Human Color Vision Deficiencies and Atmospheric Twilight

T. E. Reimchen

Department of Zoology
University of Alberta
Edmonton, Alberta, Canada

ABSTRACT: Red-green color vision deficiencies in human populations are generally regarded as visual dysfunctions, the frequencies of which have increased from mutational pressure and relaxation of selection in post-Neolithic periods. However, color vision insensitivities also occur in other vertebrates and appear to be an ecological adaptation for low light levels encountered during crepuscular activity. I hypothesize a comparable function for human vision deficiencies and test this by re-examining global frequencies of deficients in relation to photic characteristics of the habitat. Data presented here show significant positive correlations with length of atmospheric twilight and increased equatorial distance. While support for the twilight hypothesis remains equivocal, the latitudinal data imply that the high frequencies of color deficiencies in North European populations represent relatively stable frequencies from Neolithic periods when these traits were at an advantage. The data and conclusions are inconsistent with the current practice of using frequencies of these polymorphisms as cultural chronometers.

Among the various classes of inherited color vision deficiencies in humans, red and green weaknesses in dichromats and anomalous trichromats are the most prevalent, affecting approximately 5 per cent of the human population. These deficiencies result from molecular changes at the loci coding for red and green visual pigments (Nathans et al., 1986). There is no consensus as to the origin of these sex-linked polymorphisms or to the evolutionary forces influencing gene frequencies. The major model that has emerged from the last century of studies is “selection relaxation” (Pickford 1958, 1963; Post 1962a, b, 1971). A low incidence (2 per cent) of red-green (R-G) deficiencies in hunting peoples and higher values (8 per cent) in agricultural and industrial societies led Post (1962a) to suggest that natural selection was still operating to remove color vision deficients in hunting populations, while in agricultural or industrial societies there has been a relaxation of selection resulting in an increase in gene frequency from mutational pressure. Post infers that the proportion of deutsans (green weakness) in Europe has increased from 1.5 per cent to 6 per cent since the Neolithic, and Kalmus (1965), in a review of inherited color deficiencies, notes that this data “provides the only clear example of a simple inherited defect, the frequency of which is significantly increased with the advent of civilization.”

Post’s model has provided a popular framework for interpreting population differences in color deficiencies (Bhasin, 1967; Dobson et al., 1967; Ray, 1969; Chattopadhyay, 1970; Post, 1971; Malhotra et al., 1974; Malhotra, 1978; Sastry, 1974, Mueller and Weiss, 1979; Salzano, 1980). Yet various aspects of the model have been criticized (Thoday, 1965; Adam et al., 1966; Adam, 1973; Halberstein and Crawford, 1974; Murty
and Vijayalaxmi, 1974; Mukherjee et al., 1979). Ford (1964) suggests that mutation is not a sufficient condition to generate such high frequencies and that this may be a balanced polymorphism. Adam (1973) has pointed out that it is primarily the deutoanomalous trait that has increased in frequency, whereas Post's model would predict that other deficiencies should increase in frequency as well. Furthermore, although the oldest cultures are predicted to have the highest frequencies of deficiencies, this is not the case (Dobson et al., 1967; Adam 1969).

The basic premise of Post's model rests on the assumption that departures from full trichromatic sensitivity represent a suboptimal condition of the visual processing system. While this may be reasonable, there are sufficient data to justify further scrutiny of this assumption. Recent studies of the visual system of the Amazonian squirrel monkey (Saimiri) and the spider monkey (Ateles) also yield evidence for a polymorphism in color vision including variations comparable to the protoanomalous form (weak red deficiency) in humans (Jacobs, 1981), suggesting either a comparable suboptimal system or ecological adaptations. Experimental observations on human vision indicate that red-green insensitive individuals can detect relatively smaller differences in brightness of colors than normals (Adam 1969), particularly at reduced light intensity (Iinuma and Handa, 1976; Hurvich, 1981). "Color blind" individuals have distinct forms of these pigments with characteristic absorption curves (Rushton et al., 1973). Recent analysis of DNA fragments of the loci coding for the visual pigments show that deficiencies have asymmetric combinations of the major gene fragments as well as increased numbers of gene copies relatives to normal observers (Nathans et al., 1986), patterns inconsistent with mutation rate as a source of these polymorphisms. Deficients occasionally show enhanced visual acuity in detecting objects which appear camouflaged to normal observers (Judd, 1943; Ford, 1964; Thoday, 1965; Roberts, 1967; Roberts and Tanner, 1967), a trait which could be particularly relevant to early hunters for detection of predators and prey. If so, increased fitness among color insensitive individuals could in part account for this widespread polymorphism. The major difficulty with this adaptive interpretation is that, according to Post (1962a), the highest frequencies occur among agricultural and industrial societies rather than among hunting peoples.

One factor which has not been previously considered in studies of human color vision polymorphisms is the potential correlations with photic characteristics of the habitat. During a recent biological field trip, I observed that a member of our group showed exceptional visual acuity at twilight. Subsequent discussion established that he was either a deuteranope or a severe deuteranomalous trichromat (green insensitive). This prompted the hypothesis that color visions deficiencies represent evolutionary adaptations among early populations for hunting and gathering activities during the low light intensities encountered at twilight. This contrasts with the current model in which anomalies in human color vision are regarded as dysfunctions and the by-product of mutational rate.

If the hypothesis is tenable, frequencies of R-G deficiencies should be highest in those cultures where activities were more often conducted at low light levels.
Comparative data on diel activity patterns are not available in ethnographic studies to examine this prediction. On the gross assumption that each indigenous culture uses available twilight to a similar extent, frequencies of R-G deficiencies should be correlated with the total amount of twilight in the habitat.

MATERIALS AND METHODS

Length of civil twilight and total daylight were tabulated from ephemerides (Anon, 1945) for latitudes 0–60° at 10° intervals and at 65°. Total hours of twilight per year increases exponentially with geographic latitude, from 132 h at the equator to 144 h at 30° and to 341 h at 60°. Data are presented as relative lengths of twilight per day (civil twilight/civil twilight + daylight). The term “daylength” is used to designate the period of daylight plus civil twilight. Civil twilight is defined as the period between sunset (or sunrise) and the time when the true position of the sun is 6° beneath the horizon and coincides with the period when first magnitude stars are just visible.

I have compiled published frequencies of R-G deficiencies in males from 45 countries (380 populations) and tested for relationships with both twilight and latitude. Attempts were made to locate and include all published data on frequencies for indigenous peoples (those who have historically resided in one geographical area). Although R-G deficiencies represent at least six classes including protanopes, deuteranopes, protanomalous, and deuteranomalous classes, much of the literature reports only two categories, protan (red deficient) and deutan (green deficient) or combines all phenotypes into a single category as a general measure of color-deficiencies in the population. I have employed the latter method in the subsequent review. Grouping of protan and deutan is reasonable as the frequencies are positively correlated ($r = 0.24$, $P < 0.02$).

These data points represent primarily Eurasia, Africa, and South America. Indigenous people from North America and the U.S.S.R., which occupy a wide latitudinal range, have not been surveyed extensively for visual traits.

RESULTS

Relative length of twilight per day ranges from 6 per cent in equatorial regions to 25 per cent in northern latitudes (60° N) (Figure 1). Frequency of R-G insensitive individuals and the amount of twilight are highly correlated ($r = 0.55$, $P < 0.001$) as described by the power curve $y = 1.55 x^{0.85}$ where $y$ is the frequency of R-G insensitive individuals and $x$ the relative length of twilight per day (Figure 2). A stronger correlation is obtained with latitude ($x$), yielding the linear regression $y = 2.46 + 0.089x$ ($r = 0.69$, $P < 0.001$). Partitioning for the two major categories of color deficiencies yields a similar latitudinal trend for both groups but with a stronger association for deutans (protan, $r = 0.20$, $P < 0.05$; deutan, $r = 0.51$, $P < 0.001$).

Equatorial peoples (+/- 20°) from Central and South America, Africa, and Asia generally have low values (0 per cent to 4 per cent); those from midlatitudes (20–40°) including North Africa, the Middle East, India, Tibet, China, Korea, and Japan have intermediate values (3 per cent to 6 per cent); and those from northern latitudes (40–65°), mainly from Western Europe, have higher frequencies (6 per cent to 10 per
Fig. 1.—Relationship between geographical latitude and length of twilight. Per cent twilight is (civil twilight) / (civil twilight + daylight).

Fig. 2.—Geographical distribution of color vision deficiencies.
cent). Highest values occurred in Finland (Kajanoja, 1972) (14.2 per cent) at 63°N. The limited number of data points from North America show no evidence of a frequency cline with the northern Aleuts exhibiting low frequencies.

There were no consistent clines over small geographical regions such as in the United Kingdom, India, and Japan where large segments of the population have been surveyed. A weak N/S cline was described (Adam et al., 1970) for the Bantus of South Africa, where the frequencies increased from 1.5 per cent in the northern provinces (22° S) to 4.3 per cent near the southern tip of the continent (33° S).

**DISCUSSION**

The above data, which provide equivocal support for the twilight hypothesis, clearly indicate an African and Eurasian north/south cline in frequencies of color vision polymorphisms which span different cultures and continents. Thus, at the equator, where twilight is only a minor component to total daylight, color vision deficiencies are generally rare, yet in northern regions, where twilight represents up to 25 per cent of daylength, color vision frequencies are common. While long summer daylengths in northern latitudes may have offset the requirement for crepus-

---

cular activity of early hunters, short winter days could have resulted in substantial use of twilight for hunting activity and strong selective pressures for the appropriate visual system.

Such correlations between color deficiencies and twilight do not imply causation since latitude is a complex variable associated with daylength, temperature, snow cover, vegetation and other biota, to name a few. Yet there is corollary evidence to suggest that photic regime may be an important factor in this correlation. Visual pigments in other vertebrate species show that maximum absorption curves generally correspond to the most common wavelengths of light found in the habitat while minimum sensitivity correspond to segments of the spectrum which are rare or absent (Lythgoe 1979). Atmospheric spectra during twilight show loss of green wavelengths (McFarland and Munz 1975) which is the most common sensitivity loss in humans color vision deficiencies. In other vertebrate taxa, trichromatic vision is largely restricted to diurnally active forms, while reduced color discrimination, merging into scotopic vision, is prevalent among crepuscular and nocturnal taxa (McFarland and Munz, 1975; Charles-Dominique, 1977; Lythgoe, 1979; Jacobs, 1981). For a species such as Homo sapien, which exhibits substantial behavioral and ecological diversity, it is realistic to assume that selective forces on visual acuity would have operated among early hunters and gatherers in the same direction as that observed in other vertebrates.

Exceptions are evident to the broad latitudinal trends of color vision deficiency frequencies, particularly among small indigenous populations. Aleuts and Eskimos, at high latitudes, have a low incidence of deficients while among equatorial peoples, high frequencies are present in the Jivaros tribe from Ecuador (7 per cent) (Sunderland and Rosa, 1976), in one village of the Yanomama tribe in Venezuela (25 per cent) (Salzano and Neel, 1976) and in the Niue Polynesians from an island near Samoa (10 per cent) (Grosvenor, 1970). Either the twilight model is fundamentally incorrect or these require ad hoc explanations such as postulating unusual photic characteristics of the habitat. For example, the extensive use of marine resources by Aleuts and Eskimos may favor full trichromatic sensitivity. The dense canopy in Amazonian forests would greatly restrict understory illumination possibly producing conditions comparable to twilight where deficients would have an advantage. Alternatively, since these exceptions involve small populations and small sample sizes, stochastic factors may suffice as an explanation for these anomalies. Critical tests of these concepts require detailed information on subsistence activities, present and historical, and empirical data on sensitivity to movement and contrast during the transition from photopic to scotopic vision. Twilight produces a distinct bimodal spectral irradiance (McFarland and Munz 1975) which may be integral to assessment of visual sensitivity of color deficients.

The high frequencies of color deficients in Europe have been previously regarded as the consequence of mutation rate and the best example in human populations of an increase in gene frequency following the Neolithic (Neel and Post, 1963; Kalmus, 1965; Post, 1971). This thesis is based upon the assumption that color deficient frequencies in modern hunters and gatherers are representative of Neolithic frequencies.
Post (1971) estimated a frequency of 1.25 per cent for the deuteranomalous trait in Neolithic Europe based on data from modern Ugandan Bantus and Papuan New Guineans; both cultures occur within 10° of the equator. In view of latitudinal correlations described above, frequencies of color deficiencies in equatorial hunters (e.g., Post, 1963, 1971; Pickford, 1963) do not offer significant insight into present or past European frequencies. Comparisons of populations in similar geographical regions, race, and photic habitats, yet differing in the extent of hunting and agriculture, would allow a more meaningful assessment of Post's model. In northern latitudes, Lapps provide the closest comparison for industrial Europeans. This indigenous group has a color deficient frequency of 6.2 per cent, comparable to the 8.0 per cent predicted on the latitudinal regression. Comparisons of gatherers and nomads with urban groups in Asia yield equivocal support for the model (Kang et al., 1967; Ray, 1969; Malhotra et al., 1974; Malhotra, 1978; Sastry, 1974; Naidu et al., 1978); yet other studies from the same region describe either no relationship or the inverse to that predicted from Post's model (Tiwari, 1965; Bansal, 1967; Ray, 1969; Parikh et al., 1970; Murty and Vijayalaxmi, 1974; Mukerjee et al., 1979). Thus, when standardized for latitude and ethnicity, there is no compelling evidence to indicate a change in gene frequency among groups recently removed from hunting and gathering. These data, in conjunction with the latitudinal cline, greatly weaken the suggestion of a post-Neolithic increase in frequency and largely exclude the use of color deficient frequencies as cultural chronometers. Rather, the available data indicate that these human genetic polymorphisms may have an extended paleohistory and that we are observing the general gene frequencies from earlier periods when these characters were subject to positive and negative selective pressures.

What factors could be operating such that color deficient frequencies rarely exceed 10 per cent in human populations? If the twilight hypothesis is correct, it seems plausible that some indigenous populations would concentrate on crepuscular foraging leading to elevated frequencies of deficiencies. Color-vision anomalies have been investigated for possible genetic associations with syndromes such as bipolar manic-depressive illness, alcoholism, and glucose-6-phosphate dehydrogenase deficiency (Adam, 1961; Cruz-Coke and Varela, 1966; Mendelewicz et al., 1972; Smith, 1972; Gershon et al., 1980). If present, these factors could contribute to reduced fitness and reduced frequencies of color deficient alleles. Yet, it follows that when the advantage to color deficiency was lost, perhaps at the end of the Neolithic, frequencies would have declined leading to the conclusion that deficiencies were more frequent in Neolithic Europe than at present.

In contrast, there may have been an ecological context to counter selection against deficiencies which would not implicate post-Neolithic frequency changes. Major color vision anomalies are sex-linked and expressed primarily in males with female frequencies representing the square of the male frequencies (Inuma and Handa, 1976). Thoday (1965) suggests that such a genetic linkage could develop if the trait was advantageous in males for hunting but disadvantageous in females during food gathering. Color insensitive females who are rearing young may be at a particular disadvantage in their inability to
recognize fruit quality or diurnally active threats on the young. Calculation of equilibrium frequencies for sex-linked traits (Wright, 1969) for these conditions indicates that a stable polymorphism is readily achieved. For example, if female deficients had 5 per cent reduction in fitness and male deficients had < 1.0 per cent advantage, balanced frequencies (> 0.1 per cent and < 10.0 per cent) would result. If fitness of male deficients is unity at the equator but increases 0.1 per cent for each 10° of latitude, genotype frequency would increase from 1.0 per cent at the equator to 5.7 per cent at 60°N. If one considers higher fitness for male deficients, for example, 10 per cent, a 5 per cent equilibrium would be maintained only with a 95 per cent reduction in fitness of female deficients. Such manipulations of selective coefficients are merely of heuristic value but serve to indicate that even a modest reduction in fitness of female color deficients will generate low and stable frequencies of color deficiencies in populations. Additional selective parameters, such as frequency dependence, would further relax the range of fitness values necessary for equilibrium.

In summary, the correlation between length of twilight and red-green color sensitivity, the enhanced visual sensitivity to small differences in color, the presence of unique spectral sensitivities in color deficients, the increased numbers of gene copies coding for the pigment fragments, and the general relationship between color-vision deficiencies and crepuscular activities in other vertebrates constitute plausible support for the twilight hypothesis. It offers an alternative interpretation to Post's model for population differences in color deficiencies. Principally, it allows high predictability of frequencies with few inferences of cultural history, mode of survival, ethnicity, and mutation rate and offers a mechanism for the evolution of sex-linkage of the traits.

ACKNOWLEDGMENTS

I thank D. D. Beatty, S. Douglas, G. H. Jacobs, W. M. McFarland, A. R. Palmer, and G. Smith for comments on the manuscript and I. Inuma, S. Karlson, R. L. Kirk, R. Lakowski, F. A. Milan, M. G. Norman, J. R. Nursall, F. M. Salzano, and V. G. Zakharova for correspondence. This study was initiated following a field trip with N. Panter to whom I remain grateful. The work was supported in part by the National Research Council of Canada.

REFERENCES


KOLIPOULOS, J., P. IORDANIDES, G. PALIMERIS, and E. CHIMONIDOU. 1976. Data concerning


