantitative PCR (RT-qPCR) without the need for prior removal of DNA. Sci Rep. 2023;13:11470.
40. Choo K, Vissel B, Nagy A, et al. A survey of the genomic distribution of alpha satellite DNA on all the human chromosomes, and derivation of a new consensus sequence. Nucleic Acids Res. 1991;19:1179–82.
41. Aerts JL, Gonzales MI, Topalian SL. Selection of appropriate control genes to assess expression of tumor antigens using real-time RT-PCR. Biotechniques. 2004;36:84–91.
42. Ruijter JM, Ramakers C, Hoogaars WMH, et al. Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. Nucleic Acids Res. 2009;37:e45.
43. Ruijter JM, Pfa MW, Zhao S, et al. Evaluation of qPCR curve analysis methods for reliable biomarker discovery: bias, resolution, precision, and implications. Methods. 2013;59:32–46.
44. Schneider C, Rasband W, Eliceiri K. NIH image to imageJ: 25 years of image analysis. Nat Methods. 2012;9:671–5.

Publisher's note

Figure S1. Consensus sequence of 171 bp alpha satellite monomer (Choo et al. 1991) and the annealing positions of primers used in RT-qPCR.

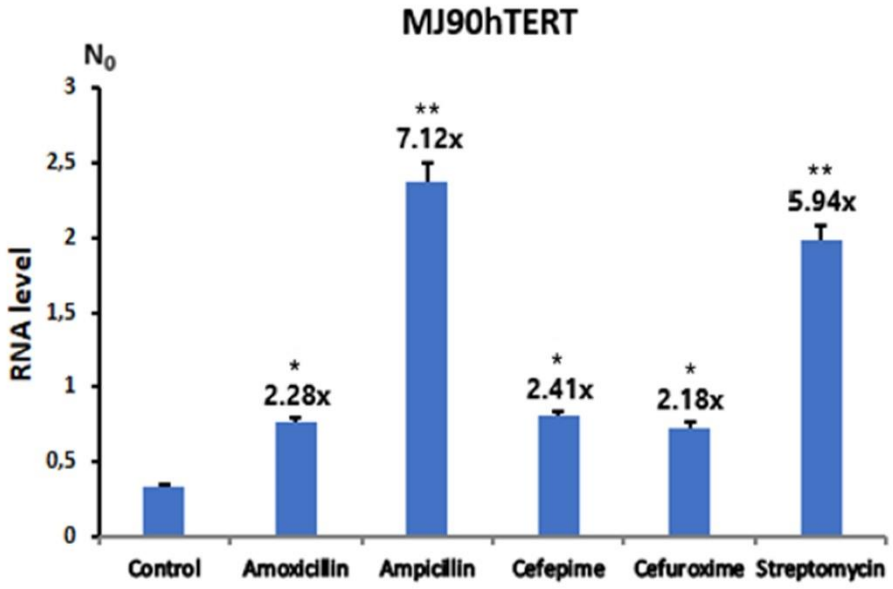
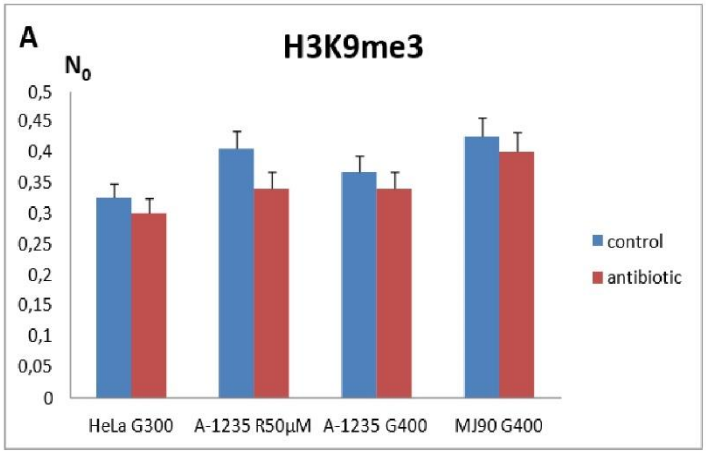


Figure S2. Transcription of alpha satellite DNA in MJ90hTERT cell line after treatment with prokaryote-specific antibiotics for 48 hrs. Two independent experiments were performed. N_0 represents normalized average N_0 value. Columns show averages of two different RT-qPCR experiments performed in triplicates and error bars represent standard deviations. Rates of alpha satellite overexpression (fold) in treated samples compared to control are displayed above error bars. Statistical significance between control and treated samples was calculated using the Student's t-test and is indicated by stars (** $p < 10–2$, * $p < 0.05$, ns - no significant difference).



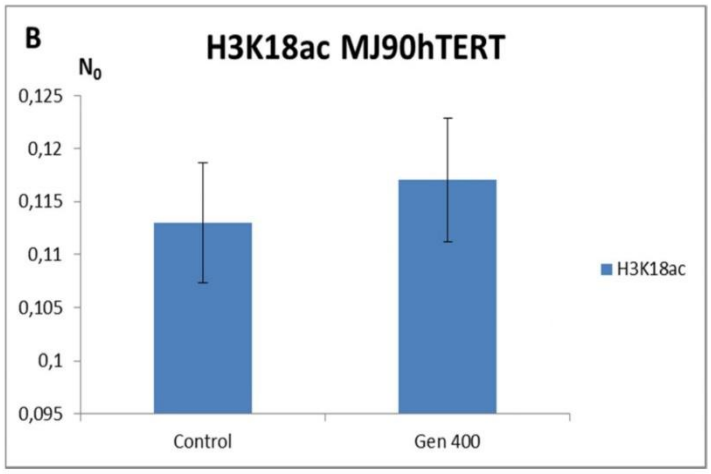


Figure S3. (A) H3K9me3 level at tandemly arranged alpha satellite arrays in: HeLa cells after treatment with geneticin (300 µg/ml), A-1235 cells after rifampicin (50 µM) treatment, A-1235 and MJ90hTERT cells after geneticin treatment (400 µg/ml). (B) H3K18Ac level at tandemly arranged alpha satellite arrays in MJ90hTERT cells after geneticin treatment (400 µg/ml). All treatments were 48 hrs.

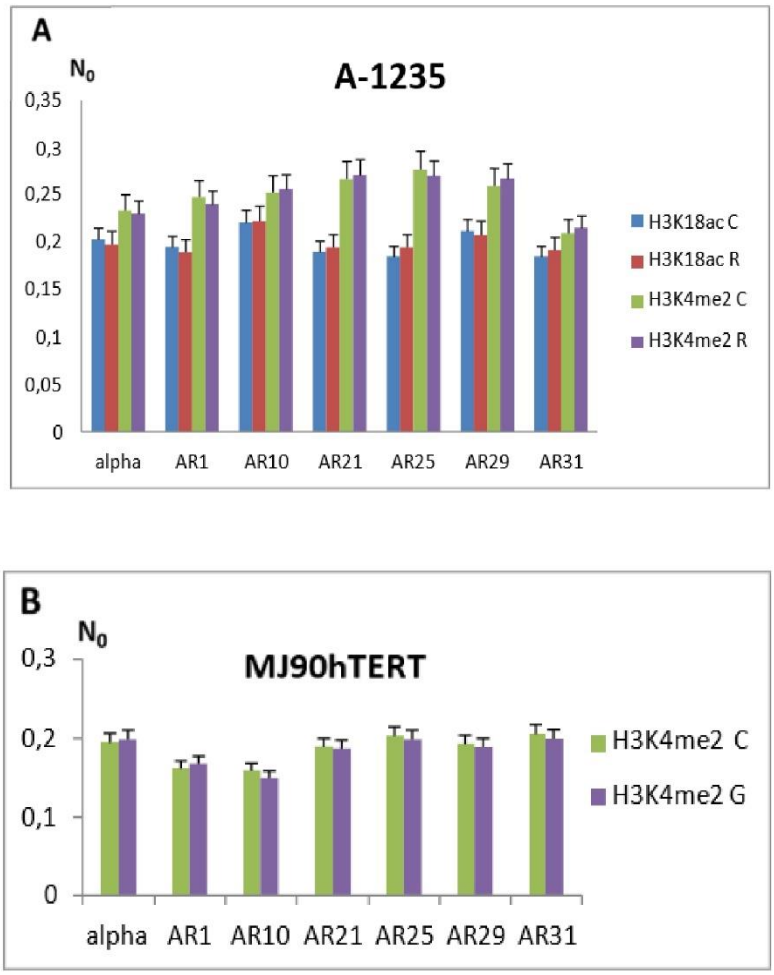


Figure S4. (A) H3K18Ac and H3K4me2 levels at tandemly arranged alpha satellite repeats and dispersed alpha repeats (AR) after treatment of A-1235 cells with rifampicin (82 µg/ml); (B) H3K4me2 levels of MJ90hTERT cells treated with geneticin (600 µg/ml). All treatments were 48 hrs. C denotes control.

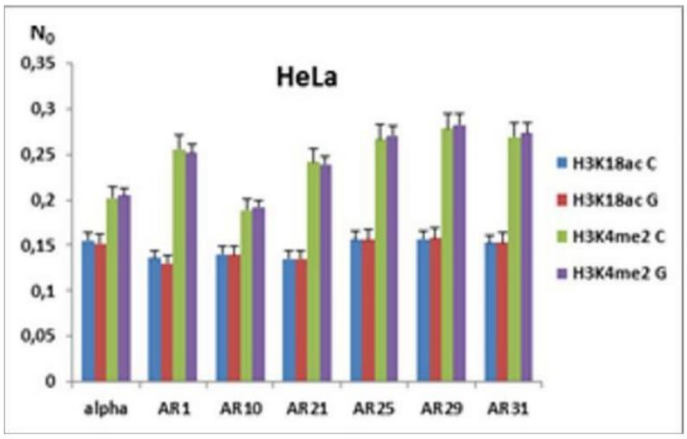


Figure S5. H3K18ac and H3K4me2 levels in HeLa cells after treatment with geneticin (400 µg/ml).

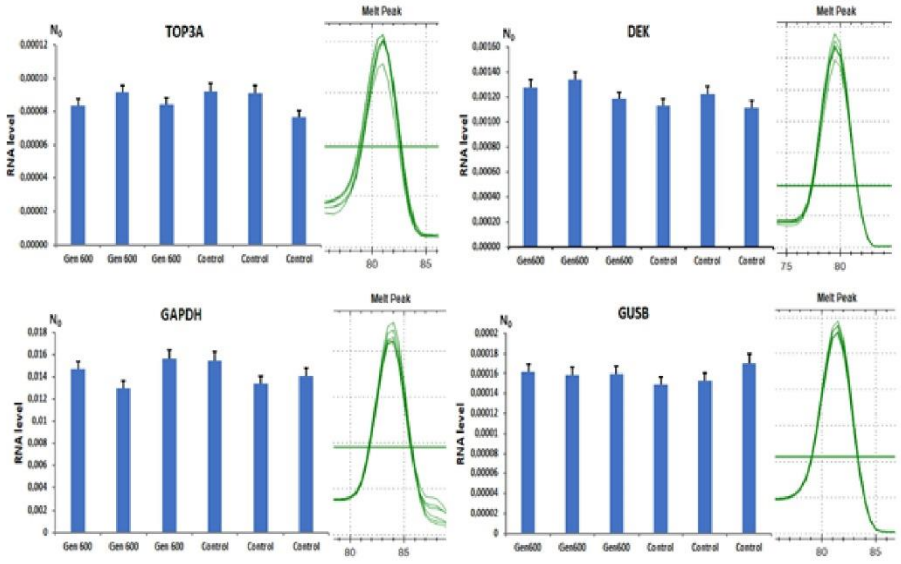


Figure S6. Expression profiles of four housekeeping genes in MJ90hTERT cell line after 48 hr treatment with geneticin (600 µg/ml) and untreated controls with associated melting curves. Three independent RT-qPCR experiments were performed. Error bars represent standard deviations and averaged N_0 values are expressed in arbitrary fluorescence units. No significant differences in gene expression were observed between treated samples and controls in all cases (Student's t-test, P>0.1).

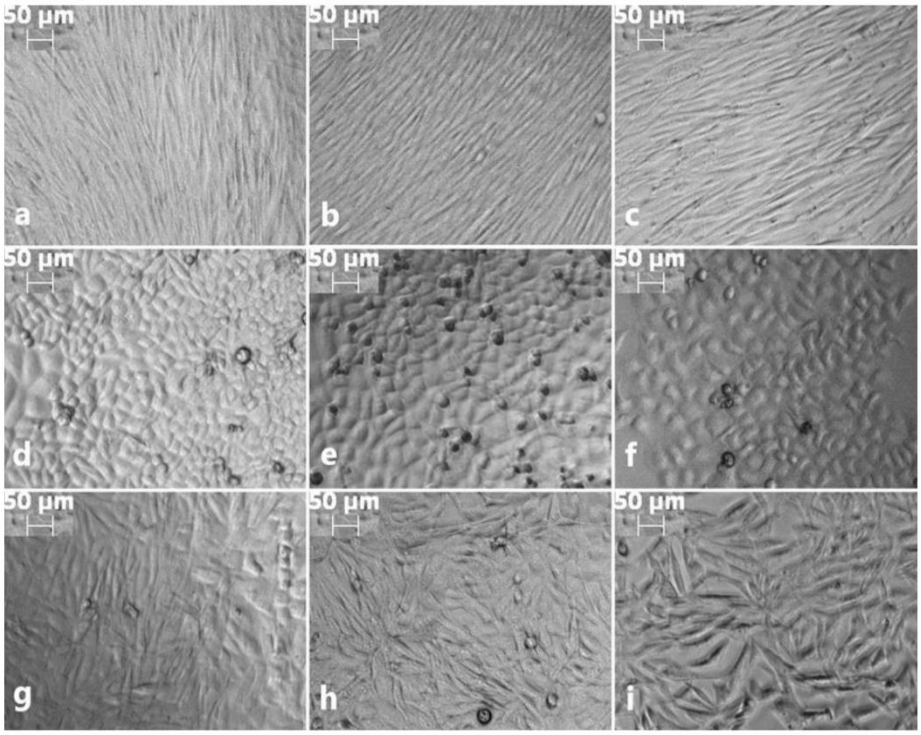


Figure S7. Microscopic imaging of investigated cell lines after treatment with antibiotics as well as of untreated controls. (a-c) MJ90hTERT cells. Untreated controls, cells treated with geneticin 600 µg/ml and hygromycin B 100 µg/ml, respectively. (d-f) HeLa cells. Untreated controls, cells treated with geneticin 400 µg/ml and hygromycin B 50 µg/ml, respectively. (g-i) A1235 cells. Untreated controls, cells treated with geneticin 400 µg/ml and rifampicin 82 µg/ml, respectively.

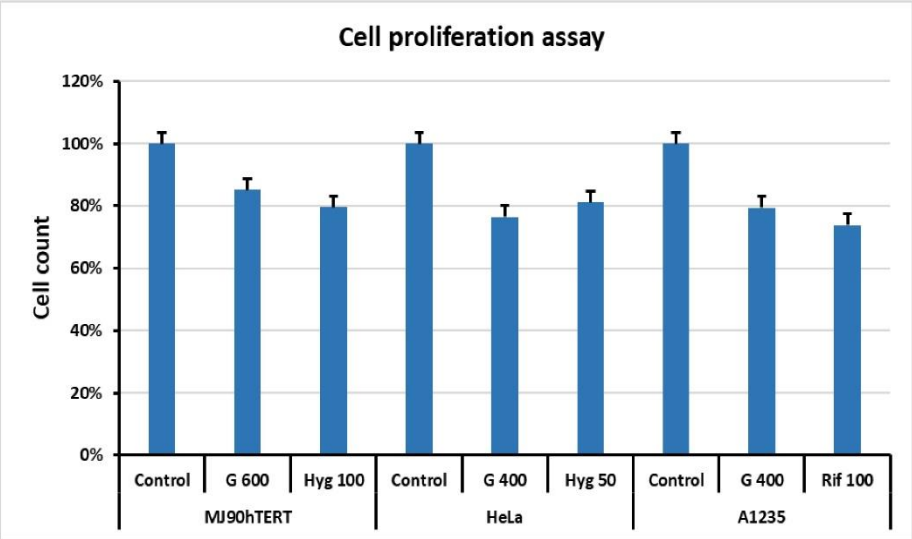


Figure S8. Cell counts after treatment of 3 tested cell lines with various antibiotic concentrations, expressed as percentages of untreated controls. Error bars represent standard deviations after 3 independent quantifications. G 600 - geneticin 600 µg/ml; G 400 - geneticin 400 µg/ml; Hyg 100 - hygromycin B 100 µg/ml; Hyg 50 - hygromycin B 50 µg/ml; Rif 100 - rifampicin 82 µg/ml.

Table S1. List of primers used in ChIP-qPCR experiments. Sequence of alpha satellite modified primer used for transcriptional analysis of alpha satellite DNA is also included and modified bases are marked in bold.

Alpha repeat No	Primers fw	Primers rev
1	GCCTCCTGAGTTCAAGCAAC	GCATGGTGCCTCATTCCTAT
10	ACAGGGATATGAAACTCTAC	AAGGAAGAACTGGATGGG
21	GCACCACACTGGTCTTCC	CTGTATTTCAGCATGAGTGACAG
25	CGACATGGGTCCAGTCTGAT	AGGAATCTGGATATGTCTTCCA
29	GCTTACCTTGGCTTCTCAAAGT	TAGTAGACACCGGGTTTCACC
31	CGCTAATGCTGAAGACATGC	GAAGTGGAATGGGATCTGAAA
Alpha satellite	CACTCTTTTTGTAGAATCTGC	AATGCACACATCACAAAGAAG
Alpha sat modified		AATGCACATATCAC**T**AT**G**TA**C**

Table S2. List of housekeeping genes used as endogenous controls for gene expression normalization with corresponding primer pairs for amplification.

Gene	Primers fw	Primers rev
GUSB	GAAAATACGTGGTTGGAGAGCTCATT	CCGAGTGAAGATCCCCTTTTTA
GAPDH	CCACTCCTCCACCTTTGAC	ACCCTGTTGCTGTAGCCA
TOP3A	TCGACTCTTTAACCACACGG	AGATCTGACCTCTACCACAG
DEK	ATGTGGGTCAGTTCAGTGGC	CCAGAAGGCTTTGGATGCAT

PUBLICATION III.

Ljubić S, Sermek A, Prgomet Sečan A, Prpić M, Jakšić B, Murgić J, Fröbe A, Ugarković Đ, Feliciello I (2022) Alpha Satellite RNA Levels Are Upregulated in the Blood of Patients with Metastatic Castration-Resistant Prostate Cancer. Genes. 13:383.

genes

MDPI

Article

Alpha Satellite RNA Levels Are Upregulated in the Blood of Patients with Metastatic Castration-Resistant Prostate Cancer

Sven Ljubić [1], Antonio Sermek [1], Angela Prgomet Sečan [2], Marin Prpić [2,3], Blanka Jakšić [2], Jure Murgić [2], Ana Fröbe [2,3,*], Đurđica Ugarković [1] and Isidoro Feliciello [1,4,*]

[1] Department of Molecular Biology, Ruđer Bošković Institute, Bijenička 54, HR-10000 Zagreb, Croatia; sven.ljubic@irb.hr (S.L.); antonio.sermek@irb.hr (A.S.); ugarkov@irb.hr (Đ.U.)
[2] Department of Oncology and Nuclear Medicine, University Clinical Hospital Centre (UHC) Sestre Milosrdnice, Vinogradska cesta 29, HR-10000 Zagreb, Croatia; angela.prgomet.secan@kbcsm.hr (A.P.S.); marin.prpic@kbcsm.hr (M.P.); blanka.jaksic@kbcsm.hr (B.J.); jure.murgic@kbcsm.hr (J.M.)
[3] School of Dental Medicine, University of Zagreb, Gundulićeva 5, HR-10000 Zagreb, Croatia
[4] Dipartimento di Medicina Clinica e Chirurgia, Universita' degli Studi di Napoli Federico II, Via Pansini 5, 80131 Napoli, Italy
* Correspondence: afrobe@irb.hr (A.F.); ifelicie@unina.it (I.F.); Tel.: +385-1-378-7510 (A.F.); +39-081-746-4317 (I.F.)

Abstract: The aberrant overexpression of alpha satellite DNA is characteristic of many human cancers including prostate cancer; however, it is not known whether the change in the alpha satellite RNA amount occurs in the peripheral tissues of cancer patients, such as blood. Here, we analyse the level of intracellular alpha satellite RNA in the whole blood of cancer prostate patients at different stages of disease and compare it with the levels found in healthy controls. Our results reveal a significantly increased level of intracellular alpha satellite RNA in the blood of metastatic cancers patients, particularly those with metastatic castration-resistant prostate cancer relative to controls. In the blood of patients with localised tumour, no significant change relative to the controls was detected. Our results show a link between prostate cancer pathogenesis and blood intracellular alpha satellite RNA levels. We discuss the possible mechanism which could lead to the increased level of blood intracellular alpha satellite RNA at a specific metastatic stage of prostate cancer. Additionally, we analyse the clinically accepted prostate cancer biomarker PSA in all samples and discuss the possibility that alpha satellite RNA can serve as a novel prostate cancer diagnostic blood biomarker.

Keywords: alpha satellite DNA; transcription; alpha satellite RNA; prostate cancer; blood biomarker

Citation: Ljubić, S.; Sermek, A.; Prgomet Sečan, A.; Prpić, M.; Jakšić, B.; Murgić, J.; Fröbe, A.; Ugarković, Đ.; Feliciello, I. Alpha Satellite RNA Levels Are Upregulated in the Blood of Patients with Metastatic Castration-Resistant Prostate Cancer. *Genes* **2022**, *13*, 383. https://doi.org/10.3390/genes13020383

Academic Editor: Selvarangan Ponnazhagan

Received: 20 January 2022
Accepted: 18 February 2022
Published: 20 February 2022

1. Introduction

Satellite DNAs are tandemly repeated sequences predominantly located within constitutive heterochromatin which is positioned in the (peri)centromeric and subtelomeric regions of chromosomes. Alpha satellite DNA is the major human satellite DNA that makes up 3–5% of each human chromosome and is composed of basic units based on divergent, 171 bp-long monomers [1]. Alpha satellite DNA contributes to essential chromosomal functions such as the formation of the centromere/kinetochore as well as of constitutive heterochromatin [2]. The heterochromatin structure is defined by epigenetic modifications, in particular by H3K9me3, which is catalysed by the histone methyltransferase SUV39H1 [3]. Of particular importance is the role of alpha satellite DNA transcripts whose interaction with the enzyme SUV39H1 is necessary for the proper formation and regulation of heterochromatin not only in the pericentromeric regions [4], but also at alpha satellite repeats dispersed within euchromatin [5].

Numerous diseases including cancers can result from deregulated epigenetic mechanisms, which also affect the structure of pericentromeric heterochromatin and the expression of sequences located therein. For instance, the lysine-specific demethylase 2A

(KDM2A)—which is specific for H3K36—is a tumour suppressor and is downregulated in prostate cancer [6]. The lower the level of *KDM2A* expression, the higher pericentromeric heterochromatin transcription and the more severe the tumour grade in prostate cancer [6]. The aberrant overexpression of sequences within pericentromeric heterochromatin in which the satellite DNAs including alpha satellite predominate is not only characteristic of prostate cancer but also of many other epithelial cancers such as those of the pancreas, lung, kidney, and colon [7]. The expression of satellite DNA in tumours can also occur due to a deficiency of other tumour suppressors such as p53 or BRCA1 which disrupts the integrity of constitutive heterochromatin and results in an extremely high expression of satellite DNAs [8,9]. Increased levels of satellite RNA destabilise the replication fork and genome integrity and further promote tumour transformation [10]. It should be mentioned that the overexpression of human satellite DNAs is not only characteristic of tumours but has also been observed in replicative senescence and aging correlating with the loss of H3K9me3 at the pericentromeric heterochromatin [11,12]. In addition, the strong upregulation of satellite DNA expression occurs upon heat stress [5,13–16].

Although the aberrant overexpression of alpha satellite DNA in many human cancers including prostate cancer has been described, it is not known whether changes in alpha satellite RNA amounts occur in the peripheral tissues of cancer patients, such as blood, and whether they could be used as potential cancer biomarkers. For the diagnosis of prostate cancer, the most clinically accepted biomarker is prostate-specific antigen (PSA). PSA is a kallikrein-related serine protease produced by prostate epithelial cells, whose levels are usually elevated in prostate cancer patients. The introduction of PSA testing in asymptomatic men has resulted in the earlier detection of the disease, with a reduction in the percentage of men with metastatic prostate cancer. PSA also allows for the early detection of latent prostate cancer that often does not develop into a significant disease and is often elevated under benign conditions such as inflammation or hyperplasia, and this lack of PSA specificity results in over-diagnosis [17]. PSA is also used to monitor the development of the disease, although often its level, especially in patients with metastatic prostate cancer, neither closely correlates with the stage of the disease nor with hormone-sensitivity. Due to the insufficient specificity of PSA as a diagnostic and prognostic marker, additional efforts are being made to find alternative biomarkers for prostate cancer [18].

To gain insight into the possible use of alpha satellite RNA as a prostate cancer biomarker, we analysed the level of intracellular alpha satellite RNA in the blood cells of cancer prostate patients and compared it with the levels of healthy controls. Our results reveal that the level of alpha satellite RNA is significantly upregulated in the blood of patients with metastatic, castration-resistant prostate cancer relative to healthy controls, indicating a link between prostate cancer pathogenesis and intracellular alpha satellite RNA levels in blood. In the blood of patients with metastatic hormone-sensitive prostate cancer, an increase in alpha satellite RNA levels was detected but not statistically significant. Testing at other stages of disease revealed no significant change of alpha satellite levels relative to controls. The clinically accepted prostate cancer biomarker PSA was also determined in all blood samples and obtained values compared to those of alpha satellite RNA levels. We discuss a possible mechanism which could be related to the increased level of intracellular alpha satellite RNA in the blood of patients at a specific stage of prostate cancer as well as the possibility that alpha satellite RNA can serve as an indicator of a particular pathophysiological state or as a potential novel prostate cancer diagnostic blood biomarker.

2. Materials and Methods

2.1. Sample Collection

The blood samples of prostate cancer patients were provided and collected by Prof. Dr. Sc. Ana Fröbe from the University Clinical Hospital Centre (UHC) Sestre Milosrdnice. The blood samples of healthy controls were provided by the Croatian Institute for Transfusion Medicine (HZTM). Informed consent was acquired from each participating individual

before blood collection. Ethical approval was obtained from the Medical Ethical Committees of the UHC Sestre Milosrdnice and of the Croatian Institute for Transfusion Medicine.

We collected 2.5 mL of blood from each cancer prostate patient and healthy control in a "PAXgene Blood RNA Tube" (Qiagen) according to the instructions of the manufacturer which was stored at −20 °C. This study comprised a total of 94 patients with prostate cancer diagnosis and 27 healthy controls. Since 50% of men older than 50 years have benign prostate hyperplasia [19], which can result in an increased PSA level, we used men with an average age of 39.4 years as the healthy control group. Enrolled patients belong to different groups according to the stage of disease: A—metastatic hormone-sensitive; B—metastatic castration-resistant; and C—localised hormone-sensitive. All patients in groups A–C received androgen-deprivation therapy by luteinising hormone-releasing hormone (LHRH) agonists. Group D included patients with newly diagnosed localised prostate cancer before receiving a hormone or any local treatment.

Control subjects were male blood donors with no personal history of prostate cancer or any other chronic disease and at the time of blood collection—they were not on a drug therapy. Other patients' and controls' characteristics that were documented included age. The characteristics of the patients and controls are presented at Table 1.

Table 1. Groups of individuals used in this study with the following characteristics: number of samples of each group (n); average age of patients (y); and the minimal and maximal age of individuals of each group.

Group	Characteristics
Healthy controls	
n	27
average age/y	39.4
age min–max/y	19–59
Group A patients	
n	19
average age/y	74.5
age min–max/y	62–85
Group B patients	
n	20
average age/y	67.4
age min–max/y	51–87
Group C patients	
n	34
average age/y	69.9
age min–max/y	57–83
Group D patients	
n	21
average age/y	71.9
age min–max/y	47–83

2.2. RNA Isolation and Reverse Transcription

Intracellular RNA from whole blood collected in the PAXgene Blood RNA Tubes was isolated using the PAXgene Blood RNA Kit (Qiagen, Hilden, Germany) according to the instructions of the manufacturer. RNA was quantified with the Quant-IT RNA assay kit using a Qubit fluorometer (Invitrogen, Waltham, MA, USA)). Approximately 1 µg of RNA was reverse transcribed using the PrimeScript RT reagent Kit with gDNA Eraser (Takara, Dalian, China) in 10 µL reaction using the specific primer ALPHrev AATGCACACATCACAAAGAAG. For all samples, negative controls without reverse transcriptase were used.

2.3. Quantitative Real-Time PCR (qPCR) Analysis

qPCR analysis was performed according to the previously published protocol [5,15]. Primers for the expression analysis of human alpha satellite DNA were: ALPHAfw CACTCTTTTTGTAGAATCTGC and ALPHrev AATGCACACATCACAAAGAAG, which were constructed according to the alpha satellite consensus sequence [20]. *Glucuronidase β (GUSB)* [21] was used as an endogenous control for normalisation in human samples and the primers used were: GUSfw GAAAATATGTGGTTGGAGAGAGCATT, GUSrev CCGAGTGAAGATCCCCTTTTTA. *GUSB* gene (Gene ID: 2990) was stably expressed without any variation among samples. The following thermal cycling conditions were used: 50 °C 2 min; 95 °C 7 min; 95 °C 15 s; 60 °C 1 min for 40 cycles followed by dissociation stage: 95 °C for 15 s; 60 °C for 1 min; 95 °C for 15 s; and 60 °C for 15 s. Amplification specificity was confirmed by dissociation curve analysis and the specificity of amplified products was tested on agarose gel. Control without template (NTC) was included in each run. Post-run data were analysed using LinRegPCR software v.11.1. [22,23] which enables the calculation of the starting concentration of amplicon ("no value"). No value is expressed in arbitrary fluorescence units and is calculated by considering PCR efficiency and baseline fluorescence. "No value" determined for each technical replicate was averaged and the averaged "no values" were divided by the "no values" of the endogenous control. The statistical analysis of qPCR data was performed using GraphPad v.6.01.

2.4. Determination of PSA Values in Blood

Peripheral blood samples were collected by the venepuncture of cubital vein in the sitting position punctured by one laboratory staff person, in 6 mL Vaccuete® serum tubes with clot activator (red cap), (Greiner Bio-One, Kremsmünster, Austria) according to national recommendations for venous blood sampling. Serum samples were centrifuged within 4 h and the analyses of PSA were made within 2 h of centrifugation. Blood for serum testing was centrifuged for 10 min at 2150× *g* at 4 °C on Hettich ROTINA35 centrifuge (Hettich, Germany). Sera samples were measured on Roche Cobas e601 (Roche Diagnostics GmbH, Mannheim, Germany) automated immunochemistry analyser with analytical principle of electrochemiluminescence reaction, using original Roche assays, calibrators, and controls.

2.5. Statistical Analyses

The Shapiro–Wilk test was used to test data normality. Statistical differences in PSA values as well as in the alpha RNA level in four groups of patients and controls were tested using Kruskal–Wallis test. The alpha satellite RNA level of patients belonging to four groups and of the control group were displayed in boxplots and were tested for statistical significance using the parametric 2-tailed Welch's *t* test if the data had normal distribution (controls, groups A and B) and non-parametric Mann–Whitney test if the data had non-normal distribution (groups C and D). The correlation between the alpha satellite RNA and PSA levels was assessed using Spearman's rank correlation. The diagnostic potential of the alpha satellite RNA level for distinguishing metastatic prostate cancer patients and controls were evaluated by computing receiver operating characteristic (ROC) curves and the results were quantified by the area under the curve (AUC) in the pROC [24] package. The statistical analyses were performed using R software [25] and graphs were created using ggplot2 [26] package. *p*-values less than 0.05 were considered statistically significant.

3. Results

3.1. Alpha Satellite RNAs Level in the Blood of Prostate Cancer Patients—qPCR Analysis

We isolated intracellular RNA from whole blood which was collected from prostate cancer patients belonging to four groups representing different stages of disease: A—representing patients with metastatic hormone-sensitive prostate cancer; B—patients with metastatic castration-resistant cancer; C—patients with localised hormone-sensitive prostate cancer; and D which included patients with newly diagnosed localised prostate cancer before receiving hormone or any other treatment. In contrast to group D, patients from groups A–C

were all under hormone treatment (LHRH agonists). In parallel, we isolated intracellular RNA from the control groups which included 27 healthy male individuals. The number of samples of each group, average age as well as age range is shown in Table 1.

To measure the level of alpha satellite RNA in the total intracellular RNA isolated from the whole blood of patients from four different groups (A–D) as well as from healthy controls, we used quantitative real-time PCR (qPCR) analysis. The obtained qPCR results (Supplementary File S1) were analysed by Kruskal–Wallis test which is used to analyse the differences among multiple groups of samples: control and four groups of patients, and it revealed significant difference ($p = 1.4 \times 10^{-4}$). The 2-tailed Welch's t test and Mann–Whitney test were used to see the difference among the pairs of samples and its statistical significance. The results reveal the increased level of alpha satellite RNA in two groups of metastatic prostate cancer patients relative to the control group (Figure 1). An increase of 2.8× with significant statistical support ($p = 2.7 \times 10^{-4}$) is characteristic of group B—with metastatic castration-resistant prostate cancer—while for group A—with metastatic hormone-sensitive prostate cancer—the increase is 1.4× but not statistically significant ($p = 0.11$). Within the other two groups of localised prostate cancers, namely C and D, there is no statistically significant difference in the level of alpha satellite RNA relative to the control ($p > 0.05$). The similarity in the alpha satellite RNA levels in groups C and D suggest that the level is not affected by drug treatment. There is also a significant increase in alpha satellite RNA level in group B relative to group A—metastatic hormone-sensitive prostate cancer of 2.0× ($p = 4 \times 10^{-3}$) as well as to groups C and D of 2.9× ($p = 4 \times 10^{-6}$) and 1.7× ($p = 0.017$), respectively.

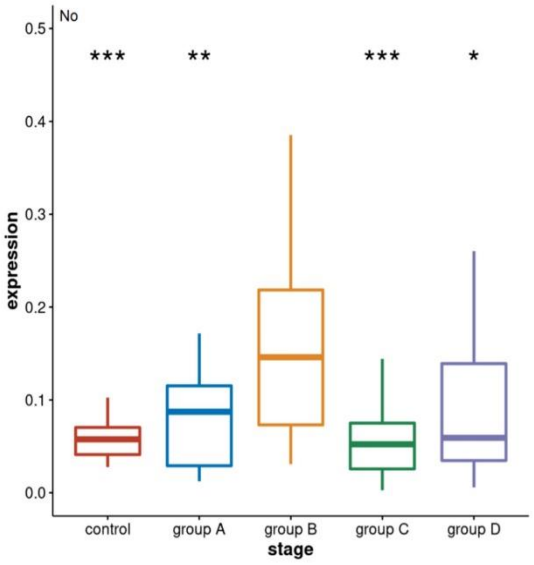


Figure 1. The level of intracellular alpha satellite RNA from the whole blood of control healthy individuals and of prostate cancer patients belonging to groups: A—metastatic hormone-sensitive on treatment; B—metastatic castration-resistant on treatment; C—localised hormone-sensitive on treatment; D—localised hormone-sensitive before any treatment. RNA level is obtained by RT-qPCR and the normalised average no value for each sample was used. Differences are analysed by 2-tailed Welch's t test for groups A, B, and the control, and Mann–Whitney for group C and group D; median values are indicated and error bars represent standard deviations. Statistically significant differences of group B relative to groups A, C, D, and the controls, respectively, are indicated by stars (*** denotes $p < 10^{-3}$, ** $p < 10^{-2}$, * $p < 0.05$).

The results show that the level of alpha satellite RNA can be used to distinguish between different stages of disease: metastatic castration-resistant relative to the metastatic castration-sensitive as well as metastatic castration-resistant relative to localised prostate cancer and to the healthy controls, respectively, and could serve as a potential diagnostic biomarker of metastatic state, particularly of castration-resistant metastatic prostate cancer. ROC curves and the calculation of AUC values (Figure 2) reveals that based on alpha satellite RNA levels, metastatic castration-resistant prostate cancer can be discriminated with high accuracy from primary localised tumours (AUC 0.85) and controls (AUC 0.85). Discrimination between metastatic castration-resistant and metastatic hormone-sensitive prostate cancer (B vs. A) is also acceptable (AUC 0.74).

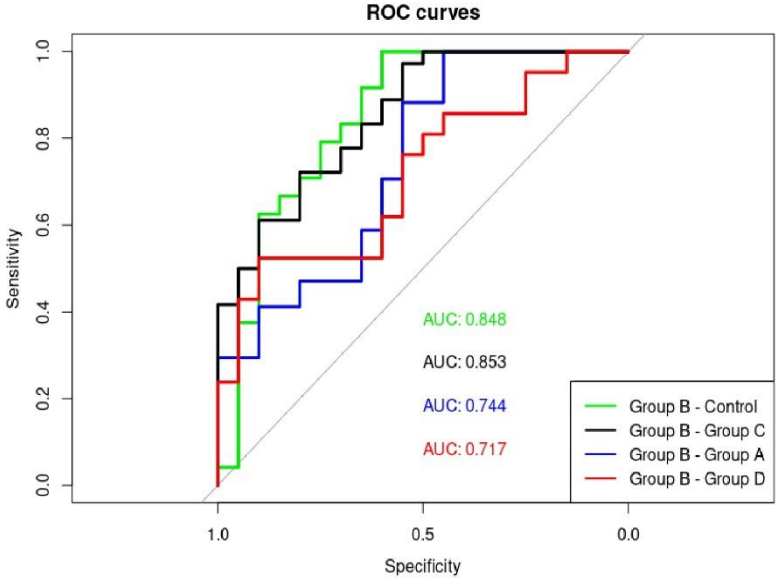


Figure 2. The diagnostic potential of alpha satellite RNA levels is determined by computing ROC curves and quantifying AUC values. The alpha RNA level shows the highest discriminatory power for distinguishing group B metastatic castration-resistant prostate cancer from: controls (AUC 0.848); group C (AUC 0.853); group A (AUC 0.744); and group D (AUC 0.717).

3.2. PSA Values in Four Groups of Prostate Cancer Patients

The most clinically accepted biomarker for prostate cancer is prostate-specific antigen (PSA) which is a kallikrein-related serine protease produced by prostate epithelial cells. PSA levels are usually elevated in prostate cancer patients. We checked PSA values in the blood of four groups of patients (A–D) as well as controls and performed statistical analyses of log PSA values by Kruskal–Wallis test which is used to analyse differences among multiple groups of samples: control and four groups of patients, and it revealed significant difference ($p = 1.7 \times 10^{-11}$). The Mann–Whitney test revealed no significant difference between the PSA level in group B relative to group A ($p = 0.9074$) and to group D ($p = 0.5063$), while the PSA level in the three groups A, B, and D is significantly increased relative to the control and to group C ($p < 10^{-4}$). PSA values between control and group C are not significantly different ($p = 0.300$; Figure 3).

The ROC curve analysis of PSA levels revealed discrimination between the controls and group D corresponding to an AUC value of 0.912 and between controls and two groups of metastatic cancers A and B with AUC values of 0.8052 and 0.9256, respectively. However, discrimination between metastatic hormone-sensitive (group A) and metastatic castration-resistant (group B) was low with an AUC of 0.512, revealing the much better performance

of alpha satellite RNA (AUC 0.744) than PSA in discriminating two stages of metastatic prostate cancer. The correlation between alpha satellite RNA level and PSA level was also assessed in each group of patients (A–D) using Spearman's rank correlation but no statistically significant correlation was found in any of the group (group A: r = −0.098, p = 0.7084; group B: r = 0.3978, p = 0.1602; group C: r = −0.2643, p = 0.1659; group D: r = −0.1056, p = 0.6968).

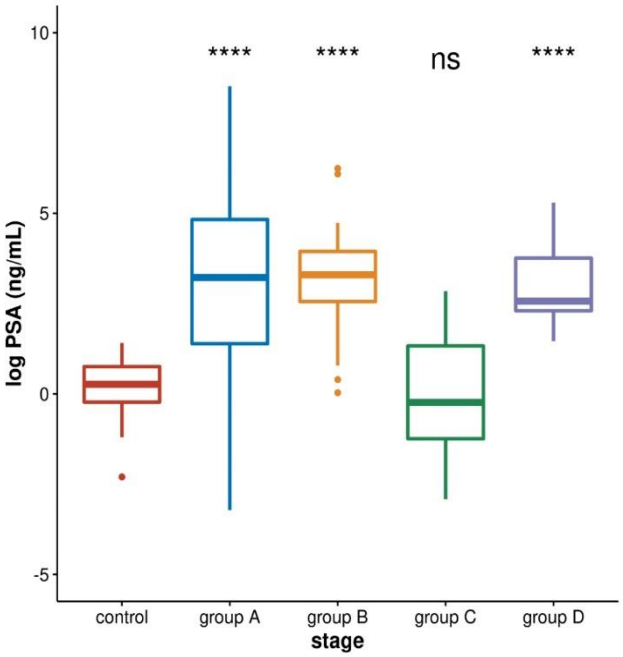


Figure 3. LogPSA values in the blood of control healthy individuals and of prostate cancer patients belonging to groups: A—metastatic hormone-sensitive on treatment; B—metastatic castration-resistant on treatment; C—localised hormone-sensitive on treatment; D—localised hormone-sensitive before any treatment. Differences between groups are analysed by Mann–Whitney and Kruskal–Wallis statistical tests, and median values are indicated and error bars represent standard deviations. Statistically significant differences of control relative to groups A, B. and D, respectively, are indicated by stars (**** denotes $p < 10^{-4}$, ns means not significant).

4. Discussion

Metastatic castration-resistant prostate cancer (mCRPC) is a clinical state in the trajectory of prostate cancer evolution characterised by disease progression despite patient castration of testosterone. Fundamentally, prostate cancer cells are exquisitely sensitive to testosterone suppression achieved by androgen-deprivation therapy via LHRH agonists. The hormone-sensitive state can last several years even in the presence of metastatic disease and is characterised by decreased or stable PSA and the resolution of metastatic lesions on imaging studies. However, castration resistance ultimately emerges as a consequence of the strong selective pressure of hormonal therapy exerted on prostate cancer cells [27]. In the spectrum of prostate cancer clinical course, mCRPC represents the final and incurable stage of disease with a survival rate of less than 3 years despite recent therapeutic breakthroughs. The key underlying mechanism of castration resistance is centred around the preserved activity of androgen receptor (AR) signalling, a crucial target of novel therapies for mCRPC. Different means of AR autonomy and therapeutic escape in mCRPC include AR protein overexpression, gene amplification, and/or AR mutations which all lead to

the uninterrupted transduction of the AR signal and the intratumoural production of androgens [28–30].

At the genomics level, there are a number of differences between mCRPC and localised or hormone-sensitive metastatic prostate cancer. In addition to prevalent AR mutations, tumour samples from patients with mCRPC are enriched for mutations in *TP53*, DNA damage repair genes, *RB1* and *PTEN*, contributing to their loss-of-function and overall genomic instability [31,32]. Following that path, the discovery of targetable genetic defects in mCRPC opened a new avenue of targeted treatment, i.e., the use of PARP inhibitors for BRCA-mutated prostate cancer which is the most prevalent genomic event in mCRPC [33]. Recently, the co-existence of androgen-dependent and androgen-independent pathways was discovered in mCRPC, explaining the limited therapeutic efficacy of novel androgen pathway-targeted therapies in the general population of patients with mCRPC [34]. Finally, mCRPC is characterised by a large number of circulating tumour cells (CTCs) in the patient's blood. Conversely, the presence of CTC in localised prostate cancer is considered to be an exceptional event. A less cohesive microenvironment of metastatic deposits in mCRPC facilitates the shedding of both individual tumour cells and cell-free DNA in the blood stream. Similar mechanisms may apply to the identification of alpha satellite RNA in the blood of patients with mCRPC.

The increased expression of pericentromeric satellite DNAs such as satellite II and alpha satellite DNA is characteristic of epithelial cancers including prostate cancer [7,35] as well as for hematopoietic malignancies [36]. Here, we observe significantly increased levels of alpha satellite RNA in the blood of patients with metastatic castration-resistant prostate cancer as well as an increase in patients with metastatic hormone-sensitive prostate cancer, while in the blood of patients with primary localised prostate cancer, no significant change relative to the controls was detected. One of the possible explanations for the increase in the alpha satellite RNA level in the blood of prostate cancer patients in metastatic stages of disease might be related to the transfer of alpha satellite RNA from prostate cancer cells to blood cells mediated by exosomes. The exosomes are a class of extracellular vesicles released by all cells, often detected in tumour microenvironments, which remove excess and/or unnecessary constituents from cells including harmful RNA and DNA [37,38]. Exosomes can transfer RNA or DNA which they contain to other cells, and in addition they can activate various signalling pathways in cells they fuse or interact with [38,39]. We propose that excess satellite RNA from prostate cancer can be transferred and delivered by exosomes to blood cells resulting in an increase in the total RNA level in blood cells. In addition, the interaction of exosomes with blood cells might activate some signalling pathways which could affect the heterochromatin structure and expression of satellite sequences located therein. In addition to the lysine-specific demethylase 2A (KDM2A) whose downregulation affects chromatin in prostate cancer [6], there are other (hetero)chromatin modifiers such as sirtuins, a family of NAD^+-dependent deacetylases which coordinate cellular responses to different types of stress and are key players in the protection and maintenance of genomic integrity [40]. Of particular importance for the preservation of constitutive heterochromatin structure are the sirtuins SIRT1 and SIRT6 which maintain the epigenetic silencing of repetitive elements [41] by regulating the activity of the histone methyl-transferase SUV39H1 responsible for the spreading of the silencing mark H3K9me3 [40,42]. In addition, SIRT6 also deacetylates histone H3K18ac in pericentric heterochromatin and its depletion results in the overexpression of pericentromeric repeats [43]. It could be proposed that changes in the activation of KDM2A, sirtuins SIRT1 and SIRT6 or of some other (hetero)chromatin modifiers might arise not only in the prostate cancer cells [44] but possibly mediated by exosomes in the blood of prostate cancer patients at specific stages, affecting heterochromatin structure and the expression of repeats located therein. It is also possible that, as mentioned above, circulating tumour cells (CTCs) which are present in the blood of patients with metastatic prostate cancer [45] could precipitate with blood cells, contributing to the increased level of alpha satellite RNA, although due to the low number of CTCs relative to blood cells, this contribution is probably not significant.

As revealed by our results, the alpha satellite RNA level was able to discriminate metastatic hormone-sensitive from metastatic castration-resistant prostate cancer (groups A and B) as well metastatic castration-resistant cancer under treatment (B) from newly diagnosed localised prostate cancer before receiving a hormone or any local treatment (group D), from localised prostate cancer under treatment (group C) and from controls. On the other hand, PSA has high discrimination power for distinguishing controls from localised cancer before treatment (group D) as well as from metastatic cancers under treatment but cannot distinguish between metastatic hormone-sensitive and metastatic castration-resistant prostate cancers (groups A and B). Based on our investigation, the alpha satellite RNA level can complement PSA as a biomarker for monitoring the progression of metastatic prostate cancer and for the diagnosis of metastatic castration-resistant stage of disease. Considering the possible use of satellite RNA as a cancer biomarker, a circulating satellite RNA level in blood serum quantified by the sensitive method of tandem repeat amplification by nuclease protection (TRAP) combined with droplet digital PCR (ddPCR) enabled the discrimination of patients with pancreatic ductal carcinoma (PDAC) from healthy controls [46]. Increased levels of circulating human satellite II in the plasma of breast, gastric, lung and bile cancers as well as sarcoma and Hodgkin's lymphoma was detected [47]. The present study reveals for the first time that not only serum or plasma-circulating satellite RNA but also alpha satellite RNA in blood cells could possibly serve as an indicator of a specific stage of cancer. In all these studies, the satellite RNA was used as a biomarker because its level is significantly increased in different cancers [7] and can be tested by quantitative real-time PCR or droplet digital PCR. Considering satellite DNA as a cancer biomarker, satellite copy number variation is characteristic for some cancers [48]; however, its detection is more complex and often requires a development of new assays [49] and technologies such as nanoplate-based digital PCR.

Further studies are necessary to explain the observed upregulation of intracellular alpha satellite RNA levels in the blood of prostate cancer patients at a specific metastatic stage and to disclose whether this phenomenon is specific to this pathological condition only.

5. Conclusions

The overexpression of satellite DNA is characteristic of many human cancers; however, it has not been investigated whether the satellite RNA level is changed in the whole blood of cancer patients. In this study, we analysed alpha satellite RNA level in the whole blood of prostate cancer patients at different stages of disease. The results reveal that the alpha satellite RNA level in the whole blood cells can discriminate castration-resistant metastatic prostate cancer from localised primary tumours and healthy controls as well as from metastatic hormone-sensitive prostate cancer with high accuracy. We discuss the possible mechanism which could result in the increased level of blood intracellular alpha satellite RNA at a specific metastatic stage of prostate cancer and propose alpha satellite RNA as a potential prostate cancer diagnostic blood biomarker.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/genes13020383/s1. File S1. Normalised no values representing the level of alpha satellite RNA obtained by RT-qPCR in blood samples belonging to groups A–D and in healthy controls.

Author Contributions: I.F., Đ.U. and A.F. conceived and designed the study. A.P.S., M.P., B.J. and J.M. performed the selection and classification of patients into different groups and the PSA assay. S.L. performed the RT-qPCR experiments and A.S. performed the statistical analyses. All authors participated in the interpretation of the results and the paper was written with the input of all authors. All authors have read and agreed to the published version of the manuscript.

Funding: The research was funded by the donation of Adris Foundation to Đ. Ugarković for the project "Alpha satellite RNA-candidate for prognostic/diagnostic prostate cancer blood biomarker" and by the Italian Ministry of Education, University and Research (MIUR), fund for Investments on

Basic Research (FIRB) and the International Staff Mobility Program of University of Naples Federico II to I. Feliciello.

Institutional Review Board Statement: This study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Committees of University Clinical Hospital Centre (UHC) Sestre Milosrdnice (protocol code EP-9251/17-8, date of approval 2 June 2017) and of Croatian Institute for Transfusion Medicine (protocol code: 003-06/21-04/01; date of approval: 22 February 2021).

Informed Consent Statement: Written informed consent was obtained from the patient(s) to publish this paper.

Data Availability Statement: Normalised no values representing the alpha satellite RNA expression level in patients belonging to four groups A–D and in healthy controls are presented in Supplementary File S1 of this paper.

Acknowledgments: We acknowledge the Croatian Institute for Transfusion Medicine for providing blood samples for healthy individuals. We also thank Mary Sopta for critical reading of the manuscript and Maria Chiara Feliciello for critical support in statistical analysis.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lee, C.; Wevrick, R.; Fisher, R.B.; Ferguson-Smith, M.A.; Lin, C.C. Human centromeric DNAs. *Hum. Genet.* **1997**, *100*, 291–304. [CrossRef] [PubMed]
2. McNulty, S.M.; Sullivan, B.A. Alpha satellite DNA biology: Finding function in the recesses of the genome. *Chromosome Res.* **2018**, *26*, 115–138. [CrossRef] [PubMed]
3. Saksouk, N.; Simboeck, E.; Déjardin, J. Constitutive heterochromatin formation and transcription in mammals. *Epigenet. Chromatin* **2015**, *8*, 3. [CrossRef] [PubMed]
4. Johnson, W.L.; Yewdell, W.T.; Bell, J.C.; McNulty, S.M.; Duda, Z.; O'Neill, R.J.; Sullivan, B.A.; Straight, A.F. RNA-dependent stabilization of SUV39H1 at constitutive heterochromatin. *Elife* **2017**, *6*, e25299. [CrossRef]
5. Feliciello, I.; Sermek, A.; Pezer, Ž.; Matulić, M.; Ugarković, Đ. Heat stress affects H3K9me3 level at human alpha satellite DNA repeats. *Genes* **2020**, *11*, 663. [CrossRef]
6. Frescas, D.; Guardavaccaro, D.; Kuchay, S.M.; Kato, H.; Poleshko, A.; Basrur, V.; Elenitoba-Johnson, K.S.; Katz, R.A.; Pagano, M. KDM2A represses transcription of centromeric satellite repeats and maintains the heterochromatic state. *Cell Cycle* **2008**, *7*, 3539–3547. [CrossRef]
7. Ting, D.T.; Lipson, D.; Paul, S.; Brannigan, B.W.; Akhavanfard, S.; Coffman, E.J.; Contino, G.; Deshpande, V.; Iafrate, A.J.; Letovsky, S.; et al. Aberrant overexpression of satellite repeats in pancreatic and other epithelial cancers. *Science* **2011**, *331*, 593–596. [CrossRef]
8. Wylie, A.; Jones, A.E.; D'Brot, A.; Lu, W.J.; Kurtz, P.; Moran, J.V.; Rakheja, D.; Chen, K.S.; Hammer, R.E.; Comerford, S.A.; et al. p53 genes function to restrain mobile elements. *Genes Dev.* **2016**, *30*, 64–77. [CrossRef]
9. Zhu, Q.; Pao, G.M.; Huynh, A.M.; Suh, H.; Tonnu, N.; Nederlof, P.M.; Gage, F.H.; Verma, I.M. BRCA1 tumour suppression occurs via heterochromatin-mediated silencing. *Nature* **2011**, *477*, 179–184. [CrossRef]
10. Zhu, Q.; Hoong, N.; Aslanian, A.; Hara, T.; Benner, C.; Heinz, S.; Miga, K.H.; Ke, E.; Verma, S.; Soroczynski, J.; et al. Heterochromatin-Encoded Satellite RNAs Induce Breast Cancer. *Mol. Cell* **2018**, *70*, 842–853. [CrossRef]
11. Enukashvily, N.I.; Donev, R.; Waisertreiger, I.S.; Podgornaya, O.I. Human chromosome 1 satellite 3 DNA is decondensed, demethylated and transcribed in senescent cells and in A431 epithelial carcinoma cells. *Cytogenet. Genome Res.* **2007**, *118*, 42–54. [CrossRef] [PubMed]
12. Smurova, K.; De Wulf, P. Centromere and Pericentromere Transcription: Roles and Regulation ... in Sickness and in Health. *Front. Genet.* **2018**, *9*, 674. [CrossRef] [PubMed]
13. Vourc'h, C.; Biamonti, G. Transcription of Satellite DNAs in Mammals. *Prog. Mol. Subcell. Biol.* **2011**, *51*, 95–118. [PubMed]
14. Pezer, Ž.; Ugarković, Đ. Satellite DNA-associated siRNAs as mediators of heat shock response in insects. *RNA Biol.* **2012**, *9*, 587–595. [CrossRef]
15. Feliciello, I.; Akrap, I.; Ugarković, Đ. Satellite DNA Modulates Gene Expression in the Beetle *Tribolium castaneum* after Heat Stress. *PLoS Genet.* **2015**, *11*, e1005466.
16. Sermek, A.; Feliciello, I.; Ugarković, Đ. Distinct Regulation of the Expression of Satellite DNAs in the Beetle *Tribolium castaneum*. *Int. J. Mol. Sci.* **2021**, *22*, 296. [CrossRef]
17. Prensner, J.R.; Rubin, M.A.; Wei, J.T.; Chinnaiyan, A.M. Beyond PSA: The next generation of prostate cancer biomarkers. *Sci. Transl. Med.* **2012**, *4*, 127rv3. [CrossRef]
18. Saini, S. PSA and beyond: Alternative prostate cancer biomarkers. *Cell Oncol.* **2016**, *39*, 97–106. [CrossRef]

19. Berry, S.J.; Coffey, D.S.; Walsh, P.C.; Ewing, L.L. The development of human benign prostatic hyperplasia with age. *J. Urol.* **1984**, *132*, 474–479. [CrossRef]
20. Choo, K.H.; Vissel, B.; Nagy, A.; Earle, E.; Kalitsis, P. A survey of the genomic distribution of alpha satellite DNA on all the human chromosomes, and derivation of a new consensus sequence. *Nucleic Acids Res.* **1991**, *19*, 1179–1182. [CrossRef]
21. Aerts, J.L.; Gonzales, M.I.; Topalian, S.L. Selection of appropriate control genes to assess expression of tumor antigens using real-time RT-PCR. *BioTechniques* **2004**, *36*, 84–91. [CrossRef] [PubMed]
22. Ruijter, J.M.; Ramakers, C.; Hoogaars, W.M.; Karlen, Y.; Bakker, O.; van den Hoff, M.J.; Moorman, A.F. Amplification efficiency: Linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Res.* **2009**, *37*, e45. [CrossRef] [PubMed]
23. Ruijter, J.M.; Pfaffl, M.W.; Zhao, S.; Spiess, A.N.; Boggy, G.; Blom, J.; Rutledge, R.G.; Sisti, D.; Lievens, A.; De Preter, K.; et al. Evaluation of qPCR curve analysis methods for reliable biomarker discovery: Bias, resolution, precision and implications. *Methods* **2013**, *59*, 32–46. [CrossRef] [PubMed]
24. Robin, X.; Turck, N.; Hainard, A.; Tiberti, N.; Lisacek, F.; Sanchez, J.; Müller, M. pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinform.* **2011**, *12*, 77. [CrossRef]
25. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2021; Available online: https://www.R-project.org/ (accessed on 20 December 2021).
26. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*; Springer: New York, NY, USA, 2016; ISBN 978-3-319-24277-4. Available online: https://ggplot2.tidyverse.org (accessed on 11 January 2022).
27. Karantanos, T.; Corn, P.G.; Thompson, T.C. Prostate cancer progression after androgen deprivation therapy: Mechanisms of castrate resistance and novel therapeutic approaches. *Oncogene* **2013**, *32*, 5501–5511. [CrossRef] [PubMed]
28. Feldman, B.J.; Feldman, D. The development of androgen-independent prostate cancer. *Nat. Rev. Cancer* **2001**, *1*, 34–45. [CrossRef]
29. Debes, J.D.; Tindal, J.D. Mechanisms of androgen-refractory prostate cancer. *N. Engl. J. Med.* **2004**, *351*, 1488–1490. [CrossRef]
30. Zhu, M.L.; Kyprianou, N. Androgen receptor and growth factor signaling crosstalk in prostate cancer cells. *Endocr. Relat. Cancer* **2008**, *15*, 841–849. [CrossRef]
31. Cancer Genome Atlas Research Network. The Molecular Taxonomy of Primary Prostate Cancer. *Cell* **2015**, *163*, 1011–1025. [CrossRef]
32. Mateo, J.; Seed, G.; Bertan, C.; Rescigno, P.; Dolling, D.; Figueiredo, I.; Miranda, S.; Rodrigues, D.N.; Gurel, B.; Clarke, M.; et al. Genomics of lethal prostate cancer at diagnosis and castration resistance. *J. Clin. Investig.* **2020**, *130*, 1743–1751. [CrossRef]
33. Pritchard, C.C.; Mateo, J.; Walsh, M.F.; De Sarkar, N.; Abida, W.; Beltran, H.; Garofalo, A.; Gulati, R.; Carreira, S.; Eeles, R.; et al. Inherited DNA-repair gene mutations in men with metastatic prostate cancer. *N. Engl. J. Med.* **2016**, *375*, 443–453. [CrossRef] [PubMed]
34. Bluemn, E.G.; Coleman, I.M.; Lucas, J.M.; Coleman, R.T.; Hernandez-Lopez, S.; Tharakan, R.; Bianchi-Frias, D.; Dumpit, R.F.; Kaipainen, A.; Corella, A.N.; et al. Androgen receptor pathway-independent prostate cancer is sustained through FGF signaling. *Cancer Cell* **2017**, *32*, 474–489. [CrossRef] [PubMed]
35. Feliciello, I.; Pezer, Ž.; Sermek, A.; Bruvo Mađarić, B.; Ljubić, S.; Ugarković, Đ. Satellite DNA-mediated gene expression regulation: Physiological and evolutionary implication. *Prog. Mol. Subcell. Biol.* **2021**, *60*, 145–168. [PubMed]
36. Onishi-Seebacher, M.; Erikson, G.; Sawitzki, Z.; Ryan, D.; Greve, G.; Lübbert, M.; Jenuwein, T. Repeat to gene expression ratios in leukemic blast cells can stratify risk prediction in acute myeloid leukemia. *BMC Med. Genom.* **2021**, *14*, 166. [CrossRef]
37. Kalluri, R. The biology and function of exosomes in cancer. *J. Clin. Investig.* **2016**, *126*, 1208–1215. [CrossRef]
38. Takahashi, A.; Okada, R.; Nagao, K.; Kawamata, Y.; Hanyu, A.; Yoshimoto, S.; Takasugi, M.; Watanabe, S.; Kanemaki, M.T.; Obuse, C.; et al. Exosomes maintain cellular homeostasis by excreting harmful DNA from cells. *Nat. Commun.* **2017**, *8*, 15287. [CrossRef]
39. Valadi, H.; Ekström, K.; Bossios, A.; Sjöstrand, M.; Lee, J.J.; Lötvall, J.O. Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nat. Cell Biol.* **2007**, *9*, 654–659. [CrossRef]
40. Bosch-Presegué, L.; Vaquero, A. Sirtuins in stress response: Guardians of the genome. *Oncogene* **2014**, *33*, 3764–3775. [CrossRef]
41. Korotkov, A.; Seluanov, A.; Gorbunova, V. Sirtuin 6: Linking longevity with genome and epigenome stability. *Trends Cell Biol.* **2021**, *31*, 994–1006. [CrossRef]
42. Santos-Barriopedro, I.; Vaquero, A. Complex role of SIRT6 in NF-κB pathway regulation. *Mol. Cell Oncol.* **2018**, *5*, e1445942. [CrossRef]
43. Tasselli, L.; Xi, Y.; Zheng, W.; Tennen, R.I.; Odrowaz, Z.; Simeoni, F.; Li, W.; Chua, K.F. SIRT6 deacetylates H3K18ac at pericentric chromatin to prevent mitotic errors and cellular senescence. *Nat. Struct. Mol. Biol.* **2016**, *23*, 434–440. [CrossRef] [PubMed]
44. Bai, L.; Lin, G.; Sun, L.; Liu, Y.; Huang, X.; Cao, C.; Guo, Y.; Xie, C. Upregulation of SIRT6 predicts poor prognosis and promotes metastasis of non-small cell lung cancer via the ERK1/2/MMP9 pathway. *Oncotarget* **2016**, *7*, 40377–40386. [CrossRef] [PubMed]
45. Galletti, G.; Portella, L.; Tagawa, S.T.; Kirby, B.J.; Giannakakou, P.; Nanus, D.M. Circulating tumor cells in prostate cancer diagnosis and monitoring: An appraisal of clinical potential. *Mol. Diagn. Ther.* **2014**, *18*, 389–402. [CrossRef] [PubMed]
46. Kishikawa, T.; Otsuka, M.; Yoshikawa, T.; Ohno, M.; Yamamoto, K.; Yamamoto, N.; Kotani, A.; Koike, K. Quantitation of circulating satellite RNAs in pancreatic cancer patients. *JCI Insight* **2016**, *1*, e86646. [CrossRef]
47. Özgür, E.; Mayer, Z.; Keskin, M.; Yörüker, E.E.; Holdenrieder, S.; Gezer, U. Satellite 2 repeat DNA in blood plasma as a candidate biomarker for the detection of cancer. *Clin. Chim. Acta* **2021**, *514*, 74–79. [CrossRef]

48. Bersani, F.; Lee, E.; Kharchenko, P.V.; Xu, A.W.; Liu, M.; Xega, K.; MacKenzie, O.C.; Brannigan, B.W.; Wittner, B.S.; Jung, H.; et al. Pericentromeric satellite repeat expansions through RNA-derived DNA intermediates in cancer. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 15148–15153. [CrossRef]
49. de Lima, L.G.; Howe, E.; Singh, V.P.; Potapova, T.; Li, H.; Xu, B.; Castle, J.; Crozier, S.; Harrison, C.J.; Clifford, S.C.; et al. PCR amplicons identify widespread copy number variation in human centromeric arrays and instability in cancer. *Cell Genom.* **2021**, *1*, 100064. [CrossRef]

PUBLICATION IV.

Đermić D, Ljubić S, Matulić M, Procino A, Feliciello MC, Ugarković Đ, Feliciello I (2023) Reverse transcription-quantitative PCR (RT-qPCR) without the need for prior removal of DNA. Scientific Reports. 13:11470.

www.nature.com/scientificreports

scientific reports

Check for updates

OPEN

Reverse transcription-quantitative PCR (RT-qPCR) without the need for prior removal of DNA

Damir Đermić[1], Sven Ljubić[1], Maja Matulić[2], Alfredo Procino[1], Maria Chiara Feliciello[3], Đurđica Ugarković[1]✉ & Isidoro Feliciello[4]✉

The procedure illustrated in this paper represents a new method for transcriptome analysis by PCR (Polymerase Chain Reaction), which circumvents the need for elimination of potential DNA contamination. Compared to the existing methodologies, our method is more precise, simpler and more reproducible because it preserves the RNA's integrity, does not require materials and/or reagents that are used for elimination of DNA and it also reduces the number of samples that should be set up as negative controls. This novel procedure involves the use of a specifically modified primer during reverse transcription step, which contains mismatched bases, thus producing cDNA molecules that differ from genomic DNA. By using the same modified primer in PCR amplification, only cDNA template is amplified since genomic DNA template is partially heterologous to the primer. In this way, amplification by PCR is unaffected by any potential DNA contamination since it is specific only for the cDNA template. Furthermore, it accurately reflects the initial RNA concentration of the sample, which is prone to changes due to various physical or enzymatic treatments commonly used by the current methodologies for DNA elimination. The method is particularly suitable for quantification of highly repetitive DNA transcripts, such as satellite DNA.

The qualitative and/or quantitative analysis of transcripts by PCR amplification is a method of choice in both molecular biology research and medical diagnostics[1]. It is widely used to detect and quantify many different types of transcripts such as messenger RNA, ribosomal RNA, non-coding RNA etc. The most common applications include gene expression analysis and precise identification of a particular microorganism.

In order to be properly analyzed, the purified RNA is subjected to a preliminary and fundamental step of reverse transcription, through which the RNA molecules are converted into cDNA (complementary DNA) by the reverse transcriptase enzyme. Indeed, the RNA itself cannot be directly amplified during the subsequent PCR steps and must necessarily be converted into cDNA[1]. The main problem of most currently used protocols lies in the often-present DNA contamination, which is impossible to chemically differentiate on a structural level from the cDNA by the polymerase enzyme during PCR amplification, thus causing false positive results[2–7]. In order to overcome this limitation, all current protocols include a couple of DNA elimination steps, both during the purification of RNA and subsequent reverse transcription step. In both cases the DNA would be eliminated, either with the aid of specific mechanical filters (silica-based columns) or through enzymatic digestion by a specific enzyme, such as DNase I (Deoxyribonuclease I)[8]. However, these treatments are not 100% effective for removal of DNA, which often remains as DNA contamination[2,3]. This is particularly the case with highly repetitive DNA which constitutes a substantial part of eukaryotic genome and is often transcribed at low levels. Furthermore, it should be emphasized that any procedure implemented to reduce the concentration of DNA in the sample certainly also causes the reduction of initial concentration of the RNA itself, a molecule which is by its very nature unstable and easily degradable.

Here we report a novel method for transcriptome analysis by PCR, wherein cDNA and DNA are differentiated and thus contamination by the latter is excluded, hence producing more precise, reliable and reproducible results.

[1]Division of Molecular Biology, Ruder Boskovic Institute, 10000 Zagreb, Croatia. [2]Division of Molecular Biology, Department of Biology, Faculty of Science, University of Zagreb, 10000 Zagreb, Croatia. [3]Department of Statistical Science, Alma Mater Studiorum, University of Bologna, 40126 Bologna, Italy. [4]Department of Clinical Medicine and Surgery, University of Naples Federico II, 80135 Naples, Italy. ✉email: ugarkov@irb.hr; isidoro.feliciello@unina.it

Materials and methods

Construction of primers. Forward, reverse and modified primers used to test the new protocol are illustrated in Table 1. The modified specific primer differs with respect to the unmodified forward and reverse primers in just four base alterations (point mutations) distributed alternatively with unchanged nucleotides and starting from the 3' end of the primer (marked in bold in Table 1).

We also tested primers that were modified in other ways (from two to six transitions/transversions at 3' end, with and without aforementioned alternation) but the best, most consistent results were obtained with 4 alternating mismatches at 3' end for 20–26 bp long primers. Of course, for longer primers more mutations can be included, but in our experience, four mismatches proved to be a necessary and sufficient number of modifications in order to achieve the expected results. In our case, the modified primers altered the complementarity between the primer and the target, while preserving the specificity and thermodynamics of primers themselves. The list of all designed primers is shown in Table 1.

RNA isolation and reverse transcription for prokaryotic gene analysis. An *Escherichia coli* strain (AB1157) used to test *ssb, sulA* and *recA* gene expression contained a multicopy plasmid, pID2, for increased expression of *ssb* gene[9]. *ssb* (coding for an essential, conserved SSB protein, which regulates DNA metabolism), *sulA* (under SOS control, coding for SulA inhibitor of cell division) and *recA* (coding for a conserved bacterial recombinase, RecA). The bacteria were grown in LB medium, with aeration, at 37 °C until they reached OD_{600} ~ 0.4 and afterwards RNA was isolated from them.

RNA was purified by RNeasy Plus Mini kit (Qiagen) which, according to the manufacturer's instructions, includes a genomic DNA elimination step by solid phase column extraction. Samples treated in such a way are indicated as + DNase I and those untreated for genomic DNA elimination are denoted as – DNase I.

Approximately 1 µg of RNA, quantified by Quant-IT RNA using a Qubit fluorometer (Invitrogen, Waltham, MA, USA), was reverse transcribed using the PrimeScript RT reagent kit without gDNA Eraser (+ RT/– DNase I) or with gDNA Eraser (+ RT/ + DNase I) (Takara, Dalian, China) using a 0.2 µM mix of specifically modified or random hexamer primers (from the kit) for *ssb*, *sulA* and *recA* expression analysis, for both the new and current method. In summary, the samples designated + DNase I had their DNA eliminated by both silica-based column during RNA isolation and by gDNA eraser (DNase I treatment) during reverse transcription step. For all samples, negative controls without reverse transcriptase enzyme (-RT) were prepared.

Nanoplate based digital PCR (dPCR) for absolute quantification of bacterial gene expression. Nanoplate based dPCR procedure was performed using the QIAcuity 2-plex instrument (Qiagen, Hilden, Germany). The dPCR reaction mixture was assembled using QIAcuity 3X Eva Green PCR Master Mix, 10X primer mix (2 µM), RNase-free water and a fixed concentration of cDNA template in a final volume of 15 µL per sample. After accurate vortexing, 12 µL of the prepared mixture was transferred onto the 24-well 8.5 K nanoplate and sealed with nanoplate rubber seal. For quality control of QIAcuity dPCR, the assay was replicated with different amounts of cDNA template input (24, 12, and 6 ng), quantified by Quant-IT ssDNA using a Qubit fluorometer (Invitrogen, Waltham, MA, USA). The sequences of primers for transcript detection of *ssb, sulA,* and *recA* genes are indicated in Table 1.

The results for all samples were obtained using 24 ng template input except for *ssb* gene expression analysis in the current method, which was performed using 60 × diluted target sample with respect to samples used for the new method due to the presence of multicopy plasmid expressing *ssb* gene (Fig. 2A).

The 8.5 K nanoplate gives rise to 8500 single partitions in which the template is distributed randomly. The QIAcuity carries out fully automated sample processing, including all necessary steps for plate priming, sealing of partitions, thermocycling and image analysis. The amplification cycling protocol consisted of 95 °C for 2 min for enzyme activation step and the following 40 cycles of 15 s. at 95 °C for denaturation, 15 s. at 60 °C for annealing, and 15 s. at 72 °C for extension, concluding with the final step at 40 °C for 5 min. Fluorescent light is emitted by positive partitions that contain a target molecule, as opposed to those without it, the negative partitions. Data were analyzed using the QIAcuity Suite Software V1.1.3 (Qiagen) and the results were expressed as copies of cDNA/µl based on Poisson distribution analyses. The partitions produced by the machine resemble Poisson process since the targets end up in different partitions independently and with a fixed rate. The Poisson distribution gives probabilities for positive integer random events. The key parameter of this distribution is the

	Forward	Reverse	PSM
ssb	GTTGTGCTGTTCGGCAAACT	GCGATCCTGACCGGATTGAT	GCGATCCTGACCG**CAATCAA**
sulA	CCTGAACCCATTCCCGACTC	GCCGGGCTTATCAGTGAAGT	CCTGAACCCATTC**GCCAGTG**
recA	AGGGCGTCACAGATTTCCAG	TTCCGGTAAAACCACGCTGA	AGGGCGTCACAGA**ATACGAC**
TCAST1	CCATAAGCGAGTTATAGAGTTGG	CTTTAGTGACTTTTATGTCTTCTCC	CCATAAGCGAGTTATA**CACTAGC**
RPS18	CGAAGAGGTCGAGAAAATCG	CGTGGTCTTGGTGTGTTGAC	CGTGGTCTTGGTG**AGATCAG**
ASAT	CACTCTTTTTGTAGAATCTGC	AATGCACACATCACAAAGAAG	AATGCACACATCAC**TATGTAC**
GUSB	GAAAATACGTGGTTGGAGAGCTCATT	CCGAGTGAAGATCCCCTTTTTA	CCGAGTGAAGATCCC**GTATATT**

Table 1. List of primers used in transcription analyses. Changes in PSM relative to reverse or forward primer are shown in bold.

expectation value for these events, which means it is the mean probability for a proportion of a counting process or the counting process per se for the dPCR analysis. Furthermore, the QIAcuity has embedded software that can quantify and produce reliable statistics. In our case the statistical measure we considered was the Poisson confidence interval at a 95% level that, when plotted (error bar), shows whether or not the events differ with 95% confidence.

RNA isolation and reverse transcription for satellite DNA analysis. Alpha satellite (ASAT) RNA was isolated from cervical cancer human cell line HeLa (obtained from ATCC (USA) using the RNeasy Plus Mini kit (Qiagen). Cells were cultivated in DMEM supplemented with 10% Fetal Bovine Serum (both from Sigma, MA USA), in humidified atmosphere of 5% CO_2 and on 37 °C.) TCAST1 satellite RNA was isolated from *Tribolium castaneum* adult beetles using the RNeasy Plus Mini kit (Qiagen) which includes a genomic DNA elimination step by solid phase column extraction. Approximately 1 µg of RNA, quantified by Quant-IT RNA using a Qubit fluorometer (Invitrogen, Waltham, MA, USA), was reverse transcribed using the PrimeScript RT reagent kit with gDNA Eraser (Takara, Dalian, China) in 10 µL reaction solution, using specifically modified primers for ASAT and TCAST1 expression analysis. For all samples, negative controls without reverse transcriptase enzyme were also prepared.

Quantitative real-time PCR (qPCR) for satellite DNA expression analysis. qPCR analysis was performed according to the previously published protocol[10,11]. Primers for the expression analysis of human alpha satellite DNA were constructed according to the alpha satellite consensus sequence[12] and TCAST1 primers according to its own satellite consensus sequence[13]. *Glucuronidase β (GUSB*-Gene ID: 2990) and ribosomal protein S18 (RPS18) were used as endogenous controls for normalisation in human and *Tribolium* samples, respectively, and were stably expressed without any significant variation among samples. The following thermal cycling conditions were used: 50 °C 2 min; 95 °C 7 min; 95 °C 15 s; 60 °C 1 min for 40 cycles followed by dissociation stage: 95 °C for 15 s; 60 °C for 1 min; 95 °C for 15 s; and 60 °C for 15 s. Amplification specificity was confirmed by dissociation curve analysis and the specificity of amplified products was tested on agarose gel. Control without template (NTC) was included in each run. Post-run data were analysed using LinRegPCR software v.11.1. which enables calculation of the starting concentration of amplicon in the sample ("N0 value"). N0 value is expressed in arbitrary fluorescence units and is calculated by considering PCR efficiency and baseline fluorescence. "N0 value" determined for each technical replicate was averaged and the averaged "N0 values" were divided by the "N0 values" of the endogenous control (Figs. 5B and 6B). In this paper we decided also to show the graphical representation of Delta Rn vs Cycle raw data amplification plots (Figs. 5A and 6A, Suppl. Fig. 1).

Results and discussion

Description of the proposed method. Our proposed method, schematically depicted in Fig. 1, takes advantage of using a modified primer (Modified Specific Primer, PSM) during the reverse transcription step of the protocol. Such a primer is specific for the RNA molecules to be quantified and its nucleotide sequence is designed to lack a perfect homology to the retro-transcribed template DNA. Generally, it is enough to add few mismatches with respect to the original sequence, preferably located in the close proximity to the 3'-OH terminal region. These modifications make the primer partially complementary to the target sequence but still able to hybridize at the temperatures of 37–42 °C used during the reverse transcription step. Nevertheless, the PSM will dissociate from the partially homologous genomic DNA sequence during the PCR step, once the operating temperature reaches around 60 °C. The aim of using such conveniently modified specific primer is to achieve amplification specifically from cDNA template while successfully avoiding genomic DNA targets. The correct number of modifications to be applied, their effectiveness and proper discriminating temperatures should be experimentally tested for each and every transcript to be analyzed, by selecting those parameters that show negative and positive amplification tendencies towards DNA and cDNA targets, respectively. This optimization phase represents a preliminary step of our method that enables the setup of negative and positive controls and, advantageously, has to be carried out only once, since it always remains valid for a specific amplicon and can be applied to a varying number of replicates under different experimental conditions. Indeed, in current protocols the negative control (NC: – RT, without reverse transcriptase) should ideally be prepared for each new sample to be tested, even though the target is the same, due to the random effectiveness of DNase I treatment. Using a PSM we are able to generate cDNA slightly different from its genomic DNA counterpart, due to the nucleotide mismatches present in the sequence.

During the phase following reverse transcription (Fig. 1B), the amplification of cDNA by PCR takes place using the same modified primer (PSM) from the previous step in addition to the unmodified specific primers (SP) starting from the opposite direction. Consequently, the resulting amplicon is a copy of the cDNA and not the DNA, due to the specifically selective annealing temperatures usually ranging from 55 °C to 62 °C. Therefore, with this procedure, there is no need to eliminate the co-purified DNA from the RNA sample since it is no longer a competing target and will not affect the final result of the assay. Indeed, in certain experimental conditions it could be useful and advantageous to have both DNA and RNA present together in the same sample if, for example, the results need to be normalized with respect to the gene copy number variation.

Validation of the method. Our proposed new method can be utilized in various experimental investigations and for the purposes of this paper, it has been tested by analyzing three bacterial *E. coli* genes: *ssb*, *sulA* and *recA* (Figs. 2, 3 and 4), and two satellite DNA transcripts: human alpha-satellite (ASAT) (Figs. 5 and 6, and Suppl. Fig. 2) and TCAST1 satellite from *Tribolium castaneum* (Suppl. Fig. 1).

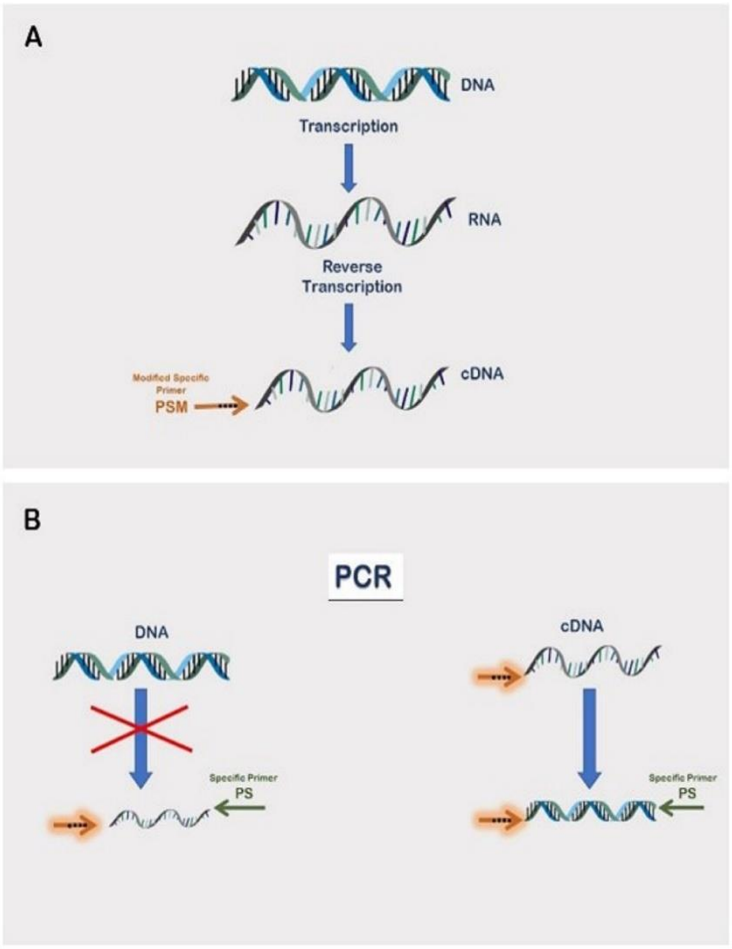


Figure 1. Schematic illustration of the new method. (**A**) Basic model of nucleic acid metabolism from DNA to cDNA. Integration of modified specific primer into cDNA by means of reverse transcription makes it a permanent part of the sequence. (**B**) Amplification of target sequence by means of polymerase chain reaction. cDNA converted by modified specific primer is properly amplified at certain discriminating temperature, while genomic DNA targets are successfully avoided.

Analysis of bacterial gene expression. Bacterial genes are a good experimental model to test our method because they do not contain introns in their coding region, removing the possibility of discriminating between transcripts and the DNA according to their different sizes. Hence, the technique could be applied to test the expression of all genes organized with a short or null intron (e.g. viral genes).

The bacterial strain used in this test was transformed with multicopy plasmid carrying a cloned *ssb* gene[9], which could compete for amplification with *ssb*-cDNA during the transcripts' quantification by PCR, unless additional DNase I treatments were implemented. The results indicated in Fig. 2 show a large difference (more than 40-fold) in *ssb* transcription levels measured by our method, as compared to the currently used method. This really high level of amplified *ssb* sequence in the latter approach, when reverse transcription was not carried out, and the DNA was eliminated in both RNA isolation and RT steps (Fig. 2A), is likely due to low efficiency of elimination of covalently closed circular plasmid DNA, meaning that it is false (i.e. it does not accurately represent the process of transcription) and is actually caused by DNA contamination.

This is likely a reason for all the observed cases of high levels of *ssb* sequence amplification using classical primers (Fig. 2A). In contrast, *ssb* sequence was amplified by our new method only in those cases when reverse transcription was performed, i.e. when cDNA was created (Fig. 2B). The level of *ssb* sequence amplification did not depend on DNA elimination (Fig. 2B), thus confirming insensibility of our method to the presence of genomic (and plasmid) DNA. Next, we quantified expression of *recA* and *sulA* genes, which are present as single copies in the *E. coli* genome. In accord with the previous assay, no *recA* sequence amplification was observed

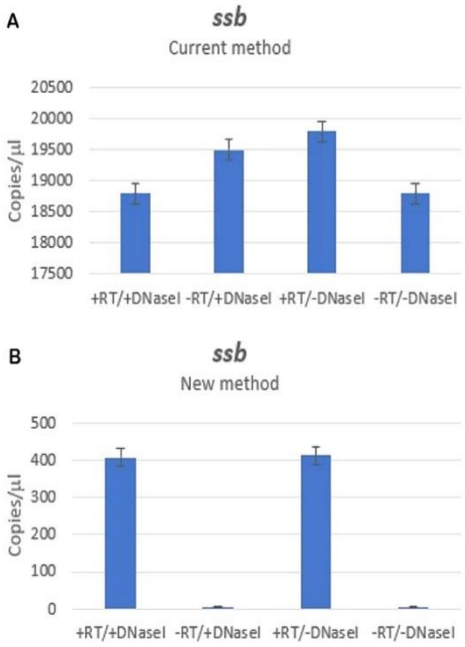


Figure 2. Transcription of *ssb* gene in exponentially growing *E. coli* cells harbouring *ssb* overexpression plasmid pID2 obtained by dPCR using current (**A**) and new method (**B**). Columns represent number of copies/µl and the plotted error bar shows whether or not the events differ with 95% Poisson confidence interval.

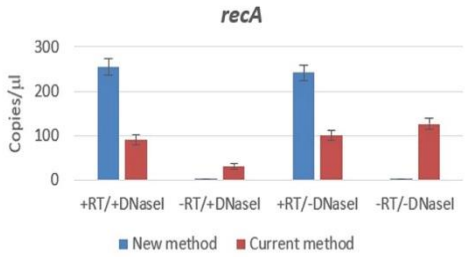


Figure 3. Transcription of *recA* gene in exponentially growing *E. coli* cells obtained by dPCR using current and new method. Columns represent number of copies/µl and the plotted error bar shows whether or not the events differ with 95% Poisson confidence interval.

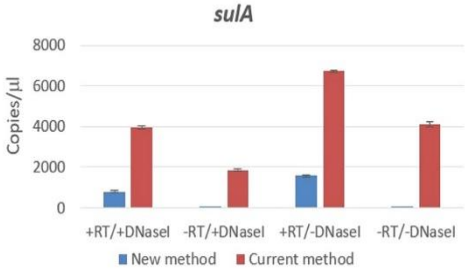


Figure 4. Transcription of *sulA* gene in exponentially growing *E. coli* cells obtained by dPCR using current and new method. Columns represent number of copies/µl and the plotted error bar shows whether or not the events differ with 95% Poisson confidence interval.

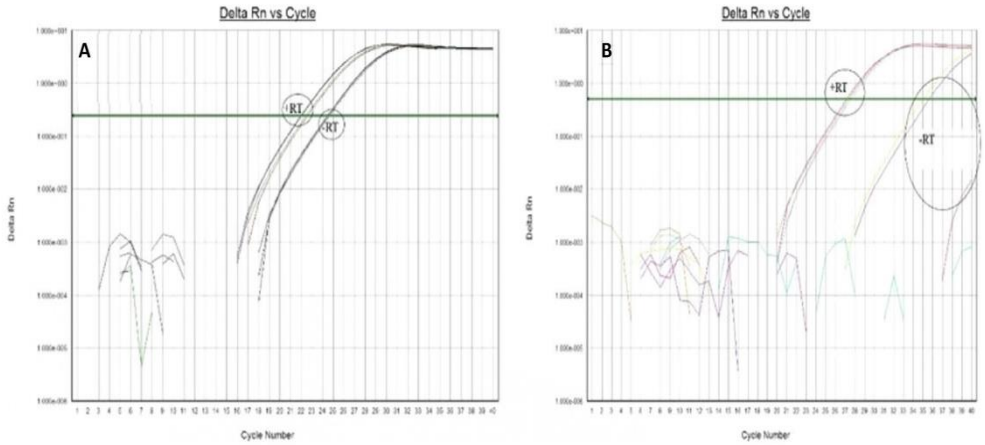


Figure 5. Delta Rn vs Cycle plot of alpha satellite DNA isolated from HeLa cells obtained by qPCR using current method (**A**) and new method (**B**). + RT and – RT represent positive and negative controls, with and without reverse transcription, respectively.

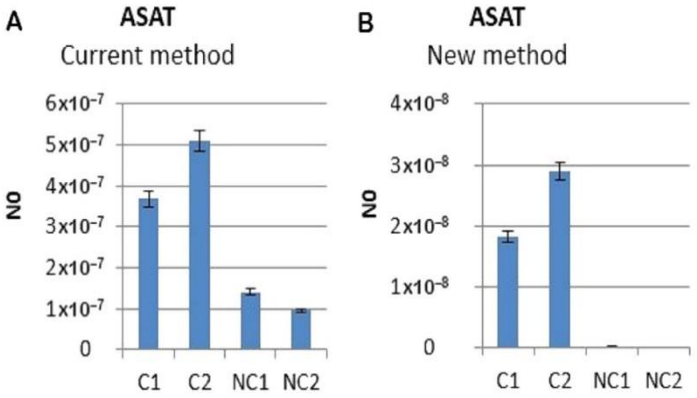


Figure 6. Transcription level of alpha satellite DNA obtained by qPCR using current method (**A**) and new method (**B**). Columns show average of 2 different loaded samples in qPCR experiments performed in triplicate. N0 represents normalized average N0 value for alpha satellite. C represent alpha samples with reverse transcription and NC represents negative controls without reverse transcription and M is 100 bp size marker.

using our method unless cDNA was created by reverse transcription (Fig. 3). The level of *recA* sequence amplification was, again, independent from genomic DNA elimination from the sample (Fig. 3). Conversely, the current method, which uses standard primers, showed a false positive signal even when reverse transcription step was skipped and the genomic DNA was (obviously incompletely) eliminated by DNase I treatment (Fig. 3).

Finally, analysis of *sulA* gene expression using a modified primer was in accord with the previous assays since amplification of *sulA* sequence occurred only after reverse transcription, i.e. it was specific for cDNA (Fig. 4). Accordingly, no effect was observed after genomic DNA elimination (Fig. 4). In contrast, amplification of *sulA* sequence using standard primers was very different, and was not abolished even in situations where genomic DNA was eliminated and reverse transcription was not performed (Fig. 4); theoretically, the – RT/ + DNase I sample should not contain any cDNA or genomic DNA.

The presented results clearly demonstrate that our method of using a modified primer during cDNA synthesis produces a cDNA-specific PCR signal, which is independent of genomic DNA, and therefore much more accurately quantifies gene expression when compared to the standard, commonly used method, which, unfortunately, does not produce real negative control since there is always possibility to have contaminating DNA in the sample.

Analysis of satellite DNA expression. Satellite DNA represents one of the best target candidates for demonstrating the effectiveness of our methodology since it is a highly repetitive non-coding genomic DNA,

ever-present in large quantities in the sample and therefore difficult, if not impossible, to remove during RNA purification.

Alpha satellite DNA is the most abundant human satellite DNA of 171 bp long, comprising up to 10% of the genome[14]. Figure 5A, shows qPCR results obtained by following the current standard protocol (old method) which implies the elimination of DNA both during the RNA purification and reverse transcription phase. In spite of that, alpha satellite DNA continues to persist in the negative control samples (– RT). Furthermore, since it is not organized into exons and introns, satellite DNA cannot be discriminated from satellite cDNA based on its length; therefore, even a slightest trace of DNA contamination often produces false-positive results. The new method, however, successfully demonstrated the disappearance of the alpha satellite DNA contamination from the qPCR amplification results (Fig. 5B, – RT), as it can be clearly seen also by loading the amplicons on agarose gel (Suppl. Fig. 2): ASAT amplicon of 126 bp long is present only in + RT samples (C: controls) respect to – RT samples (NC: negative control). The same results could be represented as in Fig. 6A (current method) and Fig. 6B (new method), where "N0 value" is the starting concentration of amplicon in the sample and columns show average of 2 different loaded samples in qPCR experiments performed in triplicate (see "Materials & method" section).

The highly abundant satellite DNA TCAST1 has previously been characterized as the major satellite that makes up to 30% of the beetle *Tribolium castaneum* genome, organizing the centromeric as well as pericentromeric regions of all 20 chromosomes[10,13]. Again, using the new method only cDNA was amplified (+RT samples) and almost nothing of genomic DNA contamination was detected in –RT samples (Suppl. Fig. 1). The results clearly show they are exactly the same as those obtained for human alpha satellite DNA.

Conclusion

In conclusion, we can affirm that the results achieved through application of our new method of quantifying different types of transcripts are certainly more precise, reproducible and affordable than those obtained by currently used protocols. This is because our method is insensitive to DNA contamination (which usually gives rise to false positive signals) and therefore there is no need for prior elimination of the template DNA. Moreover, skipping the DNA elimination step effectively preserves the RNA from degradation. In that way the two major sources of inherent inaccuracy in transcriptome analyses are avoided.

Data avalaibility

All data generated or analysed during this study are included in this published article (and its Supplementary Information files).

Received: 5 April 2023; Accepted: 7 July 2023
Published online: 15 July 2023

References

1. Foley, K. P., Leonard, M. W. & Engel, J. D. Quantitation of RNA using the polymerase chain reaction. *Trends Genet.* **9**(11), 380–385 (1993).
2. Verwilt, J. *et al.* When DNA gets in the way: A cautionary note for DNA contamination in extracellular RNA-seq studies. *Proc. Natl. Acad. Sci.* **117**(32), 18934 (2020).
3. Bustin, S. A. Quantification of mRNA using real-time reverse transcription PCR (RT-PCR): Trends and problems. *J. Mol. Endocrinol.* **29**(1), 23–39 (2002).
4. Kumar, S. V., Hurteau, G. J. & Spivack, S. D. Validity of messenger RNA expression analyses of human saliva. *Clin. Cancer Res.* **12**(17), 5033–5039 (2006).
5. Li, X., Zhang, P., Wang, H. & Yu, Y. Genes expressed at low levels raise false discovery rates in RNA samples contaminated with genomic DNA. *BMC Genom.* **23**(1), 554 (2022).
6. Padhi, B. K., Singh, M., Huang, N. & Pelletier, G. A PCR-based approach to assess genomic DNA contamination in RNA: Application to rat RNA samples. *Anal. Biochem.* **494**, 49–51 (2016).
7. Hashemipetroudi, S. H., Nematzadeh, G., Ahmadian, G., Yamchi, A. & Kuhlmann, M. Assessment of DNA contamination in RNA samples based on ribosomal DNA. *J. Vis. Exp. JoVE* https://doi.org/10.3791/55451 (2018).
8. Green, M. R. & Sambrook, J. Removing DNA contamination from RNA samples by treatment with RNase-free DNase I. *Cold Spring Harb. Protoc.* https://doi.org/10.1101/pdb.prot101725 (2019).
9. Feliciello, I. *et al.* RecF, UvrD, RecX and RecN proteins suppress DNA degradation at DNA double-strand breaks in *Escherichia coli*. *Biochimie* **148**, 116–126 (2018).
10. Feliciello, I., Akrap, I. & Ugarković, Đ. Satellite DNA modulates gene expression in the beetle *Tribolium castaneum* after heat stress. *PLoS Genet.* **11**(8), e1005466 (2015).
11. Ljubić, S. *et al.* Alpha satellite RNA levels are upregulated in the blood of patients with metastatic castration-resistant prostate cancer. *Genes* **13**, 383 (2022).
12. Choo, K. H., Vissel, B., Nagy, A., Earle, E. & Kalitsis, P. A survey of the genomic distribution of alpha satellite DNA on all the human chromosomes, and derivation of a new consensus sequence. *Nucleic Acids Res.* **19**, 1179–1182 (1991).
13. Feliciello, I., Chinali, G. & Ugarković, Đ. Structure and evolutionary dynamics of the major satellite in the red flour beetle *Tribolium castaneum*. *Genetica* **139**, 999–1008 (2011).
14. McNulty, S. M. & Sullivan, B. A. Alpha satellite DNA biology: Finding function in the recesses of the genome. *Chromosome Res.* **26**(3), 115–138 (2018).

Acknowledgements

Thanks to the students who worked in the lab, specially: Nunzia Santini and Alessandro Esposito.

Author contributions

I.F.: Conceptualization, Formal analysis, Methodology, Validation, Writing original draft. D.Đ., S.L., M.M.: sample preparation, Writing review & editing. A.P. and Đ.U.: Writing review & editing. M.C.F.: Writing review & Statistical analysis.

Funding

This work was supported by Croatian Science Foundation grants: IP-2019-04-6915 to Đ. Ugarković and IP-2019-04-3790 to D. Đermić, as well as by the Italian Ministry of Education, University and Research (MIUR), FFABR2017, fund for Investments on Basic Research (FIRB), by the International Staff Mobility Program of University of Naples Federico II to I. Feliciello and by ADRIS foundation.

Competing interests

The authors declare no competing interests.

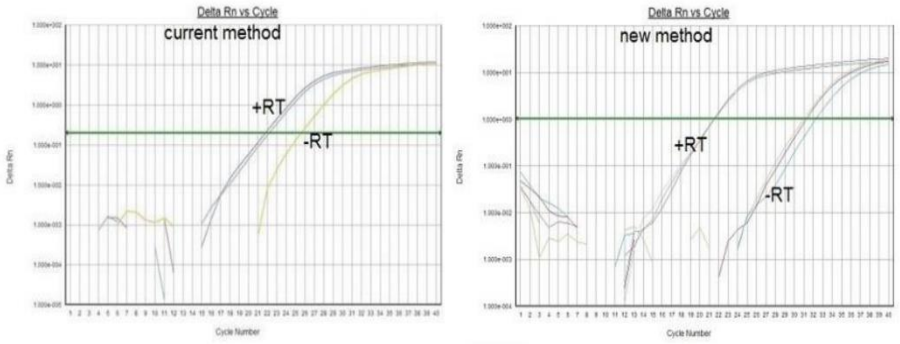
Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-023-38383-4.

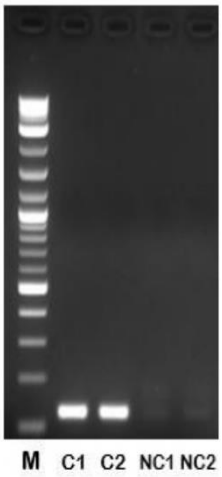
Correspondence and requests for materials should be addressed to Đ.U. or I.F.

Reprints and permissions information is available at www.nature.com/reprints.

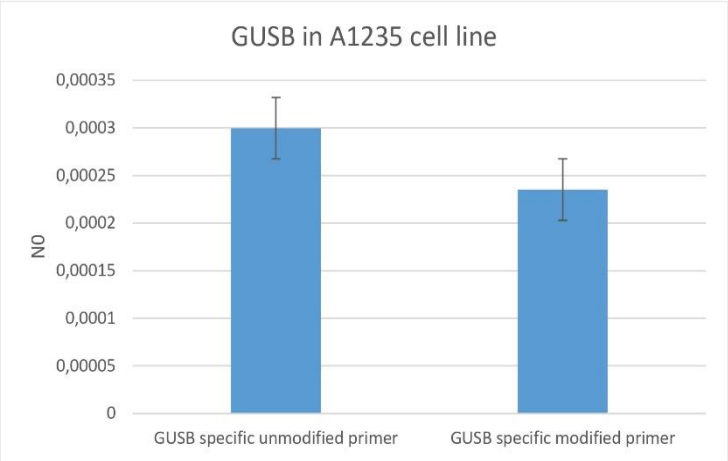
Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Supplementary Figure 1. Transcription of TCAST1 satellite DNA. Delta Rn vs Cycle obtained by quantitative real-time PCR using current method and new method. +RT and –RT represent positive and negative controls, with and without reverse transcription, respectively.



Supplementary Figure 2. Agarose gel with alpha satellite DNA amplicons of 126 bp. NC represents negative controls without reverse transcription; C represents alpha satellite samples with reverse transcription and M is molecular-weight size marker (100 bp – 10 kb).



Supplementary Figure 3. Real time PCR analysis of human GUSB gene expression in order to compare the efficiency in reverse transcription using two different specific primers: the unmodified and the modified one by inserted 4 mismatched bases. Columns show average of 2 different loaded samples in qPCR experiments performed in triplicate.

6. CURRICULUM VITAE

Sven Ljubić was born on November 29, 1986, in Zagreb, where he completed his elementary and high school education in a general gymnasium. In September 2017, he graduated *magna cum laude* from the Department of Molecular Biology at the Faculty of Science, presenting a thesis titled "Expression of semaphorins-4D, -6D, -3C and -7A in hepatitis C virus infection," supervised by Assoc. Prof. Snježana Židovec Lepej. This achievement earned him the title of Master of Molecular Biology. In October 2020, he took on the role of research assistant in the Laboratory for Evolutionary Genetics at the Department of Molecular Biology at the Ruđer Bošković Institute in Zagreb, while also starting his doctoral studies in biology at the Faculty of Science.

Throughout his doctoral journey, Sven collaborated on various projects, including "Alpha satellite DNA in the evolution of gene modulatory networks" funded by the Croatian Science Foundation, and the "Alpha satellite RNA: a candidate for diagnostic/prognostic blood biomarker of prostate cancer" project supported by the Adris Foundation, under the guidance of Prof. emer. Đurđica Ugarković and Prof. Isidoro Feliciello. He also contributed to teaching by leading laboratory exercises and seminars, assisting students with practical laboratory work, and guiding them in preparing their graduate theses. Sven has shared his research findings at several international conferences and engaged in workshops aimed at popularizing science. He is a co-author of eight scientific papers published in internationally peer-reviewed journals.

Publications:

1) Ljubić S, Matulić M, Đermić D, Feliciello MC, Procino A, Passaro F, Ugarković Đ, Feliciello I (2025) Downregulation of Gene Expression by Alpha Satellite Transcripts. International Journal of Molecular Sciences. 26:11204.

2) Ljubić S, Matulić M, Đermić D, Feliciello MC, Procino A, Ugarković Đ, Feliciello I (2025) Antibiotics induce overexpression of alpha satellite DNA accompanied with

epigenetic changes at alpha satellite arrays as well as genome-wide. Epigenetics & Chromatin. 18:62.

3) Feliciello I, Ljubić S, Đermić E, Ivanković S, Zahradka D, Đermić D (2024) Single-strand DNA-binding protein suppresses illegitimate recombination in *Escherichia coli*, acting in synergy with RecQ helicase. Scientific Reports. 14:20476.

4) Đermić D, Ljubić S, Matulić M, Procino A, Feliciello MC, Ugarković Đ, Feliciello I (2023) Reverse transcription-quantitative PCR (RT-qPCR) without the need for prior removal of DNA. Scientific Reports. 13:11470.

5) Feliciello I, Đermić E, Malović H, Ivanković S, Zahradka D, Ljubić S, Procino A, Đermić D (2022) Regulation of ssb Gene Expression in *Escherichia coli*. International Journal of Molecular Sciences. 23:10917.

6) Ugarković Đ, Sermek A, Ljubić S, Feliciello I (2022) Satellite DNAs in Health and Disease. Genes. 13:1154.

7) Ljubić S, Sermek A, Prgomet Sečan A, Prpić M, Jakšić B, Murgić J, Fröbe A, Ugarković Đ, Feliciello I (2022) Alpha Satellite RNA Levels Are Upregulated in the Blood of Patients with Metastatic Castration- Resistant Prostate Cancer. Genes. 13:383.

8) Feliciello I, Pezer Ž, Sermek A, Bruvo Mađarić B, Ljubić S, Ugarković Đ (2021) Satellite DNA-Mediated Gene Expression Regulation: Physiological and Evolutionary Implication. Progress in Molecular and Subcellular Biology. 60:145–167.