Neuron, Vol. 24, 299–312, October, 1999, Copyright ©1999 by Cell Press

The Evolution and Physiology of Human Color Vision: Insights from Molecular Genetic Studies of Visual Pigments

Review

Jeremy Nathans*
Department of Molecular Biology and Genetics
Department of Neuroscience and
Department of Ophthalmology
Howard Hughes Medical Institute
Johns Hopkins University School of Medicine
Baltimore, Maryland 21205

Nothing in biology makes sense except in light of evolution.

—T. Dobzhansky

Color vision is the process by which an organism extracts information regarding the wavelength composition of a visual stimulus (Figure 1). In its simplest form—exemplified by the wavelength-dependent phototactic responses of halobacteria—color vision is based on the relative abundances of two isoforms of a sensory pigment (Hoff et al., 1997). One isoform preferentially absorbs long wavelength light and mediates a photoattractant response, and the second isoform preferentially absorbs short wavelength light and mediates a photorepellant response. Light absorption photoconverts the first isoform into the second, and the second into the first. Thus, the steady-state ratio of pigment isoforms provides a measure of the spectral composition of the ambient light.

The halobacterial system contains the two cardinal elements of every color vision system: (1) two or more sensory pigments with different spectral sensitivities (although in most organisms these are distinct chromoproteins rather than isoforms of a single chromoprotein), and (2) a mechanism for monitoring the relative number of photons captured by the different pigments. In higher eukaryotes, this system has evolved so that, in general, each pigment resides within a distinct class of photoreceptor cells, and therefore the ratios of photoexcitation of the different pigments can be determined by assessing the ratios of activation of different photoreceptor cells.

Evolution of Vertebrate Visual Pigments

Visual pigments are G protein–coupled receptors in which a seven-transmembrane segment protein is covalently linked to a chromophore, 11-*cis* retinal. Studies of the visual pigment complement and color vision ability of different vertebrates reveal an ancient and nearly universal color vision system in which one visual pigment has an absorption maximum at <500 nm and a second visual pigment has an absorption maximum at >500 nm (Mollon, 1989; Yokoyama, 1997). Rhodopsin, a third and equally ancient pigment, has an absorption maximum at ~500 nm and plays little or no role in color vision. In general, the pigments mediating color vision reside in cone photoreceptors and are used only under bright light conditions, whereas rhodopsin resides in rod photoreceptors and is used under dim light conditions.

*E-mail: jnathans@jhmi.edu.

Present-day vertebrates vary enormously in the sophistication of their color vision, the density and spatial distribution of cone classes, and the number and absorption maxima of their cone pigments (Figure 2; Lythgoe, 1979; Jacobs, 1981, 1993; Yokoyama, 1997). At one extreme, most mammals have only three pigments: the two ancestral cone pigments and rhodopsin. At the other evolutionary extreme, chickens possess six pigments: four cone pigments, one rhodopsin, and a pineal visual pigment, pinopsin. As seen in the dendrogram in Figure 2, the chicken green pigment was derived from a duplication within the rhodopsin branch.

In this evolutionary comparison, humans and their closest primate relatives represent an intermediate level of complexity. Humans have four visual pigments: a single member of the <500 nm family of cone pigments (the blue or short-wave pigment, with an absorption maximum at ~425 nm), two highly homologous members of the >500 nm family (the green or middle-wave pigment, and red or long-wave pigment, with absorption maxima at ~530 and ~560 nm, respectively), and rhodopsin. The presence of only a single gene encoding a >500 nm pigment in almost all New World primates, and in all nonprimate mammals studied to date, places the red/green visual pigment gene duplication in the Old World primate lineage at ~30–40 million years ago, shortly after the geologic split between Africa and South America (Jacobs, 1993). Current molecular genetic evidence suggests that howler monkeys, the only known species of New World primate with two >500 nm pigment genes, have acquired a gene duplication event that is independent of the one within the Old World primate lineage (Jacobs et al., 1996; Hunt et al., 1998; Kainz et al., 1998). The relatively recent acquisition of trichromacy by Old World primates is reflected in the structure of the red and green pigment genes (Nathans et al., 1986a; Vollrath et al., 1988; Ibbotson et al., 1992). These genes reside in a head-to-tail tandem array on the X chromosome and show ~98% DNA sequence identity in the coding, intron, and 3′ flanking sequences. Whether the Old World duplication involved initially identical genes that subsequently diverged, or whether it arose from two X chromosomes that carried polymorphic variants with different absorption spectra, is not known.

Spectral Tuning: Mechanisms and Sequence Determinants

A long-standing challenge in vision research has been to explain the mechanisms by which diverse visual pigment apoproteins regulate the wavelength of absorption of the common 11-*cis* retinal chromophore, and to relate these absorption spectra to the visual abilities and evolutionary histories of each species.

The basic photochemical properties of 11-*cis* retinal are now well understood, and these constrain the strategies that a visual pigment apoprotein can use to effect an absorption shift (Ottolenghi and Sheves, 1989; Birge, 1990). In all visual pigments, 11-*cis* retinal is joined to

Figure 1. The Practical and Aesthetic Advantages of Color Vision

The figure shows the same photograph of autumn foliage with and without color. Color vision permits distinctions between objects based on differences in the chromatic composition of reflected light.

the apoprotein by a Schiff base linkage with a lysine in the center of the seventh transmembrane segment. In visual pigments with absorption maxima greater than ~440 nm, this Schiff base is protonated and therefore positively charged (Figure 3A). All vertebrate visual pigments carry a glutamate in the third transmembrane segment (corresponding to glutamate 113 in bovine rhodopsin), which serves as the counterion to the proton-

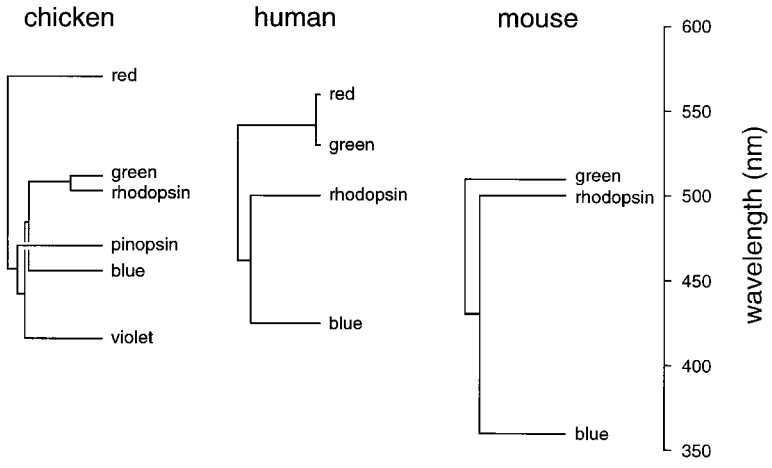

Figure 2. Dendrogram of Chicken, Mouse, and Human Visual Pigments

The branch lengths of the dendrogram correspond to percent amino acid divergence. The absorption maximum of each pigment is indicated by the vertical position of its branch relative to the wavelength axis (right).

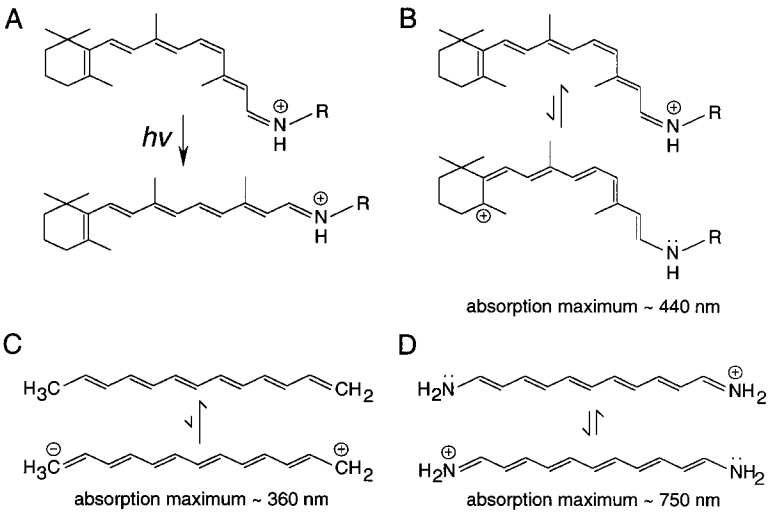

Figure 3. Retinal, Polyene, and Cyanine Structures, Each with Six Double Bonds, Showing π Electron Delocalization and Absorption Maxima

(A) Photoisomerization of a protonated Schiff base of 11-*cis* retinal to all-*trans* retinal, as occurs in vertebrate visual pigments.

(B and C) Major (upper) and minor (lower) resonance structures of (B) a protonated Schiff base of 11-*cis* retinal and (C) a polyene.

(D) Equivalent resonance structures of a cyanine.

ated retinylidene Schiff base (Sakmar et al., 1989; Zhukovsky and Oprian, 1989; Nathans, 1990). As illustrated in Figure 3B, the positive charge on the protonated Schiff base is partially delocalized by alternate resonance structures. This delocalization is relevant to spectral tuning, because the spectral consequences of any perturbation can be understood by assessing its effect on π electron delocalization within the 11-*cis* retinal chromophore: an increase in delocalization leads to a red shift and a decrease in delocalization leads to a blue shift (Kropf and Hubbard, 1958; Mathies and Stryer, 1976; Ottolenghi and Sheves, 1989). This effect is seen most dramatically in a comparison of model compounds that represent the extremes of π electron delocalization (Suzuki, 1967). A polyene with six double bonds (Figure 3C) has extreme bond length alternation, minimal π electron delocalization, and a peak absorption at ~360 nm. By contrast, the corresponding cyanine (Figure 3D) has equivalent bond lengths, maximal π electron delocalization, and a peak absorption at ~750 nm.

Current evidence suggests that vertebrate visual pigments modify the electronic environment of 11-*cis* retinal by (1) modulating the interaction between the protonated Schiff base and its counterion, and (2) modifying the dipolar environment of the conjugated chromophore backbone with neutral amino acid side chains. For example, weakening the interaction between the protonated Schiff base and its counterion promotes delocalization of the positive charge throughout the π electron system and a resulting red shift in the absorption spectrum. In contrast to the variable position of the absorption curve along the wavelength axis, the shape and width of the visual pigment absorption curve is determined by chromophore vibration and is therefore not subject to modification by the apoprotein.

Recent analyses of the absorption spectra of visual pigments expressed from cloned cDNA have allowed a precise delineation of the contribution of particular amino acids to spectral tuning. Within the >500 nm subfamily of cone pigments, five positions within the polypeptide chain have been shown to play critical roles in spectral tuning (Figure 4). Three positions with significant effects on spectral tuning among >500 nm cone pigments—180 in the fourth transmembrane segment and 277 and 285 in the sixth transmembrane segment—account for most of the spectral shift between the human red and green pigments and between the many varieties of >500 nm pigments found among New World primates (Neitz et al., 1991; Merbs and Nathans, 1992a, 1993; Williams et al., 1992; Asenjo et al., 1994). Comparisons of >500 nm pigment gene sequences from New and Old World primates indicate that the same dimorphic amino acid substitutions at these three positions—alanine/serine at 180, tyrosine/phenylalanine at 277, and alanine/threonine at 285—arose independently in multiple evolutionary lineages (Table 1; Shyue et al., 1995; Hunt et al., 1998). Within the primate >500 nm cone pigment subfamily, naturally occurring substitutions at all other sites produce a combined spectral shift of ~5 nm (Figure 4B; Merbs and Nathans, 1993; Asenjo et al., 1994; Neitz et al., 1995; Crognale et al., 1998, 1999; Sharpe et al., 1998).

The serine/alanine dimorphism at position 180 is of special interest because it exists as a common polymorphic variant in the human gene pool, with ~60% of red pigment genes coding for serine and ~40% for alanine (Nathans et al., 1986a; Winderickx et al., 1992b; Sharpe et al., 1998). Variation at position 180 is less common among human green pigment genes, as at least 90% code for alanine. The effect of this single nucleotide polymorphism can be seen in a simple color matching test, the Rayleigh test. A Rayleigh color match is established when the photon capture rates of each of the observers' cone pigments are equal for the two stimuli, and this serves as a sensitive indicator of the number and spectral distribution of the pigments in this region of the spectrum. When matching a spectrally pure yellow light to a mixture of spectrally pure red and green lights, the ~60% of human males who carry serine at position 180 in their red pigment require less red light in the mixture compared to the ~40% who carry alanine, because the serine-containing pigment is red shifted 3–4 nm relative to the alanine-containing pigment (Neitz and Jacobs, 1986, 1990; Merbs and Nathans, 1992a; Winderickx et al., 1992b; Sanocki et al., 1993, 1994; Asenjo et al., 1994; Sharpe et al., 1998).

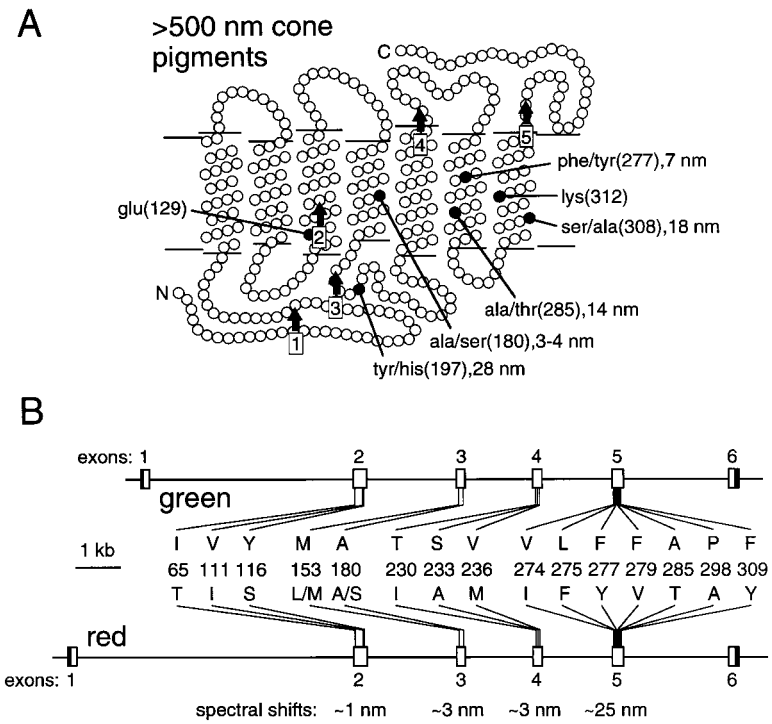

Figure 4. Amino Acids Involved in Spectral Tuning of >500 nm Cone Pigments

(A) Schematic of the human red or green pigment in the membrane. The amino terminus, N, faces the extracellular space; the carboxyl terminus, C, faces the cytosol. The five intron locations are indicated by boxed arrows above numbers. Lysine 312 and glutamate 129 (corresponding to lysine 296 and glutamate 113 in bovine rhodopsin) are the sites of covalent attachment of 11-*cis* retinal and of the Schiff base counterion, respectively. Amino acid differences at five positions that are important for spectral tuning within the >500 nm cone pigment subfamily are shown. For each pair of amino acids, the blue-shifting member of the pair is listed first, followed by the codon position in parentheses and the resulting spectral shift.

(B) Structure of the human red and green pigment genes showing the locations and identities of the 15 amino acids (in the single letter code) that distinguish the encoded pigments. Introns and exons are drawn to scale. Open boxes indicate coding regions; closed boxes indicate noncoding regions. Listed at the bottom are the average spectral shifts referable to the cumulative effects of amino acid differences within each exon.

The dimorphisms at positions 180, 277, and 285 share two attributes which suggest that they may produce spectral shifts by similar mechanisms. First, these three sites are close to the center of their respective transmembrane segments and therefore are likely to form part of the chromophore-binding pocket. Second, in each case the red-shifting member of the dimorphic pair carries a hydroxyl-bearing side chain, and the blue-shifting member carries a similarly sized aliphatic side chain (Figure 4). Most likely, the hydroxyl groups produce a red shift by promoting π electron delocalization in the 11-*cis* retinal chromophore through dipole–dipole interactions along the conjugated chain.

At present, spectral tuning within the <500 nm visual pigment family is less well defined than in the >500 nm family. One site-directed mutagenesis study of bovine rhodopsin suggests that the human blue pigment owes ~60 nm of its blue shift relative to the 500 nm absorption peak of rhodopsin to the cumulative effect of amino acid differences at nine positions in or near the chromophore binding pocket (Lin et al., 1998)

Table 1. Dimorphic Amino Acids at the Three Most Important Locations for Spectral Tuning among Different Primate >500 nm Cone Pigments

	Absorption Maximum	Amino Acids at Three Positions		
		180	277	285
Human, chimpanzee, and gorilla	560	S	Y	T
	530	A	F	A
Capuchin monkey	563	S	Y	T
	550	A	F	T
	535	A	F	A
Marmoset	563	S	Y	T
	556	A	Y	T
	543	A	Y	A

Capuchin monkeys and Marmosets are New World primates, and each species possesses three allelic variants of a >500 nm X-linked visual pigment as indicated (after Hunt et al., 1998).

Spectral Tuning: Physiologic Implications

Why do humans and other trichromatic primates have cone pigments with absorption maxima at ~425, ~530, and ~560 nm rather than at some other set of wavelengths? In particular, why are the absorption spectra of the green and red pigments so close together relative to the blue pigment spectrum? The ease with which absorption shifts can be produced by mutation implies a selective pressure for maintaining the spectral separation found in the present-day primate cone pigments. To address these questions, we need to understand the evolutionary importance of different visual tasks and the nonequivalent roles played by the different cone types in primate vision (Mollon, 1989).

Figure 5 shows a series of cone pigment spectra (left) and the chromatic discrimination curves associated with each set of spectra (right). The normal human cone pigment spectral sensitivities (Figure 5A) are compared with arrangements in which (1) the green pigment is midway between the blue and red pigments (Figure 5C), (2) the green and red pigments differ by only 5 nm as seen in humans with one type of anomalous trichromacy (Figure 5E), or (3) only a single >500 nm pigment is present, as seen in one type of human dichromacy (Figure 5G). The accompanying chromatic discrimination curves plot the threshold for distinguishing neighboring spectrally pure lights as a function of wavelength and were calculated from the amplitude of two differencing operations: (1)

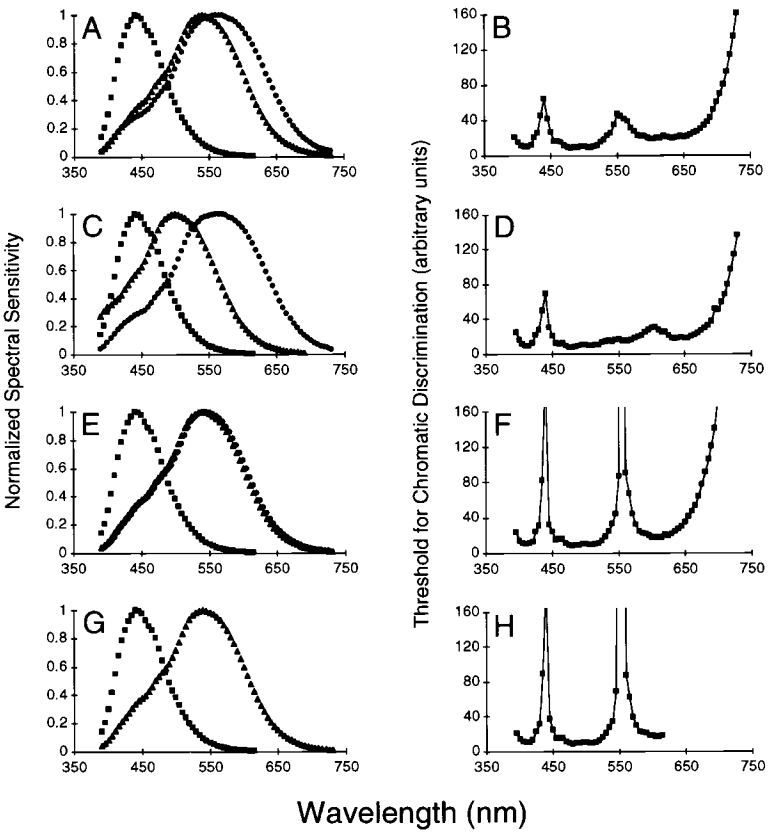

Figure 5. Relationships between Visual Pigment Absorption Spectra and Spectral Discrimination

Four sets of blue, green, and red pigment absorption curves (left) and the calculated spectral discrimination functions associated with each set (right).

(A and B) Normal human trichromat.

(C and D) Hypothetical trichromat in which the green pigment curve is shifted to a point midway between the blue and red curves.

(E and F) Anomalous trichromat in which the the normal red pigment absorption curve has been replaced by a green-like pigment that is red shifted 5 nm relative to the normal green pigment absorption curve (protanomaly).

(G and H) Dichromat missing the red pigment (protanopia). The normal spectral sensitivities (A) were determined psychophysically from color matching data (Stockman et al., 1993). For the calculation of chromatic discrimination, each cone signal was first set to unity at the point of maximal absorption. For the interval from 390–730 nm the following were calculated at intervals of 5 nm: (1) the difference between the red and green signals (R–G), and (2) the difference between the blue signal and the mean of the red and green signals (B–Y). Each of these values can be plotted in a chromatic discrimination space in which one axis represents the amplitude of the R–G signal and the other axis represents the amplitude of the B–Y signal. The distance between adjacent points in this space gives a measure of the discriminability of the two stimuli. For each pair of adjacent wavelengths, i.e., those that differ by 5 nm, the corresponding distance in chromatic discrimination space is: $\{[(R_1 - G_1) - (R_2 - G_2)]^2 + [(B_1 - Y_1) - (B_2 - Y_2)]^2\}^{1/2}$), where subscripts "1" and "2" indicate the visual pigment absorption efficiencies for the first and second of these wavelengths, respectively. The threshold for chromatic discrimination is defined as distance^{-1} in this chromatic discrimination space. Thus, a large change in cone signals between spectrally pure stimuli of X nm and X + 5 nm results in a low threshold for chromatic discrimination in the neighborhood of these wavelengths.

subtracting the green cone signal from the red cone signal, which will be referred to as the "R–G" signal, and (2) subtracting the blue cone signal from the sum of the red and green (equal to yellow) cone signals, which will be referred to as the "B–Y" signal (see legend to Figure 5). This method of calculating chromatic discriminability is referred to as a "line element" analysis. It was first introduced by Helmholtz with subsequent refinements by Schrödinger and Stiles (reviewed by Wyszecki and Stiles, 1982). The line element calculation parallels the spectral analysis performed by the retina (described more fully below), and it produces chromatic discrimination curves for normal, anomalous trichromat, and dichromat visual pigment sets that closely match those obtained psychophysically (Boynton, 1979).

The spectral positions where chromatic discrimination is poorest, indicated by maxima in the curves on the right side of Figure 5, are easily understood in terms of the underlying cone pigment absorption spectra. For the normal human cone pigment sensitivities illustrated in Figure 5A, these positions are calculated to be ~440 nm, ~550 nm, and beyond 700 nm. The first of these positions corresponds to a point at which (1) the rate of photon capture by the blue pigment changes minimally with changes in stimulus wavelength, and (2) the relative rates of photon capture by the red and green pigments change minimally with changes in stimulus wavelength. As a result, wavelength-dependent changes in both the R–G and the B–Y signals are small. Similar reasoning applies to the other points of poor chromatic discrimination.

Comparison of Figures 5B and 5D shows that shifting the green pigment spectrum to shorter wavelengths produces only small changes in the chromatic discrimination curve. However, bringing the red and green pigment spectra closer together results in a progressive loss of chromatic discrimination in two regions of the spectrum (Figures 5E and 5F). One region is near the peak of the two >500 nm cone pigments (~535 nm in this example), where the blue pigment absorption efficiency is low and the relative wavelength-dependent change in photon capture by the >500 nm pigments is minimal. The second region is at long wavelengths where chromatic discrimination relies entirely on the small differential afforded by the R–G signal. In the extreme case in which the red and green pigment spectra coincide or one of the two pigments is missing (Figures 5G and 5H), chromatic discriminability at wavelengths of >600 nm is virtually nonexistent.

The advantages conferred by well-separated red and green pigment absorption curves can be appreciated by identifying those tasks which red/green dichromat

and anomalous trichromat humans perform poorly (Mollon, 1989; Stewart and Cole, 1989). As might be expected, these individuals have special difficulty with visual tasks in which similarly sized colored objects must be discriminated under conditions where lightness varies locally in an unpredictable manner. In the laboratory or clinic, this is embodied in the Ishihara plates, a simple screening test in which a digit must be recognized as a pattern of colored dots among background dots of variable lightness and color. In the real world, identifying colored fruit among dappled foliage presents a similar challenge, and it has been known for centuries that color-defective observers have great difficulty performing this task (Huddart, 1777).

These considerations have led a number of investigators to examine the possibility that one of the principal selective pressures for primate trichromacy may have been the identification of fruit and the assessment of its state of ripeness (Polyak, 1957; Mollon, 1989; Osorio and Vorobyev, 1996; Regan et al., 1998). In support of this hypothesis, a number of tropical fruits appear to be consumed principally by primates. When ripe, these fruits are typically yellow or orange and could be easily spotted among green foliage by an animal with high acuity trichromatic vision. Importantly for the plants, the primates disperse the seeds either by ingesting them intact or by spitting them out at some distance from the tree (Gautier-Hion et al., 1985). These observations suggest the intriguing possibility that frugivorous primates and the fruit that they consume have coevolved in a relationship much like that between bees and the flowers they pollinate.

The evolutionary significance of this or other hypotheses regarding primate trichromacy might be testable in the field by taking advantage of naturally occuring color vision variation among New World primates. Most New World primates carry only a single X-linked visual pigment gene, but this gene is polymorphic within a species and the several alleles differ by 10–30 nm in their absorption spectra (Mollon et al., 1984; Jacobs, 1993). Heterozygous females have excellent trichromatic vision because X inactivation, which is known to produce a fine-grained mosaic within the retina (Reese et al., 1995), provides them with two classes of $>$500 nm cones. Males and homozygous females are dichromats. Thus, it might be possible to observe individual animals in the wild and to correlate different color vision capacities, determined by genotyping, with foraging or other behaviors.

Any discussion of visual pigment spectra would be incomplete without considering the effect of these spectra on spatial vision (Rodieck, 1998). According to the late Mathew Alpern, one of the great visual psychophysicists of the twentieth century, the real business of primate vision is the discrimination of features or objects based on their location; color vision is just a hobby (Alpern, 1982). The large fraction of the primate visual cortex devoted to foveal vision, where individual cones are connected in a one-to-one relationship with ganglion cells, attests to the importance of this 0.1% of the visual field where spatial resolution is best. Two aspects of any color vision system inevitably degrade spatial vision. First, the need to tile the retinal surface with multiple nonidentical cone types means that differences in excitation between neighboring photoreceptors can arise either from chromatic or luminance differences, an ambiguity that cannot be locally resolved. Second, color vision requires sampling of the image over a range of wavelengths, but refraction at the air–cornea interface and by the lens is accompanied by chromatic aberration. This produces an image that is in focus at one wavelength and increasingly out of focus at more distant wavelengths. The primate retina has partially solved both of these problems by sprinkling the retina only sparingly with blue cones, and, in some species, decreasing their representation in the central fovea (Curcio et al., 1991; Martin and Grunert, 1999). In contrast to the $>$500 nm cones, blue cones do not feed into ganglion cells with center–surround antagonistic receptive fields, nor do they exhibit a one-to-one connection with ganglion cells. As a result, the $>$500 nm cones, which make up $\sim$95% of the cone mosaic, appear to be the principal conveyors of spatial information. By allowing the image to be optimally focused at longer wavelengths and defocused at shorter wavelengths, this segregation of function minimizes the degradation of spatial vision caused by chromatic aberration. Among Old World primates, including humans, the $\sim$30 nm separation of red and green pigment spectral sensitivities is likely to represent a compromise between the competing requirements for chromatic and spatial discrimination.

The Arrangement and Rearrangement of Red and Green Pigment Genes

In humans (and presumably all Old World primates), the red and green pigment genes are arranged in a tandem array on the X chromosome. As shown in Figure 6, the array consists of repeat units of 39 kb, each containing a single visual pigment gene at one end (Vollrath et al., 1988). The red pigment gene is the most proximal, followed by one or more green pigment genes. As noted above, the human red and green pigment genes are remarkably similar, differing by only $\sim$2% at the nucleotide level in both coding and noncoding sequences. On the assumption that this duplication occurred $\sim$40 million years ago in the Old World lineage, it is likely that the extremely high sequence identity observed between repeat units in the human array reflects more recent gene conversions.

The similarity between repeat units predisposes the tandem array to unequal homologous recombination. Crossing-over within the $\sim$25 kb intergenic regions produces a gain or loss of one or more genes, whereas recombination within the transcription units—intragenic recombination—creates hybrid genes (Figure 6). As a consequence of the high degree of homology between red and green pigment gene introns and their large size relative to the exons, the great majority of intragenic crossovers occur within introns rather than exons, resulting in a limited number of distinct hybrid gene types. These unequal exchanges produce the common anomalies of red/green color vision that occur in $\sim$8% of Caucasian, $\sim$5% of Asian, and $\sim$3% of African males (Post, 1962; Nathans et al., 1986b; Deeb et al., 1992). They also account for the highly variable number of green pigment genes in the normal trichromat population. Approximately 25% of human X chromosomes carry two

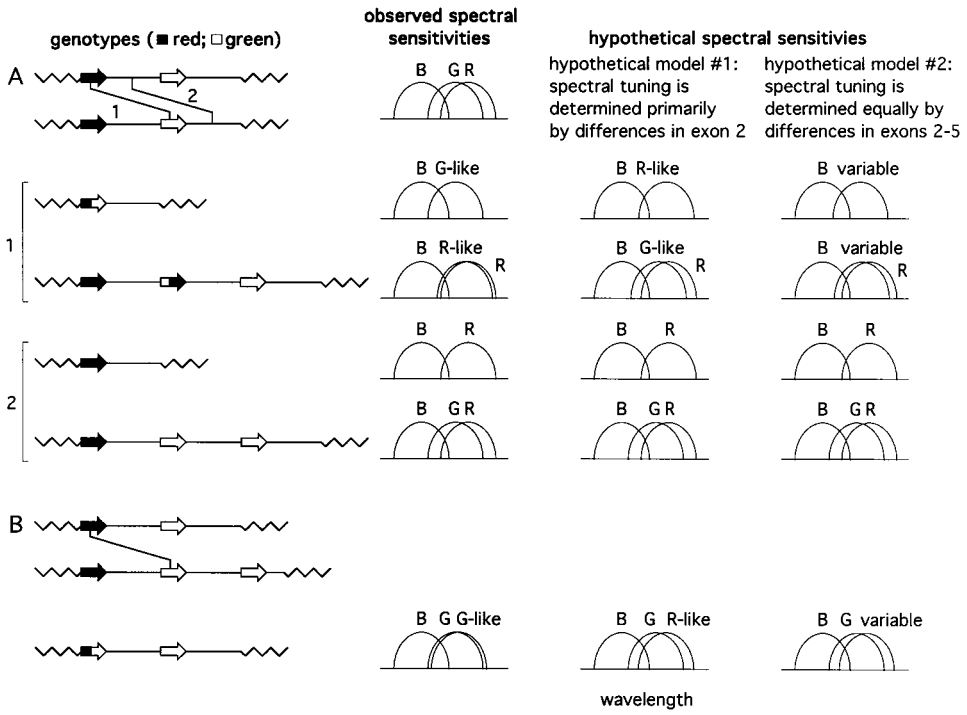

Figure 6. Phenotypic Consequences of Unequal Recombination Events within the Human Red and Green Pigment Genes

Each visual pigment gene transcription unit is represented by an arrow: the base of the arrow represents the 5′ end and the tip of the arrow represents the 3′ end. Closed symbols indicate red pigment; open symbols indicate green pigment. Straight lines represent homologous 3′ flanking sequences, and zig-zag lines represent unique flanking DNA, i.e., DNA that was not included in the primordial gene duplication event.

(A) Intragenic (indicated by "1") and intergenic (indicated by "2") unequal homologous recombination events between wild-type arrays with one red pigment gene and one green pigment gene give rise to the two pairs of reciprocal products shown beneath the parental arrays.

(B) Intragenic recombination between wild-type arrays in which one array has two green pigment genes. One recombination product (not shown) is identical to the second one in the upper set; the reciprocal product (shown) carries a 5′ red–3′ green hybrid and one green pigment gene. For males carrying the arrays illustrated at left, cone pigment spectral sensitivities are shown for the observed human pigments in which spectral tuning is determined primarily by sequence differences in exon 5 (left), a hypothetical arrangement in which spectral tuning is determined primarily by sequence differences in exon 2 (center), and a hypothetical arrangement in which spectral tuning is determined equally by sequence differences in exons 2–5 (right). In the actual human arrangement, anomalous trichromacies are characterized by pairs of closely related pigments. In the first hypothetical arrangement (center), the absorption spectra of most hybrid pigments would differ significantly from that of the remaining normal pigment. In the second hypothetical arrangement (right), the large range of hybrid pigment spectral sensitivities (indicated in the figure as "variable") would produce a correspondingly large range of anomalous trichromacies. In the schematic diagrams of visual pigment sensitivity curves, "B," "G," and "R" represent the normal blue, green, and red curves; "G-like" and "R-like" represent pigments with spectral sensitivities similar to those of the green and red pigments, respectively; and "variable" represents pigments with any of a variety of spectral sensitivities in the interval between normal green and red pigment sensitivities.

visual pigment genes, 50% carry three genes, 20% carry four, and 5% carry five or more (Nathans et al., 1986b; Drummond-Borg et al., 1989; Macke and Nathans, 1997; Yamaguchi et al., 1997; Wolf et al., 1999). As discussed more fully below, RT–PCR experiments using postmortem human retinas indicate that for those arrays with three or more genes, only the two that occupy the most 5′ positions are generally expressed (Winderickx et al., 1992a; Yamaguchi et al., 1997; Hayashi et al., 1999). This effect has been attributed to the 5′ location of a locus control region (LCR).

The amino acid differences that distinguish red and green pigments are confined to exons 2–5, with the greatest effect on spectral tuning referable to differences in exon 5 (Figure 4B). Sequence differences in exon 2 can produce shifts of at most 1–2 nm, and sequence differences in exons 3 and 4 can produce shifts of up to 6–7 nm depending upon which of several dimorphic residues are present (Merbs and Nathans, 1992a, 1992b; Asenjo et al., 1994; Sanocki et al., 1997; Crognale et al., 1998, 1999; Sharpe et al., 1998). For each exon, sequences derived from the green pigment gene produce spectral shifts toward shorter wavelengths and sequences derived from the red pigment gene produce spectral shifts toward longer wavelengths. Because exon 5 plays the largest role in spectral tuning, the hybrid genes encode pigments with absorption spectra that cluster near that of the red pigment if they carry a red pigment exon 5 or near that of the green pigment if they carry a green pigment exon 5.

Asymmetries in the Arrangement of Red and Green Pigment Genes

The asymmetric location of sequences that affect spectral tuning, together with the placement of the red pigment gene at the 5′ edge of the array, has a number of interesting consequences for the types and frequencies of color vision anomalies found in the human population.

In the sections that follow, we examine these consequences by considering the color vision phenotypes that result from the actual arrangement of red and green pigment genes and from several hypothetical arrangements.

In the naturally occuring gene arrangement, shown in Figure 6A, one consequence of placing the red pigment gene at the 5′ edge of the first repeat unit is that *intergenic* recombination events that reduce the array to a single gene inevitably leave a red pigment gene, and *intragenic* recombination events that reduce the array to a single gene inevitably leave a 5′ red–3′ green hybrid gene which encodes a pigment that is green-like in its absorption properties. In those rare cases in which an intragenic recombination event occurs close to the 5′ end of the red and green pigment genes, the resulting 5′ red–3′ green hybrid gene will encode a de facto green pigment. The reciprocal product of the intergenic exchange indicated above is an array with an additional green pigment gene, and the reciprocal product of the intragenic exchange is an array with a 5′ green–3′ red hybrid gene (Figure 6A). The latter arrangement generally confers deuteranomalous trichromacy, because the absorption spectrum of the 5′ green–3′ red hybrid pigment is red-like but distinct from that of the accompanying red pigment. Dichromacy results if the 5′ green–3′ red hybrid pigment has the same absorption spectrum as the accompanying red pigment.

It is interesting to note that because of the high frequency and large variety of distinct exon 2–4 haplotypes (Winderickx et al., 1993; Sharpe et al., 1998), in particular the high frequency of the serine/alanine polymorphism at position 180 of the red pigment gene, an array that contains a red pigment gene followed by a 5′ green–3′ red hybrid gene that encodes a de facto red pigment, will frequently produce nonidentical red pigments. If the spectral sensitivities of the two red pigments differ by a few nanometers, they will support a low level of chromatic discrimination and the carrier will be an anomalous trichromat rather than a dichromat (Merbs and Nathans, 1992b; Neitz et al., 1996; Sanocki et al., 1997), an idea originally proposed by Alpern and colleagues to explain the color matching behavior of dichromats and anomalous trichromats (Alpern and Moeller, 1977; Alpern and Pugh, 1977; Alpern and Wake, 1977). An analogous argument applies to those arrays that carry a 5′ red–3′ green hybrid gene in place of the normal red pigment gene, together with one or more green pigment genes.

Dichromacy has classically been considered a reduction of normal vision in which one of the pigments is missing and the others are unchanged (Maxwell, 1860; König and Dieterici, 1886). In fact, over 50% of dichromats carry and presumably express more than one X-linked visual pigment gene (Nathans et al., 1986b; Deeb et al., 1992; Sharpe et al., 1998). Among dichromats who carry only a single X-linked gene, those who carry a red pigment gene represent a true reduction type as originally envisioned. However, those who carry a single 5′ red–3′ green hybrid gene actually comprise a heterogeneous group, with about half possessing a hybrid pigment that is clearly different from a normal green pigment (Sharpe et al., 1998; Crognale et al., 1999). The location of the red pigment gene at the 5′ end of the array precludes the production of a single-gene array that carries only a 5′ green–3′ red hybrid gene.

The mechanics of unequal homologous recombination dictate that for each array that gains a repeat unit, there is another array that loses a repeat unit. Since ~75% of human X chromosomes carry three or more visual pigment genes and only ~1% carry a single visual pigment gene, it is likely that over time the latter class has largely been eliminated. By inference, trichromacy must have conferred a selective advantage over dichromacy during much of human existence. A variation on this line of reasoning may also explain why such a large fraction of arrays in the present-day human gene pool have “extra” visual pigment genes. These larger arrays may have a small selective advantage over the basic arrangement of one red and one green pigment gene, because a greater distance between the first and last genes should decrease the probablility of producing a recombination product that carries only one gene.

Hypothetical Arrangements of Red and Green Pigment Genes: Paths Not Taken

The dominant role of exon 5 and the subsidiary roles of exons 2–4 in spectral tuning has significant consequences for the types and frequencies of color vision anomalies. This effect can be appreciated by considering two hypothetical arrangements of amino acids involved in spectral tuning (Figure 6). In the first arrangement, spectral tuning is dominated by sequence differences in exon 2 and exons 3–5 play a subsidiary role, and in the second, spectral tuning is determined equally by sequence differences in exons 2–5. Under either scenario, no difference in outcome is to be expected following intergenic recombination. However, the consequences of intragenic recombination differ appreciably from that produced by the actual arrangement.

Under the first scenario, those intragenic recombination events in which the crossover is 3′ of exon 2 would produce hybrid pigments with absorption spectra resembling that of the pigment contributing to the 5′ part of the hybrid (Figure 6, center). Together with the location of the red pigment gene at the edge of the first repeat unit, this arrangement would result in a red-like pigment occupying the first position in the array following most intragenic recombination events. Among intragenic recombination products with two or more genes, most would produce small rather than large alterations in the complement of visual pigments because they would be exchanging only exons that play a subsidiary role in spectral tuning. Individuals carrying these arrays would have nearly normal trichromatic vision. This scenario would therefore produce a lower frequency of color anomalies than is actually observed. Most obviously, there would be a lower frequency of anomalous trichromats and a lower frequency of dichromats who lack the red pigment (protanopes).

Under the second scenario, in which spectral tuning is determined equally by sequence differences in exons 2–5, hybrid pigments would differ to greater, lesser, or intermediate extents from the parental pigments depending on the exact point of crossing over, instead of clustering into red-like and green-like groups (Figure 6, right). Under this scenario, a large variety of hybrid

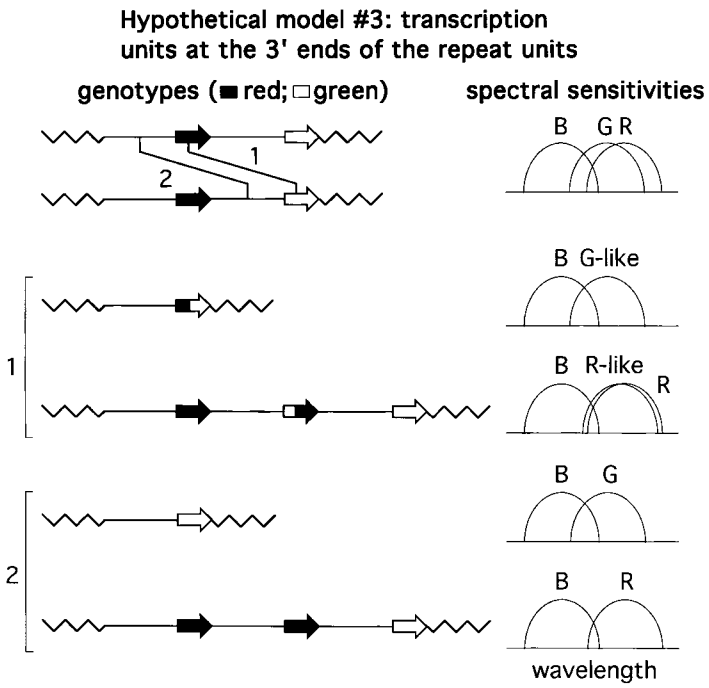

Figure 7. Schematic of Unequal Recombination Events and Their Phenotypic Consequences within a Hypothetical Human Red and Green Pigment Gene Array in which the Transcription Units Are Located at the 3′ Ends of the 39 kb Repeat Units

Intragenic (indicated by "1") and intergenic (indicated by "2") unequal homologous recombination events between wild-type arrays with one green pigment gene and one red pigment gene give rise to the two pairs of reciprocal products shown beneath the parental arrays. For males carrying the arrays illustrated at left, cone pigment spectral sensitivities are shown to the right. If expression is limited to the first two genes in the array, then any recombination product with three genes will produce either anomalous trichromacy or dichromacy.

pigments would produce a correspondingly large variety of anomalous trichromacies. Many combinations of two hybrid pigments or of a hybrid and a parental pigment would be predicted to support good chromatic discrimination and would therefore be expected to persist in the gene pool.

What would happen if the primordial duplication event had included 25 kb of DNA upstream of the ancestral visual pigment gene rather than 25 kb of DNA downstream, and as a result the visual pigment genes were located at the 3′ rather than at the 5′ end of the repeat unit (Figure 7)? (For this scenario, we assume that the LCR is located >25 kb 5′ of the nearest visual pigment gene and is therefore not duplicated.) In this case, intergenic or intragenic homologous recombination would produce single gene arrays carrying either a green pigment gene or a 5′ red–3′ green hybrid gene, respectively. Interestingly, the reciprocal product of the intergenic recombination event would carry red pigment genes in the first two positions in the array, rather than a red and a green pigment gene. If, as current evidence indicates, only the first two genes are expressed, then the phenotype produced by this array would be dichromacy. In all likelihood, selective pressure against dichromacy would eliminate most arrays with three or more genes, as described above for single-gene arrays. In the present-day human gene pool, unexpressed "extra" genes in those arrays with three or more visual pigment genes constitute a reservoir of sequence diversity that includes a variety of 5′ green–3′ red hybrid genes in ~5% of arrays (Drummond-Borg et al., 1989) and a deleterious cysteine-to-arginine substitution in ~0.5% of arrays (Winderickx et al., 1992c; Nathans et al., 1993). When these sequence variants are brought into one of the first two positions by unequal homologous exchanges, the array gives rise to variant color vision. If the ancestral gene duplication event had placed the visual pigment genes at the 3′ end rather than at the 5′ end of the repeat unit, we would see a very different distribution of variant visual pigment gene arrays in the present-day human population.

Cone Pigment Gene Expression

Psychophysical, microspectrophotometric, electrophysiological, and immunocytochemical studies indicate that in humans and in most other mammals each cone contains only a single type of visual pigment (Rushton, 1972; Dartnall et al., 1983; Schnapf et al., 1987; Curcio et al., 1991). A recent exception to this general pattern has been documented for a subset of cones in the retinas of rodents (Rohlich et al., 1994), but current evidence supports the "one cone–one visual pigment" rule in primates. How does a human cone make this all-or-none choice? For the choice between <500 nm and >500 nm pigments, the divergence of these two cone pigment families prior to the vertebrate radiation has provided more than enough time for the evolution of sophisticated mechanisms to control cell type–specific gene transcription. Intriguingly, there appears to be some similarity in the DNA sequences recognized by transcription factors in the two cone classes. Reporter transgenes driven by the mouse or human blue pigment promotor are expressed only in <500 nm cones, but similar transgenes containing the human red pigment gene promotor and LCR are expressed in both cone types (Wang et al., 1992; Chiu and Nathans, 1994a, 1994b; Shaaban et al., 1998).

The extremely recent duplication of red and green pigment genes presents an opportunity to understand what may be an unusually simple transcriptional control mechanism that determines which visual pigment gene is expressed in red and green cones. One clue to this mechanism comes from the identification of the LCR 5′ of the red and green pigment gene array (Nathans et

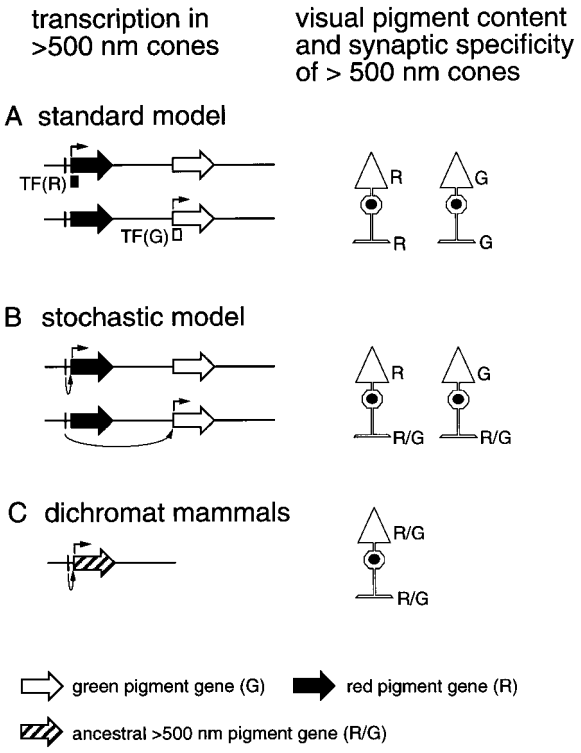

Figure 8. Models of Visual Pigment Gene Transcription and the Resulting Visual Pigment Content and Synaptic Specificity of Cones

Individual gene transcription units are represented by large rightward arrows as indicated in the key at the bottom. The LCR is shown as a thin vertical bar to the left of the visual pigment array, and hypothesized interactions between the LCR and individual visual pigment gene promotors are shown as curved arrows below the DNA. Transcriptional activation is represented by a rightward arrow above the indicated promotor region. The schematic diagrams of cone photoreceptors (right) show the outer segment (top), a cytoplasm and nucleus (center), and a pedicle/synaptic region (bottom). The identity of the visual pigment within the outer segment and the specificities, if any, of the synaptic region are indicated adjacent to those structures.

(A) The standard model in which red versus green visual pigment transcription is controlled by cell type–specific transcriptional regulatory proteins, illustrated arbitrarily as transcriptional activators, TF(R) and TF(G), present in red and greens cones, respectively. The function of the LCR in the standard model is not defined but could involve pairing with promotors and/or altering the chromatin structure in and around the visual pigment gene array. The standard model allows coordinate regulation of visual pigment type and synaptic specificity (right) and therefore red and green cones are shown with both different pigment content and synaptic specificity.

(B) The stochastic model in which a selective interaction between the LCR and one of the visual pigment gene promotors activates transcription exclusively from that gene. Red and green cones are distinguished only by pigment content, not by synaptic specificity.

(C) In dichromat mammals a single X-linked visual pigment gene is expressed in a single class of cone photoreceptors. The LCR is presumed to act on the promotor to activate transcription.

al., 1989). Humans who carry a deletion of sequences between 3.1 and 3.7 kb 5′ of the red pigment gene lack both red and green cone function despite the presence of an otherwise intact red and green pigment gene array. By analogy with the LCR adjacent to the β globin gene cluster, the visual pigment LCR is presumed to act in conjunction with individual promotors in the red and green pigment gene array to activate transcription. In support of this hypothesis, a reporter transgene driven by the human red pigment gene promotor is dependent on the presence of the LCR for its expression, and this expression is confined to cones (Wang et al., 1992). A comparison of sequences upstream of the >500 nm pigment genes in mice, humans, and cattle shows high sequence conservation only in the several hundred bases immediately 5′ of the start of transcription and in the LCR. This analysis shows no evidence for a duplication of the LCR in humans. Most likely, the LCR evolved as an enhancer to activate a single ancestral X-linked visual pigment gene.

As noted earlier, proximity to the LCR may explain why only the first two genes are expressed from the red and green pigment gene array (Winderickx et al., 1992a; Yamaguchi et al., 1997; Hayashi et al., 1999). (Rare exceptions to this pattern have been reported [Sjoberg et al., 1998].) A distance effect could plausibly arise if the LCR and its associated proteins (1) act to change the conformation or protein composition of adjacent chromatin to make the chromatin accessible for transcription, or (2) directly interact with an individual promotor to assemble active transcriptional complexes. In either case, the data indicate a steep distance dependence to the effect since transcriptional start sites ~3.5 kb and ~43 kb from the LCR are efficiently activated but a start site ~82 kb from the LCR is not. One practical consequence of this pattern of visual pigment gene expression is that predictions of color vision phenotype based on genotyping of individuals with three or more genes in the array require not only a catalog of the number and type of visual pigment genes but also a determination of their order within the array.

The possibility of a stable complex between the LCR and an individual promotor suggests a model for red versus green pigment gene expression in which the formation of this complex is the determining event in the choice of which visual pigment gene to express (Wang et al., 1992, 1999). Within this framework, two extreme versions of the model can be envisioned (Figure 8). At one extreme, the red versus green pigment choice might be orchestrated by transcription factors that are specific to either red or green cones, a scenario that will be referred to as the "standard model." At the other extreme, the red versus green pigment choice might be random, an alternative that will be referred to as the "stochastic model." In the standard model, the existence of cone type–specific transcription factors could also lead to the differential expression of other genes that might distinguish these cells, in particular genes involved in determining the specificity of synaptic connections. By contrast, the stochastic model assumes that red and green cones contain identical transcriptional regulatory proteins. In the simplest version of the stochastic model, the choice of red or green pigment gene expression is envisioned to exert no influence on the expression of other genes, implying that red and green cones possess no molecular differences other than the pigments they contain. Both the standard model and the stochastic model are compatible with the random or nearly random distribution of red and green cones within the mosaic of >500 nm cones as

measured by microspectrophotometry and high-resolution reflectometry (Mollon and Bowmaker, 1992; Roorda and Williams, 1999).

To distinguish between the two competing models, Wang and colleagues asked whether a mammal that normally possesses only a single X chromosome–linked visual pigment gene could support mutually exclusive expression of the human red and green pigment genes when these are integrated into its genome (Wang et al., 1999). The standard model predicts that only primates with trichromatic color vision have evolved the requisite transcriptional regulators to distinguish red and green pigment genes, whereas the stochastic model predicts that the red versus green choice is effected by transcriptional regulators common to all mammals and requires only the appropriate arrangement of *cis*-acting regulatory sequences. Transgenic mice were generated that carry a single copy of a minimal human X chromosome visual pigment gene array in which the red and green pigment gene transcription units were replaced, respectively, by alkaline phosphatase and β-galactosidase reporters. As determined by histochemical staining, the reporters are expressed exclusively in cone photoreceptor cells, and 63% of expressing cones were found to have alkaline phosphatase activity, 10% β-galactosidase activity, and 27% both activities. The three types of transgene-expressing cells appear to be randomly intermingled with one another. Thus, mutually exclusive expression of red and green pigment transgenes can be achieved in a large fraction of cones in a dichromat mammal, a result that strongly favors the stochastic model.

This observation also bears on the significance of X linkage of the visual pigment genes. As noted previously, the retinal mosaic produced by random X inactivation permits trichromatic color vision among heterozygous female primates in those species with a single polymorphic X-linked visual pigment gene. In primates with more than one X-linked visual pigment gene, a stochastic mechanism produces mutually exclusive expression only if, in each cone, a single visual pigment gene array is present or available for expression, a condition that is fulfilled by hemizygosity in males and X inactivation in females, respectively.

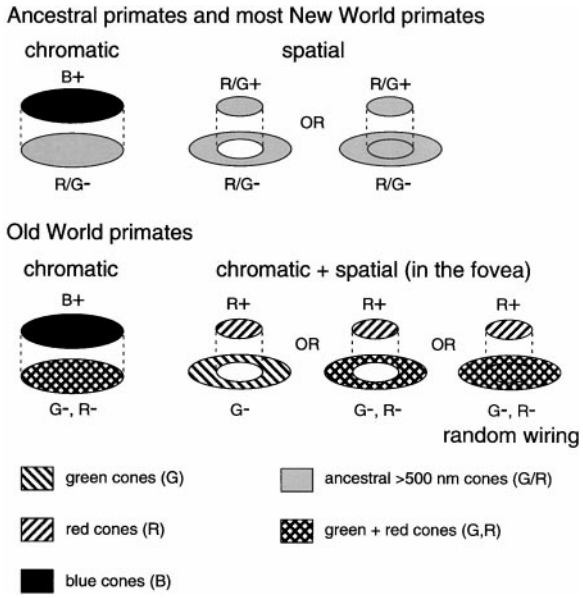

Figure 9. Spatial and Chromatic Structure of Retinal Ganglion Cell Receptive Fields in Ancestral Dichromat Primates and Present-Day New World and Old World Primates

The ancient B–Y subsystem (left) is present in all primates; the recent R–G subsystem (right) is shown as an extension of an achromatic center–surround receptive field structure to which only the >500 nm cones contribute. The ancient subsystem has both B–Y and Y–B receptive fields, but only the former is shown. For spatially opponent receptive fields, the figure shows only the subset with excitatory centers, and for Old World primates only those in which the center is driven by a single red cone. Two arrangements of spatial opponency are shown that differ depending upon whether the central input does or does not contribute to the inhibitory surround. For Old World primates, three possible arrangements of foveal receptive fields are shown in which the receptive field center is driven by a single cone. These have a chromatically uniform surround (left), a chromatically mixed but spatially distinct surround (center), or a chromatically mixed and spatially homogenous surround (right). In the simplest case, random wiring models for red and green cones predict the arrangement to the right, but under certain conditions could be compatible with the arrangement in the center. Patterns representing different cone types are indicated at the bottom.

Chromatic Processing in the Retina

The ancient dichotomy between <500 nm and >500 nm pigments and their respective cone types is reflected in a dedicated circuit within the inner retina that computes the B–Y signal (Figure 9; Dacey, 1996). A class of morphologically distinctive bistratified ganglion cells integrates chromatically antagonistic inputs from blue cone-specific ON bipolar cells and red/green cone-specific OFF bipolar cells (Dacey and Lee, 1994). The inputs to each of these ganglion cells are derived from many cones of each type and are not organized in a spatially opponent manner. These receptive field properties are in keeping with psychophysical experiments that demonstrate a minimal role for the B–Y signal in spatial vision (reviewed by Zrenner et al., 1990).

The circuitry responsible for the R–G signal has been a subject of investigation (and speculation) for decades, with different investigators reaching seemingly contradictory conclusions. Extracellular recordings from presumptive midget ganglion cells (DeMonasterio and Gouras, 1975) as well as from the parvocellular layers of the lateral geniculate nucleus, which appear to faithfully represent midget ganglion cell signals (De Valois, 1965; Wiesel and Hubel, 1966; Derrington et al., 1984), revealed center–surround receptive fields with both spatial and R–G chromatic opponency. The assumption for many years was that input to the center was drawn from one class of cones and input to an antagonistic annulus from the other class (Figure 9)—and this is how midget ganglion cell receptive fields are shown in virtually all textbooks. The simplest interpretation of this receptive field structure is that these cells do double duty by conveying both chromatic and spatial signals (Ingling and Martinez-Uriegas, 1983). However, the challenge inherent in deconvolving chromatic and spatial information has led some investigators to consider the possibility that the midget ganglion cells convey primarily or

exclusively spatial information and that there is a distinct and as yet undiscovered chromatic channel for the R–G system, a not implausible hypothesis given the large number of morphologically distinct ganglion cell types to which no function has yet been assigned (Rodieck, 1991; Dacey, 1996).

One difficulty with any model that invokes pure green or pure red cone inputs to any part of a ganglion cell receptive field is that interneurons have not yet been identified with the specificity required to carry this signal. As input to the inhibitory surround does not appear to involve direct contacts between bipolar and ganglion cells, any red versus green cone specificity in the surround would need to be reflected in the specificity of horizontal and/or amacrine cells (Rodieck, 1998). Recordings from goldfish horizontal cells beautifully revealed the first example of color-opponent responses over 40 years ago (Svaetichin and MacNichol, 1958), but primate horizontal cells are strikingly nonopponent and show no selectivity with respect to red and green cone inputs: H1 cells draw their input indiscriminantly from both red and green cones, and H2 cells draw their inputs from blue cones and to a lesser extent from red and green cones, again indiscriminantly (Dacey et al., 1996). Reconstructions from serial electron micrographs of macaque retina suggest a similar lack of red versus green specificity at the level of amacrine cells (Calkins and Sterling, 1996).

One potential solution to this conundrum is to suppose that midget ganglion cells draw their inputs *randomly* from the available pool of red and green cones within the designated regions of the receptive field (Figure 9; Paulus and Kroger-Paulus, 1983; Young and Marracco, 1989; Lennie et al., 1991). In the fovea, the receptive field center of each midget ganglion cell is driven by a single cone, whereas in the peripheral retina the receptive field centers are driven by multiple cones, and the number of cones that contribute increases with increasing eccentricity (Polyak, 1941; Rodieck, 1998). In the random wiring model, color opponency in the central retina would occur as a simple consequence of the "one cone–one ganglion cell" wiring diagram. For the peripheral retina, the random wiring model makes two related predictions regarding ganglion cell receptive fields. First, at intermediate retinal eccentricities, where a small number of cones contributes to the receptive field center, random fluctuations in the ratio of red to green cones in the receptive field center should lead to significant cell-to-cell variation in the wavelength at which the central input cancels that of the antagonistic surround when a spatially uniform stimulus is presented. Second, as the number of cones driving the receptive field center increases with increasing retinal eccentricity, the relative fluctuations in the ratio of red to green cones should decrease and the ganglion cells should become, on average, less sensitive to chromatic as opposed to spatial stimuli (Mullen and Kingdom, 1996). While these predictions are largely born out by much of the data in the literature (reviewed by Zrenner, 1983; Martin, 1998), some of the receptive field data are better fitted to a model of cone-specific surrounds (e.g., Reid and Shapley, 1992).

The random wiring model almost certainly describes the situation in those female New World primates who carry spectrally distinct X-linked visual pigments (Mollon et al., 1984). The excellent trichromatic color vision enjoyed by these heterozygous females presumably arises from color-opponent ganglion cell receptive fields that straddle the border between patches of differentially X-inactivated retina. Any hypothesis that invokes differential synaptic labeling of the X inactivation patches would need to also invoke linkage disequilibrium of visual pigment alleles and the genes encoding the labels—an unlikely scenario in an outbred population. The random wiring model is also likely to describe the state of affairs in the ancestral primate who first carried an array with duplicated red and green pigment genes. If, like present day primates, this ancestral primate possessed high-acuity foveal vision, then the transgenic mouse experiment described above suggests that he or she would have immediately acquired R–G chromatically opponent ganglion cell receptive fields. Moreover, if the stochastic model of red versus green pigment gene expression applies to present-day humans and other Old World primates, and if the choice of visual pigment is the sole difference between red and green cones, then the retina would initially develop with random wiring of red and green cones. In that case, the only way for the visual system to extract chromatic information from green versus red cone signals would be to use Hebbian mechanisms.

Acknowledgments

The author would like to thank Hui Sun, Clark Riley, and Philip Smallwood for assistance with Figures 1, 2, and 5 and Samir Deeb, John Mollon, Ed Pugh, and Ted Sharpe for helpful comments on the manuscript.

References

Alpern, M. (1982). Eye movements and strabismus. In The Senses, H.B. Barlow and J.D. Mollon, eds. (Cambridge: Cambridge University Press), pp. 201–211.

Alpern, M., and Moeller, J. (1977). The red and green cone visual pigments of deuteranomalous trichromacy. J. Physiol. (Lond.) *266*, 647–675.

Alpern, M., and Pugh, E.N. (1977). Variation in the action spectrum of erythrolabe among deuteranopes. J. Physiol. *266*, 613–646.

Alpern, M., and Wake, T. (1977). Cone pigments in human deutan colour vision defects. J. Physiol. *266*, 595–612.

Asenjo, A.B., Rim, J., and Oprian, D.D. (1994). Molecular determinants of human red/green color discrimination. Neuron *12*, 1131–1138.

Birge, R.R. (1990). Nature of the primary photochemical event in rhodopsin and bacteriorhodopsin. Biochem. Biophys. Acta *1016*, 293–327.

Boynton, R.M. (1979). Human Color Vision (New York: Holt, Rinehart, and Winston).

Calkins, D.J., and Sterling, P. (1996). Absence of spectrally specific lateral inputs to midget ganglion cells in primate retina. Nature *381*, 613–615.

Chiu, M.I., and Nathans, J. (1994a). Blue cones and cone bipolar cells share transcriptional specificity as determined by expression of human blue visual pigment-derived transgenes. J. Neurosci. *14*, 3426–3436.

Chiu, M.I., and Nathans, J. (1994b). A sequence upstream of the mouse blue visual pigment gene directs blue cone-specific transgene expression in mouse retinas. Vis. Neurosci. *11*, 773–780.

Crognale, M.A., Teller, D.Y., Motulsky, A.G., and Deeb, S.S. (1998).

Severity of color vision defects: electroretinographic (ERG), molecular and behavioral studies. Vision Res. *38*, 3377–3385.

Crognale, M.A., Teller, D.Y., Yamaguchi, T., Motulsky, A.G., and Deeb, S.S. (1999). Analysis of red/green color discrimination in subjects with a single X-linked visual pigment gene. Vision Res. *39*, 707–719.

Curcio, C.A., Allen, K.A., Sloan, K.R., Lerea, C.L., Hurley, J.B., Klock, I.B., and Milam, A.H. (1991). Distribution and morphology of human cone photoreceptors stained with anti-blue opsin. J. Comp. Neurol. *312*, 610–624.

Dacey, D.M. (1996). Circuitry of color coding in the primate retina. Proc. Natl. Acad. Sci. USA *93*, 582–588.

Dacey, D.M., and Lee, B.B. (1994). The 'blue-on' opponent pathway in primate retina originates from a distinct bistratified ganglion cell type. Nature *367*, 731–735.

Dacey, D.M., Lee, D.D., Stafford, D.K., Pokorny, J., and Smith, V.C. (1996). Horizontal cells in the primate retina: cone specificity without spectral opponency. Science *271*, 656–659.

Dartnall, H.J.A., Bowmaker, J.K., and Mollon, J.D. (1983). Human visual pigments: microspectrophotometric results from the eyes of seven persons. Proc. R. Soc. Lond. B Biol. Sci. *220*, 115–130.

Deeb, S.S., Lindsey, D.T., Sanocki, E., Winderickx, J., Teller, D.Y., and Motulsky, A.G. (1992). Genotye-phenotype relationships in human red/green color vision defects: molecular and psychophysical studies. Am. J. Hum. Genet. *51*, 687–700.

DeMonasterio, F.M., and Gouras, P. (1975). Functional properties of ganglion cells of the rhesus monkey retina. J. Physiol. *251*, 167–195.

Derrington, A.M., Krauskoff, J., and Lennie, P. (1984). Chromatic mechanisms in lateral geniculate nucleus of macaque. J. Physiol. *357*, 241–265.

De Valois, R.L. (1965). Analysis and coding of color vision in the primate visual system. Cold Spring Harbor Symp. Quant. Biol. *30*, 567–579.

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

Gautier-Hion, A., Duplantier, J.-M., Quris, R., Feer, F., Sourd, C., Decoux, J.-P., Dubost, G., Emmons, L., Erard, C., Hecketsweiler, P., et al. (1985). Fruit characteristics as a basis of fruit choice and seed dispersal in a tropical forest vertebrate community. Oecologia *65*, 324–337.

Hayashi, T., Motulsky, A.G., and Deeb, S.S. (1999). Position of a 'green-red' hybrid gene in the visual pigment array determines colour-vision phenotype. Nat. Genet. *22*, 90–93.

Hoff, W.D., Jung, K.H., and Spudich, J.L. (1997). Molecular mechanism of photosignaling by archaeal sensory rhodopsins. Annu. Rev. Biophys. Biomol. Struct. *26*, 223–258.

Huddart, J. (1777). An account of persons who could not distinguish colours. Philos. Trans. R. Soc. Lond. *67*, 260–265.

Hunt, D.M., Dulai, K.S., Cowing, J.A., Julliot, C., Mollon, J.D., Bowmaker, J.K., Li, W.-H., and Hewett-Emmett, D. (1998). Molecular evolution of trichromacy in primates. Vision Res. *38*, 3299–3306.

Ibbotson, R.E., Hunt, D.M., Bowmaker, J.K., and Mollon, J.D. (1992). Sequence divergence and copy number of the middle- and long-wave photopigment genes in Old World monkeys. Proc. R. Soc. Lond. B Biol. Sci. *247*, 145–154.

Ingling, C.R., and Martinez-Uriegas, E. (1983). The relationship between spectral sensitivity and spatial sensitivity for the primate R–G channel. Vision Res. *23*, 1495–1500.

Jacobs, G.H. (1981). Comparative Color Vision (New York: Academic Press).

Jacobs, G.H. (1993). The distribution and nature of colour vision among the mammals. Biol. Rev. *68*, 413–471.

Jacobs, G.H., Neitz, M., Deegan, J.F., and Neitz, J. (1996). Trichromatic colour vision in New World monkeys. Nature *382*, 156–158.

Kainz, P.M., Neitz, J., and Neitz, M. (1998). Recent evolution of uniform trichromacy in a New World monkey. Vision Res. *38*, 3315–3320.

König, A., and Dieterici, C. (1886). Die Grundempfindungen und ihre Intensitäts-Vertheilung im Spektrum. Sitz Akad Wiss (Berlin) *1886*, 805–829.

Kropf, A., and Hubbard, R. (1958). The mechanism of bleaching rhodopsin. Ann. NY Acad. Sci. *74*, 266–280.

Lennie, P., Haake, P.W., and Williams, D.R. (1991). The design of chromatically opponent receptive fields. In Computational Models of Visual Processing, M.S. Landy and J.A. Movshon, eds. (Cambridge, MA: MIT Press), pp. 71–82.

Lin, S.W., Kochendoerfer, G.C., Carroll, K.S., Wang, D., Mathies, R.A., and Sakmar, T.P. (1998). Mechanisms of spectral tuning in blue cone visual pigments. J. Biol. Chem. *273*, 24583–24591.

Lythgoe, J.N. (1979). The Ecology of Vision (Oxford: Clarendon Press).

Macke, J.P., and Nathans, J. (1997). Individual variation in the size of the human red and green pigment gene array. Invest. Ophthalmol. Vis. Sci. *38*, 1040–1043.

Martin, P.R. (1998). Colour processing in the primate retina: recent progress. J. Physiol. *513*, 631–638.

Martin, P.R., and Grunert, U. (1999). Analysis of the short-wavelength sensitive ("blue") cone mosaic in the primate retina: comparison of New World and Old World monkeys. J. Comp. Neurol. *406*, 1–14.

Mathies, R., and Stryer, L. (1976). Retinal has a highly dipolar vertically excited singlet state: implications for vision. Proc. Natl. Acad. Sci. USA *73*, 2169–2173.

Maxwell, J.C. (1860). On the theory of compound colours and the relations of the colours of the spectrum. Philos. Trans. R. Soc. Lond. B Biol. Sci. *150*, 57–84.

Merbs, S.L., and Nathans, J. (1992a). Absorption spectra of human cone pigments. Nature *356*, 433–435.

Merbs, S.L., and Nathans, J. (1992b). Absorption spectra of the hybrid pigments responsible for anomalous color vision. Science *258*, 464–466.

Merbs, S.L., and Nathans, J. (1993). Role of hydroxyl-bearing amino acids in differentially tuning the absorption spectra of the human red and green cone pigments. Photochem. Photobiol. *58*, 706–710.

Mollon, J.D. (1989). "Tho she kneel'd in that place where they grew..."—the uses and origins of primate colour vision. J. Exp. Biol. *146*, 21–38.

Mollon, J.D., and Bowmaker, J.K. (1992). The spatial arrangement of cones in the primate fovea. Nature *360*, 677–679.

Mollon, J.D., Bowmaker, J.K., and Jacobs, G.H. (1984). Variations of colour vision in a New World primate can be explained by polymorphism of retinal photopigments. Proc. R. Soc. Lond. B Biol. Sci *222*, 373–399.

Mullen, K.T., and Kingdom, F.A.A. (1996). Losses in peripheral color sensitivity predicted from "hit or miss" post-receptoral cone connections. Vision Res. *36*, 1995–2000.

Nathans, J. (1990). Determinants of visual pigment absorbance: identification of the retinylidene Schiff's base counterion in bovine rhodopsin. Biochemistry *29*, 9746–9752.

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., Piantanida, 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.

Nathans, J., Davenport, C.M., Maumenee, I.H., Lewis, R.A., Hejtmancik, J.F., Litt, M., Lovrien, E., Weleber, R., Bachynski, B., Zwas, F., et al. (1989). Molecular genetics of human blue cone monochromacy. Science *245*, 831–838.

Nathans, J., Maumenee, I.H., Zrenner, E., Sadowski, B., Sharpe, L.T., Lewis, R.A., Hansen, E., Rosenberg, T., Schwartz, M., Heckenlively, J.R., et al. (1993). Genetic heterogeneity among blue cone monochromats. Am. J. Hum. Genet. *53*, 987–1000.

Neitz, J., and Jacobs, G.H. (1986). Polymorphism of the long-wavelength cone in normal human colour vision. Nature *323*, 623–625.

Neitz, J., and Jacobs, G.H. (1990). Polymorphism in normal color vision and its mechanism. Vision Res. *30*, 621–636.

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

Neitz, M., Neitz, J., and Jacobs, G.H. (1995). Genetic basis of photopigment variations in human dichromats. Vision Res. *35*, 2095–2103.

Neitz, J., Neitz, M., and Kainz, P. (1996). Visual pigment gene structure and the severity of color vision defects. Science *274*, 801–804.

Osorio, D., and Vorobyev, M. (1996). Colour vision as an adaptation to frugivory in primates. Proc. R. Soc. Lond. B Biol. Sci. *263*, 593–599.

Ottolenghi, M., and Sheves, M. (1989). Synthetic retinals as probes for the binding site and photoreactions in rhodopsin. J. Membr. Biol. *112*, 193–212.

Paulus, W., and Kroger-Paulus, A.A. (1983). A new concept of retinal colour coding. Vision Res. *23*, 529–540.

Polyak, S.L. (1941). The Retina (Chicago: University of Chicago Press).

Polyak, S.L. (1957). The Vertebrate Visual System (Chicago: University of Chicago Press).

Post, R.H. (1962). Population differences in red and green color vision deficiency: a review, and a query on selection relaxation. Eugen. Quart. *9*, 131–146.

Reese, B.E., Harvey, A.R., and Tan, S.S. (1995). Radial and tangential dispersion patterns in the mouse retina are class specific. Proc. Natl. Acad. Sci. USA *92*, 2494–2498.

Regan, B.C., Julliot, C., Simmen, B., Viénot, F., Charles-Dominique, P., and Mollon, J.D. (1998). Frugivory and colour vision in *Aloutta seniculus*, a trichromatic platyrrhine monkey. Vision Res. *38*, 3321–3327.

Reid, R.C., and Shapley, R.M. (1992). Spatial structure of cone inputs to receptive fields in primate lateral geniculate nucleus. Nature *356*, 716–718.

Rodieck, R.W. (1991). Which cells code for colour? In From Pigments to Perception: Advances in Understanding Visual Processes, A. Valberg and B.B. Lee, eds. (London: Plenum Press), pp. 83–93.

Rodieck, R.W. (1998). The First Steps of Seeing (Sunderland, MA: Sinauer).

Rohlich, P., van Veen, T., and Szel, A. (1994). Two different visual pigments in one retinal cone cell. Neuron *13*, 1159–1166.

Roorda, A., and Williams, D.R. (1999). The arrangement of the three cone classes in the living human eye. Nature *397*, 520–522.

Rushton, W.A.H. (1972). Pigments and signals in colour vision. J. Physiol. *220*, 1–31.

Sakmar, T.P., Franke, R.R., and Khorana, H.G. (1989). Glutamic acid-113 serves as the retinylidene Schiff base counterion in bovine rhodopsin. Proc. Natl. Acad. Sci. USA *86*, 8309–8313.

Sanocki, E., Lindsey, D.T., Winderickx, J., Teller, D.Y., Deeb, S.S., and Motulsky, A.G. (1993). Serine/alanine amino acid polymorphism of the L and M cone pigments: effects on Rayleigh matches among deuteranopes, protanopes and color normal observers. Vision Res. *33*, 2139–2152.

Sanocki, E., Shevell, S.K., and Winderickx, J. (1994). Serine/alanine amino acid polymorphism of the L-cone photopigment assessed by Dual Rayleigh-type color matches. Vision Res. *34*, 377–382.

Sanocki, E., Teller, D.Y., and Deeb, S.S. (1997). Rayleigh match ranges of red/green color-deficient observers: psychophysical and molecular studies. Vision Res. *37*, 1897–1907.

Schnapf, J.L., Kraft, T.W., and Baylor, D.A. (1987). Spectral sensitivity of human cone photoreceptors. Nature *325*, 439–441.

Shaaban, S.A., Crognale, M.A., Calderone, J.B., Huang, J., Jacobs, G.H., Deeb, S.S. (1998). Transgenic mice expressing a functional human photopigment. Invest. Ophthalmol. Vis. Sci. *39*, 1036–1043.

Sharpe, L.T., Stockman, A., Jagle, H., Knau, H., Klausen, G., Reitner, A., and Nathans, J. (1998). Red, green, and red-green hybrid pigments in the human retina: correlations between deduced protein sequences and psychophysically-measured spectral sensitivities. J. Neurosci. *18*, 10053–10069.

Shyue, S.-K., Hewett-Emmett, D., Sperling, H.G., Hunt, D.M., Bowmaker, J.K., Mollon, J.D., and Li, W.-H. (1995). Adaptive evolution of color vision genes in higher primates. Science *269*, 1265–1267.

Sjoberg, S.A., Neitz, M., Balding, S.D., and Neitz, J. (1998). L-cone pigment genes expressed in normal color vision. Vision Res. *38*, 3213–3219.

Stewart, J.M., and Cole, L. (1989). What do colour vision defectives say about everyday tasks? J. Optom. Vis. Sci. *66*, 288–295.

Stockman, A., MacLeod, D.I.A., and Johnson, N.E. (1993). Spectral sensitivities of the human cones. J. Opt. Soc. Am. *10*, 2491–2521.

Suzuki, H. (1967). Electronic Absorption Spectra and Geometry of Organic Molecules (New York: Academic Press).

Svaetichin, G., and MacNichol, E.F. (1958). Retinal mechanisms for chromatic and achromatic vision. Ann. NY Acad. Sci. *74*, 385–404.

Vollrath, D., Nathans, J., and Davis, R.W. (1988). Tandem array of human visual pigment genes at Xq28. Science *240*, 1669–1671.

Wang, Y., Macke, J.P., Merbs, S.L., Klaunberg, B., Bennett, J., Zack, D., Gearhart, J., and Nathans, J. (1992). A locus control region adjacent to the human red and green pigment genes. Neuron *9*, 429–440.

Wang, Y., Smallwood, P.M., Cowan, M., Blesh, D., Lawler, A., and Nathans, J. (1999). Mutually exclusive expression of human red and green visual pigment-reporter transgenes occurs at high frequency in murine cone photoreceptors. Proc. Natl. Acad. Sci. USA *96*, 5251–5256.

Wiesel, T.N., and Hubel, D.H. (1966). Spatial and chromatic interactions in the lateral geniculate body of the rhesus monkey. J. Neurophysiol. *29*, 1115–1156.

Williams, A.J., Hunt, D.M., Bowmaker, J.K., and Mollon, J.D. (1992). The polymorphic photopigments of the marmoset: spectral tuning and genetic basis. EMBO J. *11*, 2039–2045.

Winderickx, J., Battisti, L., Motulsky, A., and Deeb, S.S. (1992a). Selective expression of human X-chromosome-linked green opsin genes. Proc. Natl. Acad. Sci. USA *89*, 9710–9714.

Winderickx, J., Lindsey, D.T., Sanocki, E., Teller, D.Y., Motulsky, A.R., and Deeb, S.S. (1992b). Polymorphism in red photopigment underlies variation in colour matching. Nature *356*, 431–433.

Winderickx, J., Sanocki, E., Lindsey, D.T., Teller, D.Y., Motulsky, A.R., and Deeb, S.S. (1992c). Defective colour vision associated with a missense mutation in the human green visual pigment gene. Nat. Genet. *1*, 251–256.

Winderickx, J., Battisti, L., Hibiya, Y., Motulsky, A.R., and Deeb, S.S. (1993). Haplotype diversity in the human red and green opsin genes: evidence for frequent sequence exchange in exon 3. Hum. Mol. Genet. *2*, 1413–1421.

Wolf, S., Sharpe, L.T., Schmidt, H.-J.A., Knau, H., Weitz, S., Kioschis, P., Poustka, A., Zrenner, E., Lichter, P., and Wissinger, B. (1999). Direct visual resolution of gene copy number in the human photopigment gene array. Invest. Ophthalmol. Vis. Sci. *40*, 1585–1589.

Wyszecki, G., and Stiles, W.S. (1982). Color Science (New York: Wiley).

Yamaguchi, T., Motulsky, A.G., and Deeb, S.S. (1997). Visual pigment gene structure and expression in human retinae. Hum. Mol. Genet. *6*, 981–990.

Yokoyama, S. (1997). Molecular genetic basis of adaptive selection: examples from color vision in vertebrates. Annu. Rev. Genet. *31*, 315–336.

Young, R.A., and Marracco, R.T. (1989). Predictions about chromatic receptive fields assuming random cone connections. J. Theor. Biol. *141*, 23–40.

Zhukovsky, E.A., and Oprian, D.D. (1989). Effect of carboxylic acid side chains on the absorption maximum of visual pigments. Science *246*, 928–930.

Zrenner, E. (1983). Neurophysiological Aspects of Color Vision in Primates: Comparative Studies on Simian Retinal Ganglion Cells and the Human Visual System (Berlin: Springer-Verlag).

Zrenner, E., Abramov, I., Akita, M., Cowey, A., Livingstone, M., and Valberg, A. (1990). Color perception: retina to cortex. In Visual Perception: the Neurophysiological Foundations, L. Spillmann and J.S. Werner, eds. (New York: Academic Press), pp. 163–204.