

# Different Selective Pressures Shape the Molecular Evolution of Color Vision in Chimpanzee and Human Populations

Brian C. Verrelli,\* Cecil M. Lewis Jr,\*<sup>1</sup> Anne C. Stone,† and George H. Perry†<sup>2</sup>

\*Center for Evolutionary Functional Genomics, The Biodesign Institute and School of Life Sciences, Arizona State University - Tempe; and †School of Human Evolution and Social Change, Arizona State University - Tempe

A population genetic analysis of the long-wavelength opsin (*OPNILW*, “red”) color vision gene in a global sample of 236 human nucleotide sequences had previously discovered nine amino acid replacement single nucleotide polymorphisms, which were found at high frequencies in both African and non-African populations and associated with an unusual haplotype diversity. Although this pattern of nucleotide diversity is consistent with balancing selection, it has been argued that a recombination “hot spot” or gene conversion within and between X-linked color vision genes alone may explain these patterns. The current analysis investigates a closely related primate with trichromatism to determine whether color vision gene amino acid polymorphism and signatures of adaptive evolution are characteristic of humans alone. Our population sample of 56 chimpanzee (*Pan troglodytes*) *OPNILW* sequences shows three singleton amino acid polymorphisms and no unusual recombination or linkage disequilibrium patterns across the ~5.5-kb region analyzed. Our comparative population genetic approach shows that the patterns of *OPNILW* variation in humans and chimpanzees are consistent with positive and purifying selection within the two lineages, respectively. Although the complex role of color vision has been greatly documented in primate evolution in general, it is surprising that trichromatism has followed very different selective trajectories even between humans and our closest relatives.

## Introduction

Through large-scale comparisons of the human and chimpanzee genomes, a number of studies have attempted to identify the genes underlying the evolutionary change between us and our closest primate relatives (e.g., Clark et al. 2003; Bustamante et al. 2005; Chimpanzee Sequencing and Analysis Consortium 2005). Of the genes most evolutionarily diverged, those with sensory perception functions were overrepresented, such as olfactory (Niimura and Nei 2003; Gilad et al. 2003; Nozawa et al. 2007) and bitter-taste receptor genes (Wang et al. 2004; Fischer et al. 2005; Go et al. 2005). Although these analyses capture “ancient” events, comparative population genetic approaches examining sensory perception would reveal how chimpanzee and human lineages have recently diverged (Stone and Verrelli 2006). For example, haplotype analyses of the bitter-taste receptor gene *TAS2R38* in humans are consistent with balancing selection (Wooding et al. 2004), whereas chimpanzee polymorphism shows that they have independently evolved a similar bitter-taste sensitivity (Wooding et al. 2006).

One of the most widely studied examples of adaptive diversification of sensory perception in primates is that of X-linked color vision variation. Trichromatic color vision in catarrhines (Old World monkeys and apes, including humans) is due to an autosomal-linked S-opsin (*OPNISW*, “blue”) gene and an X-linked gene duplication ~30–40 Ma that resulted in L- (*OPNILW*, “red”) and M-opsin (*OPNIMW*, “green”) genes (Nathans et al. 1986; Jacobs et al. 1996). The *OPNILW* and *OPNIMW* genes each code for a 364 amino acid protein (>98% similar). Many have estimated ranges of color vision perception from these pro-

teins’ amino acid sequences given the strong correlation between opsin wavelength absorption maxima ( $\lambda_{\max}$ ) and single amino acid variants. In fact, only ~5 amino acid residues may account for the difference in  $\lambda_{\max}$  between the L- and M-opsins (Merbs and Nathans 1992; Asenjo et al. 1994; Sharpe et al. 1998; Yokoyama and Radlwimmer 2001). The *OPNILW* and *OPNIMW* genes are separated by ~24 kb, and recombination and gene conversion in this array still frequently occurs, resulting in deletions and duplications that are responsible for “red/green” color blindness in ~8% of the male human population (Hayashi et al. 1999; Sharpe et al. 1999). Interestingly, several variants are found within the human L-opsin protein that also have functional, and even behavioral, implications (Merbs and Nathans 1992; Winderickx et al. 1992, 1993; Asenjo et al. 1994; Sharpe et al. 1998; Jameson et al. 2001).

Several studies have proposed that this L-opsin amino acid variation persists because purifying selection maintains only a few specific amino acid residues to distinguish M- and L-opsins (Merbs and Nathans 1992; Winderickx et al. 1992, 1993), and thus, these L-opsin amino acid variants are simply “neutral.” However, in the first DNA sequence analysis of *OPNILW* in a global population sample of 236 human “color normals,” Verrelli and Tishkoff (2004) found an unusual haplotype structure associated only with high-frequency amino acid polymorphisms and observed no fixed amino acid differences since our divergence from chimpanzees, which argued that positive selection has shaped human color vision variation within populations.

Opsin protein variation is found in several primates and is believed to confer color vision perception that is adaptively tuned for foraging efficiency, as in the detection of ripe fruits or immature leaves (e.g., Mollon 1989; Dominy and Lucas 2001; Martin and Ross 2005). Although studies have examined *OPNILW*-coding regions among primates (e.g., Deeb et al. 1994; Tan and Li 1999), there are very few comparative population genetic analyses of exons and introns with tests of selection at primate color vision genes. In fact, we lack these data for even another great ape to compare with the high human *OPNILW* variation. In continuing with our comparative genetic analysis

<sup>1</sup> Present address: Department of Anthropology, University of Oklahoma.

<sup>2</sup> Present address: Department of Human Genetics, University of Chicago.

Key words: opsin, long-wavelength, primates.

E-mail: brian.verrelli@asu.edu.

*Mol. Biol. Evol.* 25(12):2735–2743. 2008

doi:10.1093/molbev/msn220

Advance Access publication October 1, 2008

of primate color vision systems (Verrelli and Tishkoff 2004; Perry et al. 2007), we conducted the first DNA sequence analysis of exons and introns at *OPNILW* in a chimpanzee population to determine the selective forces shaping color vision in our closest primate relative.

## Materials and Methods

### Primate Samples

The major question of interest here involves how color vision gene variation differs between humans and chimpanzees. Thus, although several chimpanzee subspecies are recognized, *Pan troglodytes verus* appears to have nuclear genetic diversity comparable to that of humans (Chimpanzee Sequencing and Analysis Consortium 2005). We have collected a sample of population genetic data from this subspecies and, additionally, a few individuals from other chimpanzee subspecies for comparison. Nucleotide sequences were collected from DNA samples of 45 unrelated chimpanzees: 40 from the western African *P. t. verus*, 3 from the central African *Pan troglodytes troglodytes*, 1 from the eastern African *Pan troglodytes schweinfurthii*, and 1 Nigerian *Pan troglodytes vellerosus* chimpanzee purported to be an additional subspecies (Gonder et al. 1997). Samples were imported in accordance with the Convention on International Trade in Endangered Species of Wild Fauna and Flora under permit 99US013176/9. One gorilla (*Gorilla gorilla gorilla*) was also sampled for our analysis. In addition to these new data, we add 5 *P. t. verus*, 1 *P. t. troglodytes*, and 236 human *OPNILW* nucleotide sequences from Verrelli and Tishkoff (2004).

The human data set includes 163 males from 11 subpopulations within sub-Saharan Africa and 73 males from 8 groups outside of this geographic region (referred to as “African” and “non-African,” respectively). Because *OPNILW* is X linked, sampling males enables the unambiguous determination of polymorphisms (i.e., no heterozygous sites) and haplotype phase. All *Pan troglodytes* were males except for 5 *P. t. verus*. As is the case with our other analyses of these 5 females (e.g., Verrelli et al. 2006), each of these individuals had two different haplotypes that could be resolved empirically (see below). Thus, our chimpanzee analyses reflect a total of 4 *P. t. troglodytes*, 1 *P. t. schweinfurthii*, 1 *P. t. vellerosus*, and 50 *P. t. verus* *OPNILW* sequences.

### *OPNILW* Fragment Amplification and Sequencing

As in humans, the *OPNILW* and *OPNIMW* genes are situated in the same orientation and location on the X chromosome and separated by ~24 kb in chimpanzees. Exons 1 and 6 are identical in nucleotide sequence for *OPNILW* and *OPNIMW* human genes, but exons 2–5 code for amino acid residues that greatly influence  $\lambda_{\max}$  (Merbs and Nathans 1992; Asenjo et al. 1994; Sharpe et al. 1998). As in Verrelli and Tishkoff (2004) and Winderickx et al. (1992), primers in exons 2 and 5 were used to amplify single *OPNILW* copies by polymerase chain reaction (PCR). This ~5,464-bp product (a few indels vary the size), which includes all of intron 2 through a partial exon 5, was prepared for

DNA sequencing using shrimp alkaline phosphatase and exonuclease I (US Biochemicals, Cleveland, OH). All nucleotide sequence data were obtained with an ABI 3730 automated sequencer (Applied Biosystems, Foster City, CA), and sequence files were aligned using the Sequencher v. 4.5 program (Gene Codes, Ann Arbor, MI). Exon 5 codes for two major amino acid differences that account for the largest differences between the L- and M-cone opsin  $\lambda_{\max}$  and are used to predict functional absorbance from inferred amino acid sequence: the Tyr277Phe (~7 nm) and Thr285Ala (~14 nm) variants (Neitz et al. 1991; Yokoyama and Radlwimmer 2001). Expressed L-cone opsins have the Tyr277 and Thr285 residues (Deeb et al. 1992); and thus, all PCR fragments were first analyzed for their exon 5 nucleotide sequence to verify that they were only L-cone opsins. In the case of the aforementioned 5 *P. t. verus* females, the ~5.5-kb product was cloned using the Invitrogen TOPO-XL kit and a clone was sequenced along with the initial PCR product. All single nucleotide polymorphisms (SNPs) identified in the cloned sequence that did not appear in the initial PCR sequence were considered cloned artifacts and ignored.

### Data Analysis and Statistical Tests

Because our data collection reflects a random sample of chimpanzees with respect to *OPNILW* genetic variation (i.e., we have no a priori phenotypic evidence for color vision variation), all statistics and neutrality tests should reflect the frequencies of variants found in the natural population. There has been little genetic evidence to date to suggest any significant underlying demographic history within chimpanzee subspecies groups, such as population expansion or substructuring. However, human population samples from sub-Saharan Africa and outside this geographic region show markedly different genetic patterns with higher nucleotide diversity, less linkage disequilibrium (LD), and a greater effective population size ( $N_e$ ) for the former compared with the latter (e.g., Tishkoff and Verrelli 2003a). Therefore, we present all estimates of nucleotide diversity from our African and non-African humans separately to appropriately compare with our chimpanzee sample. We also present population genetic parameter estimates for *P. troglodytes* overall as well as comparisons among the subspecies.

We have used the Rozas et al. (2003) DnaSP v. 4.1 program for our population and comparative species genetic analyses. “Silent” SNPs (which refer to introns and synonymous sites in exons, throughout the study) do not alter the amino acid sequence and, here, are used to best reflect neutral evolution in all tests. Locus-specific estimates of the population parameter  $\theta = 3N_e\mu$  (for X-linked genes) were calculated using the number of SNPs as  $\theta_W$  of Watterson (1975), which is corrected by sample size through a coalescent estimation. These measures were compared with the estimate  $\theta_\pi$ , which is based on the average number of pairwise differences among all sequences, and thus, sample size is also taken into consideration. These two estimates,  $\theta_W$  and  $\theta_\pi$ , are expected to be equal under neutrality. This can be assessed using Tajima’s (1989) *D* test, which examines the SNP frequency spectrum

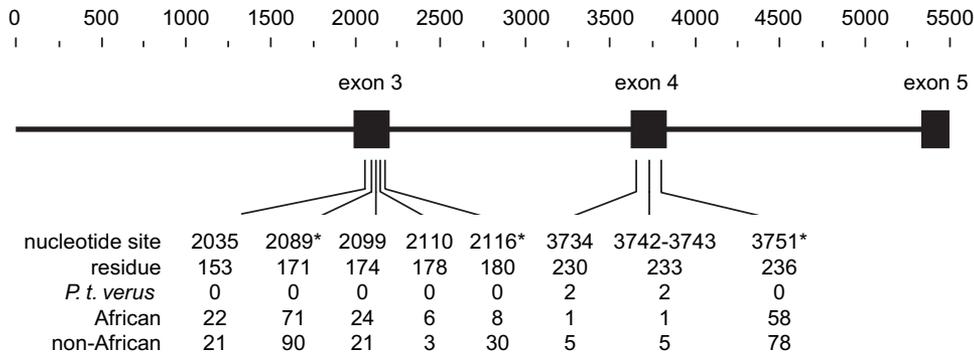


FIG. 1.—Diagram of the  $\sim 5.5$  kb of the *OPNILW* gene region analyzed in humans and chimpanzees. Exon 3 (169 bp), exon 4 (166 bp), and partial exon 5 (121 bp) are shown as black boxes with introns 2 (1,987 bp), 3 (1,467 bp), and 4 (1,554 bp) separating them. Starting at the beginning of intron 2, the nucleotide sites of the nine human amino acid replacement SNPs are shown below the gene with their respective amino acid residue positions and their frequencies (%) in population samples of *Pan troglodytes verus* chimpanzees ( $n = 50$ ), African (163), and non-African (73) humans. Asterisks denote significant differences between the latter two samples using a chi-square statistical test ( $P < 0.01$ ).

and evaluates whether the number of SNPs are higher or lower than expected, given their frequencies and the number of sequences sampled. Positive  $D$  values are consistent with balancing selection, whereas negative  $D$  values imply directional selection. Changes in demographic history, such as population structure or expansion, may also be consistent with positive or negative  $D$  values, respectively; thus, the statistical significance of all measures of diversity and neutrality tests were computed using coalescent simulations and permutations in DnaSP. For example, we simulated 1,000 genealogies with observed values of  $\theta_w$  to estimate how often we would find the observed values of  $\theta_\pi$  and Tajima's  $D$  by chance alone (i.e., neutrality), given the sample sizes of humans and chimpanzees in our analysis. In using coalescent simulations to determine statistical significance, we effectively account for all differences in sample sizes across humans and chimpanzees that may otherwise effect statistical tests of polymorphism data.

To examine the haplotype structure at *OPNILW*, we calculated LD and estimated the recombination parameter  $\rho = 3N_c c$  (where  $c$  is the recombination rate per base pair) across the  $\sim 5.5$ -kb region in chimpanzees for comparison with humans. We used the LDhat program of McVean et al. (2002), which applies the approximate likelihood method of Hudson (2001), to estimate  $\rho$ . LDhat conducts a permutation analysis to determine whether pairwise comparisons among SNPs exhibit significantly more or less LD given the locus-specific estimate of  $\rho$  and SNP frequencies. As determined by LDhat, uninformative SNPs (i.e., rare) were omitted for these analyses.

Finally, to examine older lineage-specific changes in selective pressures, we compared chimpanzee *OPNILW* polymorphism at silent and replacement sites with fixation between chimpanzee and human *OPNILW* at silent and replacement sites with the McDonald and Kreitman (1991) test of neutrality. If amino acid replacement sites are evolving according to neutrality, we may expect that the ratio of replacement polymorphism to fixation will approximate this same ratio at silent sites. The *OPNILW* sequence from the single gorilla individual was used to

facilitate the polarization of fixed differences within human and chimpanzee lineages.

## Results

### Chimpanzee *OPNILW* Diversity

In humans, single X chromosomes can have more than one *OPNILW* copy (e.g., Sjöberg et al. 1998). If there are SNPs among paralogous *OPNILW* copies, they would be easily detectable from our nucleotide sequences given that the large majority of our sampled individuals are males. For example, the presence of multiple *OPNILW* gene copies with different nucleotide sequences would be apparent from “heterozygous” sites (paralogous sequence SNPs). Such duplications were previously seen in  $\sim 5\%$  of human PCR amplifications, and these individuals were omitted from all analyses (Verrelli and Tishkoff 2004). However, in the current study, we did not observe a single instance of heterozygous sites within chimpanzee individuals.

Figure 1 shows a schematic of the *OPNILW* gene and the region sequenced for our chimpanzee and human population samples. Our analysis involves 456 bp in exons 3–5 and 5,008 bp in introns 2–4 that have a total of 5,118 effectively neutral silent sites. In our sample of 56 chimpanzee chromosomes, we find 42 variants, which include an 11-bp deletion in intron 2, a 2-bp deletion, and a 1-bp insertion in intron 3 (where insertion/deletion status was deduced from comparison with human sequence). Of the 39 SNPs, there are one silent and three replacement SNPs in exon 4, which are found as “singletons” in our population sample, and involve two amino acid changes in a single individual (Ile230Thr, site 3734 and Ala233Ser, sites 3742–3743; figs. 1 and 2). Otherwise, the remaining 35 SNPs are distributed over introns.

For silent site diversity at *OPNILW*, chimpanzees are far less variable than humans. Humans show silent site diversity ( $\pi = 0.31\%$ , after multiplying by 4/3 for it being X linked) that is over three times that found in the nuclear genome on average, and given the relatively few replacement sites in exons ( $\sim 300$ ), the nine replacement SNPs are



**Table 2**  
**Divergence at *OPNILW* Silent Sites between Chimpanzees and Humans**

Sample	<i>Pan troglodytes troglodytes</i>	<i>Pan troglodytes schweinfurthii</i>	<i>Pan troglodytes verus</i>	<i>Pan troglodytes vellerosus</i>
<i>P. t. schweinfurthii</i>	0.191			
<i>P. t. verus</i>	0.172	0.093		
<i>P. t. vellerosus</i>	0.151	0.117	0.070	
Human	1.221	1.182	1.163	1.182

NOTE.—Average pairwise SNP differences (%) among groups (see the Results section).

subspecies, but when small samples of each (e.g., one for each of *P. t. schweinfurthii* and *P. t. vellerosus*) are combined into one sample, it appears that these SNPs are singletons at low frequencies. In fact, when the *P. t. verus* sample as well as the smaller sample of *P. t. troglodytes* is analyzed alone, no excess of rare alleles is seen (table 1). When analyzed separately, African and non-African human samples also show no significant skew toward rare frequency silent SNPs. Altogether, it appears that silent site diversity in humans and chimpanzees is consistent with neutrality.

In contrast to silent site diversity, humans and chimpanzees show very different patterns of variation for amino acid replacement sites. Although both species possess replacement polymorphisms, these SNPs in *P. t. verus* are very rare in frequency (all three found as singletons in one individual, fig. 2), which results in a significantly high negative Tajima's *D* value. In contrast, human replacement SNPs are not only high in number but several are high in frequency (table 1 and fig. 1), which results in a significantly high positive Tajima's *D* under a standard neutral model of human population expansion (Verrelli and Tishkoff 2004).

#### Recombination and LD Analyses

Analyses of LD and recombination are only truly valid when conducted for within-population samples; and thus, we included only our samples of *P. t. verus* and humans for these comparisons. In addition, the amount of variation found within and between the chimpanzee subspecies at *OPNILW* also warrants that these groups be analyzed separately, as is the case for African and non-African human samples. For chimpanzee *OPNILW*, only 8 pairwise comparisons in the *P. t. verus* sample of 16 SNPs were significantly correlated given the overall maximum likelihood (ML) estimate of recombination at *OPNILW*, which finds very little evidence of historical crossing-over across the ~5.5-kb region ( $\rho \approx 0$ ). These few correlations were not clustered in any specific gene region and include SNPs across introns 2–4. Thus, no unusual pattern of LD or recombination was apparent in the chimpanzee sample.

This pattern in chimpanzees is in sharp contrast to the estimates of LD and recombination found in humans. As previously noted, overall, there is far more diversity in the human sample, and this is associated with ML estimates of recombination that are also very high in both Africans and non-Africans ( $\rho > 100$ ). Given these estimates, it is not surprising that only 56 of the >700 African pairwise comparisons and 25 of the >300 non-African pairwise comparisons show significant LD, all of which are spread across introns 2–4. Surprisingly, given the overall high levels of

gene-specific recombination, significantly high levels of localized recombination were still detected in both human population samples as a hot spot in the 169 bp of exon 3. In fact, 125 of the African and 59 of the non-African correlations had significantly less LD given the gene-specific estimate of  $\rho$ , yet 86 and 49 of these correlations, respectively, are clustered in only an 86-bp region centered on exon 3. The HOTSPOTTER program designed by Li and Stephens (2003) was used to show that this clustering was significant; the haplotype diversity in this region appears to have recombination rates that are 20 and 5 times greater in Africans and non-Africans, respectively, compared with the rest of the ~5.5 kb (Verrelli and Tishkoff 2004).

#### *OPNILW* Interspecific Analyses

Silent site fixation at *OPNILW* (45 differences, 0.9%) is similar to other nuclear gene comparisons between humans and chimpanzees (1.1%; Fischer et al. 2004; Wooding et al. 2005; Chimpanzee Sequencing and Analysis Consortium 2005; Verrelli et al. 2006), yet there is no amino acid fixation. In spite of this, both chimpanzees and humans show high levels of amino acid polymorphism, and analyses of *P. t. verus* as well as African and non-African human samples show significant McDonald–Kreitman tests of neutrality (table 1). However, measures of diversity ( $\pi$ ) and SNP frequency distributions (Tajima's *D*) of replacement SNPs in the two species indicate that, although unexpectedly high in number, these SNPs are demonstrably rare in frequency for chimpanzees but very common for human populations (table 1).

#### Discussion

Amino acid replacement polymorphisms are not common for the human genome in general, and yet when present, are often associated with low frequencies or even singletons that imply at least weak purifying selection (i.e., Ohta 1992) acting across populations (e.g., Bustamante et al. 2005; Boyko et al. 2008; Lohmueller et al. 2008). Focus remains on those genes that show different patterns of amino acid polymorphism across populations and species and thus point to phenotypes under positive selection that is unique to the human or chimpanzee lineages. One example is of the human *G6PD* locus that shows several amino acid polymorphisms that have recently risen to high frequencies with strong LD in geographic regions of malarial endemism (Tishkoff et al. 2001; Verrelli et al. 2002; Saunders et al. 2005). Yet, *G6PD* in chimpanzees and all other great apes shows strong

purifying selection and no significant LD and unusual haplotype structure, thus implying that positive selection for malarial resistance has targeted this locus recently only in humans (Verrelli et al. 2006).

In this current study, we have examined another phenotype, trichromatism, which shows very different patterns of genetic diversity between humans and chimpanzees. Some have suggested that the abundance of amino acid polymorphism at human *OPNILW* is the result of gene conversion (Winderickx et al. 1993) and not necessarily adaptive evolution. Although Verrelli and Tishkoff (2004) used human polymorphism data to address this question, the current comparative study tests a second hypothesis. If this pattern in humans is simply due to structural gene organization (i.e., gene conversion) and weak purifying selection, given that chimpanzees may also experience gene conversion in this tandem gene array (e.g., Zhou and Li 1996), we may expect to see patterns of amino acid variation in chimpanzees similar to those in humans.

First, if gene conversion is the sole explanation for *OPNILW* diversity, then we may expect that both silent and amino acid polymorphism, and fixation, show similar patterns. In spite of typical silent diversity and fixation, chimpanzee replacement SNPs are singletons found on only one haplotype (fig. 2). These two amino acid polymorphisms, at residues 230 and 233, are similarly found only together in 1% and 5% of haplotypes in African and non-African humans, respectively (fig. 1), and are identical to that found at *OPNIMW*. That is, this identical “block” of replacement SNPs found rarely in both humans and chimpanzees is clearly the result of single-gene conversion events that transport this variation from *OPNIMW* to *OPNILW*. As is true in humans, we find that chimpanzee introns at *OPNILW* are more similar to *OPNIMW* than they are to their *OPNILW* homologs in other primates, which is the most obvious impact of long-term gene conversion. However, in contrast to chimpanzees, humans have nine replacement SNPs whose derived allele frequencies range from 1% to 90% (fig. 1), showing a discrepancy in how gene conversion alters amino acid variation in the two species. It is possible that gene conversion is simply higher in humans, yet we may not expect *OPNILW* silent divergence to be higher in chimpanzees than in humans (31 vs. 14, with gorilla as outgroup). In addition, several high-frequency human *OPNILW* replacement SNPs are not found in *OPNIMW* (Winderickx et al. 1993), which is also unexpected given that gene conversion should be unbiased if it is simply a neutral mechanism that drives variation.

Another signature of gene conversion is the unusual pattern of haplotype diversity that extends across exon 3 in humans. Recombination hot spots within and between human and chimpanzee genomes have recently been identified (Wall et al. 2003; Crawford et al. 2004; Ptak et al. 2004, 2005; Myers et al. 2005; Winckler et al. 2005; Coop et al. 2008). It is possible that these hot spots over small areas (<1 kb) are the result of gene conversion that shuffles small tracks preferentially (Ardlie et al. 2001; Frisse et al. 2001; Przeworski and Wall 2001; Reich et al. 2002; Wall 2004). At *OPNILW*, the entire ~5.5-kb region shows no dramatic pattern of LD and no recombination rate variation in chimpanzees. Furthermore, the fact that recombination

rates in the Xq28 region are generally low (e.g., Ardlie et al. 2001; Frisse et al. 2001; Matisse et al. 2007) indicates that the observed pattern in humans, and not chimpanzees, is unusual.

It is difficult to reconcile how a neutral substitution model could explain the lack of amino acid fixation between the two species in the last 5–7 My since their divergence, the observation of only three replacement singletons in chimpanzees, as well as explain several high-frequency replacement SNPs in human populations. It is also difficult to employ a neutral model of gene conversion or a recombination hot spot to explain the haplotype structure associated with human replacement SNPs, given only one replacement haplotype from gene conversion and the lack of unusual LD in chimpanzees. Although it is apparent that gene conversion occurs in both species in introns and exons, replacement SNPs are found at different frequencies in human and chimpanzee *OPNILW*, many of which have measured functional affects (Merbs and Nathans 1992; Asenjo et al. 1994; Sharpe et al. 1998). Based on our data, it appears that certain residues that are implicated in changes to opsin  $\lambda_{\max}$  and spectral tuning, especially 277 and 285 (>7–14 nm) and to a lesser degree 230 and 233 (<5 nm), are under purifying selection in both species. This is unlike other amino acid sites, such as the 180 variant in humans, which although may alter the  $\lambda_{\max}$  < 5 nm, shows very different patterns of frequency and haplotype variation in the two species. Thus, several pieces of information support the conclusion that the different patterns of amino acid variation in humans and chimpanzees are the result of selective and not neutral processes alone. In fact, as our comparative population genetic studies of human and chimpanzees mount, like *G6PD* and *OPNILW*, it will be interesting to see how often unusual haplotype diversity in the genome is more often explained by different selective pressures and not by varying recombination rates across species (Tishkoff and Verrelli 2003b; Wall and Pritchard 2003; Reed and Tishkoff 2006; Verrelli et al. 2006; Clark et al. 2007).

Data collected for *OPNIMW* show very little amino acid polymorphism in humans and no amino acid fixation between humans and chimpanzees (Winderickx et al. 1993; Deeb et al. 1994; Verrelli BC, unpublished data), demonstrating the effects of strong purifying selection. However, *OPNIMW* copy number among humans is highly variable, with individuals having 1–5 copies (Sjoberg et al. 1998; Sharpe et al. 1999; Carroll et al. 2002). Thus, population genetic parameters (as estimated in the current study) would be inappropriate because assumptions of homology for *OPNIMW* gene copies among individuals would be invalidated. As previously discussed, humans rarely show evidence for *OPNILW* duplications (Sjoberg et al. 1998; Verrelli and Tishkoff 2004); yet, there has been no published study of *OPNILW* copy number variation in other primates. In general, studies of copy number variation have only recently shown the magnitude of this diversity within and between human and chimpanzee genomes (Cheng et al. 2005; Perry et al. 2006; Redon et al. 2006). In an additional study, we have examined copy number variation at the X-linked opsin region in 30 human and 30 chimpanzee males with a high-resolution microarray-based comparative

genomic hybridization platform comprised of ~385,000 X-chromosome oligonucleotide probes (median spacing = 347 bp; Perry GH, unpublished data). These data suggest that single X chromosomes in humans are variable not only for the number of *OPN1MW* gene copies but are also consistent with some individuals having more than one *OPN1LW* copy. In contrast, there was not a single chimpanzee X chromosome with copy number variation of either opsin gene, which confirms our PCR results in showing no *OPN1LW* heterozygous sites within chimpanzees. Interestingly, this contrasting copy number variation between humans and chimpanzees may indicate yet another level upon which selection differentially acts within species at X-linked opsin genes.

The global human analysis by Verrelli and Tishkoff (2004) first proposed that human *OPN1LW* replacement SNPs may be explained by historical balancing selection for spectral tuning. For example, the amino acid variant at residue 180 in humans that reaches frequencies ~30% in Africans (fig. 1) but which is not found in chimpanzees has been hypothesized to shift L-opsin  $\lambda_{\max}$  further into the “red/orange” portion of visible light (Asenjo et al. 1994; Sharpe et al. 1998; Carroll et al. 2002). As *OPN1LW* is X linked, balancing selection for heterozygotes would have only occurred in females and lead to “tetrachromacy” (blue, green, red, and red/orange opsins) as some have speculated (e.g., Jameson et al. 2001). In New World monkeys, where only one X-linked opsin gene exists and males are dichromatic, balancing selection favors heterozygous females that carry two different functional alleles because it effectively makes them trichromatic. Therefore, although it is possible that further variation may be under balancing selection in female humans, it is interesting to question why this variation has not been adaptive in chimpanzees. It is possible that this color vision variation only recently became adaptive for sex-specific behavior in human hunter-gatherer societies, especially for the need to distinguish among colored fruits and other food items on shaded backgrounds (Mollon et al. 1984; Dominy and Lucas 2001; Regan et al. 2001; Lucas et al. 2003; Surridge et al. 2003).

Finally, it has been proposed that sexual selection and mate recognition may have further shaped red color vision intraspecific variation (Fernandez and Morris 2007). Thus, specifically L-opsin spectral tuning may be responsible for adaptive changes in behavior among populations. This hypothesis is strengthened by the functional opsin variation between African and non-African populations as evidenced by the significant differences in L-opsin amino acid haplotype frequencies (Verrelli and Tishkoff 2004) and L- to M-opsin expression ratios (McMahon et al. 2008) found between these two groups. As our molecular population genetic analyses of other closely and distantly related primates continue (e.g., Perry et al. 2007), we can reveal how color vision became phenotypically unique during human and primate evolution in general.

## Acknowledgments

Chimpanzee samples were provided by H. Vasken Aposhian, the Detroit Zoological Institute (Detroit, MI),

the Lincoln Park Zoo (Chicago, IL), the Primate Foundation of Arizona (Mesa, AZ), the Riverside Zoo (Scottsbluff, NE), the Southwest Foundation for Biomedical Research (San Antonio, TX), the Sunset Zoo (Manhattan, KS), the Welsh Mountain Zoo (North Wales, UK), the Yerkes National Primate Research Center (Atlanta, GA), and the New Iberia Research Center (Lafayette, LA), where samples were supported by a National Institutes of Health–National Center for Research Resources grant U42-RR015087 to the University of Louisiana at Lafayette. This work was supported by the Center for Evolutionary Functional Genomics in The Biodesign Institute and the School of Life Sciences at Arizona State University and National Science Foundation grant BCS-0715972 (to B.C.V. and A.C.S.).

## Literature Cited

- Ardlie K, Liu-Cordero SN, Eberle MA, Daly M, Barrett J, Winchester E, Lander ES, Kruglyak L. 2001. Lower-than-expected linkage disequilibrium between tightly linked markers in humans suggests a role for gene conversion. *Am J Hum Genet.* 69:582–589.
- Asenjo AB, Rim J, Oprian DD. 1994. Molecular determinants of human red/green color discrimination. *Neuron.* 12:1131–1138.
- Boyko AR, Williamson SH, Indap AR, et al. (14 co-authors). 2008. Assessing the evolutionary impact of amino acid mutations in the human genome. *PLoS Genet.* 4(5):e1000083.
- Bustamante CD, Fledel-Alon A, Williamson S, et al. (14 co-authors). 2005. Natural selection on protein-coding genes in the human genome. *Nature.* 437:1153–1157.
- Carroll J, Neitz J, Neitz M. 2002. Estimates of L:M cone ratio from ERG flicker photometry and genetics. *J Vis.* 2:531–542.
- Cheng Z, Ventura M, She X, et al. (12 co-authors). 2005. A genome-wide comparison of recent chimpanzee and human segmental duplications. *Nature.* 437:88–93.
- Chimpanzee Sequencing and Analysis Consortium. 2005. Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature.* 437:69–87.
- Clark AG, Gnanowski S, Nielsen R, et al. (17 co-authors). 2003. Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios. *Science.* 302:1960–1963.
- Clark VJ, Ptak SE, Tiemann I, Qian Y, Coop G, Stone AC, Przeworski M, Arnheim N, Di Rienzo A. 2007. Combining sperm typing and linkage disequilibrium analyses reveals differences in selective pressures or recombination rates across human populations. *Genetics.* 175:795–804.
- Crawford DC, Bhangale T, Li N, Hellenthal G, Rieder MJ, Nickerson DA, Stephens M. 2004. Evidence for substantial fine-scale variation in recombination rates across the human genome. *Nat Genet.* 36:700–706.
- Coop G, Wen X, Ober C, Pritchard JK, Przeworski M. 2008. High-resolution mapping of crossovers reveals extensive variation in fine-scale recombination patterns among humans. *Science.* 319:1395–1398.
- Deeb SS, Jorgensen AL, Battisti L, Iwasaki L, Motulsky AG. 1994. Sequence divergence of the red and green visual pigments in great apes and humans. *Proc Natl Acad Sci USA.* 91:7262–7266.
- Deeb SS, Lindsey DT, Hibiya Y, Sanocki E, Winderickx J, Teller DY, Motulsky AG. 1992. Genotype-phenotype relationships in human red/green color-vision defects: molecular and psychophysical studies. *Am J Hum Genet.* 51:687–700.

- Deinard A, Kidd K. 1999. Evolution of a HOXB6 intergenic region within the great apes and humans. *J Hum Evol.* 36:687–703.
- Dominy NJ, Lucas PW. 2001. Ecological importance of trichromatic vision to primates. *Nature.* 410:363–366.
- Fernandez AA, Morris MR. 2007. Sexual selection and trichromatic color vision in primates: statistical support for the preexisting-bias hypothesis. *Am Nat.* 170:10–20.
- Fischer A, Gilad Y, Man O, Pääbo S. 2005. Evolution of bitter taste receptors in humans and apes. *Mol Biol Evol.* 22:432–436.
- Fischer A, Wiebe V, Paabo S, Przeworski M. 2004. Evidence for a complex demographic history of chimpanzees. *Mol Biol Evol.* 21:799–808.
- Frisse L, Hudson RR, Bartoszewicz A, Wall JD, Donfack J, Di Rienzo A. 2001. Gene conversion and different population histories may explain the contrast between polymorphism and linkage disequilibrium levels. *Am J Hum Genet.* 69:831–843.
- Gilad Y, Bustamante CD, Lancet D, Paabo S. 2003. Natural selection on the olfactory receptor gene family in humans and chimpanzees. *Am J Hum Genet.* 73:489–501.
- Go Y, Satta Y, Takenaka O, Takahata N. 2005. Lineage-specific loss of function of bitter taste receptor genes in humans and nonhuman primates. *Genetics.* 170:313–326.
- Gonder MK, Oates JF, Disotell TR, Forstner MRJ, Morales JC, Melnick DJ. 1997. A new west African chimpanzee subspecies? *Nature.* 388:337.
- Hayashi T, Motulsky AG, Deeb SS. 1999. Position of a 'green-red' hybrid gene in the visual pigment array determines colour-vision phenotype. *Nat Genet.* 22:90–93.
- Hudson RR. 2001. Two-locus sampling distributions and their application. *Genetics.* 159:1805–1817.
- Jacobs GH, Neitz M, Deegan JF, Neitz J. 1996. Trichromatic colour vision in New World monkeys. *Nature.* 382:156–158.
- Jameson KA, Highnote SM, Wasserman LM. 2001. Richer color experience in observers with multiple photopigment opsin genes. *Psychon Bull Rev.* 8:244–261.
- Kaessmann H, Wiebe V, Paabo S. 1999. Extensive nuclear DNA sequence diversity among chimpanzees. *Science.* 286:1159–1162.
- Li N, Stephens M. 2003. Modeling linkage disequilibrium and identifying recombination hotspots using single-nucleotide polymorphism data. *Genetics.* 165:2213–2233.
- Lohmueller KE, Indap AR, Schmidt S, et al. (12 co-authors). 2008. Proportionally more deleterious genetic variation in European than in African populations. *Nature.* 451:994–997.
- Lucas PW, Dominy NJ, Riba-Hernandez P, et al. (12 co-authors). 2003. Evolution and function of routine trichromatic vision in primates. *Evolution.* 57:2636–2643.
- Martin RD, Ross CF. 2005. The evolutionary and ecological context of primate vision. In: Kremers J, Silveira L, Martin P, editors. *Structure, function, and evolution of the primate visual system.* New York: John Wiley. p. 1–36.
- Matise TC, Chen F, Chen W, et al. (13 co-authors). 2007. A second-generation combined linkage physical map of the human genome. *Genome Res.* 17:1783–1786.
- Merbs SL, Nathans J. 1992. Absorption spectra of human cone pigments. *Nature.* 356:433–435.
- McDonald JH, Kreitman M. 1991. Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature.* 351:652–654.
- McMahon C, Carroll J, Awua S, Neitz J, Neitz M. 2008. The L:M cone ratio in males of African descent with normal color vision. *J Vis.* 8:1–9.
- McVean G, Awadalla P, Fearnhead P. 2002. A coalescent-based method for detecting and estimating recombination from gene sequences. *Genetics.* 160:1231–1241.
- Mollon JD. 1989. "Tho' she kneel'd in that place where they grew...": The uses and origins of primate colour vision. *J Exp Biol.* 146:21–38.
- Mollon JD, Bowmaker JK, Jacobs GH. 1984. Variations of color-vision in a New World primate can be explained by polymorphism of retinal photopigments. *Proc R Soc Lond B Biol Sci.* 222:373–399.
- Myers S, Bottolo L, Freeman C, McVean G, Donnelly P. 2005. A fine-scale map of recombination rates and hotspots across the human genome. *Science.* 310:321–324.
- Nathans J, Thomas D, Hogness DS. 1986. Molecular genetics of human color vision: the genes encoding blue, green, and red pigments. *Science.* 232:193–202.
- Neitz M, Neitz J, Jacobs GH. 1991. Spectral tuning of pigments underlying red-green color vision. *Science.* 252:971–974.
- Niimura Y, Nei M. 2003. Evolution of olfactory receptor genes in the human genome. *Proc Natl Acad Sci USA.* 100:12235–12240.
- Nozawa M, Kawahara Y, Nei M. 2007. Genomic drift and copy number variation of sensory receptor genes in humans. *Proc Natl Acad Sci USA.* 104:20421–20426.
- Ohta T. 1992. The nearly neutral theory of molecular evolution. *Annu Rev Ecol Syst.* 23:263–286.
- Perry GH, Martin RD, Verrelli BC. 2007. Signatures of functional constraint at aye-aye opsin genes: the potential of adaptive color vision in a nocturnal primate. *Mol Biol Evol.* 24:1963–1970.
- Perry GH, Tchinda J, McGrath SD, et al. (12 co-authors). 2006. Hotspots for copy number variation in chimpanzees and humans. *Proc Natl Acad Sci USA.* 103:8006–8011.
- Przeworski M, Wall JD. 2001. Why is there so little intragenic linkage disequilibrium in humans? *Genet Res.* 77:143–151.
- Ptak SE, Hinds DA, Koehler K, Nickel B, Patil N, Ballinger DG, Przeworski M, Frazer KA, Pääbo S. 2005. Fine scale recombination patterns differ between chimpanzees and humans. *Nat Genet.* 37:429–434.
- Ptak SE, Roeder AD, Stephens M, Gilad Y, Pääbo S, Przeworski M. 2004. Absence of the TAP2 human recombination hotspot in chimpanzees. *PLoS Biol.* 2:e155.
- Redon R, Ishikawa S, Fitch KR, et al. (43 co-authors). 2006. Global variation in copy number in the human genome. *Nature.* 444:444–454.
- Reed FA, Tishkoff SA. 2006. Positive selection can create false hotspots of recombination. *Genetics.* 172:2011–2014.
- Regan BC, Julliot C, Simmen B, Vienot F, Charles-Dominique P, Mollon JD. 2001. Fruits, foliage and the evolution of primate colour vision. *Philos Trans R Soc Lond B Biol Sci.* 356:229–283.
- Reich DE, Schaffner SF, Daly MJ, McVean G, Mullikin JC, Higgins JM, Richter DJ, Lander ES, Altshuler D. 2002. Human genome sequence variation and the influence of gene history, mutation and recombination. *Nat Genet.* 32:135–142.
- Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics.* 19:2496–2497.
- Saunders MA, Slatkin M, Garner C, Hammer MF, Nachman MW. 2005. The extent of linkage disequilibrium caused by selection on G6PD in humans. *Genetics.* 171:1219–1229.
- Sharpe LT, Stockman A, Jagle H, Knau H, Klausen G, Reitner A, Nathans J. 1998. Red, green, and red-green hybrid pigments in the human retina: correlations between deduced protein sequences and psychophysically measured spectral sensitivities. *J Neurosci.* 18:10053–10069.

- Sharpe LT, Stockman A, Jagle H, Nathans J. 1999. Opsin genes, cone photopigments, color vision, and color blindness. In: Gegenfurtner KR, Sharpe LT, editors. *Color vision: from genes to perception*. Cambridge. p. 3–51.
- Sjoberg SA, Neitz M, Balding SD, Neitz J. 1998. L-cone pigment genes expressed in normal colour vision. *Vision Res.* 38:3213–3219.
- Stone AC, Griffiths RC, Zegura SL, Hammer MF. 2002. High levels of Y-chromosome nucleotide diversity in the genus Pan. *Proc Natl Acad Sci USA.* 99:43–48.
- Stone AC, Verrelli BC. 2006. Focusing on comparative ape population genetics in the post-genomic age. *Curr Opin Genet Dev.* 16:586–591.
- SurrIDGE AK, Osorio D, Mundy NI. 2003. Evolution and selection of trichromatic vision in primates. *Trends Ecol Evol.* 18:198–205.
- Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics.* 123:585–595.
- Tan Y, Li W-H. 1999. Trichromatic vision in prosimians. *Nature.* 402:36.
- Tishkoff SA, Varkonyi R, Cahinhinan N, et al. (17 co-authors). 2001. Haplotype diversity and linkage disequilibrium at human G6PD: recent origin of alleles that confer malarial resistance. *Science.* 293:455–462.
- Tishkoff SA, Verrelli BC. 2003a. Patterns of human genetic diversity: implications for human evolutionary history and disease. *Annu Rev Genomics Hum Genet.* 4:293–340.
- Tishkoff SA, Verrelli BC. 2003b. Role of evolutionary history on haplotype block structure in the human genome: implications for disease mapping. *Curr Opin Genet Dev.* 13:569–575.
- Verrelli BC, McDonald JH, Argyropoulos G, Destro-Bisol G, Froment A, Drousiotou A, Lefranc G, Helal AN, Loiselet J, Tishkoff SA. 2002. Evidence for balancing selection from nucleotide sequence analyses of human *G6PD*. *Am J Hum Genet.* 71:1112–1128.
- Verrelli BC, Tishkoff SA. 2004. Signatures of selection and gene conversion associated with human color vision variation. *Am J Hum Genet.* 75:363–375.
- Verrelli BC, Tishkoff SA, Stone AC, Touchman JW. 2006. Contrasting histories of G6PD molecular evolution and malarial resistance in humans and chimpanzees. *Mol Biol Evol.* 23:1592–1601.
- Wall JD. 2004. Close look at gene conversion hot spots. *Nat Genet.* 36:114–115.
- Wall JD, Frisse LA, Hudson RR, Di Rienzo A. 2003. Comparative linkage-disequilibrium analysis of the beta-globin hotspot in primates. *Am J Hum Genet.* 73:1330–1340.
- Wall JD, Pritchard JK. 2003. Haplotype blocks and linkage disequilibrium in the human genome. *Nat Rev Genet.* 4:587–597.
- Wang X, Thomas SD, Zhang J. 2004. Relaxation of selective constraint and loss of function in the evolution of human bitter taste receptor genes. *Hum Mol Genet.* 13:2671–2678.
- Watterson GA. 1975. On the number of segregating sites in genetic models without recombination. *Theor Popul Biol.* 7:256–276.
- Winckler W, Myers SR, Richter DJ, et al. (11 co-authors). 2005. Comparison of fine-scale recombination rates in humans and chimpanzees. *Science.* 308:107–111.
- Winderickx J, Battisti L, Hibiya Y, Motulsky AG, Deeb SS. 1993. Haplotype diversity in the human red and green opsin genes: evidence for frequent sequence exchange in exon 3. *Hum Mol Genet.* 2:1413–1421.
- Winderickx J, Lindsey DT, Sanocki E, Teller DY, Motulsky AG, Deeb SS. 1992. Polymorphism in red photopigment underlies variation in colour matching. *Nature.* 356:431–433.
- Wooding S, Bufe B, Grassi C, Howard MT, Stone AC, Vazquez M, Dunn DM, Meyerhof W, Weiss RB, Bamshad MJ. 2006. Independent evolution of bitter-taste sensitivity in humans and chimpanzees. *Nature.* 440:930–934.
- Wooding S, Kim UK, Bamshad MJ, Larsen J, Jorde LB, Drayna D. 2004. Natural selection and molecular evolution in PTC, a bitter-taste receptor gene. *Am J Hum Genet.* 74:637–646.
- Wooding S, Stone AC, Dunn DM, Mummidi S, Jorde LB, Weiss RK, Ahuja S, Bamshad MJ. 2005. Contrasting effects of natural selection on human and chimpanzee CC chemokine receptor 5. *Am J Hum Genet.* 76:291–301.
- Yokoyama S, Radlwimmer FB. 2001. The molecular genetics and evolution of red and green color vision in vertebrates. *Genetics.* 158:1697–1710.
- Yu N, Jensen-Seaman MI, Chemnick L, Kidd JR, Deinard AS, Ryder O, Kidd KK, Li WH. 2003. Low nucleotide diversity in chimpanzees and bonobos. *Genetics.* 164:1511–1518.
- Zhou YH, Li WH. 1996. Gene conversion and natural selection in the evolution of X-linked color vision genes in higher primates. *Mol Biol Evol.* 13:780–783.

Adriana Briscoe, Associate Editor

Accepted September 25, 2008