Annu. Rev. Genom. Human Genet. 2006.7:301-313. Downloaded from arjournals.annualreviews.org
by DUKE UNIVERSITY on 09/18/06. For personal use only.

Structural and Functional Dynamics of Human Centromeric Chromatin

Mary G. Schueler[1] and Beth A. Sullivan[2]

[1]Genome Technology Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland 20892

[2]Institute for Genome Sciences & Policy, Duke University, Durham, North Carolina 27708; email: beth.sullivan@duke.edu

Annu. Rev. Genomics Hum. Genet. 2006. 7:301–13

First published online as a Review in Advance on June 6, 2006

The *Annual Review of Genomics and Human Genetics* is online at genom.annualreviews.org

This article's doi:
10.1146/annurev.genom.7.080505.115613

Copyright © 2006 by Annual Reviews.
All rights reserved

1527-8204/06/0922-0301$20.00

Key Words

centromere, kinetochore, α-satellite, histone, heterochromatin, methylation, cohesion, epigenetic

Abstract

Centromeres are the elements of chromosomes that assemble the proteinaceous kinetochore, maintain sister chromatid cohesion, regulate chromosome attachment to the spindle, and direct chromosome movement during cell division. Although the functions of centromeres and the proteins that contribute to their complex structure and function are conserved in eukaryotes, centromeric DNA diverges rapidly. Human centromeres are particularly complicated. Here, we review studies on the organization of homogeneous arrays of chromosome-specific α-satellite repeats and evolutionary links among eukaryotic centromeric sequences. We also discuss epigenetic mechanisms of centromere identity that confer structural and functional features of the centromere through DNA-protein interactions and post-translational modifications, producing centromere-specific chromatin signatures. The assembly and organization of human centromeres, the contributions of satellite DNA to centromere identity and diversity, and the mechanism whereby centromeres are distinguished from the rest of the genome reflect ongoing puzzles in chromosome biology.

Annu. Rev. Genom. Human Genet. 2006.7:301-313. Downloaded from arjournals.annualreviews.org
by DUKE UNIVERSITY on 09/18/06. For personal use only.

Centromere: a locus, comprised of chromatin, required in *cis* for proper segregation of chromosomes

Kinetochore: the proteinaceous structure that is assembled on CEN chromatin, linking DNA of the centromere with proteins that direct chromosome movement and monitor chromosomal attachment to spindle microtubules

Centromere region: the various domains and functions present in the vicinity of the regional centromere locus

α-satellite: a primate-specific satellite family based on a 171-bp repeat and found at the centromere of all primates examined

Monomer: the smallest unit of repeat in a satellite family

Higher-order repeat unit: a unit of repeat comprised of multiple copies of a satellite monomer

Centromeric (CEN) chromatin: the specialized chromatin at the centromere region that is the foundation of the kinetochore

INTRODUCTION

Chromosomes of species as diverse as humans and yeast, flies, and plants all accomplish chromosome segregation via the centromere locus (65, 72). The centromere is the assembly site for the kinetochore, the proteinaceous structure on each chromosome that coordinates attachment to and travel along microtubules (46). High-resolution optical studies and elegant functional studies in model organisms show that the centromere region is structurally and functionally complex, existing as a multidomain locus (67). Multiple protein components within centromere regions are highly conserved and contribute to a largely universal arrangement of structural and functional domains. The centromere coordinates chromosome movement in mitosis and meiosis and synchronizes aspects of chromosome structure, such as heterochromatin formation, sister chromatid cohesion, and chromosome condensation.

Remarkably, whereas the proteins responsible for forming the kinetochore are conserved throughout the evolutionary spectrum, the DNA sequences found at centromeres are not (52, 65). It is not clear why each species reinvents centromeric (CEN) DNA despite the conserved structure and mechanism of the kinetochore. One common DNA feature of centromeres across the evolutionary spectrum is tandemly repetitive sequence families called satellites. α-satellite DNA is primate specific and present at the centromere of all typical human chromosomes (36, 38). Proteins of the kinetochore assemble onto chromosome-specific arrays of α-satellite (37). Artificial chromosome assays provide direct evidence of a functional role for α-satellite in human chromosome stability and segregation (9). Advances in the genetic and genomic definition of human centromere regions reveal molecular signatures of CEN DNA expansion throughout primate evolution (55).

This chapter presents an overview of the structure and evolution of human CEN DNA, epigenetic factors that specify centromere identity, and protein components required for centromere function. We focus primarily on the proteins that are intimately associated with CEN DNA, and that organize the specialized type of chromatin that is the foundation for kinetochore assembly. The reader is directed to several excellent reviews describing the composition of the outer kinetochore and role of proteins involved in chromosome movement, checkpoint functions and spindle interactions, neocentromeres, and human artificial chromosomes (5, 9, 14, 16, 75, 82).

CENTROMERIC DNA

All primate chromosomes studied to date contain α-satellite DNA (2, 74, 81). α-satellite was originally characterized in the African Green Monkey genome as a satellite DNA family based on divergent 170-bp monomers arranged in a tandem, head-to-tail fashion resulting in an overall directionality (35, 47). This type of α-satellite is termed monomeric and has been identified within the pericentromeric regions of 21 of the 24 human chromosomes (**Figure 1**) (3, 50). Each human chromosome is also characterized by a chromosome-specific higher-order array of α-satellite (**Figure 1**) (79). Each higher-order array is composed of a tandemly arranged repeat unit consisting of an integral number of α-satellite monomers. Higher-order arrays can span over 3–5 Mb and are highly homogenous, consisting of the same higher-order repeat unit occurring hundreds or thousands of times within a given centromere locus (33, 80).

Despite the human genome sequencing effort (32, 73), sequence contigs spanning the junction at the edge of the chromosome-specific array exist for only chromosomes 8, 17, and X (41, 48–50). Such contigs provide the precise physical context in which to analyze CEN DNA sequences (51, 55, 56). These sequences reveal a stretch of monomeric α-satellite and other satellites joining arm sequences with the higher-order array (see

Figure 1). The stretch of monomeric α-satellite is frequently interrupted by interspersed elements [long interspersed element (LINE), short interspersed element (SINE), long terminal repeat (LTR) retrotransposons] (see **Figure 1**) and has disrupted directionality such that blocks of monomers of common directionality are defined by changes in orientation relative to adjacent blocks (55).

There is an abrupt transition from monomeric to higher-order α-satellite at the array transition; however, some monomeric α-satellite monomers near the array edge contain centromere protein (CENP)-B boxes (M.G. Schueler, unpublished data), in contrast to α-satellite currently identified in early primates (2). Comparisons of higher-order repeat units within genomic sequence from chromosomes 8, 17, and X agree with early reports that typical higher-order repeat units of a chromosome-specific array share 97–100% identity (M.G. Schueler, unpublished data) (17, 49, 54, 56). At both the p and q arm array junctions on chromosomes 8 and X, these typical pairwise identities are reached within one or two units of the transition from monomeric α-satellite and, importantly, repeat units are in the same orientation on both the p and q arm sides, indicating a continuous directionality across the array (M.G. Schueler, unpublished data) (41, 48, 56).

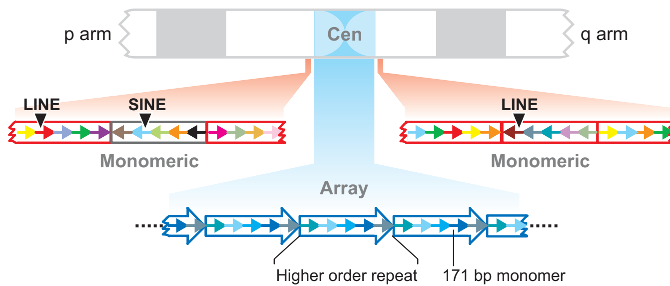

Figure 1

Centromeric (CEN) DNA organization. A typical human chromosome is schematically depicted, emphasizing the pericentromeric and CEN (*blue*) satellites in the ideogram. Each small arrow represents a single satellite monomer. In the pericentromeric regions, blocks of tandem satellite monomers from a single family (indicated by *red* versus *gray boxes*) occasionally contain embedded interspersed repetitive elements [e.g., long interspersed elements (LINEs) and short interspersed elements (SINEs)]. Adjacent satellite blocks can exist in the same or opposite orientations. In the CEN region, higher-order repeat units of α-satellite (this unit comprised of five monomers) are indicated with large blue arrows. Image is adapted from Reference 55.

α-SATELLITE EVOLUTION

The primary sequence structure of α-satellite DNA varies from chromosome to chromosome in the human genome and from genome to genome among primates. The vast majority of data describing human α-satellite arrays implicate unequal crossover between sister chromatids as the primary mode of change (74). Unequal exchange is driven by sequence identity and results from crossover between misaligned sequences. Successive unequal exchanges lead to a high degree of sequence homogeneity, to concerted evolution yielding chromosome-specific arrays, and to the observed four- to sixfold length variation between homologous arrays (51, 74).

Common restriction enzyme periodicities and sequence features of α-satellite arrays have allowed the grouping of human chromosomes into five families (33). These relationships argue for a common progenitor of each higher-order repeat family that was amplified once and then spread by transposition or conversion (4, 78). Interestingly, although α-satellite arrays sharing the characteristics of the human higher-order repeat families exist in the great apes, these are only rarely located on orthologous chromosomes (7, 18, 23, 53, 76). α-satellite of lower primates lacks CENP-B boxes and, although some early primate α-satellite has a dimeric organization, these early arrays are typically not chromosome or even species specific (2). These observations indicate a shift from predominately interchromosomal homogenization events in early primates to intrachromosomal events in great ape genomes (74).

An exhaustive account of all the sequence elements within a pericentromeric region and knowledge of their physical relationships can now provide a unified view of the events

LINE: long interspersed element

SINE: short interspersed element

CENP: centromere protein

Homogenization: the result of sequence conversion mechanisms that, through direct molecular interaction, make two independent loci more similar to each other in DNA sequence

Annu. Rev. Genom. Human Genet. 2006.7:301-313. Downloaded from arjournals.annualreviews.org by DUKE UNIVERSITY on 09/18/06. For personal use only.

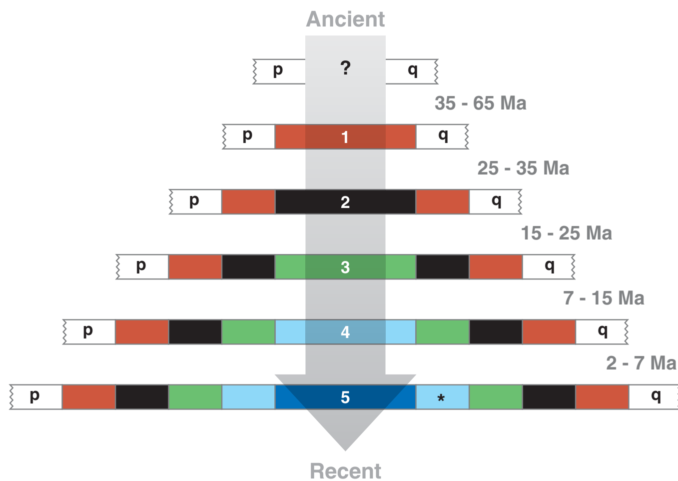

Figure 2

α-satellite DNA evolution: progressive proximal expansion. Successive additions (*colored rectangles*) to the centromere throughout primate evolution begin at the top of the image with an ancestral chromosome prior to α-satellite emergence and proceeding down through the image to the current human centromere. Each addition of new material moves previous centromeric DNA outward. An asterisk (*) indicates the region of monomeric α-satellite with centromere protein (CENP)-B boxes. The image is based on data from the human X chromosome (55). Approximate dates [millions of years ago (mya)] are derived from the primate tree (21) and from the phylogeny of L1 elements observed within the α-satellite sequences (55, 59).

Annu. Rev. Genom. Human Genet. 2006.7:301-313. Downloaded from arjournals.annualreviews.org
by DUKE UNIVERSITY on 09/18/06. For personal use only.

leading to current centromere content and organization (**Figure 2**) (55). Particularly, phylogenetic relationships of α-satellite monomers on the X chromosome and the evolutionary age of interspersed elements found embedded within those monomers indicate addition of new material into CEN regions in a series of punctuated events occurring at the central domain. This sequential addition can be seen on the X chromosome with the addition of three independent domains of monomeric α-satellite prior to the emergence of monomers containing CENP-B boxes (M.G. Schueler, unpublished data) and followed, most recently, by higher-order α-satellite (**Figure 2**).

Physical symmetry of phylogenetic domains on the X chromosome provides an essential model in defining early evolutionary events. Very recent phylogenetic α-satellite domains display symmetry at the chromosome 8 centromere (additions 4 and 5 in **Figure 2**); however, in contrast to the X, chromosome 8 monomeric α-satellite domains do not (M.G. Schueler, unpublished data). This could indicate a greater number of additions onto the X, more frequent homogenization or rearrangement of monomeric α-satellite on chromosome 8, and/or loss of material from chromosome 8. Recent data suggest that monomeric α-satellite displays chromosome specificity (51), indicating that interchromosomal exchanges in early primates may have, at one time, been limited to homologs. Preservation of the physical hallmarks of successive additions on the X chromosome may indicate that interhomolog exchanges occurred at a time in development when the X had no pairing partner, perhaps during male meiosis. Comparative sequence analysis of centromeres from additional human autosomes and of nonhuman primate centromeres is necessary to further test the validity of this model.

THE CENTROMERE REGION IS ORGANIZED INTO FUNCTIONAL SUBDOMAINS

The centromere region is broadly classified into two functional protein domains that encode (*a*) CEN chromatin/kinetochore functions and (*b*) heterochromatin assembly. In eukaryotes, linear DNA is packaged into chromatin, a higher-order structure containing histones and nonhistone proteins. Approximately 150 bp of DNA is wrapped around two copies of the core histones H2A, H2B, H3, and H4 to form a nucleosome. Human centromeres assemble from CEN (α-satellite) DNA that is packaged with core histones, variant histones, and other proteins into a specialized type of chromatin, called CEN chromatin. CEN chromatin and the kinetochore each contain distinctive DNA and proteins involved in centromere assembly and structural aspects of the kinetochore. CENP-A, CENP-B, and CENP-C are constitutive components of the kinetochore that form the prekinetochore complex, the

structural and functional precursor of the mature metaphase kinetochore. Blocks of heterochromatin flank one or both sides of CEN chromatin. Although CEN chromatin and heterochromatin are assembled independently, each is important for chromosome segregation and genome stability (10, 15).

CEN chromatin containing a specialized histone is the foundation of the kinetochore. Centromere formation in larger eukaryotes involves epigenetic (sequence-independent) mechanisms, but the specifics of these processes remain unclear. Despite divergent CEN sequences, CENPs are widely distributed in evolution (24, 68). CENP-A is central to kinetochore assembly, initiating centromere assembly. It was originally identified as an antigen recognized by human autoimmune antisera (19). Biochemical and molecular analyses confirmed that CENP-A is a protein that is related to histone H3 through a common histone-fold domain (43, 44). CENP-A proteins, also referred to as cenH3s, are present in yeasts (*S. cerevisiae* and *S. pombe*), *C. elegans* and *Drosophila*, *A. thaliana*, Oryzis (rice) and *Z. meays*, representing an evolutionary link between seemingly widely divergent CEN DNA sequences (26, 65, 68). Nucleosomes can be assembled in vitro from purified CENP-A and from histones H2A, H2B, and H4 (83), indicating that CENP-A can replace both copies of H3 in CEN nucleosomes (58). All centromeres are associated with CENP-A, and it is required to recruit many other centromere and kinetochore proteins, with the exception of proteins located in the adjacent heterochromatin domain, such as heterochromatin protein 1 (HP1) (12).

How is the CENP-A mark maintained through cell divisions during which core histones are typically dispersed? During replication, dynamic chromatin changes occur, and centromeres are exposed to microtubule tension when chromosomes segregate at anaphase. Unlike H3, which is incorporated into chromatin in early S-phase during replication, existing CENP-A nucleosomes are not removed and replaced, but are inherited semiconservatively, such that "old" CENP-A is divided between daughter strands (64, 67). Newly synthesized CENP-A is loaded onto centromeres in G2 by a replication-independent mechanism (29, 57). One model suggests that during early S-phase, H3 is placed into the sites on daughter strands in which CENP-A has oppositely distributed. In G2, when chromatids begin to condense, the "placeholder" H3 is replaced by CENP-A, with "old" CENP-A acting as a flag or marker for inserting new CENP-A. Core histone replacement involving a histone variant also occurs at sites of specialized active chromatin, when H3 is replaced by H3.3 (1); thus, a similar mechanism involving CENP-A and H3 seems reasonable.

H3K4me2: H3 dimethylated at lysine 4

MODIFICATIONS OF CORE HISTONES CONTRIBUTE TO CENTROMERE SPECIFICITY

Modifications (methylation, acetylation, phosphorylation) of amino acids in N-terminal tails of core histone integrate a complex system of epigenetic modifications and chromosomal proteins that establish and maintain distinctive types of chromatin. Distinctive combinations of histone modifications, known as the "histone code," partition the genome into functional domains, such as transcriptionally silent heterochromatin and transcriptionally active euchromatin. Perhaps a unique chromatin signature plays a role in directing CENP-A to centromeres and establishing CEN chromatin. In fact, H3 nucleosomes interspersed with subdomains of CENP-A nucleosomes within CEN chromatin are methylated at K4 (**H3K4me2**) (**Figure 3**). Histone methylation is a relatively stable epigenetic mark that is maintained through many cell divisions. A model in which H3K4me2 that surrounds unmodified H3 that has been placed into gaps left by CENP-A during S-phase replication could provide the framework for removing H3 and loading new CENP-A while maintaining the location of the centromere.

Annu. Rev. Genom. Human Genet. 2006.7:301-313. Downloaded from arjournals.annualreviews.org by DUKE UNIVERSITY on 09/18/06. For personal use only.

Annu. Rev. Genom. Human Genet. 2006.7:301-313. Downloaded from arjournals.annualreviews.org
by DUKE UNIVERSITY on 09/18/06. For personal use only.

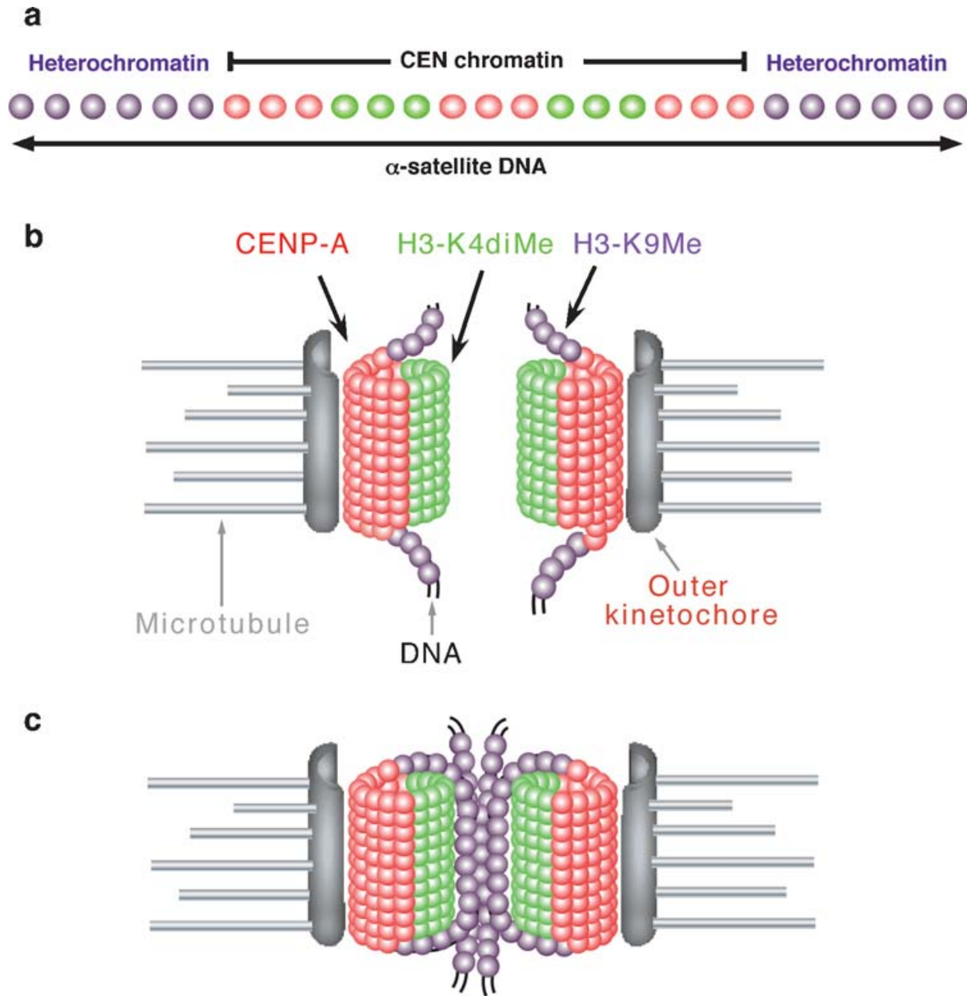

Figure 3

Unique organization of centromere regions in humans. (*a*) On linear, two-dimensional chromatin fibers, subdomains of nucleosomes containing centromeric (CEN) histone CENP-A (*red*) are interspersed with H3 dimethylated at lysine 4 (H3K4me2) (*green*) to form a domain of CEN chromatin on a fraction of the megabase regions of human α-satellite DNA. The remainder of the α-satellite DNA is assembled into heterochromatin (*purple*) that flanks one or both sides of CEN chromatin domain. (*b*) At metaphase, when mitotic chromosomes condense, the interspersed domains promote coiling of the DNA so that stacks of CENP-A nucleosomes are presented to the poleward face of the chromosome where they can interact with other kinetochore proteins. H3-containing nucleosomes are oriented between sister kinetochores. (*c*) Heterochromatin defined by nucleosomes containing H3-K9 methylation (*purple*) is assembled into a domain that is distinct from CEN chromatin. Higher-order packaging of heterochromatin between sister kinetochores may promote orientation of CENP-A, pushing it toward the outside of the chromosome. Heterochromatin in this region is also important for recruiting cohesion proteins that are sustained at the centromere until chromatid separation at anaphase.

Annu. Rev. Genom. Human Genet. 2006.7:301-313. Downloaded from arjournals.annualreviews.org
by DUKE UNIVERSITY on 09/18/06. For personal use only.

THE PREKINETOCHORE COMPLEX: CHROMATIN CONTAINING CENP-A, CENP-B, AND CENP-C PROMOTES UNIQUE STRUCTURAL FEATURES AT THE CENTROMERE

Structural features of CENP-A also contribute to centromere identity. Nucleosomes containing CENP-A and H4 are biochemically and structurally compact and conformationally more rigid than H3-H4 nucleosomes (11). Two domains within the CENP-A amino acid sequence are responsible for selective CEN targeting, and provoke a conformation change that also involves H4, restricting CENP-A from other genomic regions and creating a rigid platform for assembling downstream kinetochore components.

Higher-order packaging also distinguishes the centromere from the rest of the genome. In humans, the prekinetochore chromatin complex containing CENPs-A, -B, and -C is preferentially organized on higher-order α-satellite DNA containing the CENP-B protein binding motif (CENP-B box) that is present in every other 171-bp monomer (27, 42). CENP-B contains a dimerization domain, so that two CENP-B boxes are brought into proximity (30). In doing so, CENP-B binding induces nonrandom positioning, or phasing, of nucleosomes (69) between CENP-B boxes. CENP-A and core histones are assembled into nucleosomes between the CENP-B binding regions, and the combination of the two proteins at different regions of the α-satellite DNA establishes centromere-specific higher-order structures (**Figure 4**). Unlike CENP-A and CENP-C, which only bind along a small portion of α-satellite, CENP-B is located more extensively along α-satellite, binding at each CENP-B box within arrays that can range up to 4 Mbs (6, 22). The functional significance of this broad binding domain is unclear; however, CENP-B homologs in yeast promote histone modifications that nucleate heterochromatin assembly (39). The binding of CENP-B through the entire array of human α-satellite may be important for not only achieving CEN structure through nucleosomal spacing, but may also act as a bridge that connects uninterrupted domains of CEN chromatin and heterochromatin.

Assembly of prekinetochore chromatin containing CENP-A and CENP-B is completed by recruitment of additional CENPs. CENP-A, CENP-H, and CENP-I are all required to recruit CENP-C, a constitutive centromere protein that is part of the prekinetochore chromatin and is a structural component of the three-dimensional inner kinetochore at metaphase (20, 40, 63). CENP-C has several domains that target it to the centromere, promote protein-protein interactions and dimerization, and confer DNA binding activity (45, 60, 70, 71). The DNA binding activity of CENP-C is nonspecific, but anchors CENP-C to the centromere and incorporates it into chromatin once it is recruited by other components (6). CENP-C monomers or dimers are distributed across repetitive units of CENP-A/CENP-B nucleosome complexes (see above). The assembly of this chromatin confers a conformational change that is the platform for recruiting outer kinetochore proteins.

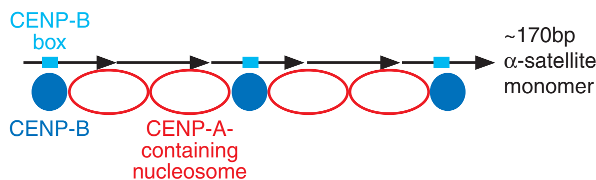

Figure 4

Arrangement of constitutive proteins on higher-order α-satellite DNA. Centromere protein (CENP)-B, a DNA binding protein, binds specifically to a 17-bp sequence motif (CENP-B box) that is present within every other ~170-bp monomer of a higher-order array. Binding of CENP-B promotes nucleosome phasing so that DNA between CENP-B boxes is wound around nucleosomes that contain CENP-A and core histones. Together with the tighter compaction of CENP-A nucleosomes, CENP-B may create the specialized higher-order structure of the centromere that visually appears as a primary constriction.

Annu. Rev. Genom. Human Genet. 2006.7:301-313. Downloaded from arjournals.annualreviews.org
by DUKE UNIVERSITY on 09/18/06. For personal use only.

topo II: topoisomerase II

Finally, the overall organization of the centromere region, coordinated by assembly of CEN chromatin and heterochromatin, may uniquely denote the centromere. Interspersion of subdomains of CENP-A and H3 nucleosomes in two dimensions (**Figure 3*a***) creates a unique three-dimensional structure in which CENP-A nucleosomes and H3 nucleosomes oppositely coalesce (**Figure 3*b***). Stacks of CENP-A nucleosomes are positioned toward the outward face of the chromosome to recruit additional kinetochore proteins and bi-orient sister kinetochores toward opposite spindle poles. Conversely, H3 subdomains are positioned toward the interior of the chromosome to establish a platform for heterochromatin assembly and CEN cohesion protein recruitment (**Figure 3*c***). Most centromeres are located in or near heterochromatin, suggesting that a repressive or distinctive surrounding chromatin environment may distinguish the centromere from the rest of the genome and/or aid in its function.

SEPARATION OF α-SATELLITE INTO DISTINCTIVE CHROMATIN DOMAINS

Epigenetic Mechanisms

At normal human centromeres, a fraction of the α-satellite arrays are assembled into CEN chromatin. In humans, the kinetochore domain comprises one half to two thirds of the satellite DNA found at eukaryotic centromeres (13) (C.D. Boivin & B.A. Sullivan, unpublished data). The remainder of the satellite DNA contributes to heterochromatin formation and chromatid cohesion (66) (A. Lam & B.A. Sullivan, unpublished data). Heterochromatin is cytologically dense material that forms predominantly at repetitive DNA sequences such as centromeres and telomeres. Its most notable property is its ability to silence euchromatic gene expression (28, 31). Centromeres in yeast, fruit flies, and mammals are flanked by heterochromatin, suggesting heterochromatin may have an important, conserved function in centromere structure and function. Heterochromatin formation is linked to chromatin regulation. It is triggered by noncoding CEN transcripts that initiate an assembly cascade that includes histone deacetylation, methylation of histone H3 at amino acid residue lysine 9, and recruitment of heterochromatin proteins, such as HP1, and cohesion and condensation proteins (31).

How are two distinct chromatin domains, CEN chromatin and heterochromatin, epigenetically assembled and maintained on the same continuous DNA array? The domains themselves are likely not fixed by specific boundaries within the α-satellite array, but may exhibit relative flexibility over the entire array, as long as each domain is present in the appropriate relative spatial location and functional proportion. Alternatively, heterochromatin may be anchored within divergent monomeric α-satellite flanking higher-order α-satellite. It may then spread into higher-order α-satellite to the edge of CEN chromatin containing CENP-A and H3K4me2, constraining the spread of CEN chromatin into non-CEN DNA. This latter model is supported by several findings. First, it has been shown that constitutive heterochromatin restricts the spread of CEN chromatin when CENP-A is overexpressed (25) (A. Lam & B.A. Sullivan, unpublished data). Furthermore, when centric heterochromatin is removed, CENP-A spreads into flanking sequences (34) (I.Y. Song & B.A. Sullivan, unpublished data).

Enzymatic Mechanisms

Enzymatic cleavage sites may also structurally define CEN domains. Topoisomerase II (topo II) is a component of the chromosome scaffold, creating transient double-strand breaks that allow decatenation (unwinding) of DNA during replication and chromosome segregation. Topo II cleavage sites are concentrated 150 kb–1 Mb from the short-arm edge of α-satellite arrays on at least three different human chromosomes

(61, 62). Topo II levels increase during the cell cycle and are concentrated at centromeres in G2/M, when metaphase kinetochore assembly and chromosome condensation occur. Topo II becomes sumoylated by PIAS, the enzyme protein inhibitor of activated signal transducers and activators of transcription proteins (STATs), promoting mitotic chromosome assembly, chromatid cohesion, and chromosome segregation during anaphase (8). Sumoylated Topo II is concentrated at the inner centromere region, where stacks of H3 nucleosomes that recruit heterochromatin and cohesion proteins are oriented. It is tempting to speculate that accumulation of topo II at centromeres in G2/M and topo II cleavage sites within α-satellite arrays are important for centromere and chromosome assembly at metaphase and centromere separation and anaphase. However, answers to these questions await more detailed functional studies.

PIAS: protein inhibitor of activated STATs

STAT: signal transducer and activator of transcription

CLOSING REMARKS

Despite substantial progress in identifying CEN components, much remains to be understood about centromere biology. Establishment and maintenance of human centromeres is likely dictated by both sequence characteristics of human α-satellite DNA and epigenetic mechanisms, such as unique structural features within nucleosomes containing the histone variant CENP-A. As genomic studies of α-satellite regions and functional studies of centromeres improve, we will better understand the processes involved in assembling this essential and complex locus.

SUMMARY POINTS

1. Despite the essential function of the centromere, the DNA that comprises it is rapidly evolving and varies from species to species.
2. Unequal crossover between sister chromatids is largely responsible for the continuing homogenization of higher-order α-satellite.
3. Primate centromere regions have undergone a series of events resulting in the addition of new material.
4. Centromere regions are subdivided into functional domains that function in kinetochore assembly and heterochromatin-mediated sister chromatid cohesion.
5. Post-translational modifications of core and variant histones within the centromere region distinguish CEN chromatin from heterochromatin and may contribute to proper kinetochore assembly.
6. Constitutive CENPs, such as CENP-A, -B, and -C, interact with α-satellite DNA to create a specialized, higher-order chromatin conformation that is important for centromere identity and kinetochore assembly and function.

ACKNOWLEDGMENTS

The authors wish to thank H. Willard and E. Green for insightful discussions.

LITERATURE CITED

1. Ahmad K, Henikoff S. 2002. The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Mol. Cell* 9:1191–200

Annu. Rev. Genom. Human Genet. 2006.7:301-313. Downloaded from arjournals.annualreviews.org by DUKE UNIVERSITY on 09/18/06. For personal use only.

2. Alexandrov I, Kazakov A, Tumeneva I, Shepelev V, Yurov Y. 2001. α-satellite DNA of primates: old and new families. *Chromosoma* 110:253–66
3. Alexandrov IA, Medvedev LI, Mashkova TD, Kisselev LL, Romanova LY, Yurov YB. 1993. Definition of a new α satellite suprachromosomal family characterized by monomeric organization. *Nucleic Acids Res.* 21:2209–15
4. Alkan C, Eichler EE, Bailey JA, Sahinalp SC, Tuzun E. 2004. The role of unequal crossover in α-satellite DNA evolution: a computational analysis. *J. Comput. Biol.* 11:933–44
5. Amor DJ, Choo KH. 2002. Neocentromeres: role in human disease, evolution, and centromere study. *Am. J. Hum. Genet.* 71:695–714
6. Ando S, Yang H, Nozaki N, Okazaki T, Yoda K. 2002. CENP-A, -B, and -C chromatin complex that contains the I-type α-satellite array constitutes the prekinetochore in HeLa cells. *Mol. Cell Biol.* 22:2229–41
7. Archidiacono N, Antonacci R, Marzella R, Finelli P, Lonoce A, Rocchi M. 1995. Comparative mapping of human alphoid sequences in great apes using fluorescence in situ hybridization. *Genomics* 25:477–84
8. Azuma Y, Arnaoutov A, Anan T, Dasso M. 2005. PIASy mediates SUMO-2 conjugation of Topoisomerase-II on mitotic chromosomes. *EMBO J.* 24:2172–82
9. Basu J, Willard HF. 2005. Artificial and engineered chromosomes: non-integrating vectors for gene therapy. *Trends Mol. Med.* 11:251–58
10. Bernard P, Allshire R. 2002. Centromeres become unstuck without heterochromatin. *Trends Cell Biol.* 12:419–24
11. Black BE, Foltz DR, Chakravarthy S, Luger K, Woods VL Jr, Cleveland DW. 2004. Structural determinants for generating centromeric chromatin. *Nature* 430:578–82
12. Blower MD, Karpen GH. 2001. The role of Drosophila CID in kinetochore formation, cell-cycle progression and heterochromatin interactions. *Nat. Cell Biol.* 3:730–39
13. Blower MD, Sullivan BA, Karpen GH. 2002. Conserved organization of centromeric chromatin in flies and humans. *Dev. Cell* 2:319–30
14. Brunet S, Vernos I. 2001. Chromosome motors on the move. From motion to spindle checkpoint activity. *EMBO Rep.* 2:669–73
15. Choo KH. 2001. Domain organization at the centromere and neocentromere. *Dev. Cell* 1:165–77
16. Cleveland DW, Mao Y, Sullivan KF. 2003. Centromeres and kinetochores: from epigenetics to mitotic checkpoint signaling. *Cell* 112:407–21
17. Durfy SJ, Willard HF. 1989. Patterns of intra- and interarray sequence variation in α satellite from the human X chromosome: evidence for short-range homogenization of tandemly repeated DNA sequences. *Genomics* 5:810–21
18. Durfy SJ, Willard HF. 1990. Concerted evolution of primate α satellite DNA. Evidence for an ancestral sequence shared by gorilla and human X chromosome α satellite. *J. Mol. Biol.* 216:555–66
19. Earnshaw WC, Rothfield N. 1985. Identification of a family of human centromere proteins using autoimmune sera from patients with scleroderma. *Chromosoma* 91:313–21
20. Fukagawa T, Mikami Y, Nishihashi A, Regnier V, Haraguchi T, et al. 2001. CENP-H, a constitutive centromere component, is required for centromere targeting of CENP-C in vertebrate cells. *EMBO J.* 20:4603–17
21. Goodman M. 1999. The genomic record of Humankind's evolutionary roots. *Am. J. Hum. Genet.* 64:31–39
22. Haaf T, Ward DC. 1994. Structural analysis of α-satellite DNA and centromere proteins using extended chromatin and chromosomes. *Hum. Mol. Genet.* 3:697–709

Annu. Rev. Genom. Human Genet. 2006.7:301-313. Downloaded from arjournals.annualreviews.org
by DUKE UNIVERSITY on 09/18/06. For personal use only.

23. Haaf T, Willard HF. 1997. Chromosome-specific α-satellite DNA from the centromere of chimpanzee chromosome 4. *Chromosoma* 106:226–32
24. Henikoff S, Ahmad K, Malik HS. 2001. The centromere paradox: stable inheritance with rapidly evolving DNA. *Science* 293:1098–102
25. Henikoff S, Ahmad K, Platero JS, van Steensel B. 2000. Heterochromatic deposition of centromeric histone H3-like proteins. *Proc. Natl. Acad. Sci. USA* 97:716–21
26. Henikoff S, Dalal Y. 2005. Centromeric chromatin: what makes it unique? *Curr. Opin. Genet. Dev.* 15:177–84
27. Ikeno M, Masumoto H, Okazaki T. 1994. Distribution of CENP-B boxes reflected in CREST centromere antigenic sites on long-range α-satellite DNA arrays of human chromosome 21. *Hum. Mol. Genet.* 3:1245–57
28. Karpen G. 1994. Position-effect variegation and the new biology of heterochromatin. *Curr. Biol.* 4:281–91
29. Kimura H. 2005. Histone dynamics in living cells revealed by photobleaching. *DNA Repair (Amsterdam)* 4:939–50
30. Kitagawa K, Masumoto H, Ikeda M, Okazaki T. 1995. Analysis of protein-DNA and protein-protein interactions of centromere protein B (CENP-B) and properties of the DNA-CENP-B complex in the cell cycle. *Mol. Cell Biol.* 15:1602–12
31. Lam AL, Pazin DE, Sullivan BA. 2005. Control of gene expression and assembly of chromosomal subdomains by chromatin regulators with antagonistic functions. *Chromosoma* 114:242–51
32. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, et al. 2001. Initial sequencing and analysis of the human genome. *Nature* 409:860–921
33. Lee C, Wevrick R, Fisher RB, Ferguson-Smith MA, Lin CC. 1997. Human centromeric DNAs. *Hum. Genet.* 100:291–304
34. Maggert KA, Karpen GH. 2001. The activation of a neocentromere in Drosophila requires proximity to an endogenous centromere. *Genetics* 158:1615–28
35. Maio JJ. 1971. DNA strand reassociation and polyribonucleotide binding in the African green monkey, *Cercopithecus aethiops*. *J. Mol. Biol.* 56:579–95
36. Manuelidis L. 1978. Chromosomal localization of complex and simple repeated human DNAs. *Chromosoma* 66:23–32
37. Masumoto H, Sugimoto K, Okazaki T. 1989. Alphoid satellite DNA is tightly associated with centromere antigens in human chromosomes throughout the cell cycle. *Exp. Cell Res.* 181:181–96
38. Mitchell AR, Gosden JR, Miller DA. 1985. A cloned sequence, p82H, of the alphoid repeated DNA family found at the centromeres of all human chromosomes. *Chromosoma* 92:369–77
39. Nakagawa H, Lee JK, Hurwitz J, Allshire RC, Nakayama J, et al. 2002. Fission yeast CENP-B homologs nucleate centromeric heterochromatin by promoting heterochromatin-specific histone tail modifications. *Genes. Dev.* 16:1766–78
40. Nishihashi A, Haraguchi T, Hiraoka Y, Ikemura T, Regnier V, et al. 2002. CENP-I is essential for centromere function in vertebrate cells. *Dev. Cell* 2:463–76
41. Nusbaum C, Mikkelsen TS, Zody MC, Asakawa S, Taudien S, et al. 2006. DNA sequence and analysis of human chromosome 8. *Nature* 439:331–35
42. Ohzeki J, Nakano M, Okada T, Masumoto H. 2002. CENP-B box is required for de novo centromere chromatin assembly on human alphoid DNA. *J. Cell Biol.* 159:765–75
43. Palmer DK, O'Day K, Margolis RL. 1989. Biochemical analysis of CENP-A, a centromeric protein with histone-like properties. *Prog. Clin. Biol. Res.* 318:61–72

Annu. Rev. Genom. Human Genet. 2006.7:301-313. Downloaded from arjournals.annualreviews.org
by DUKE UNIVERSITY on 09/18/06. For personal use only.

44. 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. *Proc. Natl. Acad. Sci. USA* 88:3734–38
45. Politi V, Perini G, Trazzi S, Pliss A, Raska I, et al. 2002. CENP-C binds the α-satellite DNA in vivo at specific centromere domains. *J. Cell Sci.* 115:2317–27
46. Rieder CL. 1982. The formation, structure, and composition of the mammalian kinetochore and kinetochore fiber. *Int. Rev. Cytol.* 79:1–58
47. Rosenberg H, Singer M, Rosenberg M. 1978. Highly reiterated sequences of SIMIANSIMIANSIMIANSIMIANSIMIAN. *Science* 200:394–402
48. Ross MT, Grafham DV, Coffey AJ, Scherer S, McLay K, et al. 2005. The DNA sequence of the human X chromosome. *Nature* 434:325–37
49. Rudd MK, Schueler MG, Willard HF. 2003. Sequence organization and functional annotation of human centromeres. *Cold Spring Harb. Symp. Quant. Biol.* 68:141–49
50. Rudd MK, Willard HF. 2004. Analysis of the centromeric regions of the human genome assembly. *Trends Genet.* 20:529–33
51. Rudd MK, Wray GA, Willard HF. 2006. The evolutionary dynamics of α-satellite. *Genome Res*. 116:89–96
52. Saffery R, Irvine DV, Griffiths B, Kalitsis P, Wordeman L, Choo KH. 2000. Human centromeres and neocentromeres show identical distribution patterns of >20 functionally important kinetochore-associated proteins. *Hum. Mol. Genet.* 9:175–85
53. Samonte RV, Ramesh KH, Verma RS. 1997. Comparative mapping of human alphoid satellite DNA repeat sequences in the great apes. *Genetica* 101:97–104
54. Schindelhauer D, Schwarz T. 2002. Evidence for a fast, intrachromosomal conversion mechanism from mapping of nucleotide variants within a homogeneous α-satellite DNA array. *Genome Res.* 12:1815–26
55. Schueler MG, Dunn JM, Bird CP, Ross MT, Viggiano L, et al. 2005. Progressive proximal expansion of the primate X chromosome centromere. *Proc. Natl. Acad. Sci. USA* 102:10563–68
56. Schueler MG, Higgins AW, Rudd MK, Gustashaw K, Willard HF. 2001. Genomic and genetic definition of a functional human centromere. *Science* 294:109–15
57. Shelby RD, Monier K, Sullivan KF. 2000. Chromatin assembly at kinetochores is uncoupled from DNA replication. *J. Cell Biol.* 151:1113–18
58. Shelby RD, Vafa O, Sullivan KF. 1997. Assembly of CENP-A into centromeric chromatin requires a cooperative array of nucleosomal DNA contact sites. *J. Cell Biol.* 136:501–13
59. Smit AF, Toth G, Riggs AD, Jurka J. 1995. Ancestral, mammalian-wide subfamilies of LINE-1 repetitive sequences. *J. Mol. Biol.* 246:401–17
60. Song K, Gronemeyer B, Lu W, Eugster E, Tomkiel JE. 2002. Mutational analysis of the central centromere targeting domain of human centromere protein C, (CENP-C). *Exp. Cell Res.* 275:81–91
61. Spence JM, Critcher R, Ebersole TA, Valdivia MM, Earnshaw WC, et al. 2002. Colocalization of centromere activity, proteins and topoisomerase II within a subdomain of the major human X α-satellite array. *EMBO J.* 21:5269–80
62. Spence JM, Fournier RE, Oshimura M, Regnier V, Farr CJ. 2005. Topoisomerase II cleavage activity within the human D11Z1 and DXZ1 α-satellite arrays. *Chromosome Res.* 13:637–48
63. Sugata N, Li S, Earnshaw WC, Yen TJ, Yoda K, et al. 2000. Human CENP-H multimers colocalize with CENP-A and CENP-C at active centromere–kinetochore complexes. *Hum. Mol. Genet.* 9:2919–26

Annu. Rev. Genom. Human Genet. 2006.7:301-313. Downloaded from arjournals.annualreviews.org
by DUKE UNIVERSITY on 09/18/06. For personal use only.

64. Sullivan B, Karpen G. 2001. Centromere identity in Drosophila is not determined in vivo by replication timing. *J. Cell Biol.* 154:683–90
65. Sullivan BA, Blower MD, Karpen GH. 2001. Determining centromere identity: cyclical stories and forking paths. *Nat. Rev. Genet.* 2:584–96
66. Sullivan BA, Karpen GH. 2004. Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. *Nat. Struct. Mol. Biol.* 11:1076–83
67. Sullivan KF. 2001. A solid foundation: functional specialization of centromeric chromatin. *Curr. Opin. Genet. Dev.* 11:182–88
68. Talbert PB, Bryson TD, Henikoff S. 2004. Adaptive evolution of centromere proteins in plants and animals. *J. Biol.* 3:18
69. Tanaka Y, Tachiwana H, Yoda K, Masumoto H, Okazaki T, et al. 2005. Human centromere protein B induces translational positioning of nucleosomes on α-satellite sequences. *J. Biol. Chem.* 280:41609–18
70. Tomkiel J, Cooke CA, Saitoh H, Bernat RL, Earnshaw WC. 1994. CENP-C is required for maintaining proper kinetochore size and for a timely transition to anaphase. *J. Cell Biol.* 125:531–45
71. Trazzi S, Bernardoni R, Diolaiti D, Politi V, Earnshaw WC, et al. 2002. In vivo functional dissection of human inner kinetochore protein CENP-C. *J. Struct. Biol.* 140:39–48
72. Tyler-Smith C, Floridia G. 2000. Many paths to the top of the mountain: diverse evolutionary solutions to centromere structure. *Cell* 102:5–8
73. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, et al. 2001. The sequence of the human genome. *Science* 291:1304–51
74. Warburton P, Willard H. 1996. Evolution of centromeric α satellite DNA: molecular organization within and between human and primate chromosomes. In *Human Genome Evolution*, ed. M Jackson, T Strachan, G Dover, pp. 121–45. Oxford: BIOS Sci. Publ.
75. Warburton PE. 2004. Chromosomal dynamics of human neocentromere formation. *Chromosome Res.* 12:617–26
76. Warburton PE, Haaf T, Gosden J, Lawson D, Willard HF. 1996. Characterization of a chromosome-specific chimpanzee α satellite subset: evolutionary relationship to subsets on human chromosomes. *Genomics* 33:220–28
77. Warburton PE, Willard HF. 1995. Interhomologue sequence variation of α satellite DNA from human chromosome 17: evidence for concerted evolution along haplotypic lineages. *J. Mol. Evol.* 41:1006–15
78. Waye JS, England SB, Willard HF. 1987. Genomic organization of α satellite DNA on human chromosome 7: evidence for two distinct alphoid domains on a single chromosome. *Mol. Cell. Biol.* 7:349–56
79. Willard HF. 1985. Chromosome-specific organization of human α satellite DNA. *Am. J. Hum. Genet.* 37:524–32
80. Willard HF. 1989. The genomics of long tandem arrays of satellite DNA in the human genome. *Genome* 31:737–44
81. Willard HF. 1991. Evolution of α satellite. *Curr. Opin. Genet. Dev.* 1:509–14
82. Yen TJ, Schaar BT. 1996. Kinetochore function: molecular motors, switches and gates. *Curr. Opin. Cell Biol.* 8:381–88
83. Yoda K, Ando S, Morishita S, Houmura K, Hashimoto K, et al. 2000. Human centromere protein A (CENP-A) can replace histone H3 in nucleosome reconstitution in vitro. *Proc. Natl. Acad. Sci. USA* 97:7266–71

Annu. Rev. Genom. Human Genet. 2006.7:301-313. Downloaded from arjournals.annualreviews.org
by DUKE UNIVERSITY on 09/18/06. For personal use only.

Annual Review of Genomics and Human Genetics

Volume 7, 2006

Contents

Annu. Rev. Genom. Human Genet. 2006.7:301-313. Downloaded from arjournals.annualreviews.org
by DUKE UNIVERSITY on 09/18/06. For personal use only.

A 60-Year Tale of Spots, Maps, and Genes
Victor A. McKusick 1

Transcriptional Regulatory Elements in the Human Genome
Glenn A. Maston, Sara K. Evans, and Michael R. Green 29

Predicting the Effects of Amino Acid Substitutions on Protein Function
Pauline C. Ng and Steven Henikoff 61

Genome-Wide Analysis of Protein-DNA Interactions
Tae Hoon Kim and Bing Ren 81

Protein Misfolding and Human Disease
Niels Gregersen, Peter Bross, Søren Vang, and Jane H. Christensen 103

The Ciliopathies: An Emerging Class of Human Genetic Disorders
Jose L. Badano, Norimasa Mitsuma, Phil L. Beales, and Nicholas Katsanis 125

The Evolutionary Dynamics of Human Endogenous Retroviral Families
Norbert Bannert and Reinhard Kurth 149

Genetic Disorders of Adipose Tissue Development, Differentiation, and Death
Anil K. Agarwal and Abhimanyu Garg 175

Preimplantation Genetic Diagnosis: An Overview of Socio-Ethical and Legal Considerations
Bartha M. Knoppers, Sylvie Bordet, and Rosario M. Isasi 201

Pharmacogenetics and Pharmacogenomics: Development, Science, and Translation
Richard M. Weinshilboum and Liewei Wang 223

Mouse Chromosome Engineering for Modeling Human Disease
Louise van der Weyden and Allan Bradley 247

The Killer Immunoglobulin-Like Receptor Gene Cluster: Tuning the Genome for Defense
Arman A. Bashirova, Maureen P. Martin, Daniel W. McVicar, and Mary Carrington 277

Structural and Functional Dynamics of Human Centromeric Chromatin
Mary G. Schueler and Beth A. Sullivan 301

Prediction of Genomic Functional Elements
Steven J.M. Jones 315

Of Flies and Man: *Drosophila* as a Model for Human Complex Traits
Trudy F.C. Mackay and Robert R.H. Anholt 339

The Laminopathies: The Functional Architecture of the Nucleus and Its Contribution to Disease
Brian Burke and Colin L. Stewart 369

Structural Variation of the Human Genome
Andrew J. Sharp, Ze Cheng, and Evan E. Eichler 407

Resources for Genetic Variation Studies
David Serre and Thomas J. Hudson 443

Indexes

Subject Index 459

Cumulative Index of Contributing Authors, Volumes 1–7 477

Cumulative Index of Chapter Titles, Volumes 1–7 480

Errata

An online log of corrections to *Annual Review of Genomics and Human Genetics* chapters may be found at http://genom.annualreviews.org/

Annu. Rev. Genom. Human Genet. 2006.7:301-313. Downloaded from arjournals.annualreviews.org by DUKE UNIVERSITY on 09/18/06. For personal use only.