Testing Hypotheses About Visual Pigments Underlying Deutan Color Vision

Joseph Carroll, Maureen Neitz, Jay Neitz*

Department of Ophthalmology and Department of Cell Biology, Neurobiology and Anatomy, Medical College of Wisconsin, 8701 Watertown Plank Road, Milwaukee, WI 53226-0509

Received 15 October 1999; accepted 9 November 1999

Abstract: A classical theory held that deuteranomaly was caused by the replacement of the normal M cone pigment by an anomalous M pigment, which was spectrally red-shifted. Two assumptions associated with that theory were: (1) that there was a single, fixed L pigment, shared by deuteranomals and normals; (2) and that everyone had two genes for cone pigments on the X-chromosome. Those two assumptions are now known to be invalid and, with regard to red-green color vision, an alternate theory, introduced by Alpern et al. *in the 1970s, has enjoyed renewed attention. It holds that photopigment spectra are inherently variable and form "clusters." Those with normal color vision are supposed to have drawn one pigment from each cluster, but deuteranomaly arises when what would normally be the L and M pigments are both drawn from the L cluster. Both the Alpern theory and the classical one continue to play important roles in current thinking. However, we present evidence here from ERG flicker photometry that observed similarities and differences between the distributions of spectral sensitivity functions of deuteranopes and deuteranomalous trichromats are intermediate between what might be predicted by Alpern's theory and what would be predicted by the classical theory.* © 2000 John Wiley & Sons, Inc. Col Res Appl, 26, S106–S111, 2001

Key words: cone photoreceptors; photopigments; deuteranopia; blue cone monochromat; genetics; color vision; color vision deficiency

* Correspondence to: Jay Neitz, Department of Ophthalmology and Department of Cell Biology, Neurobiology and Anatomy, Medical College of Wisconsin, 8701 Watertown Plank Road, Milwaukee, WI 53226-0509 (e-mail: jneitz@mcw.edu)

Contract grant sponsors: NIH; Contract grant numbers: EY09303, EY09620, EY01921; Research to Prevent Blindness

INTRODUCTION

During the 1970s and early 1980s, a generally accepted view was that X-linked color vision defects were mediated by two genes, each with multiple alleles.[1,2] One gene encoded the L-cone photopigment, and the other encoded the M-cone photopigment. It was supposed that the predominant allele at the L cone pigment gene locus encoded the stereotyped L pigment of normal color vision. Likewise, there was a presumed single normal allele of the M pigment. Thus, all color-normals shared the same L- and M-cone photopigments. According to the theory, color vision defects were the result of the inheritance of a mutant allele. In deuteranomaly, a mutant allele of the M-cone pigment gene replaced the normal M. Within each class of color vision deficiency, individuals vary in the severity of color vision loss. Two broad subcategories of deuteranomalous observers have been differentiated according to their behavior on the Rayleigh color match.[3] Individuals with less severe deficiencies, simple deuteranomalous trichromats, require a higher proportion of the green primary in the Rayleigh match and usually accept a matching range that is somewhat wider than normal. Individuals in the group of more severe defects, extreme deuteranomalous trichromats, have wide matching ranges that include both the normal and the simple deuteranomalous match. To account for different degrees of severity of color vision defect, a series of mutant alleles of the M-cone photopigment were proposed. These formed a series of "anomalous" pigments, which were proposed to be progressively spectrally shifted toward the normal L photopigment. The anomalous pigments encoded by the mutant genes were consequently different from the pigments of both normal observers and dichromats.

In the mid 1970s, at the time when the theory outlined above enjoyed popularity, Alpern and colleagues performed a series of experiments, which they interpreted as suggesting that the cone photopigments varied in λ_{max} within each

class of dichromat.[4,5] They reasoned that the normal L and M pigments must vary in spectral peak as well. This led them to propose an alternative to the favored view about the mechanisms underlying color vision defects. Their idea was that there are three groups, or "clusters," of visual pigments that are common to all normal eyes as well as those of people with red-green color vision defects. Normal color vision was proposed to result when the pigment in the L cones was drawn from somewhere (anywhere) in the L cluster, the M-cone pigment was drawn from somewhere within the M cluster, and the S-cone pigment was drawn from within the S cluster. Color vision defects were proposed to occur when the pigments in both the L and M cones were drawn from within the same cluster. Thus, *deuteranomaly* occurred when the pigments in both the L and M cones were drawn from within the L cluster. However, the L pigment in the M cones had to be slightly different in λ_{max} from the L pigment in the L cones to confer anomalous trichromacy rather than dichromacy.

The results of Alpern and colleagues were very influential in stimulating research into the issues of normal variation in cone pigment spectra; on the other hand, the theoretical approach to color vision defects that evolved from their empirical work was not widely accepted at the time. However, over the last dozen years it has become clear that two of the most important assumptions associated with the classical theory are invalid: (1) the assumption that there is a single, fixed L pigment, shared by deuteranomals and normals,[6-8] and (2) the assumption that everyone has two genes for cone pigments on the X-chromosome.[9] As a consequence of findings that the basic premises of the classical theory are untrue, there has been renewed interest in the ideas of Alpern and his collaborators. At the same time, work on characterization of the pigments underlying anomalous trichromacy has continued.[10-12] Nonetheless, no new, completely comprehensive theory of the cone pigments underlying anomalous trichromacy, their origins, and their relationships to the pigments of normal and dichromatic observers has been formulated. Thus, the classical theoretical approaches, as well as those of Alpern and colleagues, continue to be important in thinking about the pigments of color-vision defects.

As a step toward ultimately developing a comprehensive theory, the goal of the present experiments was to evaluate the relative merits of Alpern's theory and the classical theory by examining the spectral sensitivity functions obtained from ERG flicker photometry of deuteranomalous trichromats and deuteranopes. The two theories make distinctly different predictions about the distribution of spectral sensitivity function in the two groups.

METHODS

Subjects

The subjects used in the study were males with deutan color vision deficiencies ($n = 25$). They were screened for participation based upon their performance on diagnostic color vision tests (Ishihara, AO-HRR, Dvorine, D-15). Subjects were then classified as deuteranomalous or deuteranopic based on their Rayleigh matches using a Nagle anomaloscope. The subjects' ages ranged from 16–52 years, with a mean of 32 years.

Procedure

The flicker-photometric ERG has been previously described as a method to accurately measure spectral sensitivities.[13-15] The subjects' right pupil was dilated (Tropicamide 0.5%) and the ERGs were recorded using fiber from the DTL Plus™ electrode as an active corneal electrode. Two channels of a 3-channel Maxwellian view optical system were used to produce the stimuli. An achromatic reference light was alternated with a monochromatic test light at a rate of 31.25 Hz; the two lights illuminated a 70° portion of the retina, centered on the fovea. A period of darkness was interposed between each presentation of test and reference light. The wavelength of the test light was controlled with a Varispec™ liquid-crystal electronically tunable filter (half-band pass = 7 nm at 550 nm). The ERG signal resulting from each stimulus train was filtered, and the responses from the test and reference light were subtracted from one another electronically. Spectral sensitivities were recorded by measuring the null point, where the respective ERG signals cancelled one another. The null point was found by adjusting a neutral density wedge (which adjusted the intensity of the test light), while the reference intensity remained fixed. For all subjects, the null point was determined at 10 nm increments over a range of 480–680 nm. Four additional recordings were made at 5-nm increments on either side of the approximate peak of 560 nm. Spectral sensitivities were determined by the average of two complete runs, with the final spectral sensitivity values reported as quantal intensities. Intensities measured at the cornea were corrected for absorption by the lens using the age dependent lens correction of Xu, Pokorny, and Smith.[16]

Classification of Spectral Sensitivity Functions

The spectral sensitivity function of each observer was best fit to a newly formulated wavelength-shiftable visual-pigment template function. The template expression was designed to fit well near the peak, display the correct asymptotic behavior at long wavelengths, and accurately depict wavelengths significantly shorter than the peak.[15] The mathematical expression used for the template function is not given here, but it can be found at http://www.mcw.edu/cellbio/colorvision. In the curve fitting, the theoretical optical density (O.D.) of the template was fixed at 0.35 and the λ_{max} value (to the nearest hundredth of a nanometer) that produced the least mean squared error was recorded. In a preliminary analysis of the deuteranope data, we fit the spectral sensitivity curves allowing both the λ_{max} and the O.D. of the template curve to vary. On average, the O.D. values that best fit the deuteranope data were relatively high,

and a correspondingly high fixed value (0.35) was chosen. A comparison with some earlier measurements[8,17] suggests that our procedure may overestimate the O.D. of the peripheral retina covered by our stimulus; however, our O.D. value seems consistent with some recent measurements, which indicated higher O.D.s for the cones.[18] In any case, the template curve with a fixed O.D. of 0.35 provided a reasonable mathematical description of the spectral sensitivity functions measured for the subjects tested here (Fig. 2). By fixing O.D., it was possible to characterize the observed individual differences with a single continuous variable (λ_{max}).

RESULTS

All the X-encoded photopigments in the retinas of deuteranopes and deuteranomalous trichromats can be classified as L pigments. In the deuteranopes, the best fitting template curve is a direct assessment of the spectral sensitivity of their single L photopigment. The deuteranomalous observers have two (or more) spectral types of L pigment. The resulting spectral sensitivity function for the deuteranomalous observers is derived from contributions from both cone types. Nevertheless, we best-fit each deuteranomalous sensitivity function to a single photopigment template under the assumption that the single λ_{max} value that results represents approximately the weighted average of the values for the underlying photopigments in the deuteranomalous eye.[†]

Using a single λ_{max} value provides a convenient metric for comparing the results from different subjects. The values for the deuteranopes ($n = 14$) had a range of 8.1 nm and an average of 557.2 nm. When the deuteranomalous curves ($n = 11$) were fit with a single pigment template, the average λ_{max} was 554.9 nm, with a range of 6.1 nm. Distributions for the two groups are shown in Fig. 1, bottom panels. The means of the two groups were significantly different ($p = 0.02$).

DISCUSSION

The interpretation of the results from the deuteranomalous and deuteranopic subjects critically depends on being able to parse out and eliminate measurement variability that is a consequence of experimental error and preretinal absorption and thus isolate individual variations in the absorption spectra of the visual pigments. To evaluate instrumental error, we recently tested several deuteranopes repeatedly on different days.[15] Figure 2(A) shows the average spectral sensitivity for a deuteranope who was tested previously on 5 trials over a 2-month test period.[15] The average λ_{max} measured for this subject was 552.55 nm. The average absolute deviation of the measurements from individual days compared to the average was 0.15 nm. Similar reliability measurements on other deuteranopes, when examined using the same procedure as the deuteranope in Fig. 2(A), give an overall average absolute deviation of 0.27 nm for single determinations compared to the average of repeated mea-

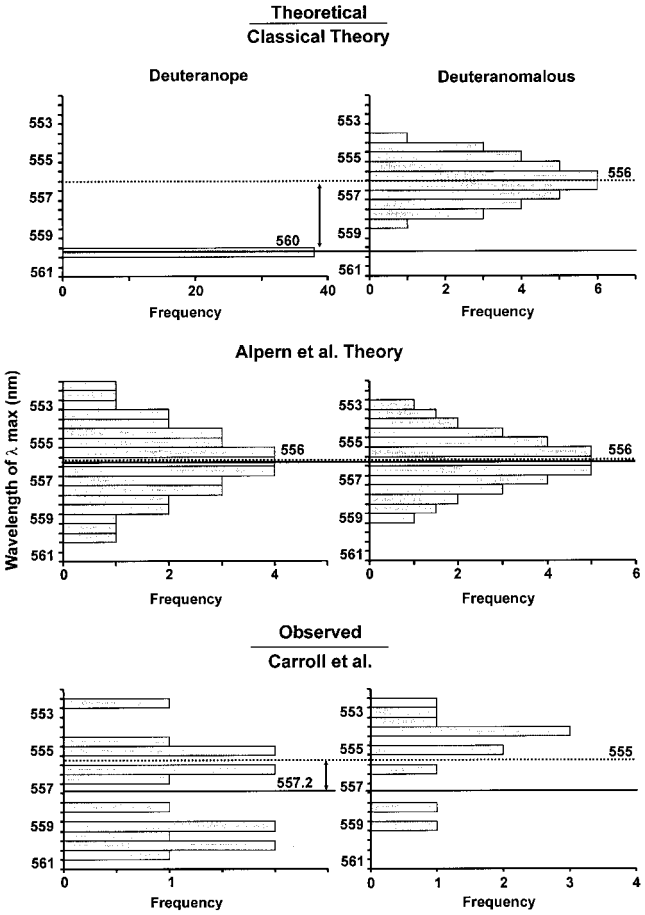

FIG. 1. Distributions of λ_{max} values for (left column) deuteranopes and (right column) deuteranomalous trichromats: (top four panels) theoretical and (bottom panels) observed. Top-most panels: The classical theory predicted no variation in L pigments, hence no variation in the λ_{max} of deuteranopes. However, all deuteranomalous observers have an added contribution from an anomalous M pigment with a spectral sensitivity shorter than the normal L. Thus, deuteranomalous spectral sensitivity curves should always peak at shorter wavelengths than the deuteranopes. Middle panels: Alpern's theory predicts that there would be no difference in the position of deuteranope compared to deuteranomalous spectral sensitivity functions, because both groups draw from the same "cluster" of L pigments. Bottom panels: Our observations do not match either theory. There is a large range of λ_{max} values for the deuteranopes, but the average spectral sensitivity of the deuteranomalous subjects is shifted significantly shorter.

[†] The meaning of the single value should not be over-interpreted. It is true that if the two pigments in a deuteranomalous subject contribute equally in linear combination to the spectral sensitivity function, the single λ_{max} value would provide a reasonably close approximation of the average of the two λ_{max} values; for example, two theoretical pigments with peaks at 554 and 560 nm yield a curve that peaks at 557.2 nm (compared to the average of the two λ_{max} values, 557 nm). However, the assumption of equal contributions of the two pigments is almost certainly not valid. There is evidence that protanomals express the first gene in the X-chromosome pigment gene array in a larger number of cones[19] in the same way that most normal trichromats have more L than M cones. Thus, the single λ_{max} value confers little information about the sensitivities of the two underlying pigments.

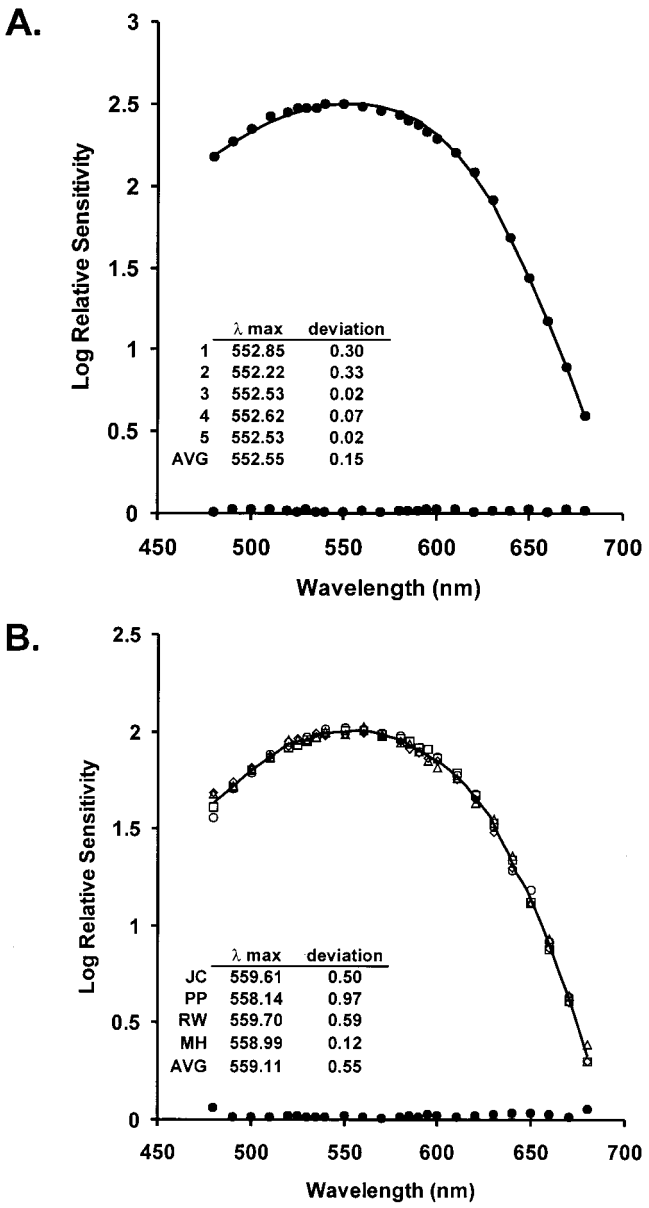

FIG. 2. Reliability of ERG spectral sensitivity functions. (A) Spectral sensitivity function of a single deuteranope. Filled circles represent the average of five independent measurements of spectral sensitivity, and the solid line is the best fitting template curve. The filled circles that plot near the X-axis show the standard deviation for each sensitivity value. The λ_{max} values determined for the 5 individual measurements are shown in the inset, as is the absolute deviation of each measurement from the mean of the five. (B) Spectral sensitivity functions for four single gene deuteranopes (open squares, circles, triangles, and diamonds). The solid line represents the average sensitivity function of these subjects with identical L pigment sequences. Standard deviations are plotted as filled circles. Each subject was measured once, and his spectral sensitivity function was fit to the template curve. The resulting λ_{max} values are given in the inset along with the average, and the absolute deviation of each individual measurement from the average.

sures. These results indicate that a single measurement on a deuteranope yields a λ_{max} value that is within about 0.3 nm from the value that would be obtained from the same subject, if he were run repeatedly and an average was taken. These experiments also provide a way to access the adequacy of our photopigment template curve as a description of the ERG spectral sensitivity data. The average data for the 5 runs of one deuteranope are shown in Fig. 2(A), compared to the best fitting template curve. The curve appears to provide an acceptable way to describe the data using a single variable, the λ_{max} value specifying the best fitting template curve. In the example of Fig. 2(A), the average absolute error per data point = 0.014 log units; this is typical of how closely the experimental results fit the theoretical curve.

Interpretation of the results also depends on our ability to minimize contributions by factors that influence spectral sensitivity but are unrelated to the extinction spectrum of the photopigment. Examples of such variables included individual differences in preretinal absorption and in effective photopigment optical density (O.D.) as might arise from differences in outer segment length or directionality. To assess the contribution of these variables, we measured individual deuteranopes who all had a single X-chromosome photopigment gene with the identical sequence[15]; thus, the encoded L-cone photopigment molecule in each subject was identical. Previous results for 4 different deuteranopes with the same sequence[15] are shown in Fig. 2(B). These indicate that a single recording on one subject can represent the spectral sensitivity function of his pigment type (as specified by the gene) with an average absolute deviation of 0.55 nm. This needs to be investigated further in additional subjects, but the results so far indicate that this type of inter-subject variability contributes an average of about 0.25 nm of error over and above instrumental error. Procedural aspects of the ERG measurements that help reduce the effect of these human variables include: (1) the large visual field (70°) reduces optical density variations that might come from individual differences in foveal outer segment length or directionality; (2) the largely extrafoveal field minimizes contributions from macular pigment variations; (3) use of an age-corrected lens density function lowers contributions from lens variation due to aging.[16] These reliability measurements indicate that the measured differences shown in the bottom panels of Fig. 1 principally reflect real differences in the spectral sensitivities of the underlying visual pigments.

The results can be compared with the predictions that might be made from the classical theory of color-vision defects and predictions from the alternative hypothesis proposed by Alpern and colleagues. Classically, there was assumed to be a single stereotyped L pigment common to all deuteranopes. Deuteranomals would have shared this L pigment and had, in addition, a red-shifted (anomalous) M pigment. This theory would predict a pattern of results similar to what is shown in the top panels of Fig. 1. All deuteranopes are shown as being identical. All deuteranomalous observers have an added contribution from a pigment with a spectral sensitivity shorter that the normal L. Thus, their spectral sensitivity curves should always peak at shorter wavelengths than the deuteranopes. The exact dis-

tribution of the deuteranomalous spectral sensitivities would depend on the relative magnitudes of the contributions of the two cone types to spectral sensitivities, the number of different anomalous pigments and their distribution in the population. For simplicity we have used a normal distribution of spectral sensitivity in the illustration (Fig. 1; top right) but this is arbitrary. An important prediction of the theory is that the average spectral sensitivity of deuteranomalous observers would be shifted to significantly shorter wavelengths than deuteranopes.

Alpern proposed the idea of "clusters" of photopigments. Normals, deuteranopes, and deuteranomals all drew pigments from the same L "cluster." A possible pattern of spectral sensitivities that could be predicted from this is shown in the middle panels of Fig. 1. Deuteranopes either draw one pigment from the L cluster or they draw two that are so nearly identical as not to be sufficiently separated to provide red-green color vision. Deuteranomalous observers draw two pigments from the same cluster, but there is a sufficient spectral separation to provide the basis for anomalous trichromacy. For simplicity of illustration, we have arbitrarily drawn the spectral sensitivities as being normally distributed. An important prediction is that, since both groups are drawing from the same distribution, the average spectral sensitivities of the deuteranopes and deuteranomals should not be spectrally shifted relative to one another.

Neither prediction exactly matches our observations (bottom panels, Fig. 1). There is a wide distribution of deuteranope spectral sensitivities (range = 8.1 nm) consistent with Alpern's observations and the more recent measurements of Sharpe *et al*..[20] The deuteranope distribution shows hints of being noncontinuous rather than gaussian, presumably reflecting the fact that the spectral position of the pigments is controlled by discrete changes of a limited number of amino acids.[21-23] However, more similar to what would have been predicted by the classical theory, the average spectral sensitivity of the deuteranomalous subjects is significantly shifted toward the short wavelengths compared to the deuteranopes. This is reminiscent of the findings of Jordan and Mollon, who reported that female carriers of deuteranomaly had significantly displaced color matches compared to normals in a ratio-matching task.[24]

The results imply that deuteranopes and deuteranomalous subjects draw from overlapping but not completely identical distributions of L cone pigments. There is probably a straightforward genetic explanation. In normal individuals, an L photopigment gene is always in the most upstream position in the X-chromosome array (Fig. 3).[25,26] The most common cause of deuteranopia appears to be the deletion of all the X-chromosome photopigment genes except the first gene in array.[27] Thus, normals and deuteranopes share in common the population of L pigments encoded by the first gene in the array. Deuteranomals presumably also have an L gene in the first position, but, as shown in Fig. 3, they ALL must also have a second gene to encode an additional L- type pigment; that gene is positioned somewhere downstream in the array (a fraction of normals have this arrangement, too[26]). The short-shifted spectral sensitivities of the

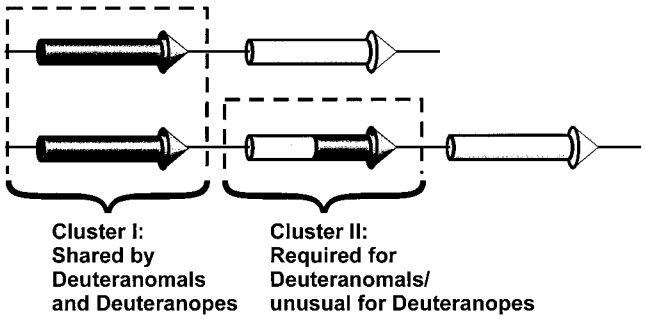

FIG. 3. Genetic model showing photopigment genes in tandem array on the X-chromosome: (black arrows) L pigment genes; (white arrows) M pigment genes. An L photopigment gene is always positioned in the most upstream position. Deuteranopes and deuteranomalous observers share the population of L genes in this first position. Deuteranomalous observers all must have an additional gene to encode a second L-type pigment; this gene must be in a downstream position. A fraction of normal observers have a downstream L gene as well. The short wavelength shifted sensitivity functions of deuteranomalous observers are proposed to be the result of a shorter average spectral sensitivity for the downstream L genes, compared to the L genes in the first position.

deuteranomals suggest that the downstream L genes have, on average, a shorter spectral peak than the L genes in the first position. It appears that over the period of human evolution, the L and M photopigment gene sequences have become very intermixed through the process of repeated recombination. As a result, a wide variety of L genes that are essentially "chimeras," with jumbled sequences from the primordial distinct L and M genes exist to provide spectral variation in the L pigments. Apparently, however, the sequences are not completely homogenized. Presumably the primordial array had an L gene first and an M gene second. Apparently, some of this is preserved in modern arrays such that L genes in downstream positions tend to contain more sequences from M genes, particularly from the 5' end of the M gene and, thus, the downstream L genes have shorter spectral peaks on average.

ACKNOWLEDGMENTS

This work was supported by NIH grants EY09303, EY09620 and EY01921, and Research to Prevent Blindness. We thank C. Bialozynski and P. M. Summerfelt for their technical expertise, and P. M. Summerfelt for her invaluable help in preparing the manuscript and figures.

1. Piantanida TP. Polymorphism of human color vision. Am J Optom Phys Opt 1976;53:647–657.
2. Pokorny J, Smith VC. Evaluation of a single pigment shift model of anomalous trichromacy. J Opt Soc Am A 1977;67:1196–1209.
3. Franceschetti A. Die bedeutung der einstellungstreite am anomaloskop fur die diagnose der einzelnen typen der farbensinnstorunger: nebst bermerkungen uber ihren vererbungsmodus. Med Wochenschr 1928; 52:1273–1278 . [As cited in Pokorny J, Smith V, Verriest G. Congenital color defects. In: Pokorny J, Smith V, Verriest G, Pinckers A,

editors. Congenital and acquired color vision defects. New York: Grune Stratton; 1979. p 183–241].
4. Alpern M, Moeller J. The red and green cone visual pigments of deuteranomalous trichromacy. J Physiol 1977;266:647–675.
5. Alpern M. Colorblind color vision. Trends Neurosci 1981;4:131–135.
6. Neitz J, Jacobs GH. Polymorphism of the long-wavelength cone in normal human color vision. Nature 1986;323:623–625.
7. Alpern M, Pugh EN. Variation in the action spectrum of erythrolabe among deuteranopes. J Physiol 1977;266:613–646.
8. Alpern M, Wake T. Cone pigments in human deutan colour defects. J Physiol 1977;266:595–612.
9. Nathans J, Thomas D, Hogness DS. Molecular genetics of human color vision: the genes encoding blue, green, and red pigments. Science 1986;232:193–202.
10. Neitz J, Neitz M, Kainz PM. Visual pigment gene structure and the severity of human color vision defects. Science 1996;274:801–804.
11. Shevell SK, He JC, Kainz PM, Neitz J, Neitz M. Relating color discrimination to photopigment genes in deutan observers. Vision Res 1998;38:3371–3376.
12. DeMarco P, Pokorny J, Smith VC. Full spectrum cone sensitivity functions for X-chromosome-linked anomalous trichromats. J Opt Soc Am A 1992;9:1465–1476.
13. Neitz J, Jacobs GH. Electroretinogram measurements of cone spectral sensitivity in dichromatic monkeys. J Opt Soc Am A 1984;1:1175–1180.
14. Jacobs GH, Neitz J, Krogh K. Electroretinogram flicker photometry and its applications. J Opt Soc Am A 1996;13:641–648.
15. Carroll J, McMahon C, Neitz M, Neitz J. Flicker-photometric ERG estimates of L:M cone photoreceptor ratio in men with photopigment spectra derived from genetics. J Opt Soc Am A 2000;17:499–509.
16. Xu J, Pokorny J, Smith VC. Optical density of the human lens. J Opt Soc Am A 1997;14:953–960.
17. Smith VC, Pokorny J. Psychophysical estimates of optical density in human cones. Vision Res 1973;13:1199–1202.
18. Burns S, Marcos S, McLellan J, Elsner A. Reflectometric measurement of cone directionality and photopigment optical density. Invest Ophth Vis Sci (Supplement) 1999;40:S365.
19. Balding SD, Sjoberg SA, Neitz J, Neitz M. Pigment gene expression in protan color vision defects. Vision Res 1998;38:3359–3364.
20. Sharpe LT, Stockman A, Jagle H, Knau H, Klausen G, Reitner A, Nathans J. Red, green, and red-green hybrid pigments in the human retina: correlations between deduced protein sequences and psychophysically measured spectral sensitivities. J Neurosci 1998;18:10053–10069.
21. Asenjo AB, Rim J, Oprian DD. Molecular determinants of human red/green color discrimination. Neuron 1994;12:1131–1138.
22. Neitz M, Neitz J, Jacobs GH. Spectral tuning of pigments underlying red-green color vision. Science 1991;252:971–974.
23. Neitz M, Neitz J. Molecular genetics and the biological basis of color vision. In: Backhaus W, Kleigl R, Werner JS, editors. Color vision — Perspectives from different disciplines. Berlin, New York: Gruyter; 1998. p 101–119.
24. Jordan G, Mollon JD. A study of women heterozygous for colour deficiencies. Vision Res 1993;33:1495–1508.
25. Kainz PM, Neitz M, Neitz J. Molecular genetic detection of female carriers of protan defects. Vision Res 1998;38:3365–3369.
26. Sjoberg SA, Neitz M, Balding SD, Neitz J. L-cone pigment genes expressed in normal colour vision. Vision Res 1998;38:3213–3219.
27. Nathans J, Piantanida TP, Eddy RL, Shows TB, Hogness DS. Molecular genetics of inherited variation in human color vision. Science 1986;232:203–210.