Heterochromatin Blocks Constituting the Entire Short Arms of Acrocentric Chromosomes of Azara’s Owl Monkey: Formation Processes Inferred From Chromosomal Locations

ORNJIRA Prakhongcheep1,2, NAMPECH Chaiprasertsri1,2, SHOKO Terada1, YURIKO Hirai1, KORNSONR Srikulnath2, HIROHISA Hirai1, and AKIHIKO Koga1,*

Primate Research Institute, Kyoto University, Inuyama City 484-8506, Japan1 and Department of Genetics, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand2

*To whom correspondence should be addressed. Tel. +81-568-63-0567. Fax. +81-568-63-0085. E-mail: koga.akihiko.5n@kyoto-u.ac.jp

Edited by Dr Mitsuo Oshimura
(Received 5 March 2013; accepted 14 May 2013)

Abstract

Centromeres and telomeres of higher eukaryotes generally contain repetitive sequences, which often form pericentric or subtelomeric heterochromatin blocks. C-banding analysis of chromosomes of Azara’s owl monkey, a primate species, showed that the short arms of acrocentric chromosomes consist mostly or solely of constitutive heterochromatin. The purpose of the present study was to determine which category, pericentric, or subtelomeric is most appropriate for this heterochromatin, and to infer its formation processes. We cloned and sequenced its DNA component, finding it to be a tandem repeat sequence comprising 187-bp repeat units, which we named OwlRep. Subsequent hybridization analyses revealed that OwlRep resides in the pericentric regions of a small number of metacentric chromosomes, in addition to the short arms of acrocentric chromosomes. Further, in the pericentric regions of the acrocentric chromosomes, OwlRep was observed on the short-arm side only. This distribution pattern of OwlRep among chromosomes can be simply and sufficiently explained by assuming (i) OwlRep was transferred from chromosome to chromosome by the interaction of pericentric heterochromatin, and (ii) it was amplified there as subtelomeric heterochromatin. OwlRep carries several direct and inverted repeats within its repeat units. This complex structure may lead to a higher frequency of chromosome scission and may thus be a factor in the unique distribution pattern among chromosomes. Neither OwlRep nor similar sequences were found in the genomes of the other New World monkey species we examined, suggesting that OwlRep underwent rapid amplification after the divergence of the owl monkey lineage from lineages of the other species.

Key words: chromosomal evolution; constitutive heterochromatin; tandem repeat; amplification; acrocentric chromosome

1. Introduction

In higher eukaryotes, centromeres and telomeres generally contain heterochromatin, the DNA component of which is usually repetitive sequences. The heterochromatin is often amplified so extensively as to form pericentric or subtelomeric heterochromatin blocks. In primates (order Primates), most of the species, including humans, carry pericentric heterochromatin blocks large enough to be visible on microscopy of metaphase chromosomes.1,2 Alpha satellite DNA is a major DNA component of these heterochromatin blocks,3–6 and other repetitive sequences are also known to be present there, such as satellite 1,6 satellite 2,7 and beta satellite.8 A well-known example of large-scale subtelomeric heterochromatin blocks in
constitutive heterochromatin blocks at chromosome arms of acrocentric chromosomes. Heterochromatin blocks at chromosomal locations similar to those of the owl monkey may be found in many other organisms. To our knowledge, the present study is the first to examine the formation processes of such heterochromatin blocks using sequence analysis of constituent repetitive DNA.

2. Materials and methods

2.1. Animals

Adult primate species bred at the authors’ institution (Primate Research Institute of Kyoto University; KUPRI) were used for this study. The main specimen, *A. azarae* (Azara’s owl monkey; female; simply called owl monkey hereafter), was the same one used in our previous study and is described therein in detail.19 *Saguinus oedipus* (cotton-top tamarin; male), *Saimiri sciureus* (squirrel monkey; male), *Cebus apella* (brown capuchin; male), *Callithrix jacchus* (common marmoset; male), *Macaca mulatta* (rhesus macaque; male), and *Hylobates lar* (white-handed gibbon; female) were used for comparison. A male adult human (*Homo sapiens*) was also a donor in the study. All experiments were conducted according to the Guidelines for Care and Use of Nonhuman Primates (version 3; June 2010) of KUPRI.

2.2. Chromosome analyses

We prepared chromosome spreads using white blood cells and subsequently performed C-band staining, 4',6-diamidino-2-phenylindole (DAPI) staining, and FISH analysis, as described previously.11,13

2.3. Genomic DNA extraction

Genomic DNA was extracted from white blood cells, or cultured cells originating from epithelial tissue, using a standard method that included cell lysis with sodium dodecyl sulphate, protein digestion with proteinase K, salt sedimentation, and DNA precipitation with isopropanol. The DNA precipitate was reconstituted in 10 mM Tris (pH 8.0).

2.4. Genomic library and screening

We used the genomic library of the owl monkey specimen that we constructed for our previous study.19 The vector was the 8.1-kb fosmid pCC1FOS, which is maintained as a single copy in bacterial cells with a negligible frequency of recombination. The insert DNA consisted of 40- to 44-kb genomic DNA fragments that had been mechanically sheared using a syringe needle and recovered from agarose gel after electrophoresis.
To screen the genomic library for repetitive sequences, we used a modified genomic hybridization method as described in our previous studies.\textsuperscript{13,20,21} Briefly, bacterial cells containing recombinant DNA were plated on agar media, individual colonies were transferred to 96-well plates containing liquid media, bacterial cell culture was blotted onto nylon membranes, and DNA was denatured using an alkali solution and then hybridized with alkaline phosphatase-labelled genomic DNA as a probe. Subsequently, we obtained candidates containing repetitive DNA by selecting clones displaying relatively strong signals. An advantage of this method is that repetitive DNA is detected irrespective of its nucleotide sequence. Another advantage is that the detection efficiency is positively correlated with the repetition scale; namely, the higher the copy number of a repetitive sequence, the easier it is to detect owing to stronger hybridization signals and a larger number of signal-producing colonies.

2.5. \textit{Other DNA manipulation}

We performed subcloning, DNA sequencing, polymerase chain reaction, colony hybridization, and Southern hybridization by using methods described previously.\textsuperscript{13} Specific experimental conditions are described in each case.

3. \textbf{Results}

3.1. \textit{Chromosome composition and cytological location of heterochromatin}

Heterochromatin appears as darkly stained portions on C-band staining of chromosomes. Figure 1 shows the C-band staining of a metaphase spread of the female owl monkey specimen. This spread contained 20 metacentric and 30 acrocentric chromosomes. The centromere regions of all chromosomes and the short arms of all acrocentric chromosomes were heavily stained. We examined more than 20 metaphase spreads, all of which exhibited these features. In addition to these heavily stained portions, lightly or faintly stained interstitial regions were observed. We did not examine the locations of these regions, however, because accurate comparison of their locations among spreads could not be expected.

3.2. \textit{Cloning of highly repetitive sequences}

The results of C-band staining showed the presence of large amounts of heterochromatin in the short arms of acrocentric chromosomes. Assuming that the heterochromatin blocks comprised similar sequences that occupied a large fraction of the genome, we screened the owl monkey genomic library for repetitive sequences. A total of 576 clones (6 plates $\times$ 96 wells) were hybridized with owl monkey genomic DNA labelled with alkaline phosphatase. The colonies displayed signals of various intensities on autoradiograms. A representative autoradiogram is shown in Fig. 2. The signal intensity reflects the repetitiveness of the respective clones in the genome. We chose the eight clones with the highest intensity signals for further study.

3.3. \textit{Identification of novel tandem repeat sequences}

We sequenced the terminal regions (600–800 nucleotides from both ends) of the eight selected clones using universal sequencing primers (M13 and reverse primers). All sequence reads were found to contain tandem repeats. Our previous study\textsuperscript{19} identified two types of alpha satellite DNA as large-scale repetitive sequences, using the method described herein. These sequences had been named OwlAlp1 and OwlAlp2, and their repeat units were 185 and 344 bp in length, respectively. Of the eight clones we selected via the library screening, two contained OwlAlp1 and four had OwlAlp2. The remaining two clones were previously unknown repeat DNA sequences comprising 187-bp repeat units. We designated the newly identified tandem repeat sequence, OwlRep.

3.4. \textit{Cloning and sequencing of an OwlRep fragment}

We determined the nucleotide sequence of an OwlRep clone. Primer walking and shotgun sequencing are the principal strategies for obtaining the sequence...
of long DNA fragments. These methods were, however, inapplicable to OwlRep because of its repetitive complexity. We therefore implemented an alternative strategy to prepare nested deletion clones of different lengths. This process involved the use of recognition sites of the restriction endonuclease \textit{AatII} (GACGTC), which were present in most of the repeat units of the sequences in the terminal regions.

We first subcloned an 8.5-kb fragment of the fosmid clone, obtained through digestion with \textit{EcoRV}, into the plasmid pBluescript II SK+ . We digested this subclone completely with \textit{BamHI} (one site in the vector, but no site in the insert DNA) and then incompletely with \textit{AatII}. After blunting the DNA ends with T4 DNA polymerase and ligation to form circular DNA molecules with T4 DNA ligase, we introduced the treated DNA into competent bacterial cells, spread the culture on agar plates, and subsequently picked up several colonies. We selected plasmid clones with insert sizes that differed by multiples of $\sim 300$ bp and sequenced them using the M13 universal primer. The sequence reads obtained were edited into a single stretch. We devised a criterion with which to exclude sequencing errors, to ensure that every nucleotide site contained a nucleotide that was (i) identical in two different clones or (ii) present in the majority of more than three different clones. A 5.1-kb region within the 8.5-kb clone fulfilled this criterion. We deposited the 5.1-kb sequence in DDBJ (accession number: AB746944) and used it for our subsequent analyses.

### 3.5. Overall structure of OwlRep

We compared the 5.1-kb sequence with itself through dot matrix analysis to identify regions that share significant similarities (Fig. 3A and B). Long diagonal lines repeated at regular intervals were observed, indicating tandem repeat structures in OwlRep. In addition, several short lines were observed within each repeat unit, suggesting that the repeat units contained short tandem repeat sequences. Another dot matrix analysis that compared OwlRep and its complementary sequence (Fig. 3C and D) suggested that the repeat units also contained inverted-repeat sequences represented by several short lines.

### 3.6. Consensus sequence

We cut the 5.1-kb OwlRep sequence into repeat units via visual examination and finally obtained 26 consecutive repeat units. These units were aligned using the ClustalW2 program (http://www.ebi.ac.uk/Tools/msa/clustalw2/), and a consensus sequence based on this alignment was drawn using the Advanced Consensus Maker program (http://www.hiv.lanl.gov/content/sequence/CONSENSUS/AdvCon.html). The alignments of the nucleotide sequences of the repeat units and the consensus sequence are shown in Fig. 4. The length of the consensus sequence was 187 bp. The distribution of the unit lengths was 187 bp (21 of 26; 81%), 182 bp (4 of 26; 15%), and 181 bp (1 of 26; 4%).

### 3.7. Sequence features of repeat units

The dot matrix analysis of the OwlRep sequence (Fig. 3) revealed that both tandem and inverted repeat structures were present within the repeat units. Notably, we identified sequence blocks exhibiting these features by manually analysing the consensus sequence.

We found several direct repeats. Among them, the repeats of the largest scale were four 17-bp blocks containing up to two mismatched nucleotides (Fig. 5A). These repeats spanned about half of the consensus sequence. With respect to inverted repeats, several pairs of blocks were found scattered throughout the consensus sequence (Fig. 5B). Of the 14 pairs identified, six were palindrome pairs (inverted repeats with no intervening nucleotides). The other eight pairs contained one nucleotide between the components, which...
could be regarded as hairpin structures. The largest inverted-repeat pair consisted of 8-bp units.

3.8. Mapping of OwlRep on chromosomes

We conducted FISH analysis of mitotic metaphase chromosomes to determine the chromosomal locations of the OwlRep sequence (Fig. 6). It was dual-color FISH assays using OwlAlp1 or OwlAlp2 as an additional probe. In each assay, we examined the hybridization patterns of more than 20 complete metaphase spreads and confirmed that they displayed identical patterns. In accordance with our previous study, the OwlAlp1 probe showed signals on the centromeres of all chromosomes (panel A, green), and OwlAlp2 showed signals in the pericentric regions of most chromosomes (panel C, green). Of 30 acrocentric chromosomes, 28 exhibited signals for OwlRep in their short arms. The other 20 chromosomes were metacentric, and 4 exhibited signals for OwlRep in the pericentric region on only one side of the centromere (panels A and C, red).

3.9. Distribution of OwlRep sequences among species

We performed a genomic Southern blot analysis to analyse the distribution of OwlRep or similar sequences among various species of the Platyrrhini clade. Analysis of the agarose gel after electrophoresis and ethidium bromide staining (Fig. 7A) revealed that both the distribution patterns and the staining intensities were similar in all lanes, indicating that the lanes contained approximately the same amount of genomic DNA, which had been completely digested with the restriction enzyme. However, the autoradiogram obtained after hybridization revealed that the signal was present only in the case of the owl monkey (Fig. 7B).

4. Discussion

4.1. Composition of short arms of acrocentric chromosomes

By applying the modified genomic hybridization method to the owl monkey genomic library, we
cloned a highly repetitive sequence (OwlRep), in addition to the two types of alpha satellite DNA (OwlAlp1 and OwlAlp2) discovered in our previous study. FISH analysis revealed that OwlRep is a component of the heterochromatin that constituted the short arms of 28 of the 30 acrocentric chromosomes. The short arms of two acrocentric chromosomes did not exhibit a detectable signal in the FISH analysis. The short arms of these two acrocentric chromosomes are much shorter than those of the other 28 acrocentric chromosomes. Thus, it is likely that OwlRep was removed from these two chromosomes or has not been introduced into them.

With respect to the 28 acrocentric chromosomes, the possibility that the short arms contained heterochromatin other than OwlRep cannot be excluded; however, OwlRep is likely the most abundant component. This likelihood can be attributed to our modified genomic hybridization method in which the cloning efficiency is expected to be positively correlated with the abundance of repeat sequences in the genome. In our initial screening, we selected eight clones exhibiting the highest signal intensity after hybridization, and these were found to be two OwlAlp1, four OwlAlp2, and two OwlRep sequences. Further, analysis of the other 32 highly positive clones via sequencing of only

Figure 4. Alignments of nucleotide sequences of repeat units. 'Con' indicates the consensus sequence assembled by collecting the most common nucleotides at respective sites. Nucleotide sites occupied by the base in the consensus sequence are indicated by dots. Nucleotide sites containing different bases are shown as the respective bases observed. The minus symbol indicates the absence of a nucleotide at that position. The asterisk shows that all 26 repeat units have identical nucleotides at that site.
1 end (both ends were not sequenced because the purpose was strictly confirmation) revealed that there were 6 OwlAlp1, 19 OwlAlp2, and 7 OwlRep sequences, confirming that no other repetitive sequences were present in a total of 40 genomic clones. These observations support the view that these three sequences comprise the great majority of the repetitive sequences in the owl monkey genome.

4.2. Duality of OwlRep blocks

Heterochromatin containing OwlRep as its main, or possibly only, DNA component constitutes the short arms of nearly all acrocentric chromosomes. These chromosomes have been categorized by cytogenetists not as telocentric, but as acrocentric, chromosomes, because the heterochromatin blocks are large enough to be called arms.\textsuperscript{16–18} These arms are, however, unlikely to contain euchromatin regions (Fig. 1). Thus, these chromosomes are more appropriately considered derivatives of telocentric chromosomes in which a satellite DNA region was elongated. Alpha satellite DNA is a major DNA component of centromeres in primates, and OwlAlp2 is the 'standard' type of alpha satellite DNA of New World monkeys.\textsuperscript{19} Because the OwlRep region resides on the acrocentric chromosomes adjacent to the OwlAlp2 region (Fig. 6), it can be regarded as a pericentric heterochromatin block. At the same time, it can be considered a subtelomeric heterochromatin block, because the proximal end of the OwlRep region of the acrocentric chromosomes corresponds to the chromosome end on the photographs of the metaphase chromosome spreads (Fig. 6).

4.3. Possible processes leading to the current OwlRep distribution on chromosomes

The transfer or exchange of sequence information most certainly occurs among the centromere regions of non-homologous chromosomes. Direct evidence in the case of the owl monkey is that OwlAlp1 is present in the centromeric constrictions of all chromosomes, as is OwlAlp2 in the pericentric regions of most chromosomes.\textsuperscript{19} Retrotransposons have recently been proposed to play at least a partial role in the mechanism for centromere homogenization among non-homologous chromosomes.\textsuperscript{22} In addition to being present in the short arms of acrocentric chromosomes, OwlRep was observed in the pericentric regions of 4 of 20 metacentric chromosomes (Fig. 6), implying that, regardless of where OwlRep was first generated, at least one transfer event of its sequence information occurred between a metacentric chromosome and an acrocentric chromosome. Thus, this distribution implicates a transfer of OwlRep to many chromosomes through the mechanism for the transfer of sequence information among centromeres. In other words, OwlRep was introduced into new chromosomes as pericentric heterochromatin blocks.

An alternative explanation for the distribution of OwlRep is that it was transferred as subtelomeric
heterochromatin blocks. The large-scale subtelomeric heterochromatin blocks of the chimpanzee, bonobo, and siamang are considered to have resulted from such sequence exchanges among non-homologous chromosomes. Nevertheless, this explanation is difficult to support for OwlRep because none of the arms of the metacentric chromosomes (40 arms) or long arms of the acrocentric chromosomes (30 arms) carries OwlRep in the subtelomeric region. For the same reason, the frequent transfer of OwlRep between the pericentric and subtelomeric regions is unlikely.

The FISH results (Fig. 6) showed that most signals for OwlRep on the acrocentric chromosomes were more intense than those on the four metacentric chromosomes, and the 15 pairs in the lower three rows are acrocentric chromosomes. The bar represents 10 μm.
had been amplified as pericentric heterochromatin blocks as OwAlp2 is (see Fig. S3 of Prakhongcheep et al. 19). Signals on the metacentric chromosomes were also observed on only one side of the centromere. This observation, however, neither supports nor denies the hypothesis because the number of OwRep-carrying metacentric chromosomes is as small as 4.

If our above hypothesis (that amplification in subtelomeric regions is more vigorous than that in pericentric regions) is true, this result may simply reflect variation in amplification efficiency owing to differences in the amplification mechanisms between the two regions. Another possible factor is the complex structures of the OwRep sequence. The repeat units of OwRep contain several tandem repeat and inverted repeat structures (Figs 3 and 5). These complex structures, especially inverted repeats, might be blocks for strand synthesis during DNA replication, interfering with the activity of DNA polymerase. If such interference occurs, the host cell may rely on a DNA repair system that includes the removal of nucleotide blocks. If this repair is incomplete, the DNA may contain unsealed nicks, which may lead to scission of the chromosome. In fact, one often encounters a situation that extension of complementary strands by DNA polymerase is halted in an inverted repeat region in reactions for DNA sequencing. It is also known that plasmid clones carrying inverted repeats are frequently rearranged unless one uses host bacterial strains that carry the uvrC and umuC mutations. These are cases observed in bacteria, but inverted repeats may have similar effects on DNA replication in eukaryotes. The harmful effect of chromosomal scission on the survival of a host cell would be more serious in the centromere region than it is in the subtelomeric region. This difference may lead to greater tolerance for vigorous amplification of OwRep in the subtelomeric region.

Heterochromatin blocks on the short arms of acrocentric chromosomes have been observed to some degree in other New World monkeys, whereas in owl monkeys, they are present on a much larger scale in all or most acrocentric chromosomes1,2. Our Southern blot analysis provided evidence against the presence of OwRep or similar sequences in the genomes of the other New World monkeys we examined (Fig. 7). Thus, the OwRep sequence was likely generated, or amplified to its currently large scale, after the owl monkey diverged from the other New World monkeys. Coupled with the rapid amplification of the new type of alpha satellite DNA we reported previously19 the owl monkey genome might have some factors leading to relatively frequent generation or amplification of tandem repeat sequences or, conversely, lack some factors to control the amount of repetitive sequences. In this regard, owl monkey is expected to help to clarify the general mechanisms of the formation and maintenance of constitutive heterochromatin blocks, of which currently little is known.

Acknowledgements: We are grateful to Drs Toshio Mouri and Masanaru Takai for helpful discussions.

Funding

This study was supported by Grants-in-Aid (23657165 to A.K., 23470098 to A.K., and 22247037 to H.H.) from the Japan Society for the Promotion of Science. O.P. and N.C. were supported by the Fellowship of Capacity Building for Kasetsart University on Internationalization.

References

10. Royle, N.J., Baird, D.M. and Jeffreys, A.J. 1994, A subterminal satellite located adjacent to telomeres in...


