

Origins and antiquity of X-linked triallelic color vision systems in New World monkeys

STEPHANE BOISSINOT*†, YING TAN*†, SONG-KUN SHYUE*, HORACIO SCHNEIDER‡, IRACILDA SAMPAIO‡, KATHERINE NEISWANGER§, DAVID HEWETT-EMMETT*, AND WEN-HSIUNG LI*¶

*Human Genetics Center, School of Public Health, University of Texas, P.O. Box 20334, 6901 Bertner Avenue, Houston, TX 77225; ‡Departamento de Genetica, Universidade Federal do Para, Belem, Para, Brazil; and §Allegheny University of Health Sciences, 4 Allegheny Center, Room 739, Pittsburgh, PA 15212

Communicated by Jiazhen Tan, Fudan University, Shanghai, Peoples Republic of China, August 19, 1998 (received for review June 15, 1998)

ABSTRACT It is known that the squirrel monkey, marmoset, and other related New World (NW) monkeys possess three high-frequency alleles at the single X-linked photopigment locus, and that the spectral sensitivity peaks of these alleles are within those delimited by the human red and green pigment genes. The three alleles in the squirrel monkey and marmoset have been sequenced previously. In this study, the three alleles were found and sequenced in the saki monkey, capuchin, and tamarin. Although the capuchin and tamarin belong to the same family as the squirrel monkey and marmoset, the saki monkey belongs to a different family and is one of the species that is most divergent from the squirrel monkey and marmoset, suggesting the presence of the triallelic system in many NW monkeys. The nucleotide sequences of these alleles from the five species studied indicate that gene conversion occurs frequently and has partially or completely homogenized intronic and exonic regions of the alleles in each species, making it appear that a triallelic system arose independently in each of the five species studied. Nevertheless, a detailed analysis suggests that the triallelic system arose only once in the NW monkey lineage, from a middle wavelength (green) opsin gene, and that the amino acid differences at functionally critical sites among alleles have been maintained by natural selection in NW monkeys for >20 million years. Moreover, the two X-linked opsin genes of howler monkeys (a NW monkey genus) were evidently derived from the incorporation of a middle (green) and a long wavelength (red) allele into one chromosome; these two genes together with the (autosomal) blue opsin gene would immediately enable even a male monkey to have trichromatic vision.

With the exception of the genus *Allouatta* (howler monkeys), which has two X-linked color photopigment (opsin) genes (1–3), all New World (NW) monkeys have only one X-linked opsin gene; all diurnal NW monkeys also have one autosomal color photopigment gene. In contrast, Old World monkeys, apes, and humans possess two X-linked genes and one autosomal photopigment gene and are therefore trichromatic. However, in some NW monkeys, such as the squirrel monkey and the marmoset (4–10), the single X-linked opsin locus has three high-frequency polymorphic alleles coding for three pigments with spectral sensitivity maxima (λ_{\max}) around 535, 550, and 562 nm in squirrel monkeys and capuchins and around 543, 556, and 562 nm in marmosets and tamarins; some of these λ_{\max} values, which are recent estimates by Jacobs (11), are slightly different from previous estimates. By convention, these alleles will be designated P535, P543, P550, P556, and P562. Because of this triallelic system, heterozygous female monkeys, like most humans, are trichromatic, although males and ho-

mozygous females are dichromatic (4–11). The triallelic system might have had a single origin, having been maintained in different NW monkey lineages by the selective advantage of heterozygous females (the single-origin hypothesis). Alternatively, it might have arisen independently in different lineages (the multiple-origin hypothesis). The intron 4 sequences and exon sequences of the squirrel monkey and marmoset favor the latter hypothesis (12). However, the issue remains unsettled because there is evidence of frequent gene conversion at X-linked opsin loci (13–19) that might have misled phylogenetic analyses. Two questions related to this issue are whether the triallelic system also exists in NW monkeys not closely related to the marmoset and squirrel monkey, and whether the common ancestor of the X-linked opsin alleles is more like a middle (17, 20) or a long wavelength opsin gene (21). Another question is whether the two X-linked duplicate opsin genes (P530 and P562) of howler monkeys (a NW monkey genus) were derived from an unequal crossing-over of two alleles very similar to present-day P535 and P562, respectively (2); such a situation would immediately make even a male howler monkey trichromatic.

To resolve these issues, we studied three other NW monkey species. The NW monkeys are divided into two families, Cebidae and Atelidae; alternatively, Atelidae is further divided into two families, Atelidae and Pitheciidae (22, 23). We selected the saki monkey (*Pithecia irrorata*), because it belongs to Atelidae (or Pitheciidae) and is one of the NW monkey species that is most divergent from the squirrel monkey and marmoset, which belong to Cebidae. If the triallelic system also exists in the saki monkey, as it does in the squirrel monkey and marmoset, then it is likely that it exists in many NW monkey species. We also selected the capuchin (*Cebus nigrivittatus*) and the tamarin (*Saguinus mystax*) because they are related to the squirrel monkey and the marmoset, respectively. We found earlier that the three alleles in the squirrel monkey were grouped together in a cluster, as were the three alleles in the marmoset, when the intron 4 sequences were used to construct a phylogenetic tree (12). The monophyly of the three alleles in each species might have been caused by allelic gene conversion, and sequencing the alleles in the capuchin and the tamarin may help clarify this issue. Previously, we have sequenced the exons of the three alleles in these two species (24). In this study, we sequenced intron 4 of each allele in both species. In addition, we found an additional allele in the capuchin and sequenced exons 3, 4, and 5, and intron 4 of this allele. A detailed analysis of the new and published sequence data was conducted.

Abbreviations: NW, New World; indel, insertion and deletion.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. AF 092850–61, AF095124–35, AF095151–53, and AF095571–73).

†S.B. and Y.T. contributed equally to this work.

¶To whom reprint requests should be addressed. e-mail: li@hgc.sph.uth.tmc.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

© 1998 by The National Academy of Sciences 0027-8424/98/9513749-6\$2.00/0 PNAS is available online at www.pnas.org.

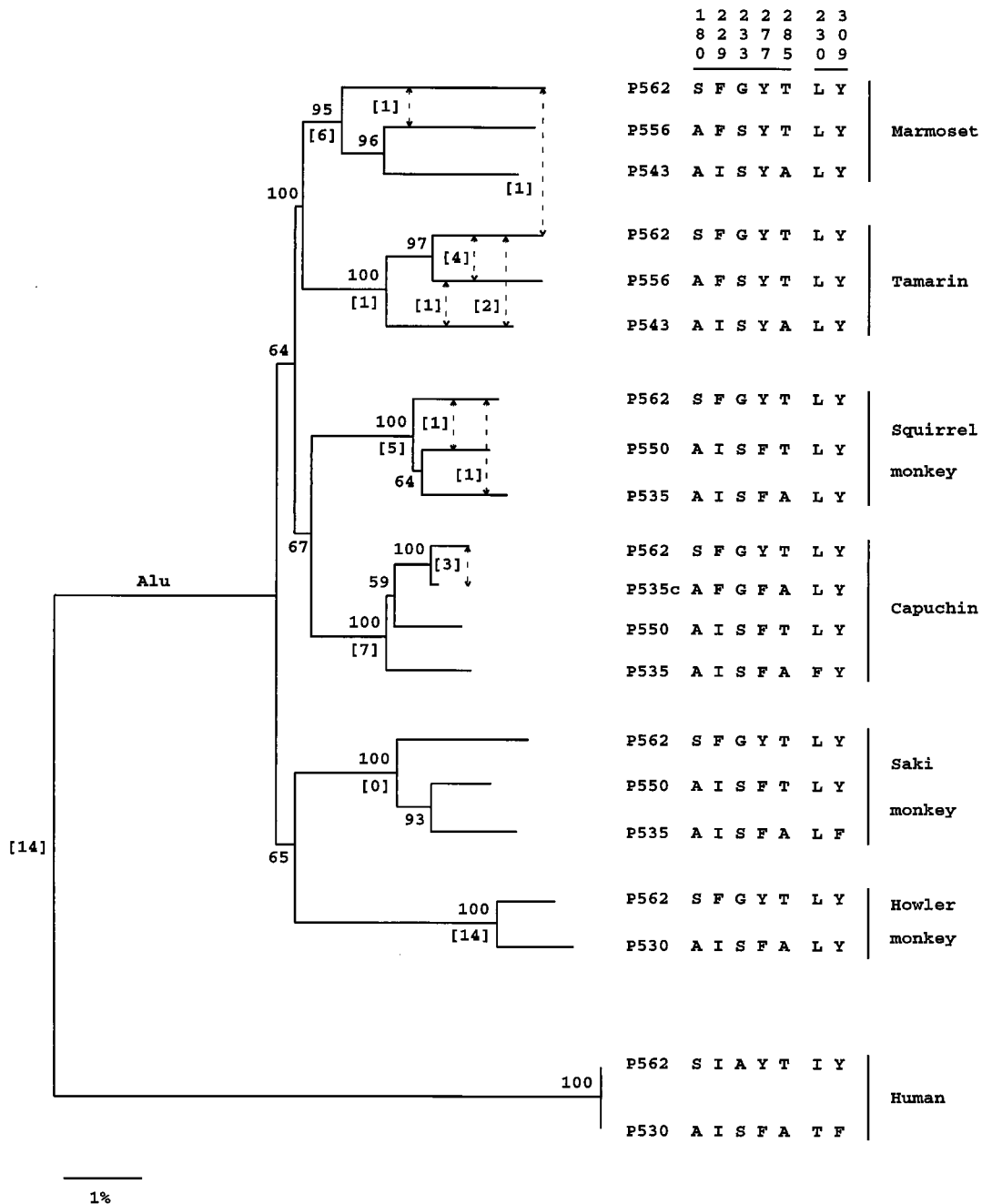


FIG. 1. Neighbor-joining tree derived from an analysis of intron 4 sequences. The distances were computed by using Kimura's (26) two-parameter method. The number at each node denotes the proportion of 500 bootstrap replicates that supported the subset of sequences. The number in brackets below each branch denotes the number of indels supporting the monophyly of the three alleles in each species. A dashed arrow with a number in brackets denotes the number of indels shared by two alleles. The seven numbers listed vertically at the upper right of the tree refer to the positions of the seven critical amino acid sites. The sequence alignment and the positions and sizes of indels are available on request from W.-H.L.

MATERIALS AND METHODS

DNA samples from 16 male saki monkeys (all from Brazil), 8 male capuchins (7 individuals from natural populations in Brazil and 1 of unknown origin), and 5 male and 1 female tamarins (all of them from natural populations in Peru) were obtained and used in this and the previous study (24).

All DNA regions under study were amplified by PCR. For saki monkeys, exons 4 and 5 and intron 4 were amplified by using two pairs of primers that amplify overlapping fragments; exon 3 was amplified separately. For the capuchin and tamarin alleles, the experimental procedure is given in Shyue *et al.* (24). The primers and the experimental conditions used are available on request. For capuchins and tamarins, the PCR prod-

ucts were purified and directly sequenced or cloned into an *EcoRV*-digested pBluescript SK+TA cloning vector. These were then transformed into competent *Escherichia coli* XL1-Blue cells. Both strands were sequenced by using a United States Biochemical Sequenase kit. For the saki monkey samples, the PCR products were purified and directly sequenced on an Applied Biosystems model 377 automatic sequencer by using the Tag DyeDeoxy Terminator Cycle Sequencing Kit (Applied Biosystems). In an initial screening for alleles, each individual was sequenced for exons 3 and 5 to determine the allele it possessed. The determination was based on three major critical sites for spectral tuning located in exon 3 (amino acid position 180) and exon 5 (positions 277 and 285); the amino acids at positions 180, 277, and 285 are, respectively,

EXON 3	EXON 4	INTRON 4			EXON 5
		11111111111111111111			
55	666677	11111122222222222222333333456799000002244446778889			888888
35	589904	2233333333334491114590223444555580225596490782236992333482082666			022235
85	256784	68901234567890193894630122149012796784929062731570481749273768237			136803
P562	TG TTTGGG	TC-----ACATCAGCTGGGGGT---GGACTGTTG-ACGC-TG-GGGCGAACAGCGGCGC			GAGAAA
P535c	GA TTTGGG	TA-----CC-TCAGCTAGAGGT---GGACTGTTGAGA-C-TG-CGGCGATCAGCGGAAC			AGAGTG
P535	GA CACAAT	CAGGGAAAGGGGTGA-TGATCAAAAACCCATAGTCACCAAGAGTACTCC-AT-GTTCTAAACGG			AGAGTG

FIG. 2. Variable sites in exons 3, 4, and 5 and intron 4 of the P562 allele and the two P535 alleles (P535 and P535c) of the capuchin. The numbers at the top refer to the positions of the sites on the complete coding sequence (for exons 3, 4, and 5) and to the positions on an alignment of intron 4 sequences (available on request). The five positions in boldface are the five critical amino acid residues, respectively, at positions 180, 229, 233, 277, and 285.

Ala, Phe, and Ala for P535; Ala, Tyr, and Thr for P543; Ala, Phe, and Thr for P550; Ala, Tyr, and Thr for P556; and Ser, Tyr, and Thr for P562 (24, 25). Once an allele was found in a species, exon 4 and intron 4 were also sequenced.

For introns, the number of nucleotide substitutions per site between sequences was computed by using Kimura's (26) two-parameter method. For exons, the numbers of substitutions per synonymous site and per nonsynonymous site between sequences were computed by using the method developed by W.-H.L. (27). For tree reconstruction, we used both the neighbor-joining method (28) and the maximum parsimony method.

RESULTS

Identification and Sequencing of Alleles. The three alleles, P535, P550, and P562, were identified from the saki monkey samples, and their exons 3, 4, and 5 and intron 4 were completely sequenced. Previously, we have identified alleles P543, P556, and P562 in the tamarin and alleles P535, P550, and P562 in the capuchin and obtained the exon sequences of these alleles (24). In this study, we sequenced the intron 4 of all these alleles in each of the two species. In addition, we found an additional allele in the capuchin, which is denoted P535c to signify the fact that although it is identical with P535 in exons 3 and 5, its intron 4 and exon 4 appear to have been converted by the sequences of P562 (see below). All of the NW monkey intron 4 sequences, including those from the howler monkey, squirrel monkey, and marmoset, contain an *Alu* repeat; this repeat is absent in the human red and green pigment genes. The sequences from the capuchin, tamarin, and saki monkey were compared with the published sequences of the squirrel monkey, marmoset, and howler monkey (2, 12), with the sequences of the human red (P562) and green (P530) pigment genes (29), and with the galago X-linked opsin sequence (30). The sequences are rather similar among species and have been aligned manually; the alignment is available on request.

Phylogenetic Tree of Intron 4 Sequences. A phylogenetic tree was inferred from the intron 4 sequences by using the neighbor-joining method (28) (Fig. 1). The branching order for the six NW monkeys is in agreement with the current view that the marmoset-tamarin and squirrel monkey-capuchin clades belong to the same family (Cebidae) and that the saki monkey and howler monkey belong to the other family (Atelidae) (22, 23). When the *Alu* sequence, which is present in all NW monkey sequences but absent in the two human sequences, is included to infer the phylogeny of the six NW monkey species, the clustering of the saki monkey with the howler monkey and that of the squirrel monkey with the capuchin are both strengthened (the bootstrap values become 94% and 88%). In Fig. 1, the alleles in each species form a monophyletic group supported by a high bootstrap value (100% or 95%). More-

over, for the parsimony analysis of insertions and deletions (indels) (Fig. 1), the monophyly of the alleles in a species is supported by six indels in the marmoset, one indel in the tamarin, five indels in the squirrel monkey, and seven indels in the capuchin. Therefore, both the neighbor-joining tree and the indel analysis strongly support the multiorigin hypothesis; in fact, they imply five independent origins for the triallelic system, one in each species (Fig. 1). Further, Fig. 1 implies that the two duplicate genes in the howler monkey had an origin independent of the triallelic system in the other NW monkeys.

Evidence of Frequent Gene Conversion. The clustering of the alleles in each species might be caused by gene conversion, and there is strong evidence supporting this interpretation. First, gene conversion events can be revealed by an analysis of the indels within each species (Fig. 1). For example, in the tamarin, two indels are shared by P543 and P562, four by P562 and P556, and one by P543 and P556. It is highly unlikely that so many indels arose independently in different alleles. Rather, this pattern strongly suggests the transfer of indels between alleles by gene conversion, partially homogenizing allelic sequences.

Second, in the capuchin, the P535 and P535c alleles are characterized by the same amino acids at the critical positions 180, 277, and 285, and their exons 3 and 5 are completely identical (Fig. 2). However, these alleles differ in exon 4 by 3.6% and in intron 4 by 2.0%. In contrast, the P535c and P562 sequences are identical in exon 4 and differ by only 0.5% in intron 4. A statistical test (31, 32) provides strong evidence ($P = 0.002$) that exon 4 and intron 4 of P535c have been converted by P562; gene conversion is preferred over recombination, because exons 3 and 5 of P535 are identical with those of P535c. As a consequence, the intron 4 sequences drastically distort the evolutionary relationships among P535, P535c, and P562 (Fig. 1).

Third, in every species, the degree of divergence between alleles is not uniform along the intron 4 sequence, suggesting that some parts of intron 4 have been converted more recently than others. For instance, in the marmoset (Fig. 3; similar patterns of distribution were found in other species), the divergence ($\approx 8.0\%$) in the first 600 bp of intron 4 is more than three times that in the rest of intron 4 ($\approx 2.5\%$) when P562 is compared with P543 or P556. In fact, the 8% divergence is even higher than the between-species divergences among the marmoset, tamarin, squirrel monkey, and capuchin (e.g., an average 5.1% divergence the marmoset P543 and squirrel monkey P535 alleles shown at the bottom of Fig. 3), indicating that marmoset P562 and P543 (or P556) diverged before the divergence of these four species. Similarly, when P543 and P556 are compared, the first 400-bp portion is more than twice as divergent as the rest of intron 4 ($\approx 6.0\%$ vs. $\approx 2.5\%$). All three comparisons are significant ($P = 0.001\%$, 0.01% , and 1% , respectively). This finding suggests that the 3' portion of

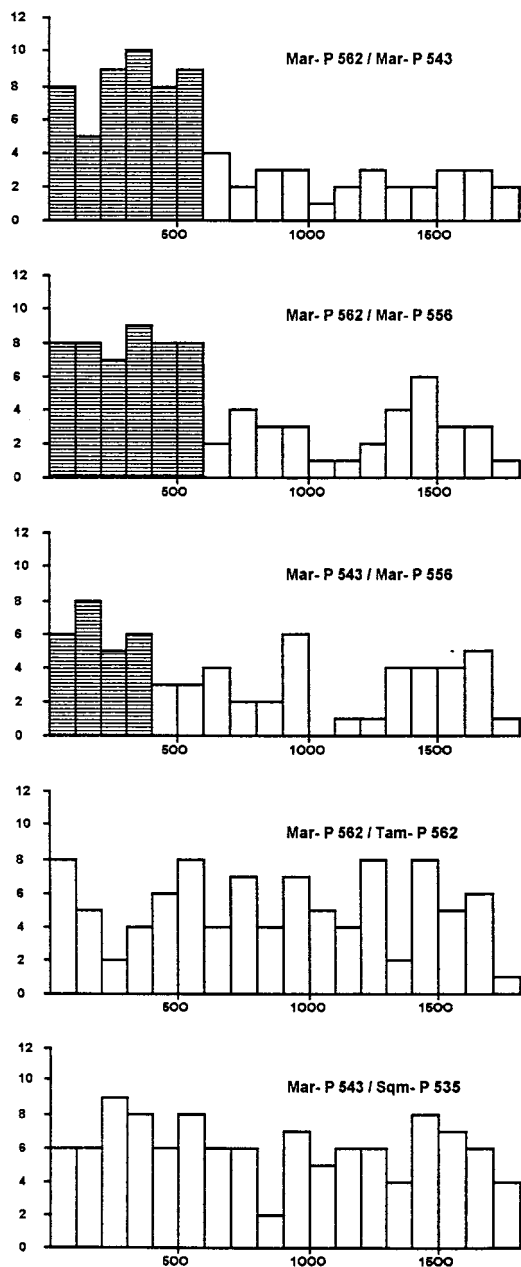


FIG. 3. Histograms showing pairwise variation in divergence as a function of position in intron 4. Horizontally, each bar represents a 100-bp segment. Vertically, each bar shows the number of nucleotide differences per segment. Three within-species and two between-species comparisons are shown (Mar, marmoset; Tam, tamarin; Sqm, squirrel monkey).

intron 4 was homogenized by recent conversion events, whereas the 5' portion was either not homogenized or was homogenized by much earlier conversion events.

Exon Tree. Fig. 4 shows the neighbor-joining tree based on the sequences of exons 3, 4, and 5. The clustering of the sequences is not in good agreement with the species tree; e.g., saki monkey P562 (SAK-P562) is clustered with squirrel monkey P562 (SQM-P562) rather than with howler monkey P562 (HOW-P562). One reason for this lack of agreement may be because these exons are also not free from the effects of gene conversion. In fact, gene conversion events between exons of X-linked duplicate opsin genes have been reported in many studies (13–19) and, as mentioned above, exon 4 of the capuchin P535c allele was derived from the P562 allele. However, these exons should be less affected by gene conver-

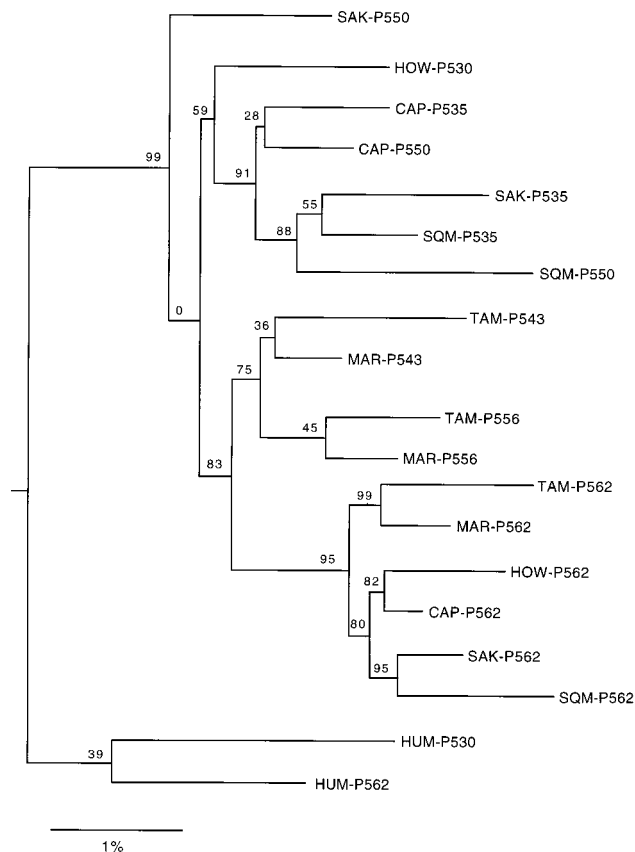


FIG. 4. Neighbor-joining tree derived from an analysis of exon 3, 4, and 5 sequences. The distances were computed by using the method of W.-H.L. (27), with weights of 80% and 20% for nonsynonymous and synonymous substitutions, respectively. The number at each node denotes the percent of 500 bootstrap replicates that supported the subset of sequences. The galago sequences of Zhou *et al.* (30) were used to root the tree (CAP, capuchin; HUM, human; HOW, howler monkey; MAR, marmoset; SAK, saki monkey; SQM, squirrel monkey; TAM, tamarin).

tion than introns because they contain critical sites for spectral tuning, and a gene conversion involving a critical site may be eliminated from the population by natural selection. Thus, Fig. 4 should be more reliable than Fig. 1 in its depiction of the origin of the alleles. Clearly, it does not support the multiorigin hypothesis of Fig. 1. In particular, the P562 alleles of the saki monkey, squirrel monkey, capuchin, marmoset, and tamarin are clustered, suggesting that these alleles were derived from a single origin. Moreover, howler monkey P562 belongs to this cluster, whereas howler monkey P530 is clustered with the P535 alleles of the saki monkey, squirrel monkey, and capuchin, suggesting that the two howler monkey duplicate genes were derived from a combination of the P562 and P535 alleles.

Maximum Parsimony Tree Based on Critical Amino Acid Residues. As gene conversion can drastically mislead phylogenetic analysis at noncritical sites for spectral tuning, the best data for inferring the evolutionary history of these X-linked color vision alleles and duplicate genes are probably the amino acid changes at critical sites. Sites 180, 277, and 285 have been identified as the major critical sites, causing λ_{\max} shifts of ≈ 5 , 8, and 15 nm, respectively, whereas sites 116, 229, 230, 233, and 309 have been suggested to have minor spectral tuning effects, causing shifts of < 3 nm (24, 25, 33–37). Sites 229 and 233 show a substitution pattern that is highly consistent with the three major sites (Fig. 1). However, sites 230 and 309 show very minor variation among sequences; therefore, they are not informative for our purpose and will not be considered further. Site 116 is in exon 2, which is not under study. In terms of

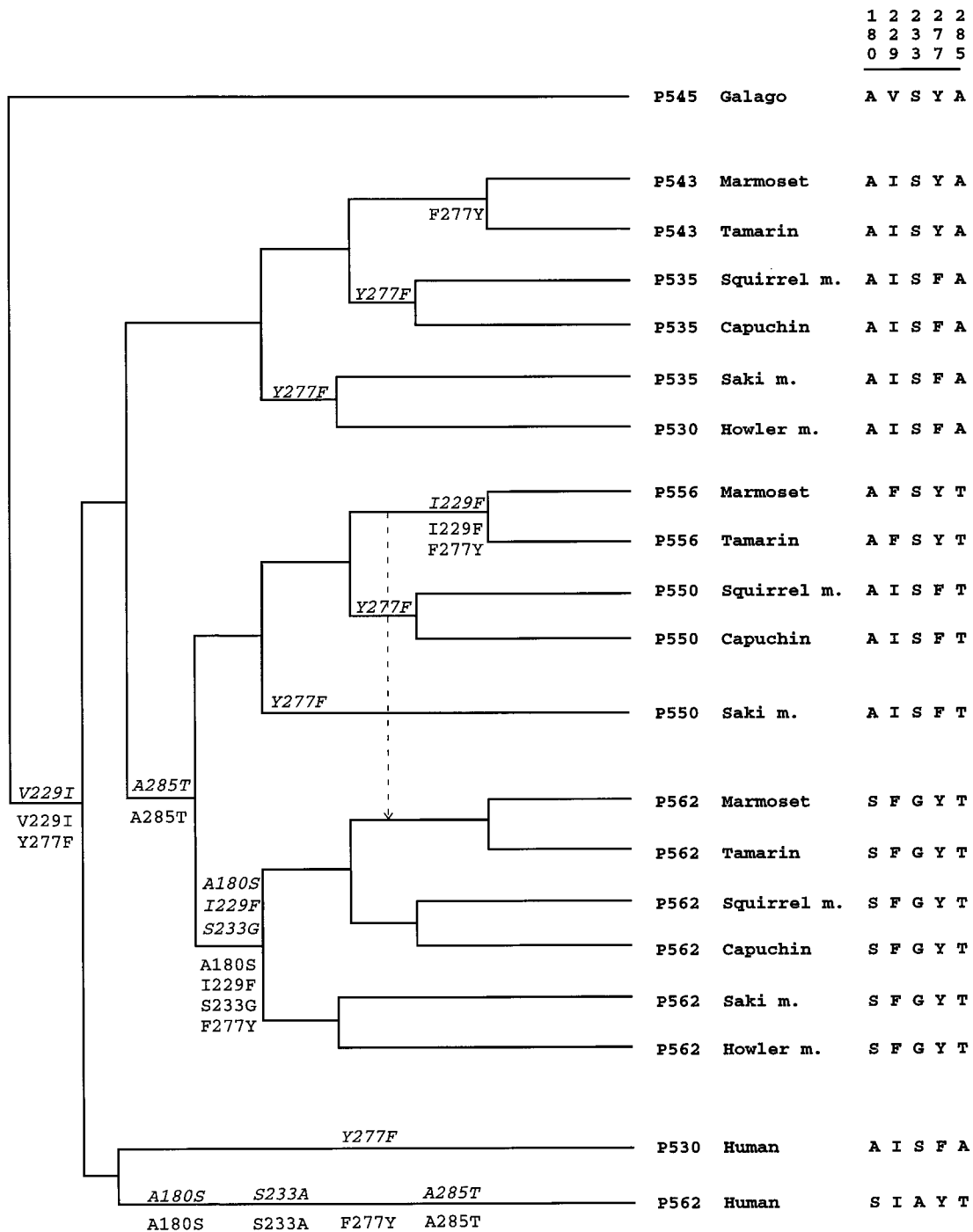


FIG. 5. Tree that requires the minimal number of amino acid changes at the five critical sites under the constraint of the species phylogeny in Fig. 1. Substitutions along branches are indicated as the ancestral amino acid followed by its position on the complete coding sequence and then followed by the new amino acid. Two equally parsimonious pathways are shown; one pathway is above branches (*italics*) and one is below branches. A shorter tree (11 steps instead of 12 steps) is obtained if the marmoset and tamarin P556 alleles have a recent origin and was derived from the P562 allele, as indicated by the arrow. However, this alternative tree is less likely (see *Results*).

parsimony, Fig. 1 is highly implausible because it requires many parallel amino acid changes at the five critical sites considered. For example, at position 180, at least seven parallel changes between serine (S) and alanine (A) are required to explain the differences among alleles and genes at this site. When all five critical sites are considered together, the minimum number of substitutions required is 37 or 38, depending on whether one assumes that the amino acids at sites 180, 229, 233, 277, and 285 in the common ancestor of all higher primates were SFGYT (i.e., as in marmoset P562) or AISFA (as in saki monkey P535, or AISYA as in marmoset P543). (Capuchin P535c is a minor

allele and, for simplicity, is not considered.) In comparison, Fig. 5 requires only 11 or 12 amino acid substitutions in the higher primate sequences, assuming that the ancestral amino acids were AISFA (as in saki monkey, squirrel monkey, and capuchin P535) or AISYA (as in marmoset and tamarin P543). Both scenarios require 13 substitutions for the entire tree (i.e., including also the two galago sequences) and are the most parsimonious scenarios subject to the constraint of the current knowledge of the phylogeny of these primates (i.e., the species phylogeny of Fig. 1). A slightly more parsimonious scenario (one fewer change) is to assume that the marmoset-tamarin

P556 was derived from the common ancestor of marmoset and tamarin P562. However, this scenario implies that the P550 allele has become lost in the common ancestor of the marmoset and tamarin (or in both species), despite the fact that it has persisted in the saki monkey, squirrel monkey, and capuchin.

CONCLUSIONS

In summary, we propose the following conclusions. First, because the tree presented in Fig. 5 requires less than one-third the number of critical amino acid changes required by the tree presented in Fig. 1 (11 or 12 vs. 37 or 38), the single-origin hypothesis is far more plausible than the multiple-origin hypothesis that was suggested by the intron 4 sequences. The single-origin hypothesis implies that the triallelic system has persisted in these NW monkeys for >20 million years, because it has been estimated that the divergence of the howler monkey lineage and the squirrel monkey–marmoset lineage occurred ≈20 million years ago (22, 23). Such antiquity of allelic lineages has been known only for major histocompatibility complex and Ig genes (38). Second, noncritical regions of the alleles in a population can be readily homogenized by allelic gene conversion, so even ancient alleles, which may have been maintained by balancing selection, may not become very divergent. In this case, sequence data may lead to erroneous phylogenetic inferences and to a severe underestimation of the antiquity of the alleles. Third, a parsimony analysis with the galago sequences as outgroups (Fig. 5) suggests that the X-linked opsin alleles and duplicate opsin genes in the higher primates (including humans) were derived from a middle wavelength (green) opsin gene similar to either P543 (20) or P535 (17), but not from a long wavelength (red) opsin gene as inferred without the galago sequences (21). Fourth, the two howler monkey duplicate genes (P530 and P562) have a separate origin from the human red and green pigment genes; in fact, unlike the latter, the howler monkey introns 4 contain the *Alu* insertion. They were evidently derived from the incorporation of the P535 and P562 alleles into one chromosome. If this suggestion is substantiated by further data, it will provide an example in which the recombination of two overdominant alleles into one chromosome provides the basis for the selective advantage of gene duplication (39). Finally, we note that even the most parsimonious tree (Fig. 5) requires several parallel substitutions at some critical sites (e.g., at least four F→Y changes at site 277). The parallel changes and the antiquity of the triallelic system strongly suggest that the system has been maintained by natural selection.

We thank W. Fitch for comments. This study was supported by National Institutes of Health Grant GM57721 and by the Betty Wheelless Trotter Professorship.

- Jacobs, G. H., Neitz, M., Deegan, J. F. & Neitz, J. (1996) *Nature (London)* **382**, 156–158.
- Boissinot, S., Zhou, Y.-H., Qiu, L., Dulai, K. S., Neiswanger, K., Schneider, H., Sampaio, I., Hunt, D. M., Hewett-Emmett, D. & Li, W.-H. (1997) *Zool. Stud.* **36**, 360–369.
- Hunt, D., Dulai, K. S., Cowing, J. A., Mollon, J. D., Bowmaker, J. K., Lee, B. B., Hewett-Emmett, D. & Li, W.-H. (1998) *Vision Res.* **38**, 3299–3306.
- Jacobs, G. H. (1984) *Vision Res.* **24**, 1267–1277.
- Mollon, J. D., Bowmaker, J. K. & Jacobs, G. H. (1984) *Proc. R. Soc. London Ser. B* **222**, 373–399.
- Jacobs, G. H. & Neitz, J. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 2545–2549.
- Jacobs, G. H. & Neitz, J. (1987) *Vision Res.* **27**, 1263–1268.
- Travis, D. S., Bowmaker, J. K. & Mollon, J. D. (1988) *Vision Res.* **28**, 481–490.
- Williams, A. J., Hunt, D. M., Bowmaker, J. K. & Mollon, J. D. (1992) *EMBO J.* **11**, 2039–2045.
- Jacobs, G. H., Neitz, J. & Neitz, M. (1993) *Vision Res.* **33**, 269–274.
- Jacobs, G. H. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 577–581.
- Shyue, S.-K., Hewett-Emmett, D., Sperling, H. G., Hunt, D. M., Bowmaker, J. K., Mollon, J. D. & Li, W.-H. (1995) *Science* **269**, 1265–1267.
- Balding, D. J., Nichols, R. A. & Hunt, D. M. (1992) *Proc. R. Soc. London Ser. B* **249**, 275–280.
- Ibbotson, R., Hunt, D. M., Bowmaker, J. K. & Mollon, J. D. (1992) *Proc. R. Soc. London Ser. B* **247**, 145–154.
- Shyue, S.-K., Li, L., Chang, B. H.-J. & Li, W.-H. (1994) *Mol. Biol. Evol.* **11**, 548–551.
- Zhou, Y.-H. & Li, W.-H. (1996) *Mol. Biol. Evol.* **13**, 780–783.
- Winderickx, J., Battisti, L., Hibiya, Y., Motulsky, A. G. & Deeb, S. S. (1993) *Hum. Mol. Genet.* **2**, 1413–1421.
- Deeb, S. S., Jorgensen, A. L., Battisti, L., Iwasaki, L. & Motulsky, A. G. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 7262–7266.
- Reyniers, E., Van Thienen, M.-N., Meire, F., De Boule, K., Devries, K., Kestelijn, P. & Willems, P. J. (1995) *Genomics* **29**, 323–328.
- Jacobs, G. H. (1993) *Biol. Rev.* **68**, 413–471.
- Nei, M., Zhang, J. & Yokoyama, S. (1997) *Mol. Biol. Evol.* **14**, 611–618.
- Schneider, H., Schneider, M. P. C., Sampaio, I., Harada, M. L., Stanhope, M., Czelusniak, J. & Goodman, M. (1993) *Mol. Phylogenet. Evol.* **2**, 225–242.
- Schneider, H., Sampaio, I., Harada, M. L., Barroso, C. M. L., Schneider, M. P. C., Czelusniak, J. & Goodman, M. (1996) *Am. J. Phys. Anthropol.* **100**, 153–179.
- Shyue, S.-K., Boissinot, S., Schneider, H., Sampaio, I., Schneider, M. P., Abee, C. R., Williams, L., Hewett-Emmett, D., Sperling, H. G., Cowing, J. A., *et al.* (1998) *J. Mol. Evol.* **46**, 697–702.
- Neitz, M., Neitz, J. & Jacobs, G. H. (1991) *Science* **252**, 971–974.
- Kimura, M. (1980) *J. Mol. Evol.* **16**, 111–120.
- Li, W.-H. (1993) *J. Mol. Evol.* **36**, 96–99.
- Saitou, N. & Nei, M. (1987) *Mol. Biol. Evol.* **4**, 406–425.
- Nathans, J., Thomas, D. & Hogness, D. S. (1986) *Science* **232**, 193–202.
- Zhou, Y.-H., Hewett-Emmett, D., Ward, J. P. & Li, W.-H. (1997) *J. Mol. Evol.* **45**, 610–618.
- Maynard Smith, J. (1992) *J. Mol. Evol.* **34**, 126–129.
- Robertson, D. L., Hahn, B. H. & Sharp, P. M. (1995) *J. Mol. Evol.* **40**, 249–259.
- Merbs, S. L. & Nathans, J. (1992) *Nature (London)* **356**, 433–435.
- Merbs, S. L. & Nathans, J. (1993) *Photochem. Photobiol.* **58**, 706–710.
- Asenjo, A. B., Rim, J. & Oprian, D. D. (1994) *Neuron* **12**, 1131–1138.
- Yokoyama, R. & Yokoyama, S. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 9315–9318.
- Yokoyama, S. & Radlwimmer, F. B. (1998) *Mol. Biol. Evol.* **15**, 560–567.
- Klein, J., Takahata, N. & Ayala, F. J. (1993) *Sci. Am.* **269**, 78–83.
- Spofford, J. B. (1972) *Brookhaven Symp. Biol.* **23**, 121–143.