The evolutionary dynamics of α-satellite

M. Katharine Rudd,[1,2,4] Gregory A. Wray,[1,3] and Huntington F. Willard[1,2,3,5]

[1]Institute for Genome Sciences & Policy, [2]Department of Molecular Genetics and Microbiology, and [3]Department of Biology, Duke University, Durham, North Carolina 27708, USA

α-Satellite is a family of tandemly repeated sequences found at all normal human centromeres. In addition to its significance for understanding centromere function, α-satellite is also a model for concerted evolution, as α-satellite repeats are more similar within a species than between species. There are two types of α-satellite in the human genome; while both are made up of ~171-bp monomers, they can be distinguished by whether monomers are arranged in extremely homogeneous higher-order, multimeric repeat units or exist as more divergent monomeric α-satellite that lacks any multimeric periodicity. In this study, as a model to examine the genomic and evolutionary relationships between these two types, we have focused on the chromosome 17 centromeric region that has reached both higher-order and monomeric α-satellite in the human genome assembly. Monomeric and higher-order α-satellites on chromosome 17 are phylogenetically distinct, consistent with a model in which higher-order evolved independently of monomeric α-satellite. Comparative analysis between human chromosome 17 and the orthologous chimpanzee chromosome indicates that monomeric α-satellite is evolving at approximately the same rate as the adjacent non-α-satellite DNA. However, higher-order α-satellite is less conserved, suggesting different evolutionary rates for the two types of α-satellite.

[Supplemental material is available online at www.genome.org.]

All α-satellite DNA is made up of tandem monomers, each ~171 bp in length (Manuelidis and Wu 1978; Willard and Waye 1987b). As revealed by patterns of monomer organization, there are two major types of α-satellite DNA in the human genome, designated higher-order and monomeric (Warburton and Willard 1996; Alexandrov et al. 2001; Rudd and Willard 2004). Higher-order α-satellite DNA is made up of monomers arranged in multimeric repeat units that are highly similar from repeat unit to repeat unit. These higher-order repeat units are positioned tandemly to make up an array of extremely homogeneous higher-order α-satellite typically several megabases in size. In contrast, monomeric α-satellite lacks detectable higher-order periodicity, and its constituent monomers are far less homogeneous than are higher-order repeat units (Rudd and Willard 2004). All normal human centromeres contain large arrays of higher-order α-satellite (Warburton and Willard 1996; Alexandrov et al. 2001), and, where investigated, these arrays have been found to be bordered by more heterogeneous monomeric α-satellite (Wevrick et al. 1992; Horvath et al. 2000; Schueler et al. 2001; Guy et al. 2003; Rudd and Willard 2004). The adjacent organization of higher-order and monomeric α-satellite, as well as the fact that lower primates have only monomeric α-satellite at their centromeres (Rosenberg et al. 1978; Musich et al. 1980; Maio et al. 1981; Thayer et al. 1981; Alves et al. 1994), has led to the hypothesis that higher-order α-satellite evolved from ancestral arrays of monomeric α-satellite and subsequently transposed to the centromeric regions of all great ape chromosomes (Warburton and Willard 1996; Alexandrov et al. 2001; Schueler et al. 2001, 2005; Kazakov et al. 2003). The relatively recent evolution of higher-order α-satellite is intriguing because centromere function is associated with higher-order and not monomeric α-satellite in the

human genome (Harrington et al. 1997; Ikeno et al. 1998; Schueler et al. 2001; Spence et al. 2002).

Like other tandem satellite families (Brown et al. 1972; Southern 1975; Coen et al. 1982), α-satellite is subject to concerted evolution, exhibiting greater sequence identity within a species than between species (Willard and Waye 1987b). For example, higher-order repeat units from an array on a particular chromosome are more similar to each other than to the orthologous repeats in other species (Jorgensen et al. 1987; Durfy and Willard 1990). The evolutionary process by which concerted evolution occurs is known as molecular drive, in which variants are able to spread quickly through a sequence family and fix in a population (Dover 1982). Molecular drive operates within and between chromosomes and includes mechanisms such as unequal crossing-over, gene conversion, and transposition (Dover 1982). Although many or all of these processes are likely to be participating in α-satellite evolution genome-wide (Alkan et al. 2004), the homogenization of tandem sequences at any given location can be largely accounted for by iterative rounds of unequal crossing-over leading to highly identical repeat units (Smith 1976; Schueler et al. 2001).

Current studies of centromere genomics and evolution rely largely on α-satellite assembled in the most recent build of the human genome assembly. However, despite its functional significance, the centromere has been largely omitted from the human genome assembly (Eichler et al. 2004; Rudd and Willard 2004). In fact, for each chromosome assembly there exists a centromere gap located at the edges of the most proximal p and q arm contigs, many of which terminate a substantial distance from the functional centromere (Rudd and Willard 2004; Rudd et al. 2004). Nonetheless, a few chromosome assemblies have reached an appreciable amount of α-satellite (International Human Genome Sequencing Consortium 2004; Rudd and Willard 2004; She et al. 2004; Ross et al. 2005) and provide a suitable resource to begin to address questions of centromere biology and evolution. We have previously characterized the types of α-satellite in the genome assembly (Rudd and Willard 2004), including their

physical relationship to extensive pericentromeric sequence duplications that are, in most cases, distal to α-satellite (Bailey et al. 2002; Eichler et al. 2004; Rudd and Willard 2004; She et al. 2004). Here, we examine in detail the α-satellite sequences to investigate the organization and evolutionary dynamics of α-satellite.

Results

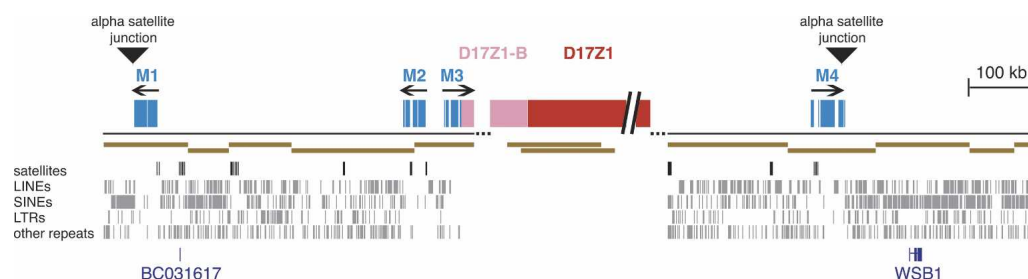### Genomic organization of the chromosome 17 centromeric region

α-Satellite DNA from chromosome 17 has been well characterized in numerous studies (Waye and Willard 1986; Wevrick and Willard 1989; Warburton and Willard 1990, 1995;) and is relatively well represented in the current genome assembly (Rudd and Willard 2004). Mapping studies have shown that the higher-order array D17Z1 ranges from ~1 to 4 Mb in size (Wevrick and Willard 1989; Warburton and Willard 1990) and is associated with the functional centromere (Harrington et al. 1997; Grimes et al. 2004). While Build 35 has not reached D17Z1 on either the p or q arm contigs of the chromosome 17 assembly, the 17p contig terminates in a related type of higher-order α-satellite, D17Z1-B (Fig. 1; Rudd et al. 2004). The D17Z1 higher-order repeat unit is 16 monomers long (Waye and Willard 1986), whereas the D17Z1-B repeat unit is comprised of 14 monomers. The two higher-order repeat units are both made up of monomers arranged in a pentameric fashion with corresponding monomers in the same order, suggesting that they diverged from a common ancestor upon a duplication or deletion event. Unlike other higher-order arrays on the same chromosome (Choo et al. 1990; Alexandrov et al. 1991; Wevrick and Willard 1991), D17Z1 and D17Z1-B are clearly part of the same family, with 92% sequence identity between colinear portions of the aligned higher-order repeat units (Rudd et al. 2004).

Based on the relative intensity of hybridization, the D17Z1-B array was estimated to be 500–900 kb among the chromosomes tested (data not shown; see Methods). FISH experiments with probes specific for D17Z1 and D17Z1-B supported the smaller size estimate for D17Z1-B and confirmed its location adjacent to D17Z1 on the p side of the centromere (Rudd et al. 2004). Although the region between D17Z1 and D17Z1-B has not been sequenced and assembled, BAC end sequencing also supports an adjacent organization of D17Z1 and D17Z1-B. Working draft quality BACs RPCI-11 5B18 (AC146710) and RPCI-11 449A3 (A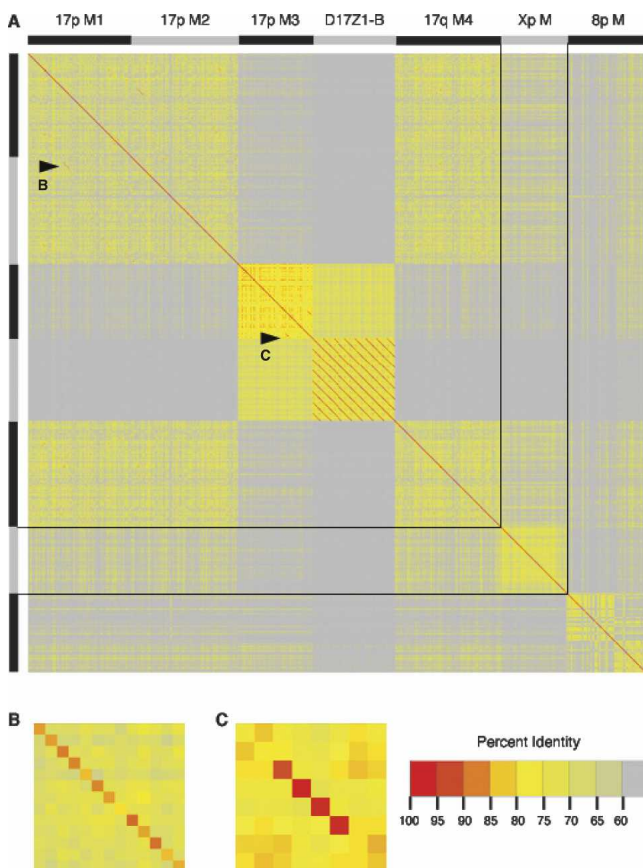C145197) both contain D17Z1 at one end and D17Z1-B at the opposite end. Furthermore, restriction digests with enzymes that release the higher-order repeat unit show both species of higher-order repeats in each BAC (Rudd et al. 2004).

Distal to the higher-order arrays, the current chromosome 17 assembly also includes four regions of monomeric α-satellite, three on the p side and one on the q side of the centromere gap (Fig. 1, M1–M4). As the chromosome 17q arm contig terminates before reaching higher-order α-satellite, there may be additional as yet undiscovered regions of monomeric α-satellite on 17q. In contrast to the two large higher-order repeat arrays, the four monomeric blocks are relatively short, each spanning 26–50 kb in length (Fig. 1). These blocks of monomeric α-satellite, as well as the regions in between monomeric blocks, are interspersed with other types of repeats. The junction between the most distal monomeric α-satellite and euchromatic sequences of the chromosome arms represents the "α-satellite junction" (Schueler et al. 2001). Like other centromeric regions (Guy et al. 2000; Schueler et al. 2001; Rudd and Willard 2004), the sequences proximal to the α-satellite junctions on chromosome 17 are enriched for other satellites as compared to the genome average. The collective concentration of non-α-satellites within the region defined by the two satellite junctions is 3.65%, >20-fold greater than the genome average (Supplemental Table 1). The concentration of other repeats such as LINEs, SINEs, LTRs, and DNA transposons proximal to the α-satellite junctions, however, is not enriched and is similar to that of the genome average. Distal to the α-satellite junctions, overall repeat content is comparable to that of the genome average (Supplemental Table 1). These data suggest a sharp demarcation between satellite-rich and euchromatic regions of the genome, as observed previously (Horvath et al. 2001; Schueler et al. 2001; She et al. 2004; Ross et al. 2005).

To search for expressed sequences close to α-satellite, we examined the Reference Sequence Collection (RefSeq; http://www.ncbi.nih.gov/RefSeq) (Fig. 1; Pruitt et al. 2000). On 17p, a sequence transcribed as a brain mRNA, BC031617, is located within the satellite-rich pericentromeric region, between the two most distal regions of monomeric α-satellite and only 33 kb from the nearest monomeric block. On 17q, the RefSeq gene *WSB1* lies within 96 kb of monomeric α-satellite. *WSB1* is also expressed in the brain and contains several WD repeats as well as a SOCS-box (Vasiliauskas et al. 1999). The genomic region between the functional centromere and the α-satellite junction (the "pericentromere"), therefore, is a complex region, containing blocks of monomeric α-satellite, other satellites, and at least some expressed sequences.

Figure 1. α-Satellite organization in the centromeric region of chromosome 17. The genomic landscape 500 kb distal of both sides of the centromere gap (dotted lines) is depicted. Blocks of monomeric α-satellite (light blue) are shown on both p and q arm contigs, and the p arm contig terminates with D17Z1-B higher-order α-satellite (pink). The proposed organization of D17Z1 (red) and D17Z1-B (pink) is shown inside the centromere gap. Arrows indicate the orientation of α-satellite monomers, and triangles show the junctions between α-satellite and non-satellite sequences (α-satellite junctions). BACs comprising the minimal tiling path are shown in brown, as are the two BACs containing D17Z1 at one end and D17Z1-B at the other. Other repeats are shown *below* the BAC contigs; from *top* to *bottom*, satellites (black), LINEs, SINEs, LTRs and other repeats (gray). The locations of RefSeq genes *BC031617* and *WSB1* are shown in dark blue at the *bottom*.

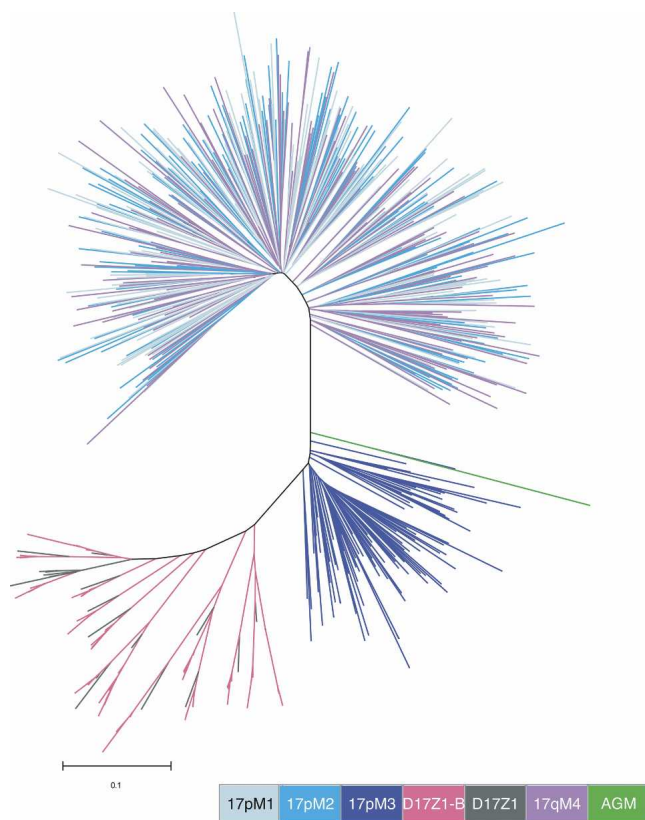cent identity on a color scale to view graphically the relationships between monomers (Fig. 2A).

Within each of the three most distal regions of monomeric α-satellite (M1, M2, M4), monomer percent identity was ~70%–73% (Table 1). Monomer percent identity was higher, however, in the most proximal region of monomeric α-satellite (M3) adjacent to D17Z1-B higher-order α-satellite, with a mean of 81.1% ± 3.2%. Short stretches of locally increased identity were apparent in proximal and distal monomeric α-satellite (Fig. 2B,C), presumably reflecting periodic short-range homogenization events; this is particularly apparent within proximal monomeric M3, evidenced by four duplicated monomers that are 97.4% identical (Fig. 2C). Notably, M3 monomers, in addition to being more homogeneous than M1, M2, and M4, were also more similar to higher-order D17Z1 and D17Z1-B monomers than were monomers from the more distal blocks of monomeric α-satellite (Fig. 2A; Table 1). These data suggest that there are two distinct classes of monomeric α-satellite in the chromosome 17 centromeric region. That M3 monomeric α-satellite is more closely related to higher-order α-satellite would be consistent with its participation in the early events that gave rise to higher-order α-satellite on chromosome 17 before becoming physically isolated from the higher-order arrays.

Interestingly, monomers within distal monomeric α-satellite are just as identical within a region of monomeric α-satellite as between regions of distal monomeric α-satellite, even between regions of distal monomeric on opposite chromosome arms (Fig. 2A; Table 1). A duplication of 13 highly similar monomers (overall 88.0% identical) present in the same order in distal monomeric regions M1 and M2 on 17p suggests that these sequences were subject to intrachromosomal homogenization mechanisms (Fig. 2B). Given the concentration of inter- and intrachromosomal segmental duplications bordering regions of distal monomeric α-satellite in the pericentromeric region of chromosome 17 (Bailey et al. 2002; Rudd and Willard 2004), blocks of monomeric α-satellite may have undergone exchanges via transposition mechanisms as well.

We also used neighbor-joining methods to examine phylogenetic relationships among chromosome 17 α-satellite monomers. In addition to the higher-order and monomeric α-satellite found in the chromosome 17 assembly, we also included monomers from D17Z1 higher-order α-satellite (Waye and Willard 1986) and a monomer from African Green Monkey α-satellite (Rosenberg et al. 1978). African Green Monkeys have only monomeric α-satellite at their centromeres (Rosenberg et al. 1978; Thayer et al. 1981; Haaf et al. 1992; Goldberg et al. 1996), and this sequence serves as an outgroup for our phylogenetic analysis. The resulting phylogenetic tree contains three major clades (Fig. 3). The related higher-order α-satellite from D17Z1 and

Figure 2. Percent identity scores for pairwise comparisons of α-satellite monomers. All pairwise comparisons were calculated for α-satellite monomers and percent identity scores were depicted according to the color scale. The chromosomal origin of α-satellite monomers is shown at the *top* of the figure in alternating black and gray bars. (*A*) Pairwise comparisons for monomers from the assemblies of chromosomes 8 and 17 and the X chromosome. Black lines indicate the boundaries of monomers from each chromosome. (*B*,*C*) Detailed versions of percent identity scores from regions indicated by arrowheads.

Higher-order and monomeric α-satellite on chromosome 17

To explore the evolutionary relationships of α-satellite monomers on chromosome 17, all α-satellite within 1 Mb of the centromere gap in the chromosome 17 assembly was broken into basic ~171-bp monomers and compared to each other. We performed CLUSTALW alignments (Thompson et al. 1994) between all possible pairwise combinations of monomers (617 monomers, 380,072 unique alignments) and expressed each alignment per-

Table 1. Mean percent identity among monomers from regions of α-satellite

	No. mons	α-satellite regions							
		17p M1	17p M2	17p M3	D17Z1-B	D17Z1	17q M4	Xp M	8p M
17p M1	141	72.2 ± 3.8	69.1 ± 4.5	63.6 ± 3.1	57.2 ± 3.1	58.0 ± 3.5	71.8 ± 4.1	65.4 ± 3.3	65.8 ± 3.6
17p M2	133		70.5 ± 4.1	62.1 ± 3.1	58.1 ± 3.1	56.3 ± 4.0	70.9 ± 4.1	62.7 ± 3.4	60.3 ± 3.8
17p M3	97			81.1 ± 3.2	70.7 ± 2.9	71.3 ± 3.5	65.8 ± 3.3	63.9 ± 3.2	69.0 ± 3.4
D17Z1-B	14				74.3 ± 5.0	76.8 ± 6.9	58.7 ± 3.5	57.2 ± 2.9	60.6 ± 3.3
D17Z1	16					75.9 ± 6.3	59.2 ± 3.9	57.6 ± 3.2	61.6 ± 3.4
17q M4	139						72.8 ± 4.3	67.1 ± 3.8	70.4 ± 3.9
Xp M	85							71.3 ± 3.8	65.1 ± 3.8
8p M	104								72.0 ± 4.7

from five diverse populations were positive for the junction PCR, and sequencing one individual from each population showed 100% identity among PCR products (see Methods). These data suggest that the junction between higher-order and monomeric α-satellite is fixed among human populations.

Monomeric α-satellite evolution

Studies involving higher-order α-satellite (Warburton and Willard 1995), as well as other satellite families (Coen and Dover 1983; Ohta and Dover 1983), have shown that intrachromosomal exchanges occur much more rapidly than do interchromosomal exchanges. To see if this was also true in monomeric α-satellite, we examined the phylogenetic relationships among higher-order and monomeric α-satellites from chromosomes 8 and 17 and the X chromosome (see Methods) using neighbor-joining methods. The resulting phylogenetic tree (Fig. 4) has a very similar topology to the chromosome 17 tree (Fig. 3). Higher-order α-satellite from the three chromosomes forms a distinct clade, and subclades are grouped by higher-order suprachromosomal subfamily. The monomers that comprise higher-order α-satellite from chromosome 17 and the X chromosome have a pentameric structure, suggesting ancient interchromosomal exchange(s) (Waye and Willard 1986). As such, the monomers from DXZ1, D17Z1, and D17Z1-B fall into one of five distinct sub-

Figure 3. Phylogenetic tree of α-satellite on chromosome 17. Neighbor-joining methods were used to generate the phylogenetic tree containing both higher-order and monomeric α-satellite from chromosome 17. The tree contains 641 monomers, including the outgroup monomer from the African green monkey. The key at the *bottom* of the figure indicates the chromosomal origin of the monomers. Monomeric α-satellite 17pM1, 17pM2, 17pM3, 17qM4; higher-order α-satellite D17Z1 and D17Z1-B; and α-satellite from the African Green Monkey (AGM) are shown.
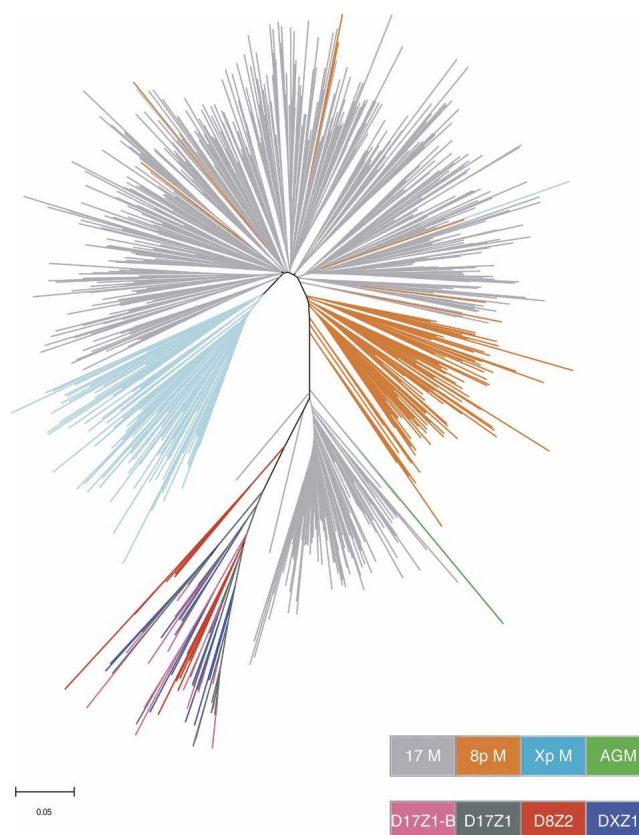
D17Z1-B forms a clade as expected, while distal monomeric α-satellite from both p and q arms forms a separate clade. Proximal monomeric α-satellite (M3) forms a third clade closest to the root as defined by African Green Monkey α-satellite. The junction between higher-order and M3 monomeric α-satellite on 17p is very distinct. An unusual 220-bp monomer clearly demarcates the division between homogeneous D17Z1-B higher-order repeat units and more divergent monomers in M3. Additionally, all monomers on either side of the junction fall into the higher-order clade or the proximal monomeric clade as predicted. This is very different from the more gradual (over 10 kb) transition described previously between higher-order and monomeric α-satellite on the X chromosome (Schueler et al. 2001; Ross et al. 2005). The 220-bp monomer on 17p may have punctuated the higher-order/monomeric junction since it would be misaligned for future unequal crossover events, while monomers in the transition regions of the X may have continued to cross over for a period of time after the fixation of DXZ1 higher-order α-satellite.

Since α-satellite is evolving rapidly and array size varies among individuals (Wevrick and Willard 1989; Mahtani and Willard 1990), we designed a PCR assay to specifically amplify the 220-bp monomer junction in order to test whether this junction was static or subject to slippage. All 30 individuals sampled

Figure 4. Neighbor-joining tree of monomers from different chromosomes. Neighbor-joining methods were used to generate the phylogenetic tree containing higher-order and monomeric α-satellite from chromosomes 8 and 17 and the X chromosome. The key at the *bottom* of the figure indicates the chromosomal origin of the monomers from monomeric α-satellite (8p M, 17 M, and Xp M) and higher-order α-satellite (D8Z2, D17Z1, D17Z1-B, and DXZ1).

clades (Willard and Waye 1987a). Higher-order α-satellite from chromosome 8 is a member of a dimeric suprachromosomal family (Ge et al. 1992), and consequently its monomers fall into two subclades that are distinct from the pentamer family subclades.

The distal monomeric α-satellites from each chromosome are present in one large clade, separate from the 17p M3 monomeric clade described above and separate from the higher-order repeat clade. Notably, however, the majority of monomers within the distal monomeric clade fall into chromosome-specific subclades (Fig. 4) that do not intermix. This finding contrasts with an earlier report based on monomers from individual BAC clones that were reported to "mix well" (Alkan et al. 2004). The results from Alkan et al. could reflect a more diverse population of monomeric monomers (from five different chromosomes) or could be due to variation in the rate of monomeric interchromosomal exchange between different chromosomes. In our study, only nine monomers (~1% of the total number of 760 monomers examined) were assigned to a chromosome-specific subclade other than the chromosome on which they are located. Given the large number of monomers in this analysis, their length of just 171 bp, and the pairwise comparisons that determine neighbor-joining trees, we interpret the placement of these few monomers to be errors of phylogenetic inference rather than evidence for monomer exchange among subfamilies. Chromosome-specific subclades within the distal monomeric clade are further supported by maximum likelihood methods (see Methods; Supplemental Fig. 1). The relationship among α-satellite monomers from different chromosomes is also evident in our sequence alignment results (Fig. 2A). Combined, these data indicate that although distal monomeric monomers are more similar to monomeric α-satellite from other chromosomes than neighboring higher-order α-satellite from the same chromosome, there has been sufficient local homogenization within regions of monomeric α-satellite on a given chromosome to drive the evolution of some chromosome specificity.

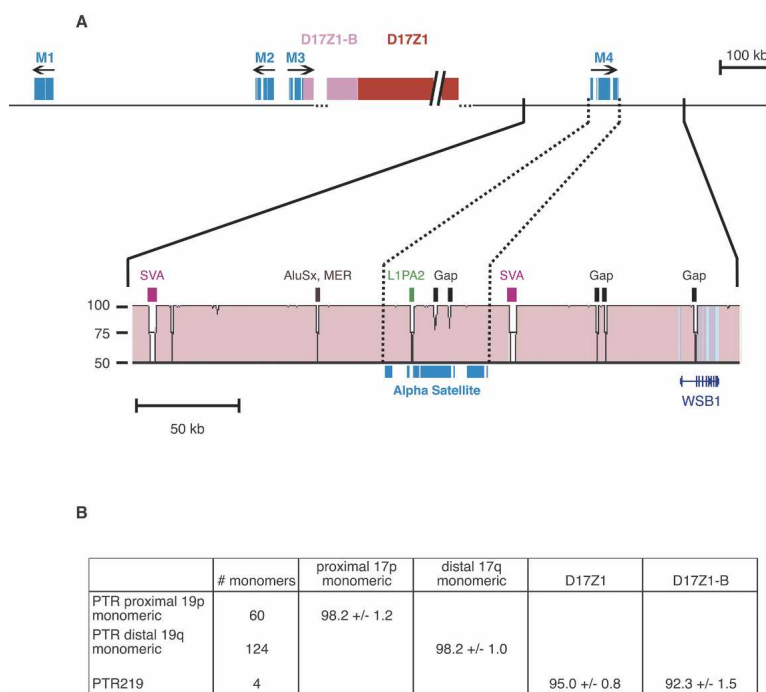## Comparative analysis of α-satellite in primates

To better understand concerted evolution of α-satellite in primates, we also examined α-satellite organization on the chimpanzee chromosome orthologous to human chromosome 17, PTR 19. The initial assembly of the *Pan troglodytes* genome (Build 1, November 2003) includes α-satellite on both the p and q arm sides of the centromere gap. This analysis was necessarily limited by the existence of several gaps in one or the other centromeric region. Nonetheless, in order to evaluate sequence conservation between monomeric α-satellite in two species, we performed VISTA alignments (http://pipeline.lbl.gov/cgi-bin/gateway2) (Couronne et al. 2003) be-

tween a 300-kb region including monomeric α-satellite on 17q (M4) and the orthologous region on PTR 19q (Fig. 5A). This region has relatively few gaps in the chimpanzee assembly, thus it provides a reasonable model to address the amount of sequence conservation between the two species.

Overall, the two sequences are 98.0% identical along the 277 kb of aligned sequence. This includes high percentage identity between the chimp and human *WSB1* orthologs. The monomeric α-satellite found in this region is also highly conserved; orthologous monomers are 98.2% ± 1.0% identical (Fig. 5B), similar to the overall sequence conservation in the region and similar to genome-wide estimates of human–chimp sequence identity (Watanabe et al. 2004; The Chimpanzee Sequencing and Analysis Consortium 2005). Although the PTR 19p assembly is not as comprehensive as the 19q side, a lesser number of chimpanzee monomers corresponding to M3 on 17p have also been assembled (data not shown). We compared orthologous monomers in this region from the two species and again found high sequence identity; the 60 aligned monomers are 98.2% ± 1.2% identical (Fig. 5B).

The PTR chromosome 19 assembly has not reached higher-order α-satellite; however, an apparently orthologous higher-order repeat, PTR219, has been reported previously (Warburton et al. 1996). We compared the sequences of PTR219, D17Z1, and D17Z1-B to determine the evolutionary relationships between the three higher-order repeats. Overall, PTR219 and D17Z1 are 95.0% identical, while PTR219 and D17Z1-B are only 92.3% iden-

	# monomers	proximal 17p monomeric	distal 17q monomeric	D17Z1	D17Z1-B
PTR proximal 19p monomeric	60	98.2 +/- 1.2			
PTR distal 19q monomeric	124		98.2 +/- 1.0		
PTR219	4			95.0 +/- 0.8	92.3 +/- 1.5

Figure 5. Genomic organization of 17q compared to the orthologous *Pan troglodytes* region. (*A*) The genomic organization of the chromosome 17 centromeric region is depicted; α-satellite is colored as described in Figure 1. A 300-kb region containing monomeric α-satellite on human chromosome arm 17q was compared to the orthologous region of chimpanzee chromosome arm 19q. A VISTA alignment of the two regions is shown in pink. Percent identity between aligning sequences is indicated by the *y*-axis (50%–100% identical). α-Satellite (blue) and the gene *WSB1* (purple) are shown. Areas of low percent identity between the two genomes can be explained by gaps in the chimpanzee assembly (black boxes) or sequences inserted in the human genome or deleted from the chimpanzee genome (colored boxes). (*B*) Mean percent identity ± one standard deviation for aligned α-satellite monomers from chimpanzees and humans.

tical (Fig. 5B). These data are consistent with divergence between chimpanzee and human higher-order α-satellite present on the X chromosome; human and chimpanzee copies of DXZ1 are 93.0% identical (Laursen et al. 1992). Thus, even though our comparative analysis of higher-order α-satellite on chromosome 17 is limited to the four monomers of PTR219 available, data from both chromosome 17 and the X chromosome support a higher rate of divergence among higher-order as compared to monomeric α-satellite.

Discussion

The evolution of α-satellite in primates gave rise to two distinct types that differ in genomic organization as well as function. Among human centromeres, higher-order arrays are several megabases in size, flanked by smaller stretches of monomeric α-satellite. Higher-order α-satellite within an array is extremely homogeneous, and few other sequences have been found embedded within higher-order α arrays (Schueler et al. 2001; Ross et al. 2005). In contrast, monomeric α-satellite is more heterogeneous in sequence and is extensively interspersed with non-α-satellite sequences (Schueler et al. 2001; Guy et al. 2003; Kazakov et al. 2003; Rudd and Willard 2004).

The evolution of higher-order from ancestral monomeric α-satellite (belonging to suprachromosomal family 4 of Alexandrov et al. 2001) can be modeled by unequal crossover events (Smith 1976; Alkan et al. 2004). After a series of mutations or duplications create homology between two previously unique sequences, an unequal crossover might occur between the two homologous sequences, creating a tandem duplication. Subsequent unequal crossovers can expand the number of tandem repeats, or monomers in the case of α-satellite. A second layer of complexity arises when a subset of monomers is multimerized into higher-order α-satellite. Unequal crossovers between misaligned higher-order repeat units will occur more frequently than between monomeric monomers because of the extremely high homology between repeat units, leading to an expansion and contraction of higher-order α-satellite (Willard and Waye 1987b). The nature of higher-order α-satellite expansion allows for the efficient spread and fixation of a sequence variant within a particular chromosomal locus. Thus, higher-order and monomeric α-satellites are predicted to evolve at different rates, causing orthologous higher-order repeat units to be less conserved than orthologous monomeric α-satellite in closely related species. Our analyses demonstrate that this is, indeed, the case for α-satellite in chimpanzees and humans (Fig. 5B).

It should be noted that the hypothesized ancestral relationship between extant monomeric and higher-order repeat families is no longer evident from examination of available sequences (Alkan et al. 2004). While it is logical that at least the initial higher-order repeats must have evolved from ancestral monomeric arrays, they must have then transposed to other centromeric sites in the genome and taken on sequence and/or epigenetic attributes that allowed them to replace monomeric α-satellite as the functional centromere (Alexandrov et al. 2001; Schueler et al. 2001, 2005; Kazakov et al. 2003). Subsequent accumulation of mutations in (nonfunctional) monomeric α-satellite and concerted evolution of the (functional) higher-order repeats over the intervening 15–25 million years may have masked many of the hypothesized ancestral relationships.

Notwithstanding these limitations, the available sequence data now permit an analysis of the relationships between monomeric α-satellite monomers on different chromosomes. Like higher-order α-satellite, monomeric α-satellite has been shaped by inter- and intrachromosomal exchange mechanisms. Interchromosomal exchange has given rise to related suprachromosomal families of higher-order α-satellite (Waye and Willard 1986; Willard and Waye 1987a; Choo et al. 1989; Alexandrov et al. 1993), while intrachromosomal exchange has produced chromosome-specific arrays of higher-order α-satellite (Willard and Waye 1987b; Durfy and Willard 1989; Warburton and Willard 1996; Schueler et al. 2001; Schindelhauer and Schwarz 2002). As shown in this study, intrachromosomal exchange creates chromosome-specific regions of monomeric α-satellite; however, in other cases, interchromosomal exchange may produce monomeric α-satellite that "mixes well" with monomeric α-satellite from other chromosomes (Alkan et al. 2004). Monomeric α-satellite appears to diverge less rapidly than does higher-order α-satellite, as evidenced by the conservation between human and likely orthologous chimpanzee α-satellites (Fig. 5). This result is itself somewhat counterintuitive, since it is higher-order α-satellite that is associated with centromere function (Harrington et al. 1997; Ikeno et al. 1998; Schueler et al. 2001; Spence et al. 2002) and thus might be predicted to be subject to selection. To explore the basis for this apparent paradox will require parallel studies of centromere function and comparative studies of the molecular evolution of both centromeric and pericentromeric genomic sequences. As part of such studies, it will be interesting to examine the relationships between higher-order and monomeric α-satellites among all chromosomes in the human genome to better model the sequence of events that created the distinct yet related organization of α-satellite on different chromosomes.

Methods

D17Z1-B array estimate

To determine the size of the D17Z1-B array relative to D17Z1, we analyzed three chromosomes 17 in which the D17Z1 array had been previously mapped. The D17Z1 arrays on chromosomes 17 in hybrid cell lines L65-14A, LT23-4C, and L745 are 3.7 Mb, 3.3 Mb, and 2.8 Mb, respectively (Warburton and Willard 1990). As D17Z1-B higher-order α-satellite is only 92% identical to D17Z1 higher-order α-satellite (Rudd et al. 2004), we used stringent Southern washing conditions (68°C, 0.5% SDS/0.1× SSC) to distinguish D17Z1 from D17Z1-B and calculate the amount of D17Z1-B relative to D17Z1.

We digested genomic DNA from hybrid cell lines L65-14A, LT23-4C, and L745 with EcoRI. Both D17Z1 and D17Z1-B contain EcoRI sites that digest each array into their respective higher-order repeat units. Digested DNA was analyzed on parallel blots, probing one with the D17Z1 higher-order repeat unit (Waye and Willard 1986) and the other with the D17Z1-B higher-order repeat unit (Rudd et al. 2004). D17Z1 higher-order repeats exist in several different lengths, ranging from 16 monomers (16-mers) to 12-mers (Warburton and Willard 1990). However, among all three chromosomes 17 analyzed here, D17Z1-B was only found as a 14-mer higher-order repeat unit. Using a PhosphorImager, we calculated the pixel intensity of all the bands. The ratio of the D17Z1-B 14-mer band to the sum of the D17Z1 bands allowed us to estimate the relative size of the D17Z1-B array. Based on the known sizes of the D17Z1 arrays from the chromosomes 17 in cell lines L65-14A, LT23-4C, and L745 (Warburton and Willard

1990), we estimated the D17Z1-B arrays to be ~930 kb, ~560 kb and ~500 kb, respectively.

Sequence alignments and phylogenetic analysis

We used the UCSC browser (http://genome.ucsc.edu) (Kent et al. 2002) to extract sequences from the May 2004 assembly (Build 35) of the human genome and the November 2003 assembly (Build 1) of the *Pan troglodytes* genome. To identify individual α-satellite monomers from the human chromosome 17 and the chimpanzee chromosome 19 centromeres, we RepeatMasked (http://repeatmasker.genome.washington.edu) the sequences 1 Mb proximal on the p and q sides of the centromere gaps using a custom RepeatMasker library containing α-satellite consensus monomers from the five suprachromosomal families (Alexandrov et al. 2001), as well as the α-satellite monomers included in the default RepeatMasker library (Smit 1999). We isolated 617 monomers from the human chromosome 17 assembly and 308 monomers from the chimpanzee chromosome 19 assembly. We also extracted monomers from the most distal regions of monomeric α-satellite on the p arms of the X chromosome (Schueler et al. 2001) and chromosome 8. Forty-one kilobases from Xp (from BAC RPCI-11 12L2, UCSC coordinates chrX:57966230–58007230) and 20 kb from 8p (from BAC RPCI-11 643N23, UCSC coordinates chr8:43546489–43566789) were RepeatMasked to isolate 85 and 104 monomers, respectively. We used CLUSTALW (Thompson et al. 1994) to compute all pairwise alignments among monomers from human chromosomes 8 and 17 and the X chromosome. Pairwise percent identity scores were translated into particular color values to generate the heat map shown in Figure 2 using Spotfire version 6.0 (Dresen et al. 2003). Monomers from chimpanzee chromosome 19 and clone PTR219 (Warburton et al. 1996) were compared to monomers from human chromosome 17 using CLUSTALW, to yield the orthologous percent identity values in Figure 5B.

To generate the chromosome 17 phylogenetic tree (Fig. 3), we isolated higher-order and monomeric monomers. The 617 monomers from the human chromosome 17 assembly (monomeric monomers as well as D17Z1-B monomers) were added to the 16 monomers making up D17Z1 higher-order α-satellite, one African Green Monkey monomer, and seven monomers from the BAC ends of BACs spanning D17Z1 and D17Z1-B (RPCI-11 5B18 and RPCI-11 449A3). The 641 total monomers were aligned using CLUSTALW, and manually examined and edited using MacClade (http://macclade.org/macclade.html). Subsequent MEGA (Molecular Evolutionary Genetic Analysis, version 2.1; http://www.megasoftware.net) phylogenetic analyses were performed (Kumar et al. 2001). Neighbor-joining methods were used with pairwise deletion parameters and 1000 bootstrap iterations.

For the interchromosomal neighbor-joining tree (Fig. 4), we used monomeric α-satellite from chromosomes 8 and 17 and the X chromosome as described above. We added monomers from D8Z2 (Ge et al. 1992), D17Z1 (Waye and Willard 1986), D17Z1-B (Rudd et al. 2004), and DXZ1 (Waye and Willard 1985) higher-order α-satellite for a total of 760 monomers. MEGA phylogenetic analysis was performed as described for the chromosome 17 tree. We confirmed the topology of the interchromosomal tree by repeating the analysis using maximum likelihood methods (Supplemental Fig. 1). Starting with the same 760 monomers, we first generated a neighbor-joining tree using PAUP 4.0 (http://paup.csit.fsu.edu/downl.html). Next, we used ModelTest (http://darwin.uvigo.es/software/modeltest.html) to perform a likelihood ratio test based on the neighbor-joining tree topology in order to find rate parameters and an appropriate evolutionary model. Based on this analysis, we estimated the maximum like-lihood tree with PAUP 4.0 using a general time-reversible (GTR) model with γ, rate matrix of (1.13943 2.43650 0.83944 1.32101 2.91431), α(G) of 4.54644, empirical nucleotide frequencies, and branch-swapping with a nearest-neighbor interchange of 3.

Junction PCR

PCR primers for the junction between D17Z1-B and proximal monomeric α-satellite were designed to amplify only the junction fragment and not other α-satellite sequences. Primers jxnF (5′-CAGATTCTACAACAAGGGTG-3′) and jxnR (5′-GATGT ATGCATTCATCACAG-3′) amplify a 298-bp product at high stringency conditions (5-min initial denaturation at 94°C followed by 30 cycles of 94°C for 30 sec, 60°C for 30 sec, and 72°C for 20 sec). Genomic DNA from individuals from five diverse populations was purchased from Coriell Cell Repositories and amplified using the junction PCR primers. DNAs from 10 Europeans, seven Africans North of the Sahara, nine Africans South of the Sahara, seven Pacific Islanders, and 10 Chinese were tested using this PCR assay, and PCR products from one individual from each population were sequenced.

References

Alexandrov, I.A., Mashkova, T.D., Akopian, T.A., Medvedev, L.I., Kisselev, L.L., Mitkevich, S.P., and Yurov, Y.B. 1991. Chromosome-specific α satellites: Two distinct families on human chromosome 18. *Genomics* **11:** 15–23.

Alexandrov, I.A., Medvedev, L.I., Mashkova, T.D., Kisselev, L.L., Romanova, L.Y., and Yurov, Y.B. 1993. Definition of a new α satellite suprachromosomal family characterized by monomeric organization. *Nucleic Acids Res.* **21:** 2209–2215.

Alexandrov, I., Kazakov, A., Tumeneva, I., Shepelev, V., and Yurov, Y. 2001. α-Satellite DNA of primates: Old and new families. *Chromosoma* **110:** 253–266.

Alkan, C., Eichler, E.E., Bailey, J.A., Sahinalp, S.C., and Tuzun, E. 2004. The role of unequal crossover in α-satellite DNA evolution: A computational analysis. *J. Comp. Biol.* **11:** 933–944.

Alves, G., Seuanez, H.N., and Fanning, T. 1994. α-Satellite DNA in neotropical primates (Platyrrhini). *Chromosoma* **103:** 262–267.

Bailey, J.A., Gu, Z., Clark, R.A., Reinert, K., Samonte, R.V., Schwartz, S., Adams, M.D., Myers, E.W., Li, P.W., and Eichler, E.E. 2002. Recent segmental duplications in the human genome. *Science* **297:** 945–947.

Brown, D.D., Wensink, P.C., and Jordan, E. 1972. A comparison of the ribosomal DNA's of *Xenopus laevis* and *Xenopus mulleri*: The evolution of tandem genes. *J. Mol. Biol.* **63:** 57–73.

The Chimpanzee Sequencing and Analysis Consortium. 2005. Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* **437:** 69–87.

Choo, K.H., Vissel, B., and Earle, E. 1989. Evolution of α DNA on human acrocentric chromosomes. *Genomics* **5:** 332–344.

Choo, K.H., Earle, E., Vissel, B., and Filby, R.G. 1990. Identification of two distinct subfamilies of α satellite DNA that are highly specific for human chromosome 15. *Genomics* **7:** 143–151.

Coen, E.S. and Dover, G.A. 1983. Unequal exchanges and the coevolution of X and Y rDNA arrays in *Drosophila melanogaster*. *Cell* **33:** 849–855.

Coen, E., Strachan, T., and Dover, G. 1982. Dynamics of concerted evolution of ribosomal DNA and histone gene families in the *melanogaster* species subgroup of *Drosophila*. *J. Mol. Biol.* **158:** 17–35.

Couronne, O., Poliakov, A., Bray, N., Ishkhanov, T., Ryaboy, D., Rubin, E., Pachter, L., and Dubchak, I. 2003. Strategies and tools for whole-genome alignments. *Genome Res.* **13:** 73–80.

Dover, G. 1982. Molecular drive: A cohesive mode of species evolution. *Nature* **299:** 111–117.

Dresen, I.M., Husing, J., Kruse, E., Boes, T., and Jockel, K.H. 2003. Software packages for quantitative microarray-based gene expression analysis. *Curr. Pharm. Biotechnol.* **4:** 417–437.

Durfy, S.J. and Willard, H.F. 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:** 810–821.

———. 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:** 555–566.

Eichler, E.E., Clark, R.A., and She, X. 2004. An assessment of the sequence gaps: Unfinished business in a finished human genome. *Nat. Rev. Genet.* **5:** 345–354.

Ge, Y., Wagner, M.J., Siciliano, M., and Wells, D.E. 1992. Sequence, higher order repeat structure, and long-range organization of α satellite DNA specific to human chromosome 8. *Genomics* **13:** 585–593.

Goldberg, I.G., Sawhney, H., Pluta, A.F., Warburton, P.E., and Earnshaw, W.C. 1996. Surprising deficiency of CENP-B binding sites in African green monkey α-satellite DNA: Implications for CENP-B function at centromeres. *Mol. Cell. Biol.* **16:** 5156–5168.

Grimes, B.R., Babcock, J., Rudd, M.K., Chadwick, B., and Willard, H.F. 2004. Assembly and characterization of heterochromatin and euchromatin on human artificial chromosomes. *Genome Biol.* **5:** R89.

Guy, J., Spalluto, C., McMurray, A., Hearn, T., Crosier, M., Viggiano, L., Miolla, V., Archidiacono, N., Rocchi, M., Scott, C., et al. 2000. Genomic sequence and transcriptional profile of the boundary between pericentromeric satellites and genes on human chromosome arm 10q. *Hum. Mol. Genet.* **9:** 2029–2042.

Guy, J., Hearn, T., Crosier, M., Mudge, J., Viggiano, L., Koczan, D., Thiesen, H.J., Bailey, J.A., Horvath, J.E., Eichler, E.E., et al. 2003. Genomic sequence and transcriptional profile of the boundary between pericentromeric satellites and genes on human chromosome arm 10p. *Genome Res.* **13:** 159–172.

Haaf, T., Warburton, P.E., and Willard, H.F. 1992. Integration of human α satellite DNA into simian chromosomes: Centromere protein binding and disruption of normal chromosome segregation. *Cell* **70:** 681–696.

Harrington, J.J., Van Bokkelen, G., Mays, R.W., Gustashaw, K., and Willard, H.F. 1997. Formation of de novo centromeres and construction of first-generation human artificial microchromosomes. *Nat. Genet.* **15:** 345–355.

Horvath, J.E., Viggiano, L., Loftus, B.J., Adams, M.D., Archidiacono, N., Rocchi, M., and Eichler, E.E. 2000. Molecular structure and evolution of an α satellite/non-α satellite junction at 16p11. *Hum. Mol. Genet.* **9:** 113–123.

Horvath, J.E., Bailey, J.A., Locke, D.P., and Eichler, E.E. 2001. Lessons from the human genome: Transitions between euchromatin and heterochromatin. *Hum. Mol. Genet.* **10:** 2215–2223.

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

International Human Genome Sequencing Consortium. 2004. Finishing the euchromatic sequence of the human genome. *Nature* **431:** 931–945.

Jorgensen, A.L., Jones, C., Bostock, C.J., and Bak, A.L. 1987. Different subfamilies of alphoid repetitive DNA are present on the human and chimpanzee homologous chromosomes 21 and 22. *EMBO J.* **6:** 1691–1696.

Kazakov, A.E., Shepelev, V.A., Tumeneva, I.G., Alexandrov, A.A., Yurov, Y.B., and Alexandrov, I.A. 2003. Interspersed repeats are found predominantly in the "old" α satellite families. *Genomics* **82:** 619–627.

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D. 2002. The human genome browser at UCSC. *Genome Res.* **12:** 996–1006.

Kumar, S., Tamura, K., Jakobsen, I.B., and Nei, M. 2001. MEGA2: Molecular evolutionary genetics analysis software. *Bioinformatics* **17:** 1244–1245.

Laursen, H.B., Jorgensen, A.L., Jones, C., and Bak, A.L. 1992. Higher rate of evolution of X chromosome α-repeat DNA in human than in the great apes. *EMBO J.* **11:** 2367–2372.

Mahtani, M.M. and Willard, H.F. 1990. Pulsed-field gel analysis of α satellite DNA at the human X chromosome centromere: High frequency polymorphisms and array size estimate. *Genomics* **7:** 607–613.

Maio, J.J., Brown, F.L., and Musich, P.R. 1981. Toward a molecular paleontology of primate genomes. I. The HindIII and EcoRI dimer families of alphoid DNAs. *Chromosoma* **83:** 103–125.

Manuelidis, L. and Wu, J.C. 1978. Homology between human and simian repeated DNA. *Nature* **276:** 92–94.

Musich, P.R., Brown, F.L., and Maio, J.J. 1980. Highly repetitive component α and related alphoid DNAs in man and monkeys. *Chromosoma* **80:** 331–348.

Ohta, T. and Dover, G.A. 1983. Population genetics of multigene families that are dispersed into two or more chromosomes. *Proc. Natl. Acad. Sci.* **80:** 4079–4083.

Pruitt, K.D., Katz, K.S., Sicotte, H., and Maglott, D.R. 2000. Introducing RefSeq and LocusLink: Curated human genome resources at the NCBI. *Trends Genet.* **16:** 44–47.

Rosenberg, H., Singer, M., and Rosenberg, M. 1978. Highly reiterated sequences of SIMIANSIMIANSIMIANSIMIANSIMIAN. *Science* **200:** 394–402.

Ross, M.T., Grafham, D.V., Coffey, A.J., Scherer, S., McLay, K., Muzny, D., Platzer, M., Howell, G.R., Burrows, C., Bird, C.P., et al. 2005. The DNA sequence of the human X chromosome. *Nature* **434:** 325–337.

Rudd, M.K. and Willard, H.F. 2004. Analysis of the centromeric regions of the human genome assembly. *Trends Genet.* **20:** 529–533.

Rudd, M.K., Schueler, M.G., and Willard, H.F. 2004. Sequence organization and functional annotation of human centromeres. *Cold Spring Harbor Symp. Quant. Biol.* **68:** 141–149.

Schindelhauer, D. and 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:** 1815–1826.

Schueler, M.G., Higgins, A.W., Rudd, M.K., Gustashaw, K., and Willard, H.F. 2001. Genomic and genetic definition of a functional human centromere. *Science* **294:** 109–115.

Schueler, M.G., Dunn, J.M., Bird, C.P., Ross, M.T., Viggiano, L., Rocchi, M., Willard, H.F., Green, E.D., and NISC Comparative Sequencing Program. 2005. Progressive proximal expansion of the primate X chromosome centromere. *Proc. Natl. Acad. Sci.* **102:** 10563–10568.

She, X., Horvath, J.E., Jiang, Z., Lui, G., Furey, T.S., Christ, L., Clark, R., Graves, T., Gulden, C.L., Alkan, C., et al. 2004. The structure and evolution of centromeric transition regions within the human genome. *Nature* **430:** 857–864.

Smit, A.F. 1999. Interspersed repeats and other mementos of transposable elements in mammalian genomes. *Curr. Opin. Genet. Dev.* **9:** 657–663.

Smith, G.P. 1976. Evolution of repeated DNA sequences by unequal crossover. *Science* **191:** 528–535.

Southern, E.M. 1975. Long range periodicities in mouse satellite DNA. *J. Mol. Biol.* **94:** 51–69.

Spence, J.M., Critcher, R., Ebersole, T.A., Valdivia, M.M., Earnshaw, W.C., Fukagawa, T., and Farr, C.J. 2002. Co-localization of centromere activity, proteins and topoisomerase II within a subdomain of the major human X α-satellite array. *EMBO J.* **21:** 5269–5280.

Thayer, R.E., Singer, M.F., and McCutchan, T.F. 1981. Sequence relationships between single repeat units of highly reiterated African Green monkey DNA. *Nucleic Acids Res.* **9:** 169–181.

Thompson, J.D., Higgins, D.G., and Gibson, T.J. 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22:** 4673–4680.

Vasiliauskas, D., Hancock, S., and Stern, C.D. 1999. SWiP-1: novel SOCS box containing WD-protein regulated by signalling centres and by Shh during development. *Mech. Dev.* **82:** 79–94.

Warburton, P.E. and Willard, H.F. 1990. Genomic analysis of sequence variation in tandemly repeated DNA: Evidence for localized homogeneous sequence domains within arrays of α satellite DNA. *J. Mol. Biol.* **216:** 3–16.

———. 1995. Interhomologue sequence variation of α satellite DNA from human chromosome 17: Evidence for concerted evolution along haplotypic lineages. *J. Mol. Evol.* **41:** 1006–1015.

———. 1996. Evolution of centromeric α satellite DNA: Molecular organization within and between human and primate chromosomes. In *Human genome evolution* (ed. S.T. Jackson and G. Dover), pp. 121–145. BIOS Scientific Publishers, Oxford.

Warburton, P.E., Haaf, T., Gosden, J., Lawson, D., and Willard, H.F. 1996. Characterization of a chromosome-specific chimpanzee α satellite subset: Evolutionary relationship to subsets on human chromosomes. *Genomics* **33:** 220–228.

Watanabe, H., Fujiyama, A., Hattori, M., Taylor, T.D., Toyoda, A., Kuroki, Y., Noguchi, H., BenKahla, A., Lehrach, H., Sudbrak, R., et al. 2004. DNA sequence and comparative analysis of chimpanzee chromosome 22. *Nature* **429:** 382–388.

Waye, J.S. and Willard, H.F. 1985. Chromosome-specific α satellite DNA:

Nucleotide sequence analysis of the 2.0 kilobasepair repeat from the human X chromosome. *Nucleic Acids Res*. **12:** 2731–2743.

———. 1986. Structure, organization, and sequence of α satellite DNA from human chromosome 17: Evidence for evolution by unequal crossing-over and an ancestral pentamer repeat shared with the human X chromosome. *Mol. Cell. Biol.* **6:** 3156–3165.

Wevrick, R. and Willard, H.F. 1989. Long-range organization of tandem arrays of α satellite DNA at the centromeres of human chromosomes: High frequency array-length polymorphism and meiotic stability. *Proc. Natl. Acad. Sci.* **86:** 9394–9398.

———. 1991. Physical map of the centromeric region of human chromosome 7: Relationship between two distinct α satellite arrays. *Nucleic Acids Res*. **19:** 2295–2301.

Wevrick, R., Willard, V.P., and Willard, H.F. 1992. Structure of DNA near long tandem arrays of α satellite DNA at the centromere of human chromosome 7. *Genomics* **14:** 912–923.

Willard, H.F. and Waye, J.S. 1987a. Chromosome-specific subsets of human α satellite DNA: Analysis of sequence divergence within and between chromosomal subsets and evidence for an ancestral pentameric repeat. *J. Mol. Evol.* **25:** 207–214.

———. 1987b. Hierarchical order in chromosome-specific human α satellite DNA. *Trends Genet.* **3:** 192–198.

Web site references

http://darwin.uvigo.es/software/modeltest.html; Modeltest.
http://genome.ucsc.edu; UCSC Genome Bioinformatics site.
http://macclade.org/macclade.html; MacClade.
http://paup.csit.fsu.edu/downl.html; PAUP.
http://pipeline.lbl.gov/cgi-bin/gateway2; VISTA browser.
http://repeatmasker.genome.washington.edu; RepeatMasker.
http://www.megasoftware.net; Molecular Evolutionary Genetic Analysis version 2.1.
http://www.ncbi.nih.gov/RefSeq; NCBI Reference Sequence Collection.