Sequence Divergence and Copy Number of the Middle- and Long-Wave Photopigment Genes in Old World Monkeys

Rachel E. Ibbotson, David M. Hunt, James K. Bowmaker, John D. Mollon

Proceedings: Biological Sciences, Volume 247, Issue 1319 (Feb. 22, 1992), 145-154.

Stable URL:
http://links.jstor.org/sici?sici=0962-8452%2819920222%29247%3C145%3ASDACNO%3E2.0.CO%3B2-3

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://uk.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

Proceedings: Biological Sciences is published by The Royal Society. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at http://uk.jstor.org.

Proceedings: Biological Sciences
©1992 The Royal Society

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact jstor@mimas.ac.uk.

©2001 JSTOR

http://uk.jstor.org/
Fri Dec 28 15:23:10 2001

Sequence divergence and copy number of the middle- and long-wave photopigment genes in Old World monkeys

RACHEL E. IBBOTSON[1], DAVID M. HUNT[1], JAMES K. BOWMAKER[2] AND JOHN D. MOLLON[3]

[1] *School of Biological Sciences, Queen Mary & Westfield College, University of London, Mile End Road, London E1 4NS, U.K.*
[2] *Department of Visual Science, Institute of Ophthalmology, University of London, Judd Street, London WC1H 9QS, U.K.*
[3] *Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, U.K.*

SUMMARY

We have studied the sequence and organization of the genes for the middle-wave (MW) and long-wave (LW) cone photopigment genes in six species of Old World monkeys. Previous studies have shown that the MW and LW pigments of all six species exhibit peak sensitivities near 535 nm and 565 nm, respectively, and thus resemble the equivalent human pigments. In the case of man, the protein components of the MW and LW photopigments differ by 15 amino acids, although only seven of these differences involve non-homologous substitutions and are therefore candidates for a role in spectral tuning. Regions corresponding to exons 4 and 5 of these genes, and including five such candidate sites, were sequenced in the Old World monkeys. In contrast to the equivalent human genes, substitutions were found at two of these sites, position 233 and 309 of the MW gene in all six species. The role of amino acid substitutions in the spectral tuning of these photopigments is discussed. A comparision of the nucleotide sequences of the MW and LW genes provides evidence for sequence homogenization within species; the role of gene conversion in the evolution of these genes is discussed. The close juxtaposition and homology of the MW and LW genes on the X chromosome is thought to underlie the high frequency of colour vision defects in man and the presence in many individuals of extra copies of the MW gene. A study of a group of talapoin (*Ceropithecus talapoin*) monkeys has revealed a similar numerical polymorphism for this gene to that present in man. In contrast to the situation in man, where the MW and LW genes may contain a shortened first intron, restriction digests of genomic DNA showed that the size of this intron does not differ across the six species of Old World monkeys examined.

1. INTRODUCTION

In man, the middle-wave (MW) and long-wave (LW) photopigment genes are closely linked on the X chromosome (Feil *et al.* 1990; Vollrath *et al.* 1988). Their juxtaposition and close homology (Nathans *et al.* 1986*a*) suggest a relatively recent duplication from a single ancestral gene, and support the idea that the trichromacy of Old World primates evolved from a dichromatic form of colour vision (Ladd-Franklin 1892; Jacobs 1981; Mollon 1991). The dichromatic vision of our mammalian ancestors depended on a comparison of the signals of short-wave cones with those of a single type of cone in the red–green spectral region. In man, dichromatic vision, and anomalous forms of trichromatic vision, occur at high frequency (Pokorny *et al.* 1979): the juxtaposition and close homology of the MW and LW genes provide a mechanism for the production of such defects by gene loss or gene hybridization as a result of illegitimate pairing followed by crossing-over in this gene region (Nathans *et al.* 1986*b*). The same mechanism is also responsible for the numerical polymorphism of MW genes in humans (Nathans *et al.* 1986*b*) where haplotypes may have up to six copies (Drummond-Borg *et al.* 1989).

A similar pair of photopigments have been identified in a number of species of Old World monkeys by Bowmaker *et al.* (1991), suggesting that these species possess a form of trichromatic colour vision similar to that present in man. The present study examines the corresponding genes in six species of Old World monkeys. Four of the species are from the *Cercopithecus* genus: the diana monkey (*Cercopithecus diana*) from the Gabon, Cameroon, and Congo basin, where it occupies the canopy and upper levels of the rain forest; the talapoin monkey (*Cercopithecus talapoin*) from the swamp forest of the Atlantic coast of Central Africa; the grivet or green monkey (*Cercopithecus aethiops*), a semi-terrestial species from sparsely wooded regions of East Africa; and the arboreal spot-nosed monkey (*Cercopithecus petaurista*). The final two species – the patas monkey (*Erythrocebus patas*) from open savannah of Central and East Africa, and the crab-eating

macaque (*Macaca fascicularis*) from Asia – are single members of the *Erythrocebus* and *Macaca* genera respectively. The six species originate therefore from a number of different geographical regions, and occupy a range of different ecological habitats. Nevertheless, microspectrophotometric (MSP) measurements of MW and LW cone pigments have shown that the wavelengths of maximal sensitivity ($\lambda_{max.}$) lie near 535 nm and 565 nm in all six species and resemble the values previously obtained for man (Bowmaker *et al.* 1980, 1991).

The consistency of the λ_{max} values of the two sets of pigments in these species offers the opportunity to confirm the role of certain amino acids in determining the spectral sensitivity of the pigments. In man, only 15 amino acid sites distinguish the polypeptide (or opsin) components of the MW and LW cone pigments. Of these, only seven – all in transmembrane regions – are considered likely to control the spectral differences between the MW and LW pigments (Kosower 1988). These seven sites are distributed across exons 2, 3, 4 and 5. Neitz *et al.* (1989) have determined the nucleotide sequence of these exons in a male protanope, a subject who is red–green colour blind because he has only a single pigment sensitive to the 'green' region of the spectrum. In its spectral sensitivity the pigment closely resembles the normal MW pigment, although the gene encoding it is a hybrid; Neitz *et al.* (1989) state that exons 2 and 3 are derived from a normal LW gene, and exons 4 and 5 from a normal MW gene. This suggests that the spectral difference between the LW and MW pigments is largely governed by amino acid differences encoded in exons 4 and 5†. The sequencing of the cone opsin gene was therefore limited to these two exons.

2. METHODS

(*a*) *Preparation of DNA*

DNA was isolated from liver tissue obtained from the following animals: one male patas monkey (*Erythrocebus patas*), one male talapoin monkey (*Cercopithecus talapoin*), one female diana monkey (*Cercopithecus diana*), one male grivet (*Cercopithecus aethiops*), one male spot-nosed monkey (*Cercopithecus petaurista*), and one male macaque (*Macaca fascicularis*). With one exception (*M. fascicularis*), the DNA was obtained from the same individual animals for which phenotypes had been established by MSP (Bowmaker *et al.*, 1991). Animals were narcotized with ketamine hydrochloride and then given a fatal dose of sodium pentabarbitone. Tissue was rapidly removed, frozen in liquid nitrogen and stored at −70 °C.

Liver tissue (1 g) was homogenized in 5 ml of NE buffer containing 150 mM NaCl, 25 mM EDTA, pH 8.0. After the addition of 0.1% (by volume) sodium dodecyl sulphate (SDS), the homogenate was digested overnight at 55 °C with 100 μg ml^{-1} proteinase K, and extracted with buffered phenol. Genomic DNA was precipitated with 0.4 M NaCl and two volumes ethanol, re-dissolved in water, precipitated with 5 M ammonium acetate and three volumes ethanol, and finally dissolved in water.

DNA was also isolated from small heparinized blood samples obtained from a group of five male and five female talapoin monkeys. White blood cells were isolated after overlaying the blood samples on to Histopaque™ and centrifuging at 4000 r.p.m. for 30 min at room temperature. After pelleting and re-suspending the cells in NE buffer, the method of DNA isolation followed that described above for liver tissue.

(*b*) *Amplification of opsin gene fragments*

A 154 base pair (b.p.) DNA fragment of exon 4 and a 191 b.p. DNA fragment of exon 5 were amplified by the polymerase chain reaction (PCR) using primers synthesized from the published human LW opsin gene sequence (Nathans *et al.* 1986*a*). The primers used were 20 or 21mers of the following sequences:

exon 4 +ve 5′-CACGGCCTGAAGACTTCATGC-3′
−ve 5′-CGCTCGGATGGCCAGCCACAC-3′
exon 5 +ve 5′-GAATTCCACCCAGAAGGCAGA-3′
−ve 5′-ACGGGGTTGTAGATAGTGGC-3′.

These primers were designed to amplify from the MW and LW genes.

PCRs contained 200 ng genomic DNA, 200 ng of each primer, 1 unit of *Taq* polymerase, 0.2 mM each of dATP, dCTP, dGTP and dTTP, and reaction buffer (Cetus) in a final volume of 50 μl. The following PCR parameters were used: 94 °C for 3.5 min, 30 cycles of 60 °C for 1 min, 72 °C for 1 min, 94 °C for 1 min, and a final step of 72 °C for 10 min. The PCR products were visualized by electrophoresis in a 1.5% low-melting-point agarose gel by using a 0.09 M Tris-borate, 0.002 M EDTA buffer, pH 8.0. Agarose containing the target bands was excised, melted at 68 °C for 10 min, and extracted with pre-warmed (68 °C) buffered phenol after the addition of $\frac{1}{10}$ volume of pre-warmed (68 °C) 5 M NaCl. After a brief vortex and spin, a second phenol extraction followed by two chloroform:amyl alcohol (24:1) extactions were done. DNA was then precipitated at −70 °C for 30 min by addition of $\frac{1}{2}$ volume 8 M ammonium acetate and 5 volumes ethanol. After centrifuging and vacuum drying, pellets were dissolved in 10 μl water.

(*c*) *Southern blotting and restriction digests*

Restriction digests of genomic DNA or PCR products were done using standard conditions, and the fragments were separated by agarose gel electrophoresis. After denaturation, fragments were Southern blotted on to Hybond-N™ nylon membrane and probed with either a cDNA of the human LW-sensitive opsin gene (a kind gift from Dr. J. Nathans) or a *Eco*RI-*Bam*HI fragment of this cDNA that includes exon 1 and part of exon 2. All probes were [^{32}P]labelled by nick translation or random priming. After UV bonding, the

† Very recently, and since the present work was completed, Neitz *et al.* (1991) have suggested that the amino acid at position 180 (coded by exon 3) also controls the spectral sensitivity of the pigment: they propose that a substitution of alanine for serine produces a 6 nm shift of the $\lambda_{max.}$ to shorter wavelengths. At first sight this result appears to contradict the earlier claim by Neitz *et al.* (1989) that a hybrid pigment indistingushable from the normal MW pigment is produced by a hybrid gene comprising exons 2 and 3 from a LW gene and exons 4 and 5 from a MW gene. In fact, the exact position of the breakpoint in their hybrid gene is uncertain. The hybrid gene exhibits alanine at site 180, in agreement with the MW sequence obtained by Nathans *et al.* (1986*a*) from a genomic clone, whereas Nathans' genomic LW sequence has serine at this position. However, Nathans *et al.* also obtained a LW sequence from a cDNA clone that exhibited alanine at site 180. It is therefore impossible to tell from the published sequence of the hybrid gene whether the latter part of exon 3 is drawn from a MW gene or from the type of LW gene that, by the hypothesis of Neitz *et al.* (1991), yields the LW pigment with the shorter of the two alternative λ_{max} values.

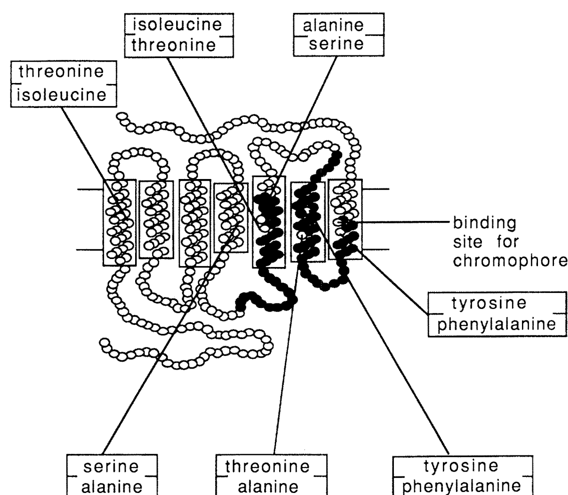

Figure 1. Molecular structure of opsin protein. The amino acids corresponding to the sequenced regions of exons 4 and 5 are indicated by filled circles. The seven candidate sites for spectral tuning are labelled and show the amino acid present in LW pigment above that present in the MW pigment.

membrane was pre-hybridized and hybridized overnight at 55 °C in a solution containing 6 × SSC, 5 × Denhardts, 60 g l^{-1} PEG, 5 g l^{-1} SDS and 1 mg ml^{-1} denatured herring testis DNA. Specific hybridization required a final wash of 0.1 × SSC at 55 °C. Autoradiography was done at −70 °C with intensifying screens. The relative amount of DNA in particular fragments was assessed by densitometry using a UVP image analysis system.

The relative amount of DNA in PCR-amplified fragments from exon 5 of the MW and LW genes was determined by two different methods. In both cases the fragment was separated by electrophoresis in 3 % low-melting-point agarose. The relative quantity of DNA in individual fragments was then determined by densitometric scanning of either the Polaroid™ negative of the ethidium bromide-stained gel, using the size-marker fragments to correct for the effect of fragment size on fluorescence, or an autoradiograph on pre-flashed film of a Southern hybridization probed at 65 °C with the [^{32}P]labelled cDNA of the human LW-sensitive opsin gene and washed at 55 °C with 0.1 × SSC.

Cloning and sequencing

The recovered PCR fragments were blunt-end ligated into the *Sma*I site of M13mp18. The ligation reaction contained approximately 7.5 ng of the fragment, 75 ng of M13 digested with *Sma*I to generate blunt ends, 0.5 mM ATP, five units of T4 DNA ligase, and buffer (Stratogene) in a total volume of 5 μl. After incubation at 10 °C overnight, competent JM101 cells were transformed and plated out on TY medium containing Xgal. White plaques were picked and checked for the presence of an inserted fragment by size comparison with single-standed M13 DNA on a 1 % agarose gel. Recombinant clones were sequenced by the dideoxy method using Sequenase™ version II and ^{35}S dATP. The products were loaded on to a 6 % polyacrylamide gel, separated at 1500 V for about 2 h, and autoradiographed overnight.

3. RESULTS

(*a*) *Amplification of exons 4 and 5 of the* MW *and* LW *photopigment genes*

The PCR (Mullis *et al.* 1986) was used to amplify a large part of exons 4 and 5 of the MW and LW genes. The selected regions code for most of the fifth transmembrane segment of the photopigment, all of the sixth, and approximately half of the seventh (to just beyond the lysine residue that is the binding site for the chromophore, retinal). The amplified regions include all those parts of exons 4 and 5 that code for sites that are thought to be candidates for the control of spectral tuning (Figure 1.

The products of an initial PCR of exons 4 and 5 are shown in Figure 2. For all six monkey samples, an intensely staining band of the expected size (154 b.p. for exon 4, 191 b.p. for exon 5) was obtained. This band was cut from the gel and the DNA eluted for cloning as described in the Methods section.

(*b*) *Identification of* PCR *fragments*

To verify that the PCR-amplified fragment had been derived from an opsin gene, the products of an exon 5

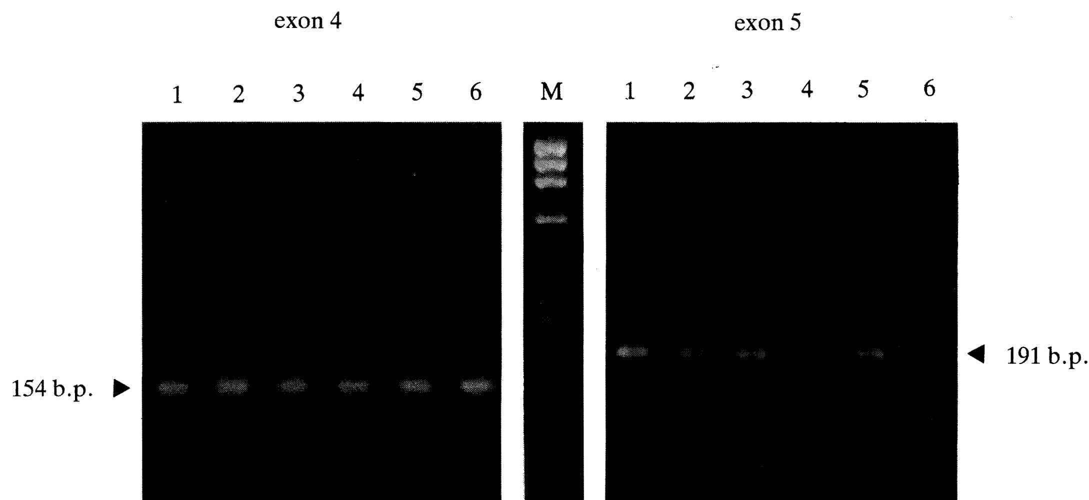

Figure 2. PCR-amplified fragments from exon 4 and exon 5 of the opsin genes. PCR amplifications were done with genomic DNA and primers homologous to exon 4 and 5 of human LW gene. The products were loaded on to a 1.5 % agarose gel and separated by electrophoresis. lane 1, *C. diana*; lane 2, *C. talapoin*; lane 3, *C. aethiops*; lane 4, *E. patas*; lane 5, *C. petaurista*; lane 6, *M. fascicularis*; M, size markers: ϕX174/*Hae* III digest.

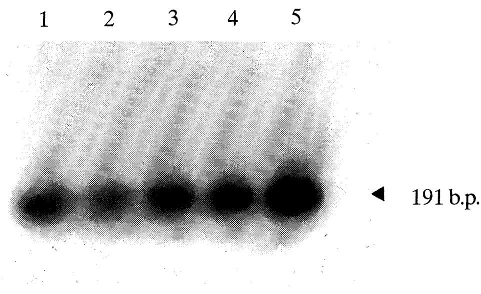

Figure 3. Hybridization of PCR fragments to human LW opsin cDNA probe. PCR fragments were separated by agarose gel electrophoresis, Southern blotted and probed with a [^{32}P]-labelled full-length human LW cDNA clone. Lane 1, *C. diana*; lane 2, *C. talapoin*; lane 3, *C. aethiops*; lane 4, *E. patas*; lane 5, *C. petaurista*.

amplification were separated by electrophoresis on a 1.5% agarose gel, Southern blotted and probed with a human LW photopigment cDNA. As shown in figure 3, after a high-stringency wash, the probe hybridizes to the 191 b.p. PCR fragment but not to the marker fragments present on the blot, suggesting that a fragment of a photopigment gene had been amplified.

(*c*) *Sequencing of exons 4 and 5*

Amplified fragments were cloned into M13 and recombinant plaques were picked and checked for an insert as described in the Methods section. For each species, at least three copies of the MW and the LW gene were sequenced.

Sequencing of clones confirmed the presence of exons from both genes in the amplified fragments. The nucleotide and amino acid sequences of exons 4 and 5 of the MW and LW genes of the six species of Old World monkey are shown in figures 4 and 5. In each case the sequences have been compared with the equivalent human sequences obtained from cloned PCR-amplified fragments produced with the same primers, and with the published sequences for the MW and LW human genes (Nathans *et al.* 1986*a*). In assigning the cloned Old World monkey sequences to MW or LW genes we have assumed close homology to the reported human sequences (Nathans *et al.* 1986*a*) and have sought to minimize the number of differences in amino acids.

The sequences of exon 4 from both the MW and LW genes remain extremely stable across the six species of catarrhine monkey. There are, however, some consistent differences compared with the human sequences, notably the silent subtitution of T for C at positions 618 and 627 in both genes. The only amino acid substitutions are threonine for serine at position 233 in the MW pigment of all six species, and valine for methionine at position 236 in the LW pigment of all species except *C. petaurista*. Rather more nucleotide and amino acid substitutions are seen in exon 5, and consistent coding differences in the MW gene compared with human are the substitution of phenylalanine for valine at position 274, alanine for proline at position 298, and tyrosine for phenylalanine at position 309. In

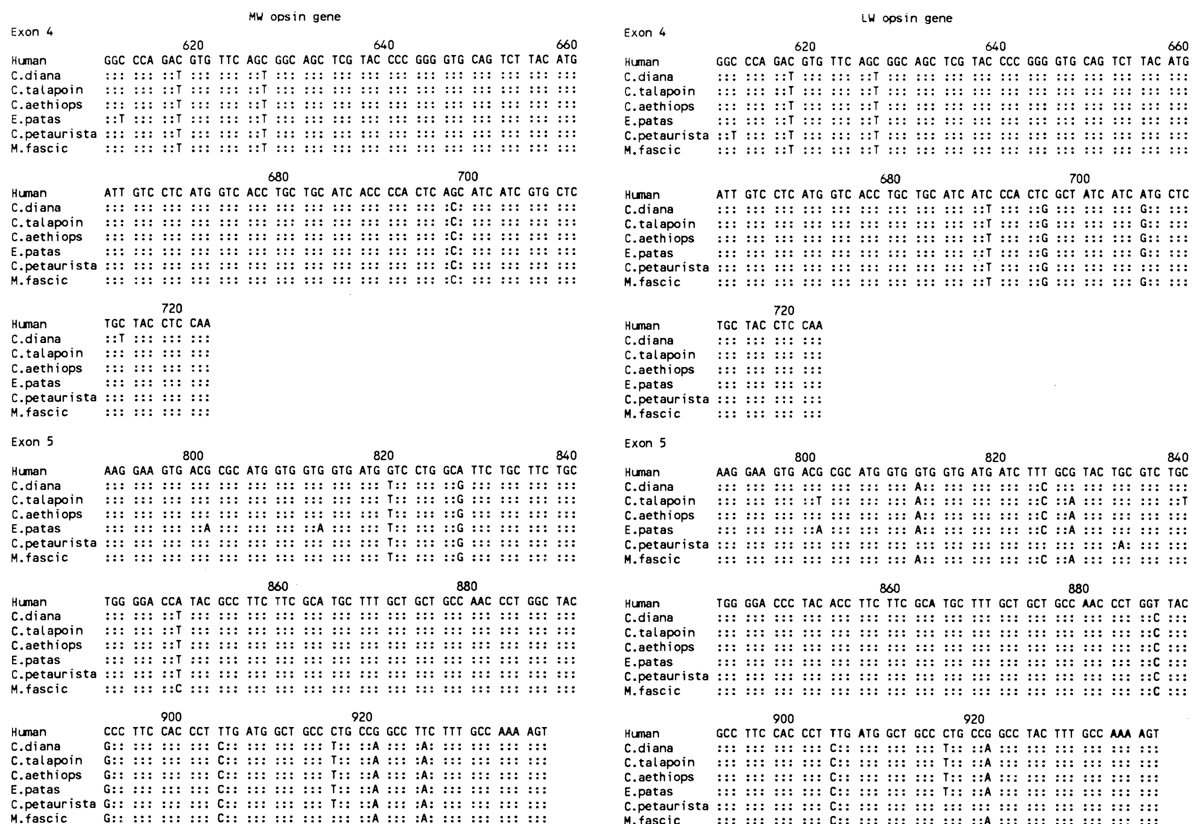

Figure 4. Nucleotide sequence of exons 4 and 5 of the MW and LW opsin genes. The Old World monkey sequences have been aligned with their equivalent human sequences. A colon indicates identity of nucleotide with the human sequence.

MW opsin gene

Exon 4

						210										220
Human	Gly	Pro	Asp	Val	Phe	Ser	Gly	Ser	Ser	Tyr	Pro	Gly	Val	Gln	Ser	Tyr	Met
C.diana	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
C.talapoin	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
C.aethiops	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
E.patas	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
C.petaurista	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M.fascic	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

									230							
Human	Ile	Val	Leu	Met	Val	Thr	Cys	Cys	Ile	Thr	Pro	Leu	Ser	Ile	Ile	Val	Leu
C.diana	-	-	-	-	-	-	-	-	-	-	-	-	Thr	-	-	-	-
C.talapoin	-	-	-	-	-	-	-	-	-	-	-	-	Thr	-	-	-	-
C.aethiops	-	-	-	-	-	-	-	-	-	-	-	-	Thr	-	-	-	-
E.patas	-	-	-	-	-	-	-	-	-	-	-	-	Thr	-	-	-	-
C.petaurista	-	-	-	-	-	-	-	-	-	-	-	-	Thr	-	-	-	-
M.fascic	-	-	-	-	-	-	-	-	-	-	-	-	Thr	-	-	-	-

			240	
Human	Cys	Tyr	Leu	Gln
C.diana	-	-	-	-
C. talapoin	-	-	-	-
C.aethiops	-	-	-	-
E.patas	-	-	-	-
C.petaurista	-	-	-	-
M.fascic	-	-	-	-

Exon 5

							270										280
Human	Lys	Glu	Val	Thr	Arg	Met	Val	Val	Val	Met	Val	Leu	Ala	Phe	Cys	Phe	Cys
C.diana	-	-	-	-	-	-	-	-	-	-	Phe	-	-	-	-	-	-
C.talapoin	-	-	-	-	-	-	-	-	-	-	Phe	-	-	-	-	-	-
C.aethiops	-	-	-	-	-	-	-	-	-	-	Phe	-	-	-	-	-	-
E.patas	-	-	-	-	-	-	-	-	-	-	Phe	-	-	-	-	-	-
C.petaurista	-	-	-	-	-	-	-	-	-	-	Phe	-	-	-	-	-	-
M.fascic	-	-	-	-	-	-	-	-	-	-	Phe	-	-	-	-	-	-

									290							
Human	Trp	Gly	Pro	Tyr	Ala	Phe	Phe	Ala	Cys	Phe	Ala	Ala	Ala	Asn	Pro	Gly	Tyr
C.diana	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
C.talapoin	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
C.aethiops	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
E.patas	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
C.petaurista	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M.fascic	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

			300										310			
Human	Pro	Phe	His	Pro	Leu	Met	Ala	Ala	Leu	Pro	Ala	Phe	Phe	Ala	Lys	Ser
C.diana	Ala	-	-	-	-	-	-	-	-	-	-	Tyr	-	-	-	-
C. talapoin	Ala	-	-	-	-	-	-	-	-	-	-	Tyr	-	-	-	-
C.aethiops	Ala	-	-	-	-	-	-	-	-	-	-	Tyr	-	-	-	-
E.patas	Ala	-	-	-	-	-	-	-	-	-	-	Tyr	-	-	-	-
C.petaurista	Ala	-	-	-	-	-	-	-	-	-	-	Tyr	-	-	-	-
M.fascic	Ala	-	-	-	-	-	-	-	-	-	-	Tyr	-	-	-	-

LW opsin gene

Exon 4

							210										220
Human	Gly	Pro	Asp	Val	Phe	Ser	Gly	Ser	Ser	Tyr	Pro	Gly	Val	Gln	Ser	Tyr	Met
C.diana	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
C.talapoin	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
C.aethiops	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
E.patas	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
C.petaurista	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M.fascic	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

									230							
Human	Ile	Val	Leu	Met	Val	Thr	Cys	Cys	Ile	Ile	Pro	Leu	Ala	Ile	Ile	Met	Leu
C.diana	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Val	-
C.talapoin	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Val	-
C.aethiops	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Val	-
E.patas	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Val	-
C.petaurista	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M.fascic	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Val	-

			240	
Human	Cys	Tyr	Leu	Gln
C.diana	-	-	-	-
C. talapoin	-	-	-	-
C.aethiops	-	-	-	-
E.patas	-	-	-	-
C.petaurista	-	-	-	-
M.fascic	-	-	-	-

Exon 5

							270										280
Human	Lys	Glu	Val	Thr	Arg	Met	Val	Val	Val	Met	Ile	Phe	Ala	Tyr	Cys	Val	Cys
C.diana	-	-	-	-	-	-	-	Met	-	-	-	-	-	-	-	-	-
C.talapoin	-	-	-	-	-	-	-	Met	-	-	-	-	-	-	-	-	-
C.aethiops	-	-	-	-	-	-	-	Met	-	-	-	-	-	-	-	-	-
E.patas	-	-	-	-	-	-	-	Met	-	-	-	-	-	-	-	-	-
C.petaurista	-	-	-	-	-	-	-	-	-	-	-	-	-	-	Tyr	-	-
M.fascic	-	-	-	-	-	-	-	Met	-	-	-	-	-	-	-	-	-

									290							
Human	Trp	Gly	Pro	Tyr	Thr	Phe	Phe	Ala	Cys	Phe	Ala	Ala	Ala	Asn	Pro	Gly	Tyr
C.diana	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
C.talapoin	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
C.aethiops	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
E.patas	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
C.petaurista	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M.fascic	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

			300										310			
Human	Ala	Phe	His	Pro	Leu	Met	Ala	Ala	Leu	Pro	Ala	Tyr	Phe	Ala	Lys	Ser
C.diana	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
C. talapoin	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
C.aethiops	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
E.patas	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
C.petaurista	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
M.fascic	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Figure 5. Deduced amino acid sequence of exons 4 and 5 of the MW and LW opsin genes. The Old World monkey sequences have been aligned with their equivalent human sequences. A dash indicates identity of amino acid with the human sequence.

contrast, exon 5 of the LW gene exhibits no amino acid substitutions that are common to all six species, although it is only *C. pertaurista* that does not have methionine in place of valine at position 271.

Evidence for gene conversion is present in the sequences of the MW and LW genes. If only silent site substitutions are considered, the average values for divergence of human and Old World monkey are 0.029 ± 0.011 for the MW gene and 0.036 ± 0.012 for the LW gene, after correction for multiple substitutions (Jukes & Cantor 1969). Remarkably, the within-species divergence of the MW and LW genes yield smaller values of 0.011 ± 0.006 and 0.014 ± 0.006 for human and Old World monkeys, respectively, suggesting that some corrective process has been acting to reduce the divergence of the two genes. A clear example of this is seen in the region of exon 5, shown in figure 6, where the MW and LW sequences are clearly more similar to each other than either is to its equivalent human sequence.

(*d*) *Restriction digests of genomic DNA*

In man, a size difference for intron 1 of the MW and LW genes has been reported (Nathans *et al.* 1986*a*). To investigate whether a difference in intron size is present in Old World monkeys, genomic DNA was digested with *Eco* RI, Southern blotted and probed with the full-length LW opsin cDNA (Fig. 7). Only one fragment is visible on the autoradiographs, suggesting that both the MW and LW opsin gene regions are cut by this enzyme into two equal-sized fragments of about 10.5 kb. There would appear to be no difference between the MW and LW genes in the size of intron 1 in any of the six species of Old World monkeys examined.

(*c*) *MW gene multiplicity*

To determine whether multiple copies of particular opsin genes are present in Old World monkeys, a group of ten talapoin monkeys was examined. In man, the MW and LW genes can be distinguished by a restriction fragment length difference, and gene copy number can then be determined by comparing the hybridization of an opsin gene probe with equivalent regions of the two genes (Nathans *et al.* 1986*b*; Drummond-Borg *et al.* 1989). Unfortunately the small samples of DNA that were obtained from the talapoin monkeys did not allow for restriction enzyme digests and Southern blotting of genomic DNA.

We adopted an alternative method that took advantage of the unique occurrence of a *Sau* 3A restriction site at position 819 in the LW gene. Digestion of PCR-amplified products with this enzyme yielded two fragments; the original 194 b.p. fragment representing the MW gene and a smaller 144 b.p. fragment representing the LW gene. The relative DNA content of each fragment was then quantified from different amplification experiments either by ethidium bromide fluorescence (Figure 8*a*) or by autoradiography after Southern hybridization with [^{32}P]labelled human LW cDNA (Figure 8*b*). As shown in table 1, the two

Old World monkey MW gene

	900 ... 920
C. diana	TAC GGC TAC GCC TTC CAC CCT CTG ATG GCT GCC TTG CCA GCC TAC TTT
C. talapoin	TAC GGC TAC GCC TTC CAC CCT CTG ATG GCT GCC TTG CCA GCC TAC TTT
C. aethiops	TAC GGC TAC GCC TTC CAC CCT CTG ATG GCT GCC TTG CCA GCC TAC TTT
E. patas	TAC GGC TAC GCC TTC CAC CCT CTG ATG GCT GCC TTG CCA GCC TAC TTT
C. petaurista	TAC GGC TAC GCC TTC CAC CCT CTG ATG GCT GCC TTG CCA GCC TAC TTT
M. fascicularis	TAC GGC TAC GCC TTC CAC CCT CTG ATG GCT GCC CTG CCA GCC TAC TTT

Old World monkey LW gene

	900 ... 920
C. diana	TAC GGC TAC GCC TTC CAC CCT CTG ATG GCT GCC TTG CCA GCC TAC TTT
C. talapoin	TAC GGC TAC GCC TTC CAC CCT CTG ATG GCT GCC TTG CCA GCC TAC TTT
C. aethiops	TAC GGC TAC GCC TTC CAC CCT CTG ATG GCT GCC TTG CCA GCC TAC TTT
E. patas	TAC GGC TAC GCC TTC CAC CCT CTG ATG GCT GCC TTG CCA GCC TAC TTT
C. petaurista	TAC GGC TAC GCC TTC CAC CCT CTG ATG GCT GCC CTG CCG GCC TAC TTT
M. fascicularis	TAC GGC TAC GCC TTC CAC CCT CTG ATG GCT GCC CTG CCA GCC TAC TTT

Human

	900 ... 920
MW	TAC GGC TAC CCC TTC CAC CCT TTG ATG GCT GCC CTG CCG GCC TTC TTT
LW	TAC GGT TAC GCC TTC CAC CCT TTG ATG GCT GCC CTG CCG GCC TAC TTT

Figure 6. Sequence homology in exon 5 of the Old World monkey MW and LW opsin genes. The regions of homology are underlined. A more restricted homology in the human genes is similarly identified.

1 2 3 4 5 6

◄ 10.5 kb

Figure 7. Southern hybridization of *Eco* RI-digested genomic DNA. Genomic DNA was digested with *Eco* RI, fractionated by agarose electrophoresis, Southern blotted and probed with a [^{32}P]labelled full-length human LW cDNA clone. Lane 1, *C. diana*; lane 2, *C. talapoin*; lane 3, *C. aethiops*; lane 4, *E. patas*; lane 5, *C. petaurista*; lane 6, *M. fascicularis*.

methods gave essentially identical results, and the calculated ratios, assuming that only the MW gene is present in multiple copies (Nathans *et al.* 1986*b*; Drummond-Borg *et al.* 1989), suggest a polymorphism in copy number ranging from one to at least four. A precise estimate of the range in copy number in talapoin monkeys is complicated by the use of females. Irrespective of sex, where a ratio of 1 is found, this suggests that a single MW gene is present on the X chromosome. However, in those females that give a higher ratio, it is not possible to be so precise because a ratio of 2 may mean the presence of three MW genes on one chromosome and one on the other, or an equal number of two on both. The highest ratio of 4 seen in female 10, places a lower limit on the MW copy number of four, and a possible upper limit of seven, given that the contribution of the two X chromosomes could be very different. The values of 3.4, 2.3 and 3.7 (corrected to 3, 2 and 4 respectively in table 1) obtained for females 6, 7 and 10 suggest the presence of an odd number of MW genes in these animals when the contribution of the two X chromosomes are totalled.

To check for contamination by any non-opsin DNA, digests with *Sph* I were also done. This enzyme will cut exon 5 of both genes into two fragments of 95 and 96 b.p. respectively. The absence of any undigested

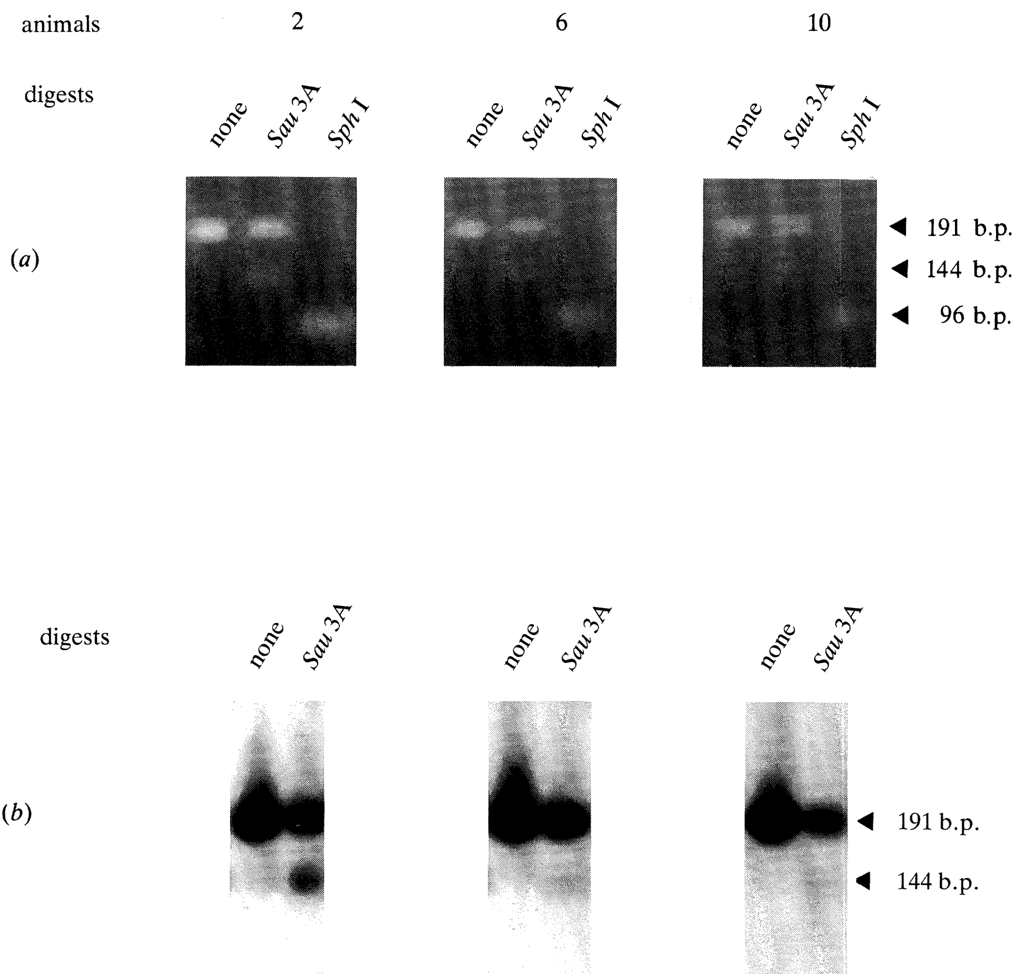

Figure 8. Determination of copy number for the MW opsin gene in talapoin monkey DNA. Exon 5 fragments of the MW and LW opsin genes were obtained by PCR amplifications from talapoin monkey DNA, and restriction enzyme digested. (*a*) Agarose gel showing *Sau* 3A and *Sph* I digests of fragments from three animals. The gel was photographed under UV light and the Polaroid(TM) negative was scanned with a UVP image analysis system to determine the relative amount of DNA in the two fragments produced after *Sau* 3A digestion. The ϕX174/*Hae* III size markers were used to correct for the effect of fragment size on fluorescence. (*b*) Autoradiograph of *Sau* 3A digests of fragments from the same three animals. The gel was Southern blotted and probed with a [^{32}P]labelled full-length human LW cDNA clone. The resulting autoradiograph was scanned with a UVP image analysis system.

Table 1. *Copy number of MW opsin genes*

(The relative amount of DNA in PCR-generated fragments was determined by densitometry after electrophoresis in 3 % agarose. The fluorescence values were obtained by scanning of UV-photographic negatives of the ethidium bromide-stained gels, and the autoradiography values by scanning autoradiographs obtained from Southern hybridizations with the [^{32}P]-labelled cDNA of the human LW-sensitive opsin gene.)

sample	sex	fluorescence MW:LW ratio	autoradiography MW:LW ratio	corrected to nearest integer
talapoin 1	male	1.0	0.9	1
talapoin 2	male	2.3	1.9	2
talapoin 3	male	1.9	1.7	2
talapoin 4	male	—	1.8	2
talapoin 5	male	—	2.0	2
talapoin 6	female	3.2	3.4	3
talapoin 7	female	2.3	2.3	2
talapoin 8	female	1.2	1.1	1
talapoin 9	female	1.0	1.1	1
talapoin 10	female	4.0	3.7	4
human	female	2.4	1.6	2

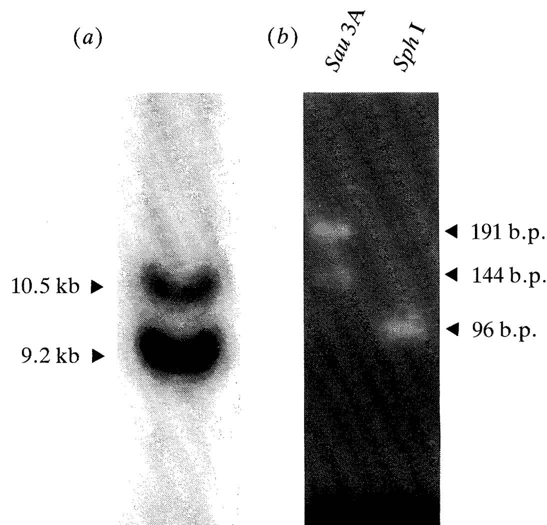

Figure 9. Determination of copy number for the MW opsin gene in a human DNA sample. DNA was isolated from a blood sample obtained from a human female subject with normal colour vision. This was used either for total genomic digests or as the template for amplification of exon 5 fragments of the MW and LW opsin gene by PCR. (*a*) Autoradiograph of *Eco*RI digest of genomic DNA. The gel was Southern blotted and hybridized with a [^{32}P]labelled probe that consisted of the 5′ *Eco*/*Bam* fragment of the human LW cDNA. The resulting autoradiograph was scanned with a UVP image analysis system. (*b*) Agarose gel showing *Sau*3A and *Sph*I digests. PCR-amplified fragments of exon 5 of the MW and LW opsin genes were restriction enzyme digested and loaded on to an agarose gel. The gel was photographed under UV light, and the Polaroid™ negative was scanned with a UVP image analysis system to determine the relative amount of DNA in the two fragments produced after *Sau*3A digestion. The ϕX174/*Hae*III size markers were used to correct for the effect of fragment size on fluorescence.

DNA shows that the PCR did not amplify any DNA that lacked *Sph*I sites (figure 8*a*).

Finally, the accuracy of this method was confirmed by reference to exon 5 fragments generated in the same way from a human female DNA sample. The MW:LW ratio had previously been established for this individual by *Eco*RI digestion of total genomic DNA followed by Southern blotting and hybridization with a probe comprising the *Eco*/*Bam* 5′ end of the human LW cDNA. The fragments shown in figure 9*a* correspond to the A_g and A_r fragments of Nathans *et al.* (1986*b*) and Drummond-Borg *et al.* (1989), and gave a MW:LW ratio of 1.5. As shown in table 1 and figure 9*b*, the PCR-based methods gave similar ratios.

4. DISCUSSION

The spectral differences of the MW and LW photopigments in man are thought to arise from the net effect of hydroxyl group differences in the vicinity of the chromophore (Kosower 1988). There are only seven differences of this kind, and five of these are coded by exons 4 and 5. In the Old World monkey MW pigment, two of these sites differ from the corresponding sites in the human MW pigment: serine is substituted for threonine at position 233, and phenylalanine for tyrosine at position 309. These differences are summarized in table 2.

The replacement of phenylalanine by tyrosine at position 309 of the MW gene is a non-homologous substitution of an amino acid that lacks an hydroxyl group for an amino acid that possesses one. It also means that the MW and LW genes of all six species of Old World monkeys are identical at this position. As the spectral properties of the two pigments in all six species are similar to those of man (Bowmaker *et al.* 1991), the presence of either tyrosine or phenylalanine at position 309 site can have little or no effect on the spectral tuning of these pigments.

The spectral tuning of cone photopigments in the middle- to long-wave region has recently been examined by Neitz *et al.* (1991). Their major conclusion is that only three substitutions are important in the case of primates: serine or alanine at position 180, tyrosine or phenylalanine at position 277, and threonine or alanine at position 285. Neitz and his colleagues propose that position 233 has no role in spectral tuning. However, our results show that the MW and LW pigments differ at this site in all the Old World monkey species examined, as they do in man. Moreover, although the human and Old World monkey MW pigments differ at site 233, the substitution that occurs – serine for threonine – is one considered to be homologous (Lehninger 1982), in that one hydroxyl-bearing amino acid is replaced by another (table 2). By contrast, the corresponding site in the LW pigment is not hydroxyl-bearing. Because there are only three hydroxyl-bearing amino acids, the conserved substitution in the MW pigment adds to our suspicion that site 233 may be important for spectral tuning. It is also very suggestive that Neitz and his colleagues themselves record that the MW and LW pigments of the neotropical squirrel monkey differ at this site by substitution of a hydroxyl-bearing amino acid for one that is not hydroxyl-bearing (Neitz *et al.* 1991), and Yokoyama & Yokoyama (1990) find a similar difference at the equivalent position in the MW- and LW-like pigments of

Table 2. *Candidate amino acid sites for spectral tuning coded by exons 4 and 5*

(Hydroxyl-bearing amino acids are indicated by a superscript OH.)

	230	233	277	285	309
Old World monkey LW	ile	ala	tyr^{OH}	thr^{OH}	tyr^{OH}
Human LW	ile	ala	tyr^{OH}	thr^{OH}	tyr^{OH}
Old World monkey MW	thr^{OH}	thr^{OH}	phe	ala	tyr^{OH}
Human MW	thr^{OH}	ser^{OH}	phe	ala	phe

the blind cave fish, *Astyanax fasciatus*. Amongst the primates, separate MW and LW genes have so far been found only in Old World monkeys; the duplication event that resulted in the production of two genes must have occurred after the separation of the New and Old World lineages. Substitution at position 233 has been retained therefore through this process of duplication and divergence. It seems unlikely that it is without effect.

Digests of genomic DNA with *Eco* RI yield a single fragment in all six species of Old World monkeys, suggesting that both the MW and LW gene regions are cut into two equal-sized fragments. A similar result has been reported by Nathans *et al.* (1986*a*) for the human LW gene. In contrast, the human MW gene yields two fragments of different sizes, the fragment from the 5′ end of the gene being shortened as a result of a reduction in the length of the first intron. A similarly shortened LW gene possessing the same deleted intron has recently been reported at high frequency in Afro-Americans by Lund Jorgensen *et al.* (1990), and a gene conversion event originating from misaligned pairing and unequal crossing-over of the two genes has been proposed as the mechanism for the transfer of this altered intron between genes. As this sample of Old World monkeys has included only a single representative of each species, such a polymorphism in intron size would not be apparent. However, there is no evidence for any variation in fragment size between the six species, suggesting that the deletion event that gave rise to the shortened intron in humans may have occurred after the separation of the human and Old World monkey lineages.

If the divergence of the MW and LW genes is combined, this gives an average value for silent sites of 0.033 ± 0.011 between this group of Old World monkeys and man. After correcting for X-linkage by multiplying by $\frac{3}{2}$ (Miyata *et al.* 1987), this gives a value of 0.050 ± 0.017. This contrasts with an average value of 0.094 ± 0.027 for sequence divergence of three non-visual pigment genes reported by Kuma *et al.* (1988), suggesting either that photopigment genes (and particularly exons 4 and 5) evolve at a slower rate compared with other genes, or that some other process is involved. Gene conversion has been suggested as an important process in equalizing the sequence of the MW and LW photopigment genes in humans (Kuma *et al.* 1988; Lund Jorgensen *et al.* 1990), and further evidence for this process is found in the divergence of the Old World monkey gene sequences. Because the duplication to give separate MW and LW genes occurred before the separation of the human and Old World monkey lineages, closer homology in silent sites would be expected between MW genes in the different species than between the MW and LW genes within species. The opposite is found, however, suggesting that some correction or homogenization of the MW and LW sequences must have occurred in each lineage, perhaps by the process of gene conversion that has been implicated in sequence correction seen in a number of other duplicated genes (Robinson & Ingram 1982; Shen *et al.* 1981; Weiss *et al.* 1983). One mechanism for achieving sequence conversion is through the production of hybrid genes by mispairing and crossing-over at meiosis, and this is known to occur at high frequency in man and to be a major factor in the development of colour vision defects (Nathans *et al.* 1986*b*). As the Old World monkey data suggest that a similar correction process has occurred, a high frequency of hybrid gene occurrence would appear likely in these species. No evidence for anomalous colour vision has, however, been found in Old World monkeys (Jacobs & Harwerth 1989); it would appear, therefore, that any hybrid genes that occur confer a strong selective disadvantage, and are rapidly eliminated from the population.

Mispairing and crossing-over between the two X-linked opsin genes will also give rise to gene duplication, and this has been extensively reported in man (Nathans *et al.* 1986*b*; Drummond-Borg *et al.* 1989) where copy number for the MW gene varies from one to at least six. Only the MW gene appears to have undergone an increase in copy number; a lack of homology immediately upstream of the LW gene is thought to preclude mispairing in this region and thereby prevent the duplication (or the complete deletion) of the LW gene. Because mispairing and hybrid gene formation between the MW and LW genes of Old World monkeys is the probable mechanism of gene conversion, we determined the copy number of the MW genes relative to the LW gene in a group of talapoin monkeys. Assuming the same constraints on LW gene duplication apply in monkeys as in man, copy number for the MW gene in these animals varies from one to at least four. Some of the ratios are derived from female monkeys, for which the contribution of the two X chromosomes cannot be precisely determined. For example, a ratio of 4 in a female could mean either an equal contribution of four MW genes from each X chromosome or unequal contributions of five and three, six and four, or seven and one MW genes. The majority of animals, however, show smaller ratios, and the modal copy number of two found for male monkeys is identical to that found in humans by Drummond-Borg *et al.* (1989).

This work was supported by a Wellcome Trust grant number 17016.

REFERENCES

Bowmaker, J. K., Dartnall, H. J. A. & Lythgoe, J. N. 1980 Microspectrophotometric demonstration of four classes of photoreceptor in an Old World primate, *Macaca fascicularis*. *J. Physiol., Lond.* **298**, 131–143.

Bowmaker, J. K., Astell, S., Hunt, D. M. & Mollon, J. D. 1991 Photosensitive and photostable pigments in the retinae of Old World monkeys. *J. exp. Biol.* **156**, 1–19.

Drummond-Borg, M., Deeb, S. S. & Motulsky, A. G. 1989 Molecular patterns of X chromosome-linked color vision genes among 134 men of European ancestry. *Proc. natn. Acad. Sci. U.S.A.* **86**, 983–987.

Feil, R., Aubourg, P., Heilig, R. & Mandel, J. L. 1990 A 195Kb cosmid walk encompassing the human Xq28 colour vision pigment genes. *Genomics* **6**, 367–373.

Jacobs, G. H. 1981 *Compaarative color vision*. New York: Academic Press.

Jacobs, G. H. & Harwerth, R. S. 1989 Color vision

variations in Old and New World primates. *Am. J. Primatol.* **18**, 35–44.

Jukes, T. H. & Cantor, C. R. 1969 Evolution of protein molecules. In *Mammalian protein metabolism*, vol. 3 (ed. H. N. Munro), pp. 21–132. New York: Academic Press.

Kosower, E. M. 1988 Assignment of groups responsible for the 'opsin-shift' and absorptions of rhodopsin and red, green and blue iodopsins (cone pigments). *Proc. Natn. Acad. Sci. U.S.A.* **85**, 1076–1080.

Kuma, K., Hayashida, H. & Miyata, T. 1988 Recent gene conversion between genes encoding human red and green visual pigments. *Jap. J. Genet.* **63**, 367–371.

Ladd-Franklin, C. 1892 A new theory of light sensation. In *International Congress of Psychology, 2nd Session,* pp. 103–108. London: Williams & Norgate.

Lehninger, A. L. 1982 *Principles of biochemistry*. New York: Worth.

Lund Jorgensen, A., Deeb, S. S. & Motulsky, A. G. 1990 Molecular genetics of X chromosome-linked color vision among populations of African and Japanese ancestry: high frequency of a shortened red pigment gene among Afro-Americans. *Proc. Natn. Acad. Sci. U.S.A.* **87**, 6512–6516.

Mollon, J. D. 1991 The uses and evolutionary origins of primate colour vision. In *Vision and visual dysfunction*, vol. 2 (ed. J. R. Conly-Dillon & R. L. Gregory), pp. 306–319. London: Macmillan Press.

Miyata, T., Hayashida, H., Kuma, K. & Yasunaga, T. 1987 Male-driven molecular evolution demonstrated by different rates of silent substitutions between autosome- and sex chromosome-linked genes. *Proc. Japan Acad.* **63B**, 327–331.

Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. & Erlich, H. 1986 Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction. *Cold Spring Harb. Symp. quant. Biol.* **51**, 263–273.

Nathans, J., Thomas, D. & Hogness, D. S. 1986*a* Molecular genetics of human color vision: the genes encoding blue, green and red pigments. *Science, Wash.* **232**, 193–202.

Nathans, J., Piantanida, T. P., Eddy, R. L., Shows, T. B. & Hogness, D. S. 1986*b* Molecular genetics of inherited variation in human color vision. *Science, Wash.* **232**, 203–210.

Neitz, J., Neitz, M. & Jacobs, G. H. 1989 Analysis of fusion gene and encoded photopigment of colour-blind humans. *Nature, Lond.* **342**, 679–682.

Neitz, M., Neitz, J. & Jacobs, G. H. 1991 Spectral tuning of pigments underlying red–green color vision. *Science, Wash.* **252**, 971–974.

Pokorny, J., Smith, V. C., Verriest, G. & Pinkers, A. J. L. G. (ed.) 1979 *Congenital and acquired color vision defects*. New York: Grune & Stratton.

Robinson, I. B. & Ingram, V. M. 1982 Gene evolution in the chicken β-globin cluster. *Cell* **28**, 515–521.

Shen, S., Slightom, J. L. & Smithies, O. 1981 A history of the human globin gene duplication. *Cell* **26**, 191–203.

Vollrath, D., Nathans, J. & Davies, R. W. 1988 Tandem array of human visual pigment genes at Xq28. *Science, Wash.* **240**, 1669–1672.

Weiss, E. H., Mellor, A., Golden, L., Fahrner, K., Simpson, E., Hurst, J. & Flavell, R. A. 1983 The structure of a mutant H-2 gene suggests that the generation of polymorphism in H-2 genes may occur by gene conversion-like events. *Nature, Lond.* **301**, 671–674.

Yokoyama, R. & Yokoyama, S. 1990 Convergent evolution of the red- and green-like visual pigment genes in fish, *Astyanax fasciatus*, and human. *Proc. Natn. Acad. Sci. U.S.A.* **87**, 9315–9318.

Received 28 October 1991; accepted 26 November 1991