Am. J. Hum. Genet. 66:1351–1361, 2000

Evidence for Variable Selective Pressures at MC1R

Rosalind M. Harding,[1] Eugene Healy,[2,*] Amanda J. Ray,[2] Nichola S. Ellis,[2] Niamh Flanagan,[2] Carol Todd,[2] Craig Dixon,[2] Antti Sajantila,[3] Ian J. Jackson,[4] Mark A. Birch-Machin,[2] and Jonathan L. Rees[2,†]

[1]Institute of Molecular Medicine, Oxford; [2]Department of Dermatology, University of Newcastle, Newcastle-upon-Tyne; [3]Department of Forensic Medicine, University of Helsinki, Helsinki; and [4]MRC Human Genetics Unit, Edinburgh

It is widely assumed that genes that influence variation in skin and hair pigmentation are under selection. To date, the melanocortin 1 receptor (MC1R) is the only gene identified that explains substantial phenotypic variance in human pigmentation. Here we investigate MC1R polymorphism in several populations, for evidence of selection. We conclude that MC1R is under strong functional constraint in Africa, where any diversion from eumelanin production (black pigmentation) appears to be evolutionarily deleterious. Although many of the MC1R amino acid variants observed in non-African populations do affect MC1R function and contribute to high levels of MC1R diversity in Europeans, we found no evidence, in either the magnitude or the patterns of diversity, for its enhancement by selection; rather, our analyses show that levels of MC1R polymorphism simply reflect neutral expectations under relaxation of strong functional constraint outside Africa.

Introduction

Variation in human skin and hair color is largely explained by the levels and ratio of the two major forms of melanin—eumelanin, which is brown-black, and pheomelanin, which is red-yellow (Quevedo and Holstein 1998). The type of melanin produced is under the control of two genes, identified initially by the mouse mutations *extension* and *non-agouti*. These correspond to the melanocortin 1 receptor (MC1R [MIM 155555]), a G protein–coupled seven-pass transmembrane receptor, expressed on cutaneous melanocytes, and an antagonist ligand, agouti (Robbins et al. 1993; Jordan and Jackson 1998; Ollmann and Barsh 1999). The MC1R gene (chromosome 16q24.3) is the only gene identified, thus far, that explains substantial phenotypic variance in human pigmentation (Rees and Flanagan 1999).

We and others have previously shown that MC1R is highly polymorphic in European populations and that, in Irish, Dutch, and Swedish populations, homozygotes or compound heterozygotes for three particular variants are associated with red hair and increased sensitivity to burning from UV radiation (Valverde et al. 1995; Box et al. 1997; Smith et al. 1998). These three variants (Arg151Cys, Arg160Trp, and Asp294His) all bind α-melanocyte–stimulating hormone but show diminished intracellular ability to activate adenylate cyclase (Frändberg et al. 1998; Schiöth et al. 1999). As a consequence, there is, within the melanocyte, a switch in production from eumelanin to pheomelanin (Cone et al. 1996). In addition, the low-frequency variant, Arg142His, and the relatively common Val60Leu variant have also been found to show a reduced ability to activate adenylate cyclase (Schiöth et al. 1999). Arg142His—but not Val60Leu—has been found in association with red hair (authors' unpublished data). Val60Leu, however, may be associated with fair or blonde and light-brown hair colors (Box et al. 1997).

MC1R has several features that make it ideal for population-genetic analysis. First, because it is a small gene, it is well suited to complete-sequence analysis. Second, the functional role of MC1R within the biochemical pathways responsible for melanogenesis is well characterized. Third, functional and epidemiological studies are providing considerable detail about genotype-phenotype associations of MC1R variants, many of which are common polymorphisms in European populations. Fourth, since it is widely assumed that genes that influence phenotypic variation in human skin and hair pigmentation are under selection (Bodmer and Cavalli-Sforza 1976), the MC1R gene is a primary candidate for variable selection intensity between human populations (Rana et al. 1999). Here we examine the role of selection on MC1R polymorphism in several European, Asian, and African populations.

Received November 8, 1999; accepted for publication January 27, 2000; electronically published March 24, 2000.

Address for correspondence and reprints: Prof. Jonathan L. Rees, Department of Dermatology, University of Edinburgh, Royal Infirmary, Lauriston Place, Edinburgh EH3 9YW, United Kingdom. E-mail: jonathan.rees@ed.ac.uk

[*] Present affiliation: Department of Dermatopharmacology, University of Southampton, Southampton, United Kingdom.

[†] Present affiliation: Department of Dermatology, University of Edinburgh, Edinburgh.

© 2000 by The American Society of Human Genetics. All rights reserved.
0002-9297/98/6303-0018$02.00

Populations and Methods

Populations

Individuals from the European sample that we studied were from England, Ireland, Sweden, Finland, and Italy (Rome and Sardinia). Some of the English, Irish, and Swedish individuals were from our previous studies (Smith et al. 1998). Hair-color phenotype was ascertained after DNA samples were collected from the subjects. For comparison, we also examined both MC1R diversity in redheads from Ireland and Sweden and a number of melanoma cases from England. Hair-color phenotype has not been ascertained for the Finnish and Italian samples that were collected for unrelated studies. The Saami sample is from Finland. The Saami and other Finnish samples were collected by A. Sajantila, and the Italian samples were given to us by A. Di Rienzo.

African diversity was represented by samples from The Gambia and the Ivory Coast and from African immigrants to the United Kingdom. Additionally, we analyzed a sample of African Americans. The Asian samples were from Japan, Papua New Guinea (both coastal and highland), and southern India and also included Indian immigrants to the United Kingdom. Included in the Asian regional grouping, for the purpose of population-structure analysis, was an Inuit sample from Igloolik', Northwest Territories, Canada. Genomic DNA samples from anonymous individuals from the African, African American, and Asian populations were given to us by A. Daly, A. Quinn, A. V. S. Hill, M. Takata, R. Krishnamoorthy, and J. B. Clegg.

MC1R Sequence Analysis

Genomic DNA from blood was amplified by PCR, and the entire coding region (317 codons) was analyzed by automated sequencing with an ABI PRISM dye primer cycle sequencer (PE Biosystems) on an Applied Biosystems 373 automated sequencer (or, in a minority of cases, by manual cycle sequencing [Smith et al. 1998]). Variants were confirmed from genomic DNA, by repeat sequencing or RFLP analysis. Haplotypes were confirmed by cloning. The Irish sample was augmented with genomic DNA from 38 individuals in whom sequencing was completed between codons 48 and 163, and the 294 variant was investigated by RFLP analysis, thus covering all sites known to be polymorphic in Europe.

We also studied DNA samples, provided to us by J. B. Clegg, of two chimpanzees. We found identical MC1R sequences that differed from the human consensus sequence at 10 nonsynonymous positions, in codons 41, 57, 116, 165, 174, 183, 186, and 199, as has been reported elsewhere (Rana et al. 1999), and in seven synonymous positions, in codons 10, 82, 88, 165, 204, 223, and 314 (fig. 1*a*). Figure 1*b* compares the human MC1R consensus sequence with the variant haplotypes found in our survey. The silent polymorphism in codon 314 distinguishes the root of human MC1R diversity vis-à-vis the human consensus sequence (fig. 2).

Tests for Selection

The 10 nonsynonymous differences between the root MC1R haplotype for human diversity and the chimpanzee sequence indicate an amino acid–substitution rate of .026/codon. This rate is higher than the average found in 46 protein-coding sequences analyzed by Eyre-Walker and Keightley (1999) but is exceeded by rates of .049–.121 for several genes that they also report. An additional six synonymous nucleotide changes in MC1R contribute to a total sequence divergence of 1.68% (16/954, including the termination codon) and a total mutation rate of 1.68×10^{-9}/site/year, under the assumption that the human-chimpanzee divergence time is 5 million years. We estimate a silent-substitution rate of 2.3×10^{-9}/effectively silent site (6/261)/year. As an initial premise for the first test for selection, we take the conclusion by Eyre-Walker and Keightley (1999)—that the contributions that generally high amino acid–substitution rates make to hominid sequence divergence are more plausibly explained by low levels of constraint than by high rates of adaptive substitution.

McDonald and Kreitman's (1991) test compares, in an interspecific comparison between polymorphism data, both the ratio of nonsynonymous to synonymous sites and the divergent sites. In application to our data, rather than calculating a G statistic and testing for significance under the assumption of a χ^2 distribution, we use an exact test that is appropriate for the small sample of segregating sites (Sokal and Rohlf 1969). Our null hypothesis is based on the 10 nonsynonymous and 6 synonymous human-chimpanzee differences; a segregating site will generate either nonsynonymous or synonymous polymorphism, with binomial probabilities of .625 and .375, respectively. We calculated the probabilities of each possible ratio of nonsynonymous to synonymous polymorphism, for a given number of segregating sites. The sum of the probabilities for the observed ratio and any other equally or less likely possible ratios provides a test of the null hypothesis.

A second method for detection of selection in polymorphism data is the Hudson/Kreitman/Aguade (HKA) test (Hudson et al. 1987). Like the McDonald-Kreitman test, expectations for polymorphism are based on the sequence divergence between two species. Instead of a comparison of numbers of nonsynonymous sites and synonymous sites within a gene, diversity over all sites can be compared between two different gene regions.

The software package ARLEQUIN, version 1.1 (Schneider et al. 1997), was used for the following anal-

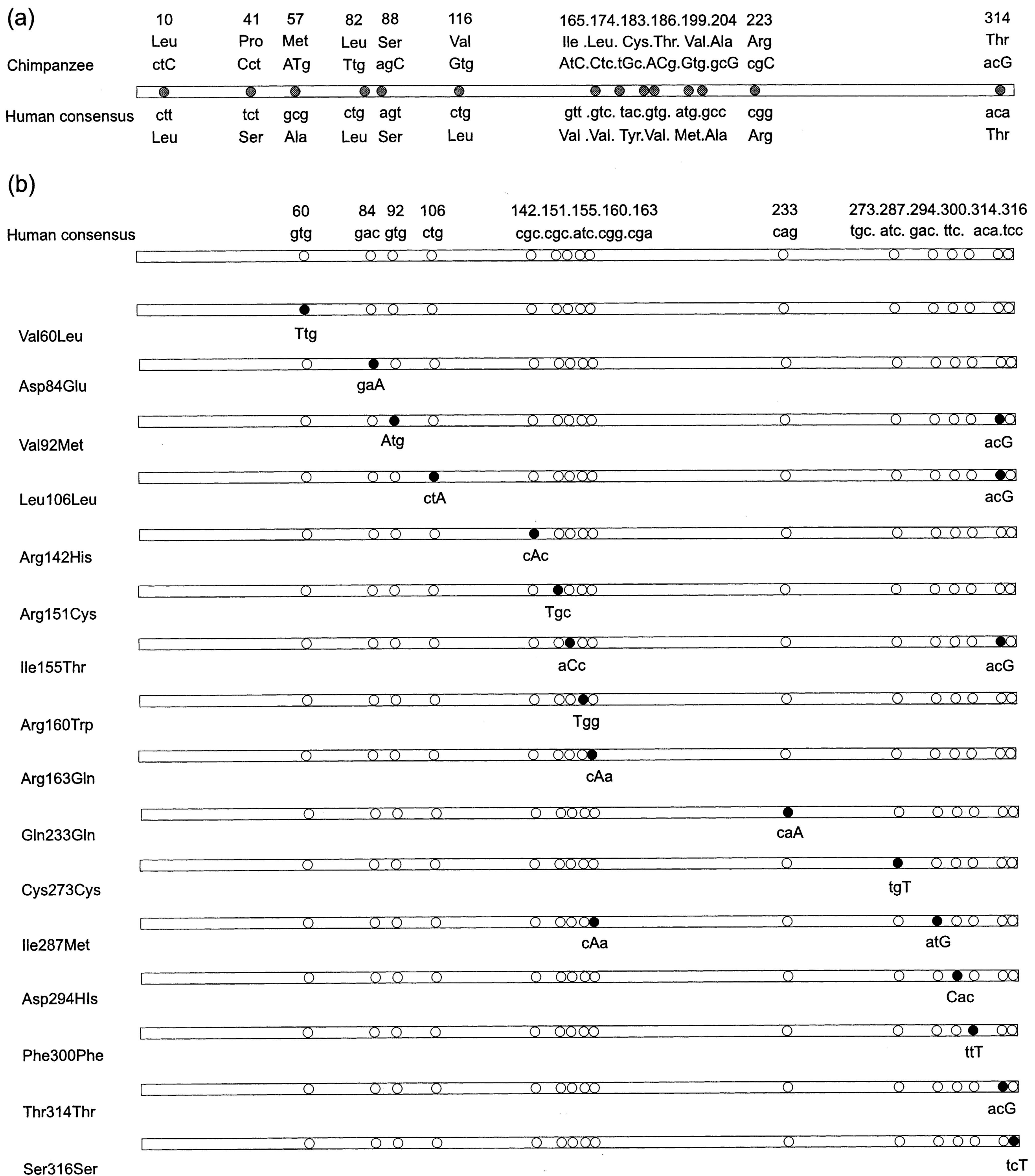

Figure 1 MC1R variation. *a,* Divergent codons and amino acids, between a chimpanzee sequence and the human consensus sequence, with the codon position indicated (*blackened circles*). Nucleotides that are different from those of the human consensus sequence are shown in uppercase letters. *b,* Variant haplotypes. The variants are labeled by a single defining amino acid change compared with the human consensus sequence. Both the codon position (*numbers*) and variant codons are indicated (*blackened circles*). Nucleotide that differ from those of the consensus sequence are shown in uppercase letters.

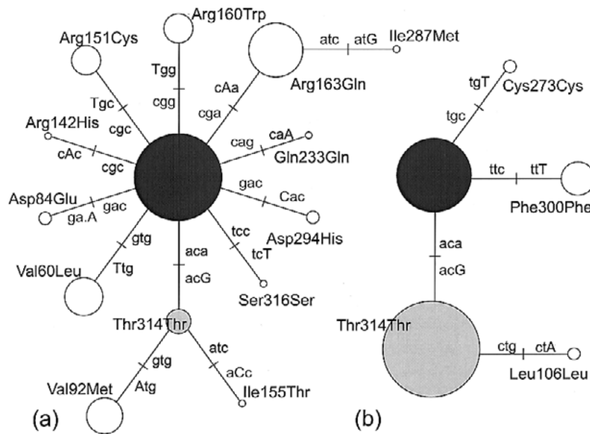

Figure 2 Gene trees connecting MC1R haplotypes through mutation events. Mutations in codon sequences, relative to the consensus sequence, are shown in uppercase letters. The MC1R consensus sequence is indicated (*blackened circles*), as is the root haplotype (*gray-shaded circles*), which is determined by comparison with a chimpanzee sequence. Variant haplotypes are labeled by a single defining amino acid change, compared with the human consensus sequence. Circled areas for haplotypes are proportional to their frequencies, not to sample numbers. *a,* European and Asian samples. *b,* African samples.

yses. We applied two further neutrality tests: the Ewens-Watterson test for equilibrium numbers of haplotypes and Tajima's D test. The exact probabilities that observed numbers of genotypes fit Hardy-Weinberg expectations were determined with a Markov chain applied to all haplotypes. Population structure was investigated by an analysis of variance of the distribution of haplotypes.

A coalescence model for the ancestral history of a sample of genes (Kingman 1982; Donnelly and Tavaré 1995) was used to estimate the time scale of polymorphic variation in MC1R. Ages were estimated, in units of $2N_e$, from gene trees for MC1R, conditional on a maximum-likelihood estimate for θ, the population mutation–rate parameter (Griffiths and Tavaré 1994, 1998; Harding et al. 1997). The gene trees represent the complete information available from the samples of sequences. None of the haplotypes that we observed gave evidence of recombination events.

Results

The most common MC1R haplotype in our survey (table 1) is the same as the consensus haplotype that has emerged from previous studies of European variation (Valverde et al. 1995; Box et al. 1997; Smith et al. 1998). The root of modern human MC1R diversity judged by comparison with a chimpanzee sequence is a haplotype common in Africa and equatorial Asia but found at low frequency elsewhere. The root haplotype differs from the consensus sequence only by a silent substitution in codon 314 (fig. 2). All five MC1R haplotypes in the African samples differ only at silent third-base-pair positions. The absence of any nonsynonymous variants in Africa suggests that strong functional constraint is acting on the consensus amino acid sequence, to maintain high levels of eumelanin pigmentation.

We first tested for goodness of fit between the ratio of synonymous to nonsynonymous variants in African polymorphism and the ratio in sites divergent between humans and chimpanzee (Sokal and Rohlf 1969); for the African sample, the probability of observing a ratio of zero nonsynonymous variants to four silent variants is $P = .0198$. Since we observed the least likely ratio, we can reject the hypothesis that it occurred as a chance observation within a standard 95% confidence interval. African MC1R polymorphism shows a significantly aberrant pattern compared with data on evolutionary divergence.

The absence of amino acid variants in Africa—as well as their low frequency in African Americans and in Asians from Papua New Guinea and India (table 1), where skin pigmentation is typically very dark—implies strong functional constraint on MC1R, probably as a means to minimize sensitivity to UV radiation. The presence of a small number of nonsynonymous variants in the African American and southern-Asian sample populations can be explained by admixture. The other non-African samples, in contrast, reveal high frequencies of a large number of nonsynonymous variants.

In our survey of European and Asian populations for MC1R variation (table 1), we found both nine nonsynonymous variants that had been reported elsewhere (Valverde et al. 1995; Box et al. 1997; Smith et al. 1998, Rana et al. 1999) and a new variant, Ile287Met (fig. 1), on the 163 haplotype common in Asia. By comparison, only three silent variants were found in these samples (table 1). However, judging this ratio of 10 nonsynonymous to 3 synonymous variants by comparing it with the ratio of 10 nonsynonymous to 6 silent substitutions that have occurred during human-chimpanzee divergence, we find no evidence for selection (McDonald and Kreitman 1991). Another nonsynonymous variant on the 163 haplotype, Arg67Gln, has been observed in random samples of Asian individuals, as reported elsewhere (Box et al. 1997; Rana et al. 1999). Including this observation does not change the conclusion that we determined on the basis of our data. Additional low-frequency amino acid variants have been found in studies for red hair, but, because of their ascertainment bias and because of a potentially incomplete reporting of silent variation, they cannot be included in this test.

The pattern of MC1R polymorphism in Europeans

Table 1

MC1R Haplotype Distributions

Population or Category	HCS[a]	Val 60 Leu	Asp 84 Glu	Val 92 Met	Leu 106 Leu	Arg 142 His	Arg 151 Cys	Ile 155 Thr	Arg 160 Trp	Arg 163 Gln	Gln 233 Gln	Cys 273 Cys	Ile 287 Met	Asp 294 His	Phe 300 Phe	Thr 314 Thr	Ser 316 Ser	AAC 29 AAAC	TTC 179 TTCC	Total
	No. (Frequency) of Haplotypes																			
Africa:																				
Ivory Coast	14	...	...	...	1	...	...	...	...	...	...	...	...	...	3	18	...	...	...	36
The Gambia	13	...	...	...	...	...	...	...	...	...	...	...	...	...	4	29	...	...	...	46
U.K. Africans	10	...	...	...	...	...	...	...	...	...	...	...	...	...	...	13	...	...	...	24
Total	37 (.35)	...	...	...	1 (.01)	...	...	...	...	...	...	1 (.01)	...	...	7 (.07)	60 (.57)	...	...	...	106
African Americans	22	...	...	...	4	...	...	...	...	2	...	...	...	1	3	10	...	...	...	42
Asia:																				
Southern India	34		...	...	...	...	...	...	...	1	...	...	...	...	...	9	...	...	...	44
U.K. Asian Indians	15	2	...	1	...	...	...	...	...	1	...	...	...	...	...	3	...	...	...	22
Papua New Guinea	20	...	...	3	...	...	...	...	...	4	...	...	...	...	...	5	...	...	...	32
Japan	4	...	...	2	...	...	...	...	...	23	...	...	1	...	...	...	...	...	...	30
Inuit	2	...	...	1	...	...	...	...	...	37	...	...	...	...	...	...	...	...	...	40
Total	75 (.45)	2 (.01)	...	7 (.04)	...	...	...	...	...	66 (.39)	...	...	1 (.01)	...	...	17 (.10)	...	...	...	168
Europe:																				
England	30	7	...	12	...	1	4	...	7	2	...	...	...	...	...	1	...	...	...	64
Ireland:																				
Complete sequence	15	7	3	6	...	...	7	...	5	3	...	...	...	...	...	...	...	...	...	46
Partial sequence	48	9	48	9	...	...	4	1	3	3	...	...	...	3	...	...	...	...	...	76
Sweden	24	11	...	2	...	...	3	1	5	4	...	...	...	...	...	1	1	...	...	52
Finland	10	1	...	2	...	...	3	...	4	9	...	...	...	...	...	1	...	...	...	30
Saami	13	...	...	6	...	...	3	...	...	4	...	...	...	...	...	...	...	...	...	26
Italy	22	7	...	1	...	...	1	...	1	1	...	...	...	...	...	1	...	...	...	34
Sardinia	19	5	...	...	...	...	...	...	...	...	1	...	...	2	...	1	...	...	...	28
Total	181 (.51)	47 (.13)	4 (.01)	33 (.09)	...	1 (.00)	25 (.07)	2 (.01)	25 (.07)	26 (.07)	1 (.00)	...	...	5 (.01)	...	5 (.01)	1 (.00)	...	...	356
	No. (Frequency) of Cases																			
Melanoma in England	34	10	...	3	...	2	3	...	1	1	...	...	...	...	...	...	...	...	...	54
Redheads:																				
Irish	1	3	...	...	...	...	6	...	4	...	...	...	...	5	...	...	...	...	1	20
Swedish	2	2	1	...	...	2	12	...	8	...	...	...	...	1	...	...	...	2	...	30
Total	3 (.06)	5 (.10)	1 (.02)	...	...	2 (.04)	18 (.36)	...	12 (.24)	...	...	...	...	6 (.12)	...	...	...	2 (.04)	1 (.02)	50

[a] HCS = human consensus sequence (other haplotypes are as labeled in figs. 1 and 2).

and northern Asians is consistent with expectations based on the evolutionary history of this gene during divergence of humans from the common ancestor that they share with chimpanzees. The explanation favored by Eyre-Walker and Keightley (1999) for this evolutionary pattern is low constraint resulting from fixation of slightly deleterious mutations when the long-term effective population size is small. One way to test the alternative explanation of a high rate of adaptive substitution is to evaluate whether MC1R diversity in Europe and northern Asia, in particular, has been selectively enhanced in excess of neutral expectations based on another gene. The distributions of MC1R variation within human populations and between humans and chimpanzee were compared, by use of the HKA model (Hudson et al. 1987), with those in β-globin, a gene that, except in areas of endemic malaria, is under strong functional constraint globally.

The HKA model was used to determine neutral expectations for African and non-African data separately. For African data, we assumed that mutation could arise only in effectively silent sites, and we estimated that there were 2,320 such sites in β-globin (Harding et al. 1997) and, likewise, 261 effectively silent sites in MC1R. We assumed no functional constraint on MC1R outside Africa, which resulted in 954 effectively silent sites. β-Globin diversity (R. M. Harding, unpublished data) was estimated from 16 segregating sites observed in 103 Africans and from 15 segregating sites observed in 350 non-Africans (Harding et al. 1997). MC1R diversity was estimated from four segregating sites observed in a sample of 106 Africans and from 13 segregating sites observed in a sample of 524 non-Africans (table 1). The numbers of sites divergent between a chimpanzee sequence and root sequences for human diversity are 27/2,320 for β-globin, 6/261 for MC1R in Africa, and 16/954 for MC1R outside Africa. HKA goodness-of-fit tests do not reject neutral expectations for either African or non-African data.

The HKA model with slightly higher mutation rates for MC1R compared with β-globin and with a divergence time of ∼5 million years (when a generation is assumed to be 20 years) gave remarkably good fits to both the African and non-African data. For the African data, the expectations for θ $4N_e\mu$ are .0027/site, for MC1R ($N_e \approx 14{,}700$), and .0013/site, for β-globin ($N_e \approx 14{,}000$). With the non-African data, we found estimates of ∼12,000 for N_e, for both genes. The expectations for θ are .0017/site, for MC1R, and .0011/site, for β-globin. These estimates for N_e are entirely consistent both between MC1R and β-globin data and, also, with estimates from other loci (Takahata et al. 1995). There is no evidence that MC1R diversity outside Africa has been enhanced by selection.

In a different approach, two neutrality tests were applied to the haplotype distributions, to detect departures from neutral expectations in the distribution pattern of diversity among individuals (table 2). It is feasible that these tests would be more sensitive to recent selection than are the tests discussed above, which evaluate the magnitude of diversity. The expectation, if selection were enhancing MC1R diversity, would be an excess of haplotypes showing intermediate frequencies. This effect is observed for the smaller sample of fully sequenced Irish alleles (table 1), but only the Ewens-Watterson test is significant; and, for the larger Irish sample, neither of the two tests is significant (table 2). No other European sample shows a similar distortion. Overall, these neutrality tests support our conclusion that the atypically high ratio of nonsynonymous to silent polymorphism in Europe and northern Asia reflects relaxed functional constraint, not positive selection. Low probabilities for goodness of fit to Hardy-Weinberg equilibrium that were found for the English and Irish samples, for Africans and Asian Indians from the United Kingdom, and for the sample of African Americans are not due to simple overall excesses of either homozygotes or heterozygotes (table 2). They may be explained by admixture, but it is feasible that assortative mating for skin and hair color also contributes.

Our analyses show that levels of both silent MC1R variation in Africa and total MC1R diversity outside Africa are compatible with neutral expectations if MC1R is under complete functional constraint in Africa and if functional constraint is relaxed outside Africa and southern Asia. These variable selective pressures between Africa and Eurasia explain the unusually high level of population structure, with 29% of total variance in MC1R haplotype distributions being apportioned among populations. In comparison, averages over multiple autosomal loci indicate that usually only 10%–15% of the total variance is explained by population structure (Bowcock et al. 1991; Jorde et al. 1998; Flint et al. 1999; Harding 1999). Relaxed functional constraint also explains the higher nucleotide diversity for Europe—at 0.12%, compared with 0.09% for Asia and 0.07% for Africa (table 2). Loci showing higher diversity in Europe or Asia, compared with Africa, when there is no reason to suppose an ascertainment bias, are atypical (Mountain and Cavalli-Sforza 1994; Harding et al. 1997; Jorde et al. 1998; Cargill et al. 1999). The more usual pattern is observed when diversity in MC1R is estimated over silent sites only: 0.08% for Europe and 0.09% for Asia, compared with 0.25% for Africa.

To further relate MC1R polymorphism to our current understanding of the demographic history of modern humans (von Haeseler et al. 1996; Harding et al. 1997; Jorde et al. 1998), we estimated the ages of MC1R alleles, assuming a coalescent model for the ancestral history of a sample of genes. For this approach, we first

Table 2

Random Mating and Neutrality Tests, and Nucleotide Diversity for Each Sample

Population or Category	No. of Haplotypes	Statistic			
		Hardy-Weinberg Equilibrium[a]	Slatkin's Exact Probability[b]	Tajima's D[c]	π[d] (%)
Africa:					
Ivory Coast	36	.22	.48	.01(52)	.08
The Gambia	46	.74	.21	.73 (.76)	.07
U.K. Africans	24	.06	.47	.27 (.61)	.06
Overall	106	.21	.70	−.28 (.41)	.07
African Americans	42	.05	.45	−.55 (.31)	.10
Asia:					
Southern India	44	.37	.65	−.33 (.39)	.04
U.K. Asian Indians	22	.07	.85	−1.10 (.14)	.07
Papua New Guinea	32	.46	.16	.13 (.56)	.08
Japan	30	.43	.77	−.90 (.20)	.07
Inuit	40	1.00 (.00)	.93	−1.43 (.07)	.03
Overall	168	.00	.45	−.07 (.49)	.09
Europe:					
England	64	.01	.42	−.40 (.37)	.13
Ireland:					
Complete sequence	46	.04	.01	−.26 (.42)	.15
Partial sequence	122	.11	.18	−.89 (.20)	.11
Sweden	52	.71	.51	−1.12 (.13)	.11
Finland	30	.08	.37	−.51 (.32)	.13
Saami	26	.45	.08	.44 (.67)	.13
Italy	34	.93	.98	−1.47 (.07)	.08
Sardinia	28	1.00 (.00)	.81	−1.13 (.13)	.06
Overall	356	.11	.58	−.97 (.17)	.11
	No. of Cases				
Melanoma in English sample	54	.15	.78	−1.29 (.10)	.08
Redheads:					
Irish	20	1.00 (.00)	.25	.29 (.62)	.16
Swedish	30	.01	.49	−.58 (.30)	.14

[a] Data are probabilities that genotypes are in Hardy-Weinberg equilibrium.
[b] From Ewens-Watterson distribution for haplotypes (outliers in both tails).
[c] Data in parentheses are probabilities from a beta distribution for Tajima's D (outliers in both tails).
[d] Average pairwise sequence difference per nucleotide.

found maximum-likelihood estimates for θ by making use of the complete information of the gene trees (fig. 2). From the 448 complete MC1R sequences (954 sites) in the European and Asian samples, we find an estimate of 2.1 for θ (2.1/954 = .0022/site) and derive an estimate of 16,400 for N_e; from the 106 MC1R sequences (261 effectively silent sites) in the African samples, we find an estimate of 0.8 for θ (0.8/261 = .0031/site) and derive an estimate of 16,700 for N_e. Again, estimates for N_e provide no evidence that MC1R diversity in Europe and northern Asia is higher than that suggested by other autosomal loci in the human genome, when neutrality of polymorphic variation has been assumed.

Both African and non-African data suggest that the time to the most recent common ancestor is ~1 million years and that the age of the global 314 variant is 650,000 years. On this time scale, ages for the Eurasian-distributed Val60Leu, Val92Met, and Arg163Gln variants are 250,000–100,000 years; the ages for African silent variants—Leu106Leu, Cys273Cys, and Phe300Phe—are 110,000–40,000 years. For the European red hair–associated Arg151Cys and Arg160Trp variants, we estimate an age of ~80,000 years; for Asp294His, and Ser316Ser, we estimate an age of ⩽30,000 years. SDs are approximately half these expectations. These ages are entirely compatible with age distributions estimated for African and non-African mutations in other nuclear genes (Harding et al. 1997; Zietkiewicz et al. 1998). The ages estimated for the Arg151Cys and Arg160Trp red hair–associated variants are consistent with a widespread European distribution, as we also observed.

Discussion

The evidence that selection eliminates nonsynonymous MC1R variation to maintain a high-eumelanin pheno-

type in African populations is consistent with two other sets of observations. First, it seems that many of the amino acid changes to the consensus sequence that have been observed to arise in non-African populations do affect MC1R function (Frändberg et al. 1998; Schiöth et al. 1999). Second, a recent study reports that, in Europeans, fairer skin and increased sensitivity to burning from UV radiation are detectable in heterozygotes for a number of MC1R variants associated with red hair. MC1R alleles (Arg151Cys, Arg160Trp, Asp294His, and Arg142His) that act as recessives in the red-hair phenotype (~11% of the British and Irish population) also contribute to UV radiation sensitivity in heterozygous genotypes (~28% of the population; J. L. Rees, unpublished data). These data indicate that heterozygotes for functional variants exhibit a phenotype that would be deleterious in Africa. In contrast to Africa, where selection imposes strong functional constraint on MC1R, loss of function is tolerated in European and eastern Asian populations.

Red hair in Europeans is almost completely accounted for by amino acid variants of MC1R, but a number of cases that remain unexplained indicate that, even for this specific phenotype, genetic determination remains incomplete. Homozygotes or compound heterozygotes for Arg151Cys, Arg160Trp, Asp294His, and Arg142His and for insertions at codons 29 and 179 account for red hair in 6 of 10 Irish redheads and in 12 of 15 Swedish redheads (table 1); among the other 7 redheads are 4 compound heterozygotes (Arg151Cys and Val60Leu, Arg160Trp and Val60Leu, Asp294His and Val60Leu, and insertion 29 and Asp84Glu) and a Val60Leu homozygote; additionally, there is one Arg151Cys heterozygote and one homozygote for the consensus sequence. Therefore, we extended our sequencing survey when a previously unidentified second exon was recently discovered by Tan et al. (1999), who report that alternative mRNA splicing adds an additional 65 amino acids at the predicted intracellular, C-terminal tail, to generate a new MC1R isoform. Neither binding of α-melanocyte–stimulating hormone nor activation of adenylate cyclase appears to be significantly altered for this isoform (Tan et al. 1999). Sequencing of the second exon has so far been completed for 40 Irish, English, and Swedish individuals (including 26 with a red-hair phenotype that cannot be explained by the standard genotypes given above), 15 black Africans or African Americans, and 5 Japanese, without detection of any polymorphism at all. These data suggest that this second exon provides no additional explanation for either the red-hair phenotype or the pattern of selection inferred from diversity in the first MC1R exon.

For the first exon of MC1R, the combined African and non-African samples revealed 16 polymorphisms, in a ratio of 10 nonsynonymous to 6 synonymous, which happens to be the same ratio as is seen for the interspecific evolutionary divergence. Consequently, both polymorphism and divergence data indicate that nonsynonymous MC1R variants survive at 63% of the rate of synonymous changes. This rate is higher than the 38% rate that was observed, as an average of many genes, both by Cargill et al. (1999), in a large single-nucleotide–polymorphism survey, and by Eyre-Walker and Keightley (1999), in a survey of nucleotide differences between the human and chimpanzee genomes. A relatively high evolutionary rate for MC1R was also detected by Rana et al. (1999), in a comparison of MC1R with the other four melanocortin receptors in this gene family. Both the relatively high evolutionary rate and the high polymorphism are likely to reflect a lower constraint on amino acid substitution, because MC1R function is less essential compared with that of these other melanocortin receptors and their ligands (Cone et al. 1996).

The possibility that both the relatively high evolutionary rate and the high European diversity are consequences of adaptation, has been discussed by Rana et al. (1999) and Owens and King (1999). However, selection on the evolutionary-divergence rate is usually inferred when the rate of nonsynonymous substitution is greater than the silent rate (Goldman and Yang 1994; Nielsen and Yang 1998). In fact, although the nonsynonymous rate in MC1R is higher than the average found over many genes, it is still lower than the silent evolutionary rate. We have also investigated the level of European polymorphism. There is a popular hypothesis that fair skin in Europeans has been positively selected to increase sensitivity to UV radiation and that, in northern latitudes, this adaptation is needed to increase UV radiation–induced vitamin D synthesis and to prevent rickets (Bodmer and Cavalli-Sforza 1976). For many European and Asian individuals, variant MC1R alleles contribute to both lighter skin color and sun sensitivity. However, we found no statistical evidence that MC1R diversity has been enhanced by selection, either in its apparently high levels or in its haplotype frequency–distribution patterns.

The statistical detection of selection from DNA sequence variation is a challenging exercise. Of the tests used, the McDonald-Kreitman test, which does indicate a pattern of differential selection on MC1R polymorphism between the African and non-African samples, should be the most powerful. This is because, for the HKA, Tajima's *D,* and the Ewens-Watterson homozygosity tests, genetic drift additionally contributes to the evolutionary variance around expectations, under neutrality. Weak selection is difficult to distinguish against a background of genetic drift. Specification of the test hypotheses is also relevant. For the McDonald-Kreitman and HKA tests, distinguishing between selective

enhancement of diversity and relaxed functional constraint probably is not feasible with rates of divergence and diversity in nonsynonymous sites that, although atypically high, are less than the rates for silent sites. Similarly, functional constraint is subsumed into the null hypothesis, for strict neutrality, in both Tajima's *D* and Ewens-Watterson tests, and, in application of the HKA test to the African data, we have adjusted neutrality expectations, for functional constraint. Of course, we also applied the HKA test to the African data without adjusting for functional constraint, but we did not detect that the diversity was significantly low. Another important factor here is the small size of the MC1R gene, limiting the number of sites that constitute the data.

To estimate the ages of MC1R alleles, we assumed that their observed frequencies simply reflect genetic drift in constant-size, randomly mating populations, after allowing for different patterns of functional constraint. These estimates suggest that the MC1R variants Val60Leu, Val92Met, and Arg163Gln may trace back to ancestors in Eurasian populations existing 250,000–100,000 years ago. Our estimates of 80,000 years for the red hair–associated Arg151Cys and Arg160Trp variants likewise suggest a distant ancestral contribution from Paleolithic Eurasians to the western European populations of today. Our estimated age of 80,000 years for the red hair–associated variants suggests that they are likely to have a wide geographic distribution, perhaps extending into central Asia. This prediction is easily testable with data from other surveys of MC1R variation.

Our age estimates for MC1R alleles sampled both within and outside Africa are compatible with age distributions for neutral allelic variation estimated for other nuclear loci (Harding et al. 1997; Zietkiewicz et al. 1998; Harris and Hey 1999). However, an incompatibility arises between estimated ages in the range of 250,000–100,000 years, for non-African MC1R allelic variation, and ages, from fossil evidence, of $\leqslant$100,000 years for the dispersal of modern humans outside Africa and the Middle East (Stringer and Andrews 1988). One possible explanation for incompatible ages from genetic data is that they have been overestimated under an assumption of levels of genetic drift that are consistent with constant population size. These age estimates would be younger under an alternative model with a high rate of population expansion out of a Eurasian founding population.

Various sources of non-African diversity data indicate a small Eurasian founding population and subsequent population expansion. A smaller size for the Eurasian population compared with that for Africa has been inferred from lower levels of diversity outside Africa (Jorde et al. 1998; Cargill et al. 1999). Collapse into a population bottleneck is one interpretation of positive values of Tajima's *D,* reported for some, mostly non-African, populations, on the basis of analyses of autosomal loci (Harding et al. 1997; Clark et al. 1998; Zietkiewicz et al. 1998), with the strongest signature found for eastern Asian populations (Fay and Wu 1999). For non-African populations, it seems difficult to detect the population-expansion phase, but data from loci with high mutation rates, including mtDNA (Marjoram and Donnelly 1994; Fay and Wu 1999) and microsatellites (Di Rienzo et al. 1998), are suggestive. A bottleneck in Asian demographic history may explain the high frequency of the Arg163Gln variant. Similarly, a single allele for an X-chromosome region has been found at high frequency in Asian populations (Kaessmann et al. 1999). Further theoretical work is needed to determine whether a specific population-expansion model that accommodates all the various genetic data would also reduce age estimates of old non-African alleles sufficiently to fit expectations based on the fossil evidence for modern human distributions.

We can speculate that we have failed to detect a selective enhancement of MC1R diversity in Eurasian populations—particularly for Arg163Gln, as suggested by Rana et al. (1999). Perhaps MC1R was the first gene to respond to selection for lighter pigmentation when modern humans moved into latitudes away from the equator, simply because, at this gene, any of a large number of knockout mutants will suffice. Thus, selection may have enhanced frequencies of Val60Leu, Val92Met, and Arg163Gln—if, indeed, these all are functional variants. There is some evidence for Val60Leu being a loss-of-function variant, but there is no evidence for Arg163Gln being such a variant, on the basis of either our statistical association studies of redheads or our assays for ability to activate adenylate cyclase (J. L. Rees and S. Phillips, unpublished data). Subsequently, a more effective mutation in another gene may have arisen, and, in sweeping through the population, it may have swamped any fitness differentials introduced by MC1R variation. However, until substantially more data describing patterns of human diversity are available, we are unable to statistically resolve the relative likelihoods of subtle selective histories.

It is also clear that phenotypic variation in pigmentation is only partially explained by the MC1R switch between eumelanin and pheomelanin and that additional loci control total levels of melanin pigmentation. Perhaps other genes will provide evidence for adaptation. If so, a problem for the future will be to explain why MC1R has escaped this selective impact on pigmentation phenotype. We conclude, on the basis of all data presently available to us, that selection on MC1R is differentiated between functional constraint within Africa, where amino acid variants of the consensus

sequence reduce fitness, and relaxation of constraint in Europe and eastern Asia.

Acknowledgments

We thank Magnus Nordborg for discussion and Nick Hastie and John Clegg for encouragement. This study was supported in part by grants from the Leech Trust (CRC, Department of Health, United Kingdom), the British Skin Foundation, and the MRC Human Genetics Unit (Edinburgh). E.H. is a MRC Clinician Scientist Fellow, and N.S.E. is a University of Newcastle Research Committee Ph.D. student.

Electronic-Database Information

Accession numbers and URLs for data in this article are as follows:

ARLEQUIN: a software for population genetic data analysis, http://anthropologie.unige.ch/arlequin

GenBank, http://www.ncbi.nlm.nih.gov/Genbank (for chimpanzee MC1R sequence [AJ245705] and human MC1R sequence [X65634])

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for MC1R [MIM 155555])

References

Bodmer WF, Cavalli-Sforza LL (1976) Genetics, evolution and man. WH Freeman, San Francisco

Bowcock AM, Kidd JR, Mountain JL, Hebert JM, Carotenuto L, Kidd KK, Cavalli-Sforza LL (1991) Drift, admixture, and selection in human evolution: a study with DNA polymorphisms. Proc Natl Acad Sci USA 88:839–843

Box NF, Wyeth JR, O'Gorman LE, Martin NG, Sturm RA (1997) Characterization of melanocyte stimulating hormone receptor variant alleles in twins with red hair. Hum Mol Genet 6:1891–1897

Cargill M, Altshuler D, Ireland J, Sklar P, Ardlie K, Patil N, Lane CR, et al (1999) Characterization of single-nucleotide polymorphisms in coding regions of human genes. Nat Genet 22:231–238

Clark AG, Weiss KM, Nickerson DA, Taylor SL, Buchanan A, Stengård J, Salomaa V, et al (1998) Haplotype structure and population genetic inferences from nucleotide-sequence variation in human lipoprotein lipase. Am J Hum Genet 63: 595–612

Cone RD, Lu D, Koppula S, Vage DI, Klungland H, Boston B, Chen W, et al (1996) The melanocortin receptors: agonists, antagonists, and the hormonal control of pigmentation. Recent Prog Horm Res 51:287–317

Di Rienzo A, Donnelly P, Toomajian C, Sisk B, Hill A, Petzl-Erler ML, Haines GK, et al (1998) Heterogeneity of microsatellite mutations within and between loci, and implications for human demographic histories. Genetics 148:1269–1284

Donnelly P, Tavaré S (1995) Coalescents and genealogical structure under neutrality. Annu Rev Genet 29:401–421

Eyre-Walker A, Keightley PD (1999) High genomic deleterious mutation rates in hominids. Nature 397:344–347

Fay JC, Wu C-I (1999) A human population bottleneck can account for the discordance between patterns of mitochondrial versus nuclear DNA variation. Mol Biol Evol 16: 1003–1005

Flint J, Bond J, Rees DC, Boyce AJ, Roberts-Thomson JM, Excoffier L, Clegg JB, et al (1999) Minisatellite mutational processes reduce F_{ST}, estimates. Hum Genet 105:567–576

Frändberg PA, Doufexis M, Kapas S, Chhajlani V (1998) Human pigmentation phenotype: a point mutation generates nonfunctional MSH receptor. Blochem Biophys Res Commun 245:490–492

Goldman N, Yang Z (1994) A codon-based model of nucleotide substitution for protein-coding DNA sequences. Mol Biol Evol 11:725–736

Griffiths RC, Tavaré S (1994) Ancestral inference in population genetics. Stat Sci 9:307–319

——— (1998) The age of a mutant in a general coalescent tree. Stochastic Models 14:273–295

Harding RM (1999) More on the X files. Proc Natl Acad Sci USA 96:2582–2584

Harding RM, Fullerton SM, Griffiths RC, Bond J, Cox MJ, Schneider JA, Moulin DS, et a1 (1997) Archaic African and Asian lineages in the genetic ancestry of modern humans. Am J Hum Genet 60:772–789

Harris EE, Hey J (1999) X chromosome evidence for ancient human histories. Proc Natl Acad Sci USA 96:3320–3324

Hudson RR, Kreitman M, Aguade M (1987) A test of neutral molecular evolution based on nucleotide data. Genetics 116: 153–159

Jordan SA, Jackson IJ (1998) Melanocortin receptors and antagonists regulate pigmentation and body weight. Bioessays 20:603–606

Jorde LB, Bamshad M, Rogers AR (1998) Using mitochondrial and nuclear DNA markers to reconstruct human evolution. Bioessays 20:126–136

Kaessmann H, Heißig F, von Haeseler A, Pääbo S (1999) DNA sequence variation in a non-coding region of low recombination on the human X chromosome. Nat Genet 22:78–81

Kingman JFC (1982) The coalescent. Stochastic Proc Appl 13: 235–248

Marjoram P, Donnelly P (1994) Pairwise comparisons of mitochondrial DNA sequences in subdivided populations and implications for early human evolution. Genetics 136: 673–683

McDonald JH, Kreitman M (1991) Adaptive protein evolution at the *Adh* locus in *Drosophila.* Nature 351:652–654

Mountain JL, Cavalli-Sforza LL (1994) Inference of human evolution through cladistic analysis of nuclear DNA restriction polymorphisms. Proc Natl Acad Sci USA 91:6515–6519

Nielsen R, Yang Z (1998) Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. Genetics 148:929–936

Ollmann MM, Barsh GS (1999) Down-regulation of melanocortin receptor signaling mediated by the amino terminus of Agouti protein in *Xenopus melanophores.* J Biol Chem 274:15837–15846

Owens K, King M-C (1999) Genomic views of human history. Science 286:451–453

Quevedo WC, Holstein TJ (1998) General biology of mammalian pigmentation. In: Nordlund JJ, Boissy RE, Hearing

VJ, King RA, Ortonne JP (eds) The pigmentary system: physiology and pathophysiology. Oxford University Press, New York, pp 43–58

Rana BK, Hewett-Emmett D, Jin L, Chang BH-J, Sambuughin N, Lin M, Watkins S, et al (1999) High polymorphism at the human melanocortin 1 receptor locus. Genetics 151: 1547–1557

Rees JL, Flanagan N (1999) Pigmentation, melanocortins and red hair. Q J Med 92:125–131

Robbins LS, Nadeau JH, Johnson KR, Kelly MA, Roselli-Rehfuss L, Baack E, Mountjoy KG, et al (1993) Pigmentation phenotypes of variant extension locus alleles result from point mutations that alter MSH receptor function. Cell 72: 827–834

Schiöth HB, Phillips SR, Rudzish R, Birch-Machin MA, Wikberg JES, Rees JL (1999) Loss of function mutations of the human melanocortin 1 receptor are common and associated with red hair. Biochem Biophys Res Commun 260:488–491

Schneider S, Kueffer J-M, Roessli D, Excoffier L (1997) ARLEQUIN: a software for population genetic data analysis, version 1.1. Genetics and Biometry Lab, Department of Anthropology, University of Geneva, Geneva

Smith R, Healy E, Siddiqui S, Flanagan N, Steijlen PM, Rosdahl I, Jacques JP, et al (1998) Melanocortin 1 receptor variants in an Irish population. J Invest Dermatol 111: 119–122

Sokal RR, Rohlf FJ (1969) Biometry. WH Freeman, San Francisco

Stringer C, Andrews P (1988) Genetic and fossil evidence for the origin of modern humans. Science 239:1263–1268

Takahata N, Satta Y, Klein J (1995) Divergence time and population size in the lineage leading to modern humans. Theor Popul Biol 48:198–221

Tan CP, McKee KK, Weinberg DH, MacNeil T, Palyha OC, Feighner SD, Hreniuk DL, et al (1999) Molecular analysis of a new splice variant of the human melanocortin-1 receptor. FEBS Lett 451:137–141

Valverde P, Healy E, Jackson I, Rees JL, Thody AJ (1995) Variants of the melanocyte-stimulating hormone receptor gene are associated with red hair and fair skin in humans. Nat Genet 11:328–330

von Haeseler A, Sajantila A, Pääbo S (1996) The genetical archaeology of the human genome. Nat Genet 14:135–140

Zietkiewicz E, Yotova V, Jarnik M, Korab-Laskowska M, Kidd KK, Modiano R, Scozzari R, et al (1998) Genetic structure of the ancestral population of modern humans. J Mol Evol 47:146–155