TESTS AND TECHNIQUES

A New Mass Screening Test for Color-Vision Deficiencies in Children

Maureen Neitz,* Jay Neitz

Department of Ophthalmology and Department of Cellular Biology, Neurobiology and Anatomy, Medical College of Wisconsin, 925 North 87th Street, Milwaukee, WI 53226-4812

Received 5 November 1999; accepted 8 February 2000

Abstract: Five thousand one hundred and twenty-nine Wisconsin children, ages 4–12 years, were tested for color-vision deficiencies using a newly devised, precisely calibrated paper-and-pencil test. The disposable 1-page test consisted of 1 demonstration and 8 test panels. Thousands of copies of the test were produced, and they were distributed and administered in classrooms by teachers. Children wrote directly on the test and were allowed to trace over the symbols with a pencil or crayon, if they had difficulty. Performance on the paper-and-pencil color vision test was compared with that on conventional tests of color vision including Ishihara's tests, the American Optical–Hardy Rand and Rittler (AO-HRR) plate test, and the APT-5 Color Vision Tester. Older children were also tested on the Nagel Anomaloscope. All the children who were classified as having a color vision deficiency by the paper-and-pencil test also failed one or more of the conventional tests. Likewise, among children who passed the paper-and-pencil test, none were classified as having a color-vision defect from the results on the conventional tests. In the sample of all males, 7.5% were classified as having a color-vision deficiency, which is consistent with what has been observed previously in large population studies. Children who were classified as having color vision deficiencies were examined further using a new minimalist genetic test that was shown to be accurate and reliable. Genetic material derived from buccal swabs was used to determine the type of deficiency, protan vs. deutan, and to provide added information about severity. Among the subjects for whom type could be determined, 27% were protans, consistent with large population studies in which approximately 25% of red-green deficiencies have been found to be of the protan type. Classification of the severity of the deficiencies determined from the paper-and-pencil test plus minimal genetics were in good agreement with classification based on a battery of conventional tests. In conclusion, we found the methods used here to be rapid, efficient, and reliable for testing color vision in children.

Key words: color vision; color vision deficiency; genetics; genetic test; deuteranopia; protanopia

INTRODUCTION

Congenital red-green color-vision deficiencies are common, affecting about 1 in 13 males, and 1 in 230 females. These occur with a broad range of severity, from nearly normal to very poor color discrimination. In early education, color-coding is increasingly used as a tool of communication and for teaching basic curriculum such as reading and math. Due to the prevalence of color-vision defects, as the use of color in education increases so does the need for color-vision testing in preschool and school-age children.

Ishihara's color-vision tests have become a standard by

* Correspondence to: Maureen Neitz, Department of Ophthalmology and Department of Cellular Biology, Neurobiology and Anatomy, Medical College of Wisconsin, 925 North 87th Street, Milwaukee, WI 53226-4812 (e-mail: mneitz@mcw.edu)

Contract grant sponsors: RPB; David and Ruth S. Coleman Foundation; Ronald McDonald Charities; Knights Templar Eye Foundation; Aaberg Retina Foundation; NEI; Contract grant numbers: EY09303, EY01931

which all other screening tests can be judged. In institutional settings such as schools and clinics, few if any copies of Ishihara's color tests are usually available. People with normal color vision often make some errors on Ishihara's tests, and a trained practitioner must determine if the pattern of errors constitute a color-vision defect. The numerical designs are difficult for young children, making the test inappropriate for children less than 8 years old. Because of these limitations, the test is not an efficient screening tool for classroom-sized groups of early- and preschool children. There does not appear to be an accepted alternative test available for such use.

Once a child is identified as having a color-vision deficiency, it is often useful to characterize the defect in terms of type and severity. Tests required to accurately define the type and severity, such as the anomaloscope, are too sophisticated for many young children to perform. However, insight has been gained into the structures of X-linked photopigment genes underlying normal and deficient red-green color vision, laying the ground-work for the development of a genetic analysis that could be used in conjunction with a simple screening test to accurately determine type and severity of abnormal color vision.

The long- (L) and middle-wavelength (M) sensitive cone pigments lie in a head-to-tail tandem array on the X-chromosome.[1,2] There is variation in the number of both L and M pigment genes per X-chromosome array.[1-4] Spectral tuning studies have demonstrated that two amino acid substitutions encoded by exon 5 of the L and M genes specify the majority of the spectral difference between L and M pigments.[5-7] Other amino acid substitutions produce relatively small spectral shifts ($\leq$ 7 nm), and can be thought of as producing spectral subtypes of L and spectral subtypes of M pigments.[5,6,8-13] Thus, the X-encoded pigments can be thought of as falling into two classes—L and M—depending on the amino acids encoded by exon 5 of the genes. The L-class has tyrosine at 277 and threonine at 285 with their peak spectral sensitivities falling in a range of about 550–563 nm. The M-class has phenylalanine and alanine at positions 277 and 285, respectively, and their peak sensitivities fall in a range of about 528–537 nm.

There is overlap in the gross structure of X-linked visual pigment gene arrays underlying normal and deficient color vision; however, within the group of males with color-vision deficiencies, there are usually distinct differences between protan and deutan gene arrays.[2,13-17] Deutan color-vision defects are characterized by the absence of M cone contribution to color vision. Accordingly, arrays containing only L genes, but lacking M genes, are associated with a fraction of deutan defects. Arrays with multiple L genes and one or more M genes are also found in individuals with deutan color defects.[1-4,15,18,19] Protan color-vision defects are characterized by the absence of M cone contribution to color vision and, accordingly, arrays containing only M genes (lacking L genes) underlie protan defects. Only rarely have individuals with a protan defect been found to have arrays containing both L and M genes.[2,20]

Here, we explored the feasibility of a 1-page disposable paper-and-pencil test for color-vision screening *en masse* in a classroom setting. We further investigated the use of the screening test results in conjunction with a newly developed simple genetic test to determine the type and severity of color-vision deficiency. A detailed report of the complete study, including all data and a comprehensive discussion of the reliability of the color vision tests will be given elsewhere. As an interim report here, we present the results of a combination of color-vision behavior and genetics to determine the type and severity of congenital color-vision defects among school-age, male children.

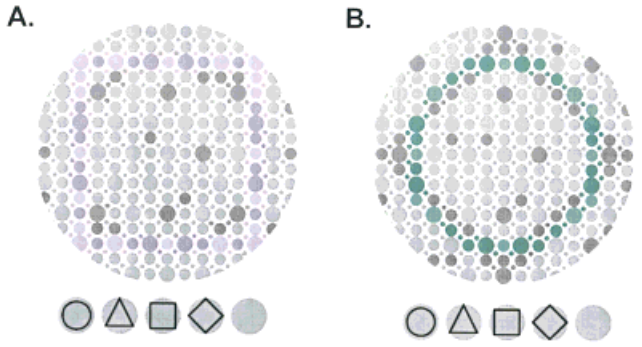


FIG. 1. Sample test plates from the Neitz paper-and-pencil test. These plates are shown as examples to illustrate how the stimuli are designed; they are not for color vision testing. (A) Example of a blue-yellow plate. A person with normal color vision reads a lavender-colored outline of a square shape. An indistinct outline of a circle formed by darker dots can also be discernable to the person with normal color vision. However, the instruction is to choose the symbol formed by the colored dots, and those with normal color vision do not choose a circle as the correct response. (B) Example of a red-green plate. With the instruction to find the colored shape, a person with normal color vision reads the plate as a circle. People with moderate to severe red-green deficiencies read the symbol as a diamond, which is made up of the darker dots.

METHODS

Paper and Pencil Test

Design. A Hewlett–Packard Color LaserJet 5 printer was used to produce thousands of copies of a precisely calibrated 1-page test. Deutan, protan, and tritan confusions were included. The test is designed to rapidly screen and differentiate red-green and blue-yellow color-vision defects and to provide limited information for estimating the degree of color-vision defect. There are 8 test plates and one demonstration plate printed on a single page designed for one-time use. In each plate, colored dots of three different sizes and two different saturations form the outline of a geometrical shape. The shapes, presented as multiple choices, are a circle, a square, a triangle, and a diamond. The shape outlines are presented against a pattern of gray dots of three sizes. A few random gray dots on each plate are darker than the majority and provide a small amount of luminance noise in the background. However, in addition,

on each plate, a set of darker dots forms an imprecise outline of an alternate shape (Fig. 1). Thus, the plate design is a hybrid between the vanishing type (in which the symbol vanishes for a defective observer) and the transformation type (in which normal and defective observers identify different figures). Any failure to identify the correct colored outline including a wrong choice of shape or the response that no shape is seen was scored as an error indicating a color-vision deficiency.

Calibration. Tests were calibrated using a Colortron II, 32-band color sensor. The instrument is factory calibrated, but the manufacturer supplies limited information about the accuracy of the calibration. To evaluate the calibration of the instrument, we used the Colortron II to measure the spectral transmission of a set of narrowband interference filters (380–700 nm at 10-nm intervals; 10-nm half-band pass). There was good correspondence between the manufacturer's specified value of the spectral transmission peak for the filters and the values measured using the Colortron II. The absolute difference between the specified value and the peak as measured by the Colortron was usually less than 1 nm and never exceeded 3 nm. Thus, the accuracy of the Colortron II was considered to be appropriate for calibration of the tests. As an additional check of the Colortron II as an appropriate calibration tool for this purpose, we used it to measure the chromaticities of the colors in the AO-HRR pseudo-isochromatic plates (Fig. 2). The original versions of these plates have been proven to be effective in the classification of color-vision deficiencies. Thus, it is expected that the chromaticities of the shapes and backgrounds as measured by the Colortron II should fall along target dichromatic confusion lines.

The Neitz test was designed to be used in natural daylight or under a variety of fluorescent lighting conditions that are encountered in institutional settings. The colors of the dots that made up the figures in the plates were chosen to fall near a confusion line on the CIE color diagram when the illuminant is either natural daylight, from a variety of fluorescent lights, or from mixtures of natural daylight and fluorescent lights. The Colortron II is supplied with software that calculates the chromaticity coordinates under different lighting conditions (illustrated in Fig. 2).

In the printing process, the calibration was checked and (if necessary) adjusted after every 50 copies were produced. Theoretically, because the colored dots were adjusted back into correspondence with the dichromatic confusion lines after every 50 copies, the largest deviations in chromaticity would be expected to occur within a print run of 50 copies. To evaluate the inter-test color variability, colors from 5 tests were measured, the 10^{th}, 20^{th}, 30^{th}, 40^{th}, and 50^{th} printed test from a single print run. Chromaticity variation from test to test was small (Fig. 2).

Administration. Teachers tested all children in a classroom simultaneously, usually between 20–30 students per class. Teachers were instructed to read verbatim a set of directions to the children. Children wrote directly on the test and were allowed to trace over the symbols with a pencil or crayon, if they had difficulty. All tests contained the same set of stimuli, but the stimuli were rearranged on different tests. Children were informed that their test would be different from their neighbor's test to discourage copying.

Subjects. Subjects were children ages 4–12 years in elementary schools in Wisconsin. In all, 5129 children were tested with the paper-and-pencil test. This study adhered to the Declaration of Helsinki. An effort was made to test every child in each of the participating schools. The only reasons for exclusion from the study were parental request or absence from school on all days of testing.

Children who made one or more errors on the paper-and-pencil test were tested a second time. The second test was administered in a separate room, to groups of 5–15 children. Children who made one or more errors on the second test were tested using Ishihara's tests (Concise Edition), the 1955 American Optical–Hardy Rand and Rittler (AO-HRR) plate test, and the APT-5 Color Vision Tester (LKC Technologies). A Nagel Anomaloscope was used to obtain Rayleigh matches from older children who made errors.

Two children from each classroom who made no errors on the test were selected as controls, and tested a second time with the Neitz test along with the children who had made errors the first time. The 145 children in the control group were also tested with the same battery of conventional color-vision tests as the potentially color-deficient children. To protect the anonymity of the children who were color deficient, teachers were informed that we were retesting normal as well as color-deficient children.

Newly Developed Genetic Test

Tissue Collection and DNA Extraction. Buccal swabs were collected from children who made errors on screening tests and whose parents agreed to have genetic testing done. Each child rinsed his/her mouth with water twice before swabs were collected. Two swabs were collected from each child, using sterile, individually wrapped, cytology brushes. Each brush was immediately swirled 10 times in a 1.5 mL microcentrifuge tube containing 200 μL of 10 mM Tris-HCl, pH 8.0; 0.1 mM EDTA pH 8.0; 1% SDS, and 4 mg/mL proteinase K, then discarded. The remainder of the DNA extraction protocol was as previously described.[21]

Determination of the Ratio of L/M Genes and Ratio of First/Downstream Genes. A newly developed real-time polymerase chain reaction (PCR) method was used to estimate the ratio of L:M cone pigment genes and their numbers. Real-time detection of PCR products in a closed tube at each cycle of the amplification process remedies limitations in the reproducibility and accuracy of quantitative PCR. The procedure is unprecedented in ease and efficiency, eliminating the post-PCR steps of DNA processing and analysis inherent to PCR-based quantitative methods used earlier. Light from an Argon laser is routed to 96 fibre-optic channels, each to a separate DNA sample well. An ABI Prism 7700 Sequence Detector measures fluorescence of labeled DNA probes, which are activated as specific sequences are amplified, providing direct quantitative information about the DNA in real-time, at each cycle in the

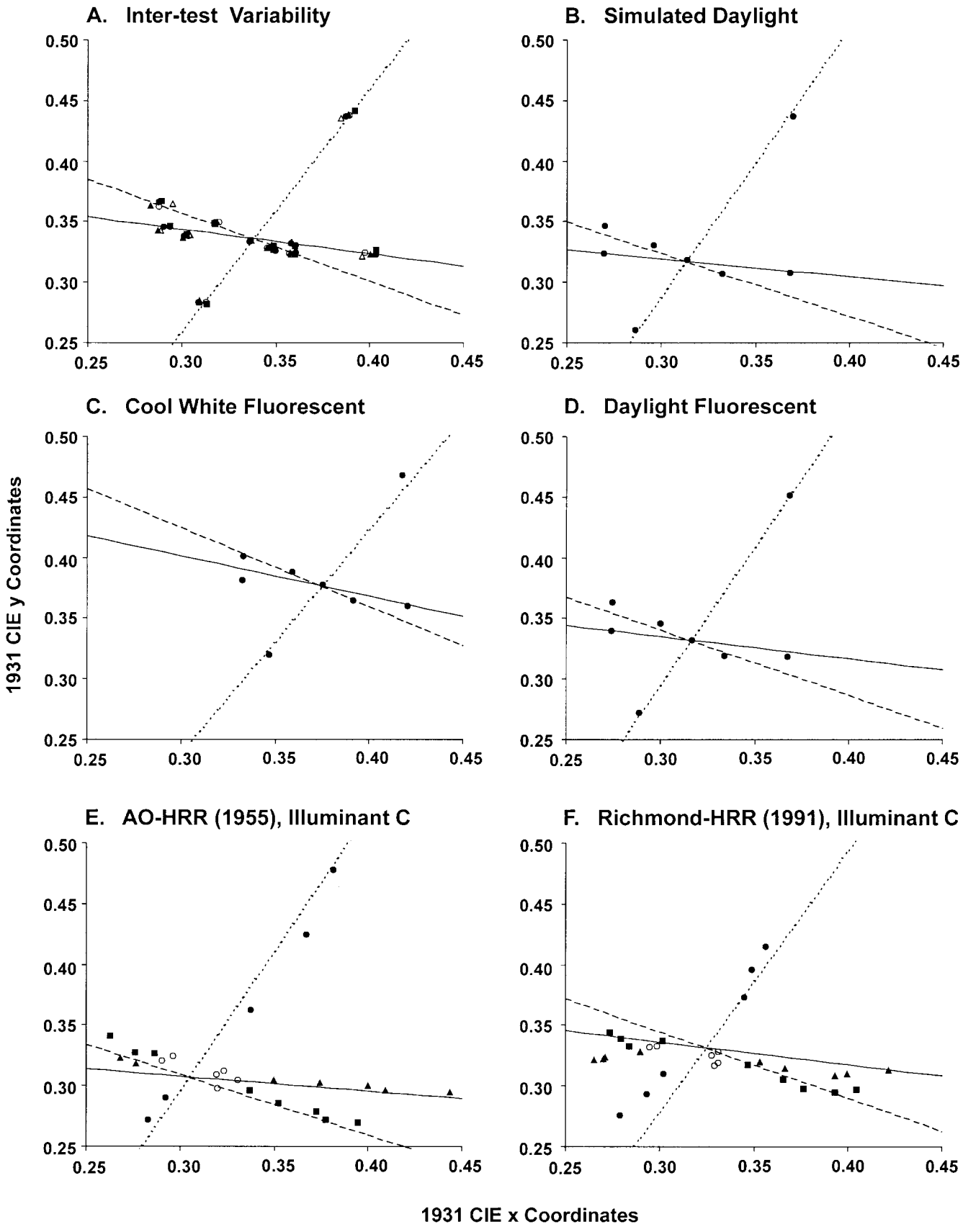

A. Inter-test Variability
B. Simulated Daylight
C. Cool White Fluorescent
D. Daylight Fluorescent
E. AO-HRR (1955), Illuminant C
F. Richmond-HRR (1991), Illuminant C
1931 CIE y Coordinates
1931 CIE x Coordinates
0.25
0.30
0.35
0.40
0.45
0.50

FIG. 2.

reaction. The assay used Taqman chemistry (Perkin–Elmer) to determine the L/M gene ratio in genomic DNA from each subject, as well as the ratio of first/downstream genes. For both assays, the TaqMan kit was used as recommended by the manufacturer (Perkin–Elmer). These recommendations included using dUTP in place of dTTP and using UNG enzyme.

For the assay to determine the L/M gene ratio, the forward primer (5'TTTGCTGCTGCCAACCCT) corresponded to sequences within exon 5, and the reverse primer (5'TTGCTTACCTGCCGGTTCATAA) spanned the junction between exon 5 and intron 6. Both primers hybridized to sequences shared by all L and M genes. Two TaqMan probes were used in each assay — one was specific for L genes (5' CCTGCCGGCCTACTTTGCCAAA) and was tagged at the 5' end with the fluorescent dye, 6-FAM, the other was specific for M genes (5'CCTGCCGGCCTTCTTTGCCAAA) and was labeled at the 5' end with the fluorescent dye, JOE. Both probes were tagged with TAMARA at the 3' end, which acts as a fluorescent quencher dye in the reaction. The single nucleotide difference between the L and M gene probes corresponded to the second nucleotide of codon 309 in the L and M genes, which specifies an amino acid difference that was shown empirically to distinguish L from M genes.[22] That is, when codons 277 and 285 specify tyrosine and threonine, respectively, 309 encodes tyrosine. Conversely, when codons 277 and 285 specify phenylalanine and alanine, 309 specifies phenylalanine. Each quantitative PCR assay was done in a final volume of 25 μL with final concentrations of 150 nM L-probe, 250 nM M-probe, 900 nM of each primer, and 5 mM $MgCl_2$. Samples were incubated first at 50°C for 2 min, then at 95°C for 10 min. The samples were then cycled 3 times at 97°C for 30 s and 67°C for 1 min, followed by 37 cycles of 95°C for 15 s and 67°C for 1 min. These reactions were done in an ABI 7700, and data were collected using the SDS Version 1.6 software.

In the assay to estimate the ratio of first/downstream genes, the forward primer (5'CCAGGCCCAATTAAGAGATCA) corresponded to sequences just upstream of exon 1, which were shared by all the X-linked visual pigment genes. The reverse primer (5'TTTGGAGGCTCCACTGCTG) corresponded to sequences within exon 1 that were also shared by L and M genes. Two probes were used. One was specific for the sequence 5' to exon 1 in the first gene in the array; the other was specific for downstream genes. The first gene-specific probe (5'ACCCTCAGGTGATGCGCCAGGG) was tagged at the 5' end with the fluorescent label, JOE. The downstream-gene specific probe (5'ACCCTCAGGTGACGCACCAGGG) was tagged at the 5' end with the fluorescent label, 6-FAM. The two nucleotide differences between these probes was shown empirically to distinguish first from downstream genes by direct DNA sequence analysis of this region of genomic DNA in a large sample of subjects. The PCRs were done in a final volume of 25 μL, and final concentrations of 300 nM forward primer, 900 nM reverse primer, 150 nM first-gene probe; 250 nM downstream-gene probe, and 5 mM $MgCl_2$. Reactions were first incubated at 50°C for 2 min followed by 10 min at 95°C. They were then cycled 3 times at 97°C for 30 s and 63°C for 1 min. This was followed by 37 cycles of 95°C for 15 s and 63°C for 1 min. Finally, reactions were incubated at 50°C for 2 min.

Standard Curves. Standards were generated by mixing cloned DNAs of known nucleotide sequence in specific ratios (Fig. 3). The reproducibility and reliability of the real-time quantitative PCR assays were determined by repeating all measurements on the children's samples at least four times and by measuring samples from a small number of subjects repeatedly dozens of times over a period of several months.

RESULTS

In all, 5129 children (2637 males and 2492 females), ages 4–12, were screened for color-vision defects using the Neitz paper-and-pencil test for color vision. Of these, 905 children (18%) made one or more errors and were retested. All the children who made one or more errors on each of two administrations were classified as having a color-vision deficiency. All children who made zero errors on at least one of the two administrations were classified as having normal color vision. One purpose of the experiments reported here was to evaluate the validity of the Neitz paper-and-pencil test for color vision as a tool for detecting the

FIG. 2. Isochromatic (confusion) lines of protanopes (solid), deuteranopes (dashed), and tritanopes (dotted) plotted in 1931 CIE *x*, *y* chromaticity coordinates compared to the measured chromaticities of colored and gray dots in color-vision test plates. The top four panels (A, B, C, and D) show examples of chromaticities of dots in the Neitz paper-and-pencil test. The data points that plot nearest to the intersection of the three isochromatic lines are average measurements of gray background dots in the plates. (A) Comparison of 5 different copies of the Neitz test under theoretical equal-energy light conditions. The different copies were drawn from a print run of 50 copies. The calibration was adjusted prior to printing the first copy of the run, and then 50 copies were made without any readjustment. Chromaticities of the dots were measured for the 10^{th}, 20^{th}, 30^{th}, 40^{th}, and 50^{th} test of the print run. Data points are plotted in different symbols (filled triangles, filled squares, open triangles, filled circles, open circles) to differentiate the measurements from the five different tests. Panels B, C, and D show sample measurements from the dots of one test page recalculated to show the relationship between the test plate chromaticities and the dichromatic confusion lines with changing illuminants: (B) simulated daylight; (C) cool white fluorescent; and (D) daylight fluorescent. For comparison with results from the Neitz test, panels E and F show a similar analysis done on two other color-vision tests. For these measurements, chromaticities are plotted for the test when administered under Standard Illuminant C as specified by the publishers. (E) AO-HRR pseudo-isochromatic plates (1955 edition). Even though the test book is 45 years old, the chromaticities of the dots are close to the target dichromatic confusion lines. (F) The recently available Richmond HRR pseudo-isochromatic plates (1991 edition). The chromaticities in this newly produced test deviate considerably from the original.

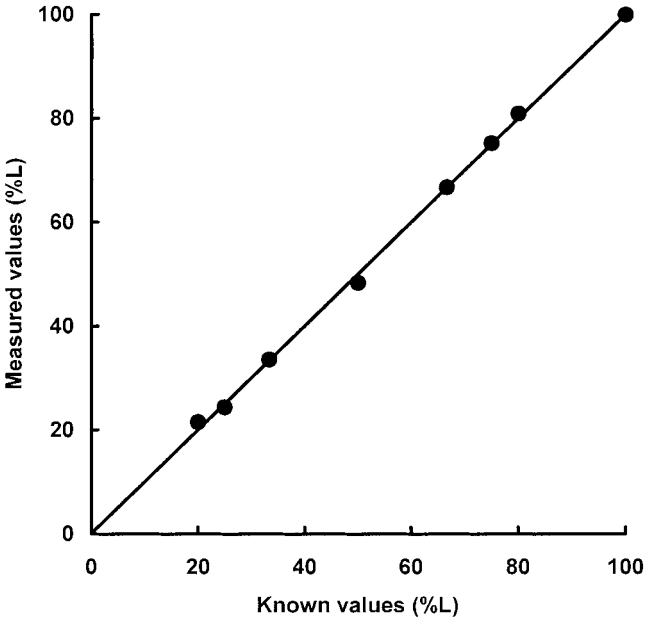


FIG. 3. Real-time quantitative PCR measurements of DNA samples with known L/M ratios. Known mixtures of L and M pigment cDNA clones were measured to evaluate the validity of the quantitative technique.

presence of color-vision deficiencies. One hundred and ninety seven males (7.5%) made errors on both the first and second test and, thus, were categorized by the Neitz paper-and-pencil test as having a color-vision deficiency.

A second goal was to evaluate the use of the combined results of the paper-and-pencil test together with genetic information about the number and ratio of pigment genes for determining the type and severity of color-vision deficiency. Genetic testing was done on 115 of the males who made errors on the Neitz test twice and were, thus, categorized as having a color-vision defect. Of the children classified as color deficient, only the results for the 115 males for whom we also obtained genetic information will be discussed here.

Detecting the Presence of Color-Vision Deficiencies. Six panels of the Neitz test contained red-green confusions and two panels contain tritan confusions. Of the 115 males for whom genetic testing was done (all of whom made errors on red-green plates on two separate tests), none made the same error twice on the tritan plates.

The children who made errors on two administrations of the Neitz test were tested individually using Ishihara's tests (Concise Edition), the American Optical–Hardy Rand and Rittler (AO-HRR) plate test, and the APT-5 Color Vision Tester (LKC Technologies). Rayleigh matches using a Nagel Anomaloscope were also obtained from the older children. All the children who were classified as color deficient by the Neitz test also failed one or more of the other tests. Thus, we found no instance of a child classified as color-vision deficient by the Neitz test who would be categorized as normal by a battery of conventional tests.

In the sample of all the males, 7.5% (197 out of 2637 males tested) were classified as having a color-vision deficiency. The 95% confidence interval for this percentage is 6.5–8.6%, consistent with what has been observed previously in large population studies. For example, Schmidt[23] observed a frequency of 7.75% color-vision deficiencies in a sample of about 7,000 German males, which may be an appropriate comparison population, because a large fraction of people in southeastern Wisconsin are descendents of German emigrants. Finding the expected number of color-deficient males is additional evidence that the Neitz test does not misclassify normal children as having a color-vision deficiency. Finding the expected percentage is also evidence that the test is effective in detecting the presence of color-vision defects. To investigate this further, we tested a group of children who passed the Neitz test using the battery of conventional tests. One hundred and forty five children who made no errors in the classroom administration of the Neitz test were randomly selected. None of these children made errors when the Neitz test was administered to them a second time. None of these children were classified as having a color-vision defect from the results on the conventional tests. Thus, we found no case of a child who passed the Neitz test but was nonetheless determined to have a color-vision defect based on results from the other tests. We note that this small *n* of 145 does not give a great deal of power statistically. A result of zero out of 145 indicates a 95% confidence interval with an upper limit of about 3%, i.e., in the population, the test could miss a 3% rate of color deficiencies, and we still could have, by chance, obtained zero instances in our sample of 145. However, this is an upper limit, and this result together with the finding of approximately the predicted frequency of color-vision deficiencies is consistent with the conclusion that color-deficient children would rarely be misclassified as normal by the Neitz test.

Type and Severity of Red-Green Color-Vision Defect. Of the six panels that are diagnostic for red-green color-vision defects, two of them contain protan confusions and four contain deutan confusions. The red-green plates occur in three different levels of saturation. In order to evaluate red-green color vision severity, performance on the Neitz test color-vision behavior was categorized as follows. Missing only the symbol in one or both of the most desaturated red-green panels was designated behavior 1 (abbreviated B1). Errors that included missing the symbol in one or both of the most saturated red-green panels was designated behavior 3 (B3), and errors that included missing the symbol in one or both of the plates that were intermediate in saturation (but not missing the symbols in the most saturated plates) was behavior 2 (B2). Eleven children were categorized as B1, 47 as B2, and 56 as B3. If a child behaved differently on each of the two tests, which 35 (out of 115) of them did, the designation assigned was the one for the worst performance.

The number and ratio of L and M genes on the X-chromosome in the 115 males was examined. Figure 3 shows the standard curve that was used to calibrate the real-time quantitative PCR measurement of L/M gene proportion. These results illustrate that the L and M gene probes are

highly specific and that this method gives an accurate measurement of the L/M proportion. Gene ratio measurements were made a minimum of four times for each child, which allowed careful monitoring of the reliability of each genetic measurement. We also evaluated the reliability of the measurements by running repeated measures extensively on a few samples over a period of months. These measurements allowed not only an estimate of the day-to-day reliability of the quantitative measurements, but also the week-to-week and month-to-month reliability. It was estimated from these measurements that, by taking the average of four repeated measurements for each sample, more than 95% of the estimated ratios for our subjects have an error of less than $\pm$ 4%. This 4% error represents total error in the measurement, including errors in sample loading, fluctuations in the reagents, and the long-term stability of the instrument.

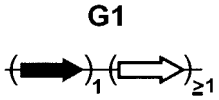
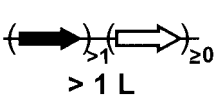
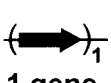
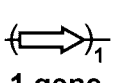
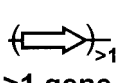
All the genetic results were expressed in terms of gene percentages. For each subject, two genetic values were obtained: (1) a percentage of downstream genes and (2) a percentage of L genes; each is compared to total number of X-chromosome pigment genes. A result of zero percent downstream genes indicates that there were no genes downstream of the first pigment gene in the array; thus, the person must be a single-gene dichromat. 50% downstream genes indicates 2 photopigment genes total; 67% downstream genes indicates 3 genes, 75% indicates 4 genes, and so on. The genetic data for the subjects were put into five categories according to the relative number and ratio of L and M genes. Arrays with 1 L gene and one or more M genes were designated G1. For example, an array with 50% downstream genes, and 50% L genes was interpreted as having two genes of which one is L; thus, he was categorized as having 1 L and one M gene. An array measured to have 67% downstream genes and 33% L genes was interpreted as having 1 L and 2 M genes. G1 arrays have been most often associated with normal color vision,[2] and only rarely with color vision defects.[16,17] An array with more than one L gene was termed G2. For example, an array with 66% L genes and 66% downstream genes was interpreted as having 2 L and 1 M genes. Arrays with multiple L genes were the most common finding among the color-defective children, and the majority of these also had one or more M genes. Arrays with more than one L gene and no M genes were also categorized as G2; however, having multiple L genes and no M genes was rare. It occurred in only four of our 115 color-vision deficient children. G2 arrays have been associated with deutan color-vision defects.[2,15,19] However, having an array with more than one L gene is not by itself diagnostic of a color vision defect, because both individuals with normal color vision and those with color-vision deficiencies can have an array with one or more M genes and more than one L gene. G3 was the designation for single-gene array with an L gene, and these have been found in deuteranopes.[2,14,24] G4 was the designation for single-gene arrays with only an M. These arrays have been found only in protanopes.[2,14,25] An array with multiple M genes, and no L genes, was designated G5. These have been found in protanopes as well as protanomalous trichromats.[2,14,25,26] Of the 115 color-deficient males for whom genetic testing was done, 14 were categorized as G1, 61 as G2, 13 as G3, 6 as G4, and 21 as G5.

Subjecting young children who passed the Neitz test to genetic testing was not judged to be justifiable. In the cases in which a child was classified as having a color-vision deficiency from the behavioral tests, the additional genetic test was justified because of the possibility that the genetic information could help to clarify the classification. However, it was difficult for us to rationalize the purpose of a genetic test to examine the photopigment genes in a child for whom there is no problem suspected. Thus, no children who were classified as having normal color vision were genetically tested. Instead, as a control for the genetic portion of the test, 91 male medical students were screened for color-vision defects using the Neitz test for color vision and then they were tested genetically. Of the 84 who made no errors on the Neitz test, the majority had G1 arrays, although a few had G2 arrays. G3, G4, and G5 array structures were never found among these 84 men categorized as having normal color vision.

For the 115 male children for which genetic test results were available, a determination of the color-vision phenotype was made by considering the genetic data together with performance on the Neitz color-vision test, Ishihara's tests, AO-HRR test, the APT 5 vision tester and, when available, Rayleigh match data. The final classifications based on all the available data for each child are displayed in Table I. Based on the number and pattern of errors for all color-vision tests and the genetic results, four levels of severity were assigned. From least to most severe they were mild (MI), moderate (MO), and severe anomalous trichromacy (S), and dichromacy (P for protanopia and D for deuteranopia). In some cases, we could not distinguish between severe anomalous trichromacy and dichromacy based on the available information. In those cases, we left the classification ambiguous; for example, a deutan subject who could either have deuteranopia or severe deuteranomaly was designated as D/SDA.

We also assigned a classification based only on the information from the Neitz paper-and-pencil test plus the minimal genetic information derived from an estimate of the number and ratio of X-linked pigment genes. Table I shows that there are 15 possible combinations of gene arrangements, and behavior on the paper-and-pencil test. It was predicted that no one whose test indicated the mildest deficiency would turn out to be a single-gene dichromat genetically. This turned out to be true; no one had test behavior B1 and genetic results G3 or G4. Thus, there were 13 remaining possible combinations of genes and paper-and-pencil test behavior. Each combination was assigned a classification, and they are shown in bold type at the top of each cell in Table I. When a classification was from only the paper-and-pencil test and the genetics, severity of the defect was estimated as follows (see Table I). Everyone with only one gene on the X-chromosome was classified as a dichromat. Individuals who made errors only on the most desaturated plates were classified as having a mild deficiency. G1

TABLE I. Color vision phenotypes predicted from genetic data and performance on the Neitz test are given in bold type at the top of the cell. Classifications made based on performance on the Neitz test, genetics, and a battery of conventional tests, are listed in each cell along with the number of subjects with that classification. Numbers in parentheses are the cell totals.

	G1  1L, ≥1 M	G2  > 1 L ≥0 M	G3  1 gene L only	G4  1 gene M only	G5  >1 gene M only	
	Designations from genetic results					
	Unspecified	Deutan		Protan		
Behavior	G1	G2	G3	G4	G5	Total
B1	**MIUS** 2 MIUS 1 MIPA	**MIDA** 7 MIDA	**N/A** 0	**N/A** 0	**MIPA** 1 MIPA	
	(3)	(7)	(0)	(0)	(1)	11
B2	**MOUS/SUS** 1 SPA/P 4 MIUS	**MODA/SDA** 5 SDA/D 9 SDA 9 MIDA 11 MODA	**D** 4 D	**P** 0	**MOPA/SPA** 3 SPA/P 1 MIPA	
	(5)	(34)	(4)	(0)	(4)	47
B3	**SUS** 2 D 4 MOUS	**D/SDA** 5 D 7 SDA/D 5 SDA 3 MODA	**D** 9 D	**P** 6 P	**P/SPA** 5 P 6 SPA/P 4 SPA 1 MOPA	
	(6)	(20)	(9)	(6)	(16)	57
Total	14	61	13	6	21	115

[a] Abbreviations: MIUS = mild, unspecified; MOUS = moderate unspecified; SUS = severe unspecified. MIDA = mild deuteranomaly; MODA = moderate deuteranomaly; SDA = severe deuteranomaly; D = deuteranopia; SDA/D = severe deuteranomaly or deuteranopia; MIPA = mild protanomaly; MOPA = moderate protanomaly; SPA = severe protanomaly.

individuals (i.e., males who have only one L gene and one or more M genes) who make errors on the most saturated plates cannot be normal (as might be presumed from the genetics alone), and they presumably cannot be conventional deuteranomalous trichromats, because deuteranomaly requires two slightly different L-type pigments. Of the usual available categories, they could be deuteranopes, protanopes, or protanomalous, or they could have something that does not fit into conventional categories; we assigned them the ambiguous classification of severe unspecified (SUS). There were a few subjects who had G1 category genes and B1 and B2 behaviors. Again, from the genes alone, normal color vision might be assumed. However, in every case these individuals not only made consistent errors on the paper-and-pencil test, but they consistently made errors on other more conventional tests. Based on the paper-and-pencil test, we categorized the B1/G1 males as mild unspecified (MIUS) and the B2/G1 males as moderate to severe unspecified (MOUS/SUS). All G2 and G3 individuals were classified as having deutan deficiencies. This is presumably definitive. The G3 males have only one L gene; they must be deuteranopes. Normal individuals can have a G2 genotype, but once someone is classified as having a red-green defect, the G2 genotype indicates the type of defect as deutan. No exception to this has been observed. G4 and G5 gene arrangement have only M genes. A classification of protan is presumably definitive.

As shown in Table I, classification as protan or deutan was made for 101 subjects, 74 were classified with deutan defects and 27 with protan defects. Thus, among the subjects for whom type could be determined, 27% were protans (95% confidence interval, 18–36%). This is consistent with large population studies in which approximately 25% of red-green deficiencies have been found to be of the protan type. This is additional evidence that the combination of the paper-and-pencil test and the genetics is an effective classification tool.

In Table I, the classification made from the paper-and-pencil test and genetic information only (bold at the top of each cell) can be compared to the classification made from the additional information from all test results. The classifications from the genes and Neitz test compare favorably with those that are based on information from considerable additional testing. The biggest discrepancy was among individuals designated with B2 behavior from the Neitz test. We designated them as having a moderate to severe color-vision deficiency. This was true for 70% of the males categorized as B2 from the Neitz test. However, 30% of the B2 males were categorized as mild after consideration of the data from the other tests. Overall, males who are catego-

rized as mild by the Neitz test are found to be mild in all tests. Those who test as severe on the Neitz test are found to be severe on other tests. Those who test as intermediate in severity (B2) on the Neitz test fall into categories from relatively mild to severe, based on results from other tests.

DISCUSSION

The goal of the present study was to determine how well color-vision phenotype could be predicted from consideration of data from a newly developed minimal genetic test and performance on a paper-and-pencil color-vision screening test. The results suggest that this works well both in discriminating people with color-vision defects from those with normal color vision and in providing information about the type and severity of defects.

A unique aspect of this study is that the genetic results come from males with color-vision deficiencies who were drawn from a larger general population in which presumably nearly all the color-defective males were identified. Thus, the study may provide unique information about how X-chromosome gene arrangements are distributed in the population of individuals with color-vision deficiencies. Previous studies, e.g., Nathans *et al*.,[2] Neitz *et al*.,[15] Deeb *et al*.,[24] Shevell *et al*.,[18] and Neitz *et al*.,[26] that have examined genes in subjects with red-green color-vision defects recruited subjects by advertising for individuals with a color-vision defect. It is hard to know exactly how this might bias the sample; however, one might predict that people who answer advertisements represent more severe cases that know for certain they have a color deficiency. Severely affected people may also be more interested in seeking information about their condition. A striking aspect of the present sample of genetically tested males is that 12% (14 out of 115; 95% confidence interval = 7–20%) had gene arrays that have not typically been associated with color vision defects (G1; Table I); the arrays are ones expected for color normals. Cases like this have been observed in the past;[2,18,20,24] however, the present study may give the best indication of the frequency of these gene arrangements in the color-defective population. The 12% observed here is higher than one might have guessed from previous studies. Two of the 14 subjects with grossly normal arrays were deuteranopes. In each of those two cases, a deleterious mutation was found to interrupt the function of the M pigment genes and, thus, cause the color-vision defect. These two cases are described in a separate report in this issue.[17] For the other 12 subjects, we do not yet know any genetic correlate to their color-vision deficiency. Two of the subjects with one L and one or more M genes were classified as protan from the color-vision test battery. They may have defects in their one L pigment gene, perhaps, analogous to the M gene defects found in the two deuteranopes who were investigated in detail.[17] The remaining males with grossly normal gene arrays have mild to moderate deficiencies. In each case, the fact that consistent errors were made on more than one test suggests that they cannot be normal subjects misclassified as having color-vision deficiencies. On the other hand, the patterns of errors indicate that it is unlikely that they are dichromats. Deuteranomalous individuals have two L type genes that encode two slightly different L pigments. The G1 individuals appear to have only one L gene. Thus, deuteranomaly seems to be excluded as a classification. One possibility is that some of them have both L and M cones with normal spectral sensitivities, but one cone type is drastically underrepresented in number, as has been suggested for some cases of color-vision deficiency.[27,28] The result of 12% of color defective males with a single L gene and normal-appearing M genes suggests that unconventional explanations, such as extreme biases in cone ratio, may be required for about 1 in 10 individuals with a color-vision deficiency.

Here we report the use of combined results from a simple 1-page, paper-and-pencil test and simplified minimalist genetic test as a highly economical and efficient method for detecting color-vision deficiencies and determining their type and severity. A practical question is, what are the prospects for developing a purely genetic test for color-vision deficiencies? Our experiences indicate that a purely genetic test for color-vision defects is not likely to be feasible or practical in the near future. The first problem is with the use of genetics to detect the presence of a defect. In the present study, 65% of the color-deficient males had either G1 or G2 arrays (75/115, Table I). Although G1 arrays are most common in individuals with normal color vision and G2 arrays are most often found in males with deutan color deficiencies, there are a significant number of exceptions for each gene arrangement. Thus, the minimal genetic information gathered here cannot be used alone to definitively separate normals from color defectives. On the other hand, detection of color-vision defects is something that the paper-and- pencil test does very efficiently and reliably.

Once a color-vision defect has been detected, the minimal genetic test provides an efficient objective method to access certain information about type and severity. There are two aspects of type and severity in which the minimal molecular genetics are unsurpassed in providing definitive information. First, individuals with single-gene arrays are expected to have the most severe loss in color vision — dichromacy. Thus, the finding of a single pigment gene on the X-chromosome should give a clear classification of dichromacy. The genetic information is also unsurpassed in providing information that can be used to separate deutan from protan defects. The only limitation is that the fraction of individuals with G1 genotypes cannot be assigned as protan or deutan type from the genetics. The only diagnostic method that gives comparably clear classification of protan vs. deutan is the anomaloscope. However, anomaloscopes can be used only with adults and older children, they are not widely available, and the use of the instrument requires a highly trained operator.

Previous studies have shown that the spectral separation between L pigments in deuteranomalous men correlates with the degree of color-vision loss.[15,18] That is, the more similar the peak sensitivities of the L pigments are in a

deuteranomalous male, the worse his color vision. Additional genetic tests could be done to deduce the spectral separation between the pigments underlying red-green color vision in each individual, and the results would provide information that would correlate with severity. However, such DNA sequencing efforts are time-consuming and expensive (hundreds of US dollars per subject) compared to the minimal genetic information obtained in this study. The finding of a strong correlation between peak separation and severity of deuteranomalous trichromacy has been important theoretically, but the examination of the gene sequences as a general screening tool for determining the severity of color-vision loss seems impractical.

The paper-and-pencil test was designed so that the information it provides would complement the information provided by an efficient and relatively economical minimalist genetic screening test. Both halves of the test are easily used with young children. Individuals with no prior training in color-vision testing can administer the paper-and-pencil test and collect buccal swabs from the fraction determined to have color-vision defects. The paper-and-pencil test is effective at detecting color-vision deficiencies and the combined information from behavior and genetics provides a good indication of type and severity, as shown in Table I. Consider the cost and effort of administering the combined test to 100 children. About 15–20 min are required to administer the test for a first time to a classroom of 25, thus about an hour of total time would be needed to test 100 children. From our experience, about 20% make errors on the first test and require a second test. It might require a second hour to reexplain the instructions and retest 20 children from the original 100 in smaller groups. A third hour might be needed for an untrained person to evaluate the test results and thus complete the paper-and-pencil segment of the testing. Thus, identification of all the color defective children would have been made with a total time investment of less than 2 min per child (180 min/100 children). This is extremely efficient when one considers that the quality of information is such that the test would need to be done only once during each person's childhood. If the 100 children included, 50 boys and 50 girls, we expect that, on average, about four would be identified as having a color-vision deficiency. Buccal swabs need only to be collected from those four and submitted for genetic analysis to complete the second segment of the test. The minimalist genetic test could theoretically be performed at a reasonably low cost (see below), and, because it is done on only about 4% of children, the cost and effort distributed over all children tested is small. Our experiences suggest that a paper-and-pencil screening test like that used here combined with subsequent minimal genetic testing on children who test as having color-vision deficiencies is an unsurpassed method in terms of ease and effectiveness for testing color vision in children.

Here we propose that a genetic test be conducted after screening with the paper-and-pencil test. Ultimately, if the proposed procedure is to be implemented on a larger national or international scale, facilities would have to become available to the general public, where someone might send samples for genetic analysis. The genetic procedures used here were designed to be simple, reliable, accurate, and easily standardized, so that they could ultimately be implemented on an appropriately large scale. Although obstacles remain, the translation between the genetic portion of the test as an experimental procedure and its ultimate implementation as part of a standard color-vision diagnostic seems theoretically feasible. One consideration is the cost per subject. In our laboratory, the full cost (labor and supplies) per child is about US $50.00 to run a full set of repeated measures of two assays, one to estimate L/M gene ratio and the other to estimate upstream/downstream gene ratio. Presumably, the cost per child could be streamlined, if implemented on a larger scale. An interim goal is to continue to test and evaluate the genetic analysis. It would be possible for other investigators to contact us about performing our minimalist genetic analysis on interesting subjects. The results from such analyses would be for experimental use only and not for diagnostic use. Our capacity to handle such submissions is limited. Up-to-date information about the availability and status of this service can be obtained at: http://www.mcw.edu/cellbio/colorvision/.

ACKNOWLEDGMENTS

This work was supported by RPB, the David and Ruth S. Coleman Foundation, the Ronald McDonald Charities, Knights Templar Eye Foundation, the Aaberg Retina Foundation, and NEI grants EY09303, EY01931. We gratefully thank P. Summerfelt, C. Bialozynski, S. Balding, and J. Besharse for their enormous help, with special thanks to S. Bence and Prevent Blindness Wisconsin for their invaluable assistance in launching this study.

1. 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.
2. Nathans J, Piantanida TP, Eddy RL, Shows TB, Hogness DS. Molecular genetics of inherited variation in human color vision. Science 1986;232:203–210.
3. Wolf S, Sharpe LT, Knau H, Wissinger B. Numbers and ratios of X-chromosomal-linked opsin genes. Vision Res 1998;38:3228–3231.
4. Neitz M, Neitz J. Numbers and ratios of visual pigment genes for normal red-green color vision. Science 1995;267:1013–1016.
5. Asenjo AB, Rim J, Oprian DD. Molecular determinants of human red/green color discrimination. Neuron 1994;12:1131–1138.
6. Neitz M, Neitz J, Jacobs GH. Spectral tuning of pigments underlying red-green color vision. Science 1991;252:971–974.
7. Neitz M, Neitz J, Jacobs GH. Analysis of fusion gene and encoded photopigment of colour-blind humans. Nature 1989;342:679–682.
8. Merbs SL, Nathans J. Role of hydroxyl-bearing amino acids in differentially tuning the absorption spectra of the human red and green cone pigments. Photochem Photobiol 1993;58:706–710.
9. Merbs SL, Nathans J. Absorption spectra of the hybrid pigments responsible for anomalous color vision. Science 1992;258:464–466.
10. Mollon JD, Bowmaker JK, Jacobs GH. Variations of colour vision in a New World primate can be explained by polymorphism of retinal photopigments. Proc R Soc Lund B 1984;222:373–399.
11. Winderickx J, Lindsey DT, Sanocki E, Teller DY, Motulsky AG, Deeb

SS. Polymorphism in red photopigment underlies variation in colour matching. Nature 1992;356:431–433.
12. Williams AJ, Hunt DM, Bowmaker JK, Mollon JD. The polymorphic photopigments of the marmoset: spectral tuning and genetic basis. EMBO J 1992;11:2039–2045.
13. Neitz M, Kraft TW, Neitz J. Expression of L cone pigment gene subtypes in females. Vision Res 1998;38:3221–3225.
14. Neitz M, Neitz J, Jacobs GA. Genetic basis of photopigment variations in human dichromats. Vision Res 1995;35:2095–2103.
15. Neitz J, Neitz M, Kainz PM. Visual pigment gene structure and the severity of human color vision defects. Science 1996;274:801–804.
16. Winderickx J, Sanocki E, Lindsey DT, Teller DY, Motulsky AG, Deeb SS. Defective colour vision associated with a missense mutation in the human green visual pigment gene. Nat Genet 1992;1:251–256.
17. Bollinger K, Bialozynski C, Neitz J, Neitz M. The importance of deleterious mutations of M pigment genes as a cause of color vision defects. Col Res Appl Suppl 2001;S100–S105.
18. 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.
19. Hayashi T, Motulsky AG, Deeb SS. Position of a 'green-red' hybrid gene in the visual pigment array determines colour-vision phenotype. Nat Genet 1999;22:90–93.
20. Crognale MA, Teller DY, Motulsky AG, Deeb SS. Severity of color vision defects: electroretinographic (ERG), molecular and behavioral studies. Vision Res 1998;38:3377–3385.
21. Hagstrom SA, Neitz J, Neitz M. Variations in cone populations for red-green color vision examined by analysis of mRNA. NeuroReport 1998;9:1963–1967.
22. Sjoberg SA, Neitz M, Balding SD, Neitz J. L-cone pigment genes expressed in normal colour vision. Vision Res 1998;38:3213–3219.
23. Schmidt I. Ergebnis einer Massenunters chung des Farbensinns mit dem Anomaloskop. Z fur Bahnharzt. 1936;2 [result cited in: Congenital and acquired color vision defects. Pokorny J, Smith VC, Verriest G, Pinckers AJLG, editors. New York: Grune Stratton; 1979].
24. Deeb SS, Lindsey DT, Hibiya Y, Sanocki E, Winderickx J, Teller DY, Motulsky AG. Genotype-phenotype relationships in human red/green color-vision defects: molecular and psychophysical studies. Am J Hum Genet 1992;51:687–700.
25. 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.
26. Neitz J, Neitz M, He JC, Shevell SK. Trichromatic color vision with only two spectrally distinct photopigments. Nat Neurosci 1999;2:884–888.
27. Regan BC, Jordan G, Mollon JD. A reinterpretation of Pigmentfarbenanomalie. Invest Ophthalmol Vis Sci (Suppl) 1999;40:S353.
28. Jacobs GH, Neitz J. Electrophysiological estimates of individual variation in the L/M cone ratio. In: Drum B, editor. Color vision deficiencies XI. Dordrecht: Kluwer Academic; 1993.