SHORT PAPER

Recent gene conversion between genes encoding human red and green visual pigments

Kei-ich KUMA, Hidenori HAYASHIDA¹⁾ and Takashi MIYATA

*Department of Biology, Faculty of Science,
Kyushu University, Fukuoka 812*

(Received May 20, 1988)

ABSTRACT

Nucleotide sequences of the human X-linked red and green pigment genes were compared, and the number of silent substitutions per site (K_s^c) between these genes was analysed in comparison with the corresponding values of primate genes. Taking the retarded mutation rate of X-linked genes into consideration (Miyata et al., 1987), the red and green pigment genes were shown to have undergone gene conversion at around the time of separation of African apes and orangutan. Thus the recent gene conversion and retarded mutation rate in these X-linked genes are probably responsible for the strong sequence similarity between these genes, which is likely to facilitate the occurrence of red-green color blindness in the human population. It was also shown that the red pigment gene evolved about five times more rapidly than the green pigment gene since the latest gene conversion.

Human color vision is mediated by three light-sensitive pigments, each composed of an integral membrane apoprotein, opsin, which is covalently linked to 11-*cis*-retinal. Recently genomic and cDNA clones of the three (blue, red, and green) pigment genes have been isolated and sequenced (Nathans et al., 1986a), providing a molecular basis for investigating underlying mechanisms of the inherited variation in human color vision. The human red and green pigment genes are tightly linked to each other on the q arm of the X chromosome. Loci for the green pigment gene vary in number among color-normal individuals, whereas only a single locus exists for the red pigment gene. The two pigment genes are structurally closely related; the degree of homology is surprisingly high (Nathans et al., 1986a, b).

We report here that this strong sequence similarity has possibly been generated by a very recent gene conversion event, which is estimated to have taken place at around the time of separation of orangutan and African apes. Contrarily, the divergence of the red and green pigment genes probably occurred at a remote

¹⁾ Present address: National Institute of Genetics, Mishima 411, Japan

time before apes and Old World monkeys split by gene duplication. In addition, the limited mutation frequency in X-linked genes (Miyata et al., 1987) would have contributed also to the sequence similarity of these genes in part.

A single nucleotide substitution in a codon results in either replacement of the encoded amino acid (replacement substitution) or alternation of synonymous codon (silent substitution). Both the silent and replacement substitutions behave like clocks; the latter is essentially identical to the protein clock (Miyata et al., 1985). The silent clock runs with similar pace for different genes, which is in sharp contrast to the protein clock running with different paces for different proteins (Miyata., 1980; 1985). This nature characteristic of the silent clock is useful for estimating the divergence time of duplicated genes.

To know the divergence time of the red and green pigment genes, nucleotide sequences of the protein coding regions were compared, and the degree of nucleotide difference K_S at the silent positions was calculated by the method described previously (Miyata and Yasunaga, 1980). The nucleotide difference was corrected for multiple substitutions by the formula: $K_S^c = -(3/4)\ln(1 - (4/3)K_S)$ (Jukes and Cantor, 1969; Kimura, 1983). The K_S^c value was shown in Table 1, in comparison with those of several primate genes. The value of K_S^c of the human red and green pigment genes apparently lies between those of human/chimpanzee pair and human, chimpanzee/orangutan pairs.

Since the red and green pigment genes are mapped on the X chromosome, a correction for their K_S^c value is required for an accurate estimate of the divergence time. We recently have shown that, in the XX female/XY male system, the mutation frequency of X-linked genes is reduced to 2/3 of that of autosome-linked genes, due to the presence of strong asymmetry between male and female in the number of germ-cell divisions (Miyata et al., 1987). This factor is expected to reduce the rate of silent substitution, and thus the value of K_S^c (red/green) (the K_S^c between the red and green pigment genes) multiplied by 3/2 should be compared with those of autosome-linked genes. This leads to an estimate that the degree of sequence divergence between the human red and green pigment genes is close to, or slightly smaller than, that between African apes and orangutan (Table 1).

If gene duplication is responsible for the sequence divergence, the present result appears to contradict the evidence that the Old World monkeys also have the red and green pigment genes on the X chromosome, whereas the New World monkeys have only a single locus for a long wavelength (red or green) visual pigment on the X chromosome (Jacobs and Neitz, 1987). The following interpretation is likely to reconcile well with the contradiction. Loci for the red and green pigments already existed before the separation of apes and Old World monkeys, but the nucleotide sequence corresponding to one of the primordial pigment genes was replaced by another through a gene conversion event that occurred at around the time of separation of African apes and orangutan.

Table 1. Comparison of the K_S^c value between red and green pigment genes with those of several primate genes compared between different species

Genes or species compared	Genes	$K_S^c \pm SD$
1) human red and green pigments		0.037 ± 0.012
the above multiplied by 3/2		0.056 ± 0.015
2) human/chimpanzee	Ig ϵ l	0.020 ± 0.008^d
3) human, chimpanzee/orangutan	3 genes ^a	0.067 ± 0.010
4) human/Old World monkeys	3 genes ^b	0.094 ± 0.027^d
5) human/New World monkey	γ -globin	0.134 ± 0.036
6) human, rabbit, bovine, dog	11 genes ^c	0.478 ± 0.048^e
7) human/bovine	rhodopsin	0.441 ± 0.043

a, genes coding for Ig ϵ l, and α 1- and α 2-globins. b, genes coding for α 1-antitrypsin, and β - and δ -globins. c, genes coding for growth hormone, prolactin, IgC γ , IgC κ , IgC α , TSH- α , α -, β - and ϵ -globins, insulin, and ACTH. d, data were taken from Sakoyama et al. (1986). e, data was taken from Kikuno et al. (1985). For 1), 2), 5) and 7), the standard error was estimated as $\sqrt{K_S^c/N_S}$, where N_S is the number of sites compared. In comparisons between mammals, rodents were excluded, because of the exceptionally rapid rate of evolution (Kikuno et al., 1985; Li, 1985). Sequence data sources used in the present work: For 3), α 1- and α 2-globin genes from human (Michelson and Orkin, 1980; 1983) and from orangutan (Marks et al., 1986); the K_S^c values of Ig ϵ l were taken from Sakoyama et al. (1986), and for 5), γ -globin genes from human (Slightom et al., 1980) and from the New World monkey (Giebel et al., 1985) and for 7), human (Nathans and Hogness, 1984) and bovine (Nathans and Hogness, 1983) rhodopsin genes.

Red-green color blindness is caused by alternations in the corresponding genes, possibly resulting from unequal crossing-over, gene conversion, or both (Nathans et al., 1986b). The frequent occurrence of this genetic disease in human population is, therefore, likely to be the result of a strong sequence similarity found between the red and green pigment genes, which would have been generated by gene conversion during recent evolution of apes, and also by the reduced rate of mutation frequency on these X-linked genes. Since a similar degree of sequence similarity is expected to exist in at least African apes, individuals of red-green color blindness would be found in the African ape population.

An examination of replacement substitutions provides additional information about evolution of the red and green pigment genes. An expected K_S^c value is obtained, when the gene encoding the fourth pigment, rhodopsin that mediates vision in dim light, is compared between human and bovine; $K_S^c=0.441$. This value is similar to the standard value between mammalian species (Table 1), implying that no unique constraint is operating at the silent positions of the rhodopsin genes. The corrected difference at the replacement site, K_A^c , on the other hand, is considerably small in value (i.e., $K_A^c=0.036$), being only about one tenth of that at the silent sites ($K_A^c/K_S^c=0.08$). This indicates the presence of strong functional constraints against amino acid replacements in rhodopsin. A

similar degree of constraints against amino acid replacements is also expected for the color pigment genes. A comparison of nucleotide sequences between the human red and green pigment genes, however, revealed an unexpectedly large K_A^c value ($=0.019$) relative to K_S^c ($=0.037$); the ratio K_A^c/K_S^c amounts to about 0.5. It is highly unlikely that the elevated value of the ratio reflects a strong constraint at the silent positions, because these positions are constrained only weakly in most mammalian genes (Miyata et al., 1980; Miyata and Yasunaga, 1981; Miyata and Hayashida, 1981). A more plausible interpretation may be that, after the latest gene conversion, one of the genes accumulated amino acid substitutions at an extraordinarily rapid rate, as often found in duplicated genes.

Using the blue pigment gene that is distantly related to the red and green pigment genes (Nathans et al., 1986a, and b), we estimated the evolutionary distance $K_A^c(o/red)$ between the common ancestor of red/green gene pair "o" and the temporal red gene by the method described previously (Miyata and Hayashida, 1982). The estimated $K_A^c(o/red)$ value with the standard error is 0.016 ± 0.004 . Similarly we have $K_A^c(o/green) = 0.003 \pm 0.002$. Surprisingly the red pigment gene has accumulated amino acid substitutions at five times more rapidly than the green gene counterpart since the latest gene conversion. Although the reason for the rapid rate of red pigment gene evolution is obscure at present, nucleotide sequence analyses including the African ape and orangutan counterparts will clarify this problem.

The nucleotide sequences of ape and the Old World and New World monkey counterparts should provide further detailed information about evolution of the color vision system in primates. It may also be interesting to know whether individuals of red-green color blindness exist in the African ape population or not.

We thank Prof. M. Hasegawa for valuable discussion, and Miss K. Mitsuyasu for technical assistance. This work was supported in part by grants from the Ministry of Education, Science and Culture of Japan.

REFERENCES

- GIEBEL, L. B., van STANTEN, V. L., SLIGHTOM, J. L. and SPRITZ, R. A. (1985) Nucleotide sequence, evolution, and expression of the fetal globin gene of the spider monkey *Ateles geoffroyi*. *Proc. Natl. Acad. Sci. USA* **82**, 6985-6989.
- JACOBS, G. H. and NEITZ, J. (1987) Inheritance of color vision in a New World monkey (*Saimiri sciureus*). *Proc. Natl. Acad. Sci. USA* **84**, 2545-2549.
- JUKES, T. H. and CANTOR, C. R. (1969) Evolution of protein molecules. In: *Mammalian Protein Metabolism III* (ed.: H. n. Munro), pp. 21-132. Academic Press, New York.
- KIKUNO, R., HAYASHIDA, H. and MIYATA, T. (1985) Rapid rate of rodent evolution. *Proc. Japan Acad.* **61(B)**, 153-156.
- KIMURA, M., (1983) *The neutral theory of molecular evolution*, pp. 55-79. Cambridge University Press, Cambridge.
- MARKS, J., SHAW, J-P. and SHEN, C-K. J. (1986) The orangutan adult alpha-globin gene locus: Dupli-

- cated functional genes and a newly detected member of the primate alpha-globin family. *Proc. Natl. Acad. Sci. USA* **83**, 1413-1417.
- MICHELSON, A. M. and ORKIN, S. H. (1983) Boundaries of gene conversion within the duplicated human alpha-gene. *J. Biol. Chem.* **258**, 15245-15254.
- MIYATA, T. and YASUNAGA, T. (1980) Molecular evolution of mRNA: A method for estimating evolutionary rates of synonymous and amino acid substitutions from homologous nucleotide sequences and its application. *J. Mol. Evol.* **16**, 23-36.
- MIYATA, T., YASUNAGA, T. and NISHIDA, T. (1980) Nucleotide sequence divergence and functional constraint in mRNA. *Proc. Natl. Acad. Sci. USA* **77**, 7328-7332.
- MIYATA, T. and YASUNAGA, T. (1981) Rapidly evolving mouse alpha-globin-related pseudo gene and its evolutionary history. *Proc. Natl. Acad. Sci. USA* **78**, 450-453.
- MIYATA, T. and HAYASHIDA, H. (1981) Extraordinarily high evolutionary rate of pseudogenes: Evidence for the presence of selective pressure against changes between synonymous codons. *Proc. Natl. Acad. Sci. USA* **78**, 5739-5743.
- MIYATA, T., and HAYASHIDA, H. (1982) Recent divergence from a common ancestor of human IFN- α genes. *Nature* **295**, 165-168.
- MIYATA, T., HAYASHIDA, H., KIKUNO, R., TOH, H. and KAWADA, Y. (1985) Evolution of interferon genes. In: *INTERFERON 6* (ed.: I. Grosser) pp. 1-30. Academic Press, London.
- MIYATA, T., HAYASHIDA, H., KUMA, K. and 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.
- NATHANS, N. and HOGNESS, D. S. (1983) Isolation, sequence analysis, and intron-exon arrangement of the gene encoding bovine rhodopsin. *Cell* **34**, 807-814.
- NATHANS, J. and HOGNESS, D. S. (1984) Isolation and nucleotide sequence of the gene encoding human rhodopsin. *Proc. Natl. Acad. Sci. USA* **81**, 4851-4855.
- NATHANS, J., THOMAS, D. and HOGNESS, D. S. (1986a) Molecular genetics of human color vision: The genes encoding blue, green, and red pigments. *Science* **232**, 193-202.
- NATHANS, J., PIAATANIDA, T. P., EDDY, R. L., SHOWS, T. B. and HOGNESS, D. S. (1986b) Molecular genetics of inherited variation in human color vision. *Science* **232**, 203-210.
- SAKOYAMA, Y., HONG, K. J., BYUN, S. M., HISAJIMA, H., UEDA, S., YAOITA, Y., HAYASHIDA, H., MIYATA, T. and HONJO, T. (1987) Nucleotide sequences of immunoglobulin epsilon genes of chimpanzee and orangutan: DNA molecular clock and hominoid evolution. *Proc. Natl. Acad. Sci. USA* **84**, 1080-1084.
- SLIGHTOM, J. L., BLECHL, A. E. and SMITHIES, O. (1980) Human fetal G gamma- and A gamma-globin genes: Complete nucleotide sequences suggest that DNA can be exchanged between these duplicated genes. *Cell* **21**, 627-638.
- WU, C. I. and LI, W. H. (1985) Evidence for higher rates of nucleotide substitution in rodents than in man. *Proc. Natl. Acad. Sci. USA* **82**, 1741-1745.