

Selection on the Human Bitter Taste Gene, TAS2R16, in Eurasian Populations

Author(s): Hui Li, Andrew J. Pakstis, Judith R. Kidd, and Kenneth K. Kidd

Source: Human Biology, 83(3):363-377. 2011.

Published By: Wayne State University Press

DOI: 10.3378/027.083.0303

URL: <http://www.bioone.org/doi/full/10.3378/027.083.0303>

BioOne (www.bioone.org) is an electronic aggregator of bioscience research content, and the online home to over 160 journals and books published by not-for-profit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Web site, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at www.bioone.org/page/terms_of_use.

Usage of BioOne content is strictly limited to personal, educational, and non-commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

Selection on the Human Bitter Taste Gene, TAS2R16, in Eurasian Populations

HUI LI,^{1,2} ANDREW J. PAKSTIS,¹ JUDITH R. KIDD,¹ AND KENNETH K. KIDD^{1*}

Abstract Bitter taste is one of the most important senses alerting humans to noxious foods. In gatherer communities, sensitivity to bitterness is presumably advantageous because of various noxious plants. *TAS2R16* is the gene coding the taste receptor molecules for some of the most common toxins in plants. A previous study of this gene indicated selection has increased the frequency of a derived allele in this gene that arose before the human expansion out of Africa. We have applied a different methodology for detecting selection, the Long Range Haplotype (LRH) analysis, to *TAS2R16* in a larger sampling of populations from around the world. The haplotype with the derived alleles at both the functional polymorphism and a polymorphism in the regulatory region of *TAS2R16* showed evidence for recent positive selection in most of the Eurasian populations, though the highest selection signal occurs in Mbuti Pygmies, an African hunter-gatherer group. In Eurasia, only populations of Mesopotamia and the southeast coast of China have no signals of selection. The evidence of recent selection found in most Eurasian populations differs from the geographic pattern seen in the earlier study of selection. One can speculate that the difference may result from a gathering lifestyle extending into the most recent 10,000 yrs and the need to recognize newly encountered bitter natural toxins as populations expanded into new environments and the biota changes with the ending of the most recent ice age. Alternatively, the promoter region variant may be a marker for altered function beyond what the derived amino acid allele conferred.

Bitter taste differs from the other four distinct basic taste qualities (Chandrasekar et al. 2006) in humans (sour, sweet, salty, umami) in that it evokes a strong negative hedonic tone (Steiner 1994). It is assumed a bitter sensation detects toxins in food, providing important protection against food poisoning (Rozin and Vollmecke 1986). Numerous chemically diverse compounds elicit a

¹Department of Genetics, School of Medicine, Yale University, New Haven CT 06520.

²MOE Key Laboratory of Contemporary Anthropology, School of Life Sciences, Fudan University, Shanghai 200433, China.

*Correspondence to: Kenneth K. Kidd. E-mail: kenneth.kidd@yale.edu.

Electronic Sources cited: ALFRED: <http://alfred.med.yale.edu>; dbSNP: www.ncbi.nlm.nih.gov/SNP; UCSC Genome Browser: www.genome.ucsc.edu; Kidd Lab: <http://info.med.yale.edu/genetics/kkidd/contents.html>; Hui Li Site: <http://comonca.org.cn/LH/Doc/A52S.htm>.

Human Biology, June 2011, v. 83, no. 3, pp. 363–377.

Copyright © 2011 Wayne State University Press, Detroit, Michigan 48201-1309

KEY WORDS: *TAS2R16*, GENETIC VARIATION, POSITIVE SELECTION, GATHERING, BITTER TASTE.

bitter taste; most arise from plants, while some also arise from animals. Some 10% of plants may contain toxic glycosides or alkaloids (Kingsbury 1964) that elicit a bitter taste sensation. The great number of bitter compounds suggested the existence of a family of bitter receptors and genes (Alder et al. 2000). The bitter taste receptor gene family, named the taste receptor type 2 (TAS2R) gene family (Adler et al. 2000), consists of 25 full-length genes and 11 pseudogenes (Go et al. 2005). These bitter receptor genes are G protein-coupled receptors. Among the TAS2R genes, *TAS2R16* encodes the receptor mediating response to various β -glucopyranosides commonly found in nature (Bufe et al. 2002; Soranzo et al. 2005). The ability to detect these substances possibly helps provide protection against food poisoning.

Presumably the species immediately ancestral to modern humans had a functional TAS2R16 receptor. To the degree that the modern human *TAS2R16* gene differs from the ancestral human form those differences either reflect neutral genetic drift or the result of selection on new mutations. The nonsynonymous K172N (Lys172Asn, rs846664) polymorphism of *TAS2R16* was examined by Soranzo et al. (2005) both for function and for a signature of natural selection. They showed in transient transfection studies that the derived 172Asn form of the receptor had greater sensitivity to several naturally occurring glycosides. By sequencing the entire coding region of *TAS2R16* they detected a signal of selection favoring the 172Asn allele using Fay and Wu's H statistics (Fay and Wu 2000) in 19 of the 60 populations they studied, primarily samples from the HGDP panel (Cann et al. 2002). The 19 populations were mostly from Africa and a few from Europe and the Americas. The derived 172Asn allele is largely fixed in the remaining populations, consistent with a selective sweep but not allowing a test using the H statistics. However, as the individual sample sizes were rather small and the short sequences studied by Soranzo et al. (2005) might not be sensitive enough to detect selection signals, such signals might be detectable with other techniques. In this paper, we tested *TAS2R16* for signatures of natural selection in multiple large population samples by Long Range Haplotype (LRH) analyses (Endo et al. 1996; Sabeti et al. 2002). As more recombinations and mutations will accumulate over a long molecular range than over a short-range, in any unit of time, our analysis was sensitive to more recent selection events and demonstrated a different geographic pattern of selection signals from what was seen by Soranzo et al. (2005).

In addition to the Soranzo et al. (2005) study, the K172N polymorphism (rs846664) has also been studied by others with respect to alcoholism. The 172Lys allele is a high-risk allele for alcoholism in the COGA study, especially in their African-American sample (Hinrichs et al. 2006; Edenberg and Faroud 2006; Wang et al. 2007). Therefore, we have examined the evidence for selection on *TAS2R16* globally using rs846664 within the gene and closely flanking SNPs to define core haplotypes clearly segregating outside of Africa and 51 additional SNPs in the region.

Subjects and Methods

Population Samples. We examined the *TAS2R16* region in the following 45 populations (abbreviations and sample sizes are in the parentheses) covering most of the regions of the world.

Sub-Saharan Africa: Mbuti (MBU, 36), Biaka (BIA, 63), African-Americans (AAM, 84), Yoruba (YOR, 73), Chagga (CGA, 44), Hausa (HAS, 34), Ibo (IBO, 46), Masai (MAS, 19), Sandawe (SND, 40), Ethiopian-Jews (ETJ, 32); Middle East: Yemenite-Jews (YMJ, 43), Ashkenazi-Jews (ASH, 80), Druze (DRU, 124), Samaritans (SAM, 39); Europe: Adygei (ADY, 54), Danes (DAN, 49), European-Americans (EAM, 84), Finns (FIN, 35), Hungarians (HGR, 145), Irish (IRI, 114), Russians-Archangel (RUA, 33), Russians-Vologda (RUV, 48), Chuvash (CHV, 40); North Asia: Khanty (KTY, 49), Komi (KMZ, 47), Yakut (YAK, 50), Koreans (KOR, 54), Japanese (JPN, 50); East Asia & Pacific: Cambodians (CBD, 25), Laotians (LAO, 119), Han-Canton (CHS, 53), Han-Minnan (CHT, 47), Han-Hakka (HKA, 41), Ami (AMI, 38), Atayal (ATL, 40), Micronesians (MCR, 34), Nasioi Melanesians (NAS, 23); the Americas: Cheyenne (CHY, 54), Pima-Arizona (PMA, 50), Pima-Mexico (PMM, 98), Maya (MAY, 46), Quechua (QUE, 23), Rondonian Surui (SUR, 43), Ticuna (TIC, 61), Karitiana (KAR, 55). All samples were collected with informed consent under protocols approved by all relevant human subjects panels. For several populations the individuals in the HGDP that were tested by Saranzo et al. (2005) are subsets of the individuals we have studied (definitions of the relationships among samples of the same population can be found in ALFRED; Rajeevan et al., 2005; <http://alfred.med.yale.edu>).

Genetic Markers. We typed 55 SNPs covering a 242,855-bp region around *TAS2R16* (Figure 1) in all 2,459 individuals in the 45 populations. Typing was done on DNA purified using standard phenol-chloroform protocols from lymphoblastoid cell lines maintained in the Kidd Lab at Yale School of Medicine. As the coding region is quite short (less than 1 kb), our markers are in the flanking regions except for rs846664, the only nonsynonymous polymorphism in *TAS2R16* that we were able to type. In addition to the functional differences known for the nonsynonymous polymorphism Lys172Asn (K172N, rs846664), we speculated that a SNP in the 5' regulatory region might possibly be functional and therefore also tested rs702424. We also chose two additional SNPs flanking these two for a total of four SNPs (rs1357949, rs846664, rs702424, rs9691345; allele information is in Table 1) covering 4.7 kb in and immediately flanking *TAS2R16* to determine haplotype diversity among populations globally.

All of the markers were typed by a custom Illumina bead array or Applied Biosystems TaqMan® assays. To determine the ancestral alleles at all the SNPs we also typed three individuals from each of five other primate groups: *Pan troglodytes*, *Pan paniscus*, *Gorilla gorilla*, *Pongo pygmaeus*, and *Hylobates sp.*

Statistical Methods. The haplotypes of 55 SNPs and of a four-SNP core were estimated from the typing results using PHASE 2.1 (Stevens and Scheet 2005)

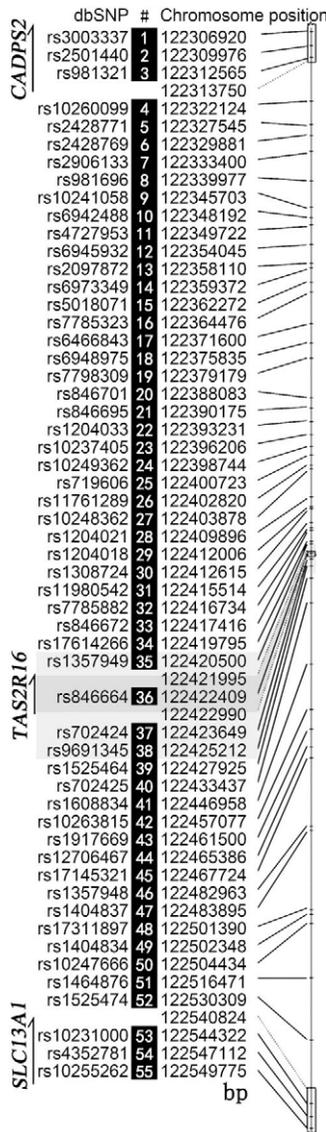


Figure 1. Spacing and molecular positions of the markers typed. The rs numbers can be used to retrieve the allele frequencies in ALFRED (<http://alfred.med.yale.edu/>). Gaps in the numbering (dotted lines) and associated positions (NCBI Build 36) indicate the borders of the genes.

assuming codominance and no null alleles. The relationships among the haplotypes of all 55 SNPs were constructed by NETWORK 4.500. The evolutionary history of the four-SNP core was constructed based on knowledge of ancestral alleles and geographic distributions of the population samples. Two LRH analyses were used to test for a signature of selection: extended haplotype homozygosity (EHH) and

Table 1. Allele Information of the Four SNPs In and Around *TAS2R16*

<i>rs Number</i>	<i>Ancestral on + Strand</i>	<i>Derived on + Strand</i>	<i>Reference on + Strand</i>	<i>Strand Reported in UCSC</i>
rs1357949	A	G	A	+, plus
rs846664	C	A	A	-, minus
rs702424	A	G	G	-, minus
rs9691345	C	T	C	+, plus

In this paper we have referred to alleles on the plus strand, but in this table we show the “reference alleles” and strand reported in dbSNP and UCSC for comparison.

relative EHH (REHH) (Sabeti et al. 2007). As the REHH value depends on the haplotype frequency, we linearized the REHH values using the curve estimation method in SPSS 13.0. The linearized REHH (IREHH) values were displayed in a contour map using the default parameters in SURFER 8.0®. Correlation between IREHH and latitude of the populations was calculated using bivariate correlation in SPSS 13.0.

Results

The Frequencies and Phylogenies of *TAS2R16* Haplotypes. Allele frequencies for all the SNPs in all of the populations are in ALFRED (http://alfred.med.yale.edu/alfred/recordinfo.asp?condition=loci.locus_uid=%27LO003182O).

The derived allele (nucleotide A; amino acid Asn) frequency of rs846664 is high to fixed in all of the populations, and the ancestral allele (nucleotide C, amino acid Lys) is seen primarily in Africa (Figure 2). The derived allele (G) frequency of rs702424 is highest in MBU, low in the other African populations, and moderate to high in all the populations of other continents except for NAS and KAR, two small and isolated populations.

The four SNPs form eight observed haplotypes, only four of which are common (Figures 2 and 3). The ancestral haplotype ACAC for all four sites occurs almost exclusively in Africa and is the only one with the ancestral Lys172 allele. The other three haplotypes, AAAC, AAGC, and GAAT, are all common throughout the world, dividing the chromosomes with the derived 172Asn into three common haplotypes. The haplotypes GAAC and AAAT, one of which is the likely intermediate evolutionary step between AAAC and GAAT (Figure 3), are quite rare. The rare occurrences of GAAC in four geographically separated populations suggest these chromosomes may be the results of recent recombination events, not the remnants of the original intermediate. This was confirmed by examination of their flanking polymorphisms. In the shortest network of the long-range haplotypes defined by all 55 SNPs, GAAC appears on three different terminals, while AAAT lies between the AAAC and GAAT groups (figure produced by Network is on the authors’ websites). In any case, the rarity of both haplotypes raises the question of why the derived GAAT haplotype is so common.

EHH and REHH. The LRH analysis was done in each population based on the haplotypes of all the markers we typed. We chose rs846664 and rs702424 to

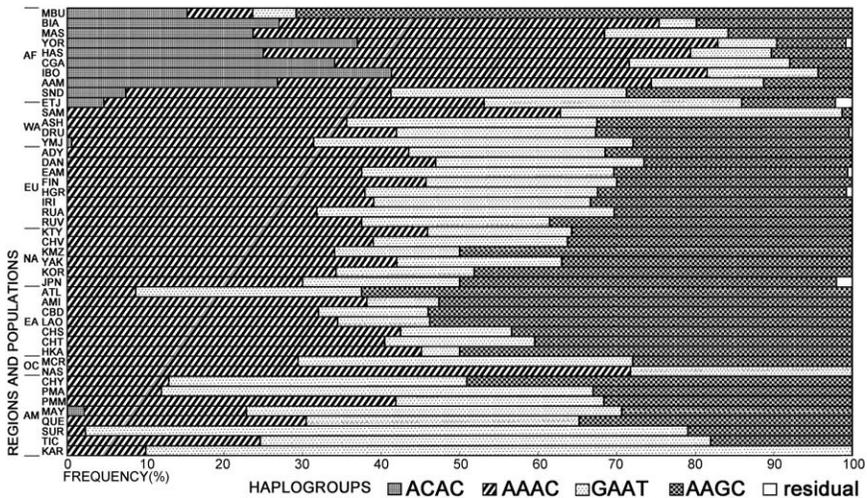


Figure 2. Haplotype diversity of TAS2R16 among the global populations. The haplotype frequencies are presented in stacked-bar format, and the 45 populations studied are grouped geographically: Africa, western Asia, Europe, northern Asia, eastern Asia, Oceania, and the Americas. The haplotype frequencies are also in ALFRED (<http://alfred.med.yale.edu>). The SNPs forming the core haplotypes are rs1357949, rs846664, rs702424, and rs9691345, in that order. Color versions of this figure and others in, and supplemental to, this paper are available on the Kidd Lab website at <http://info.med.yale.edu/genetics/kkidd/contents.html> or on Hui Li's website at <http://comonca.org.cn/LH/Doc/A52S.htm>.

form the core haplotype to calculate EHH and REHH. The haplotype AG (with derived alleles for both SNPs) tends to have higher EHH and REHH values than the other haplotypes, especially in the European, Western Asian, and Northern Asian populations (Figure 4). The REHH statistic eliminates the difference among the polymorphism diversities and therefore is more sensitive than the EHH statistic. However, the significance of the REHH statistic largely depends on the haplotype frequency. The haplotypes with a low frequency will exhibit a much higher REHH value than those with higher frequencies under the same selective stress but will be less significant. Therefore, we have transformed the REHH values to make them independent of haplotype frequency and thus informative for comparison among populations.

Linearization. We assumed that the core haplotypes other than AG have not undergone selection as their REHH and EHH values are relatively low. For a new statistic independent of the population haplotype frequencies, the REHH values of populations for the non-selected haplotypes should exhibit a flat horizontal distribution as a function of the haplotype frequencies. We chose two positions, 100 kb downstream and 120 kb upstream from the core haplotype, to estimate the distribution of REHH values of the 45 populations (Table 2; Figure 5A). For both locations, the distributions of population-specific REHH values most closely fit

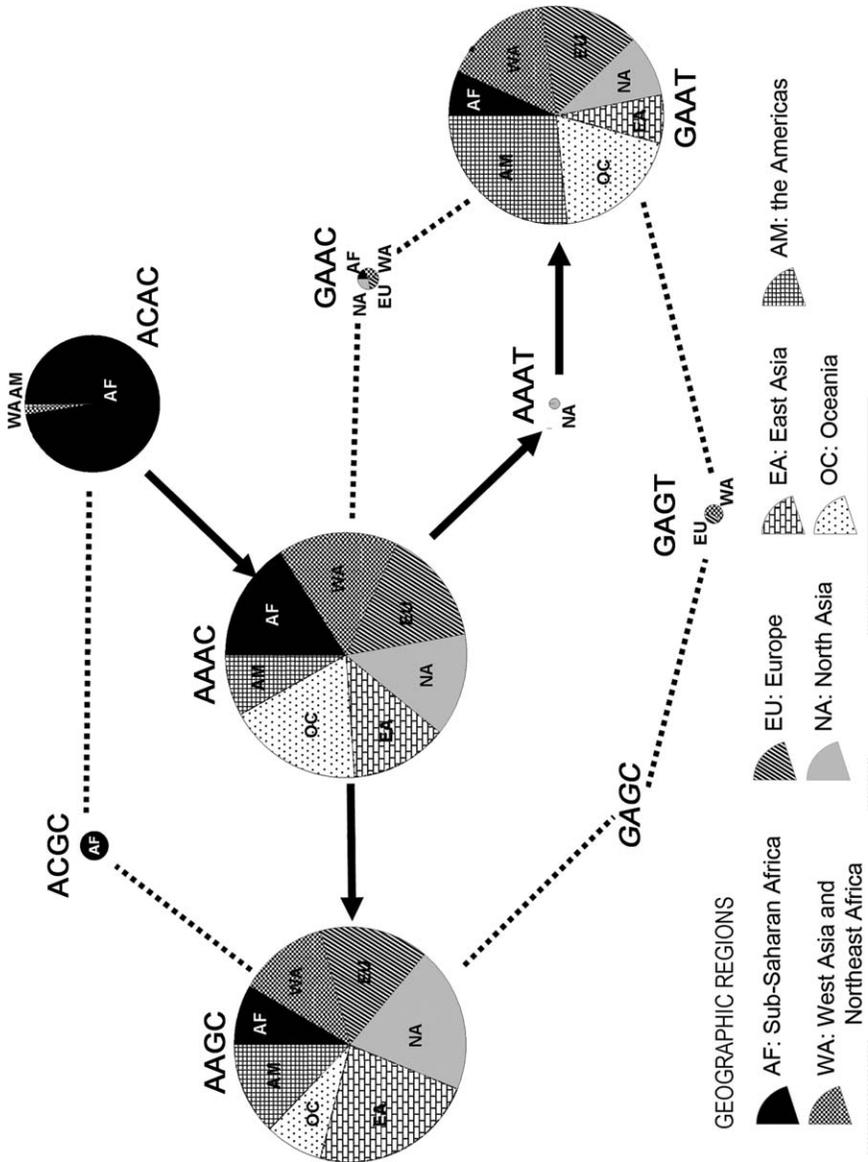


Figure 3. Evolutionary relationships among the core four-SNP haplotypes in Figure 2. The solid arrows indicate the likely historical accumulation of mutations, and the dotted lines indicate possible recombinations. The four SNPs are those in Figure 2, in the same order, with ACAC the ancestral haplotype.

logarithmic curves among the eleven curve types tested using SPSS. For 100 kb downstream, the curve is $y = 0.059 - 0.812 \ln x$ ($R^2 = 0.312$, $F = 27.605$, $P < 0.001$), yielding the transformation, $IREHH = (0.059 - REHH)/0.812 \ln(\text{Freq})$ to

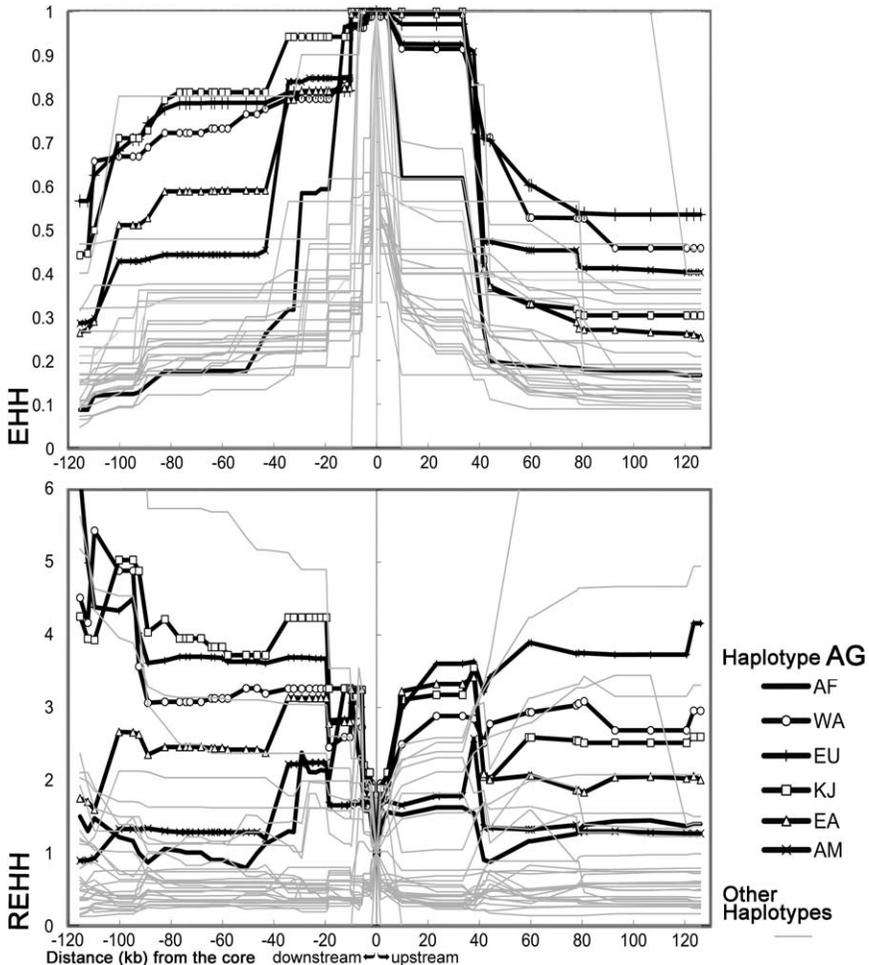


Figure 4. Graphs of EHH and REHH values for the core haplotypes formed by rs846664 and rs702424. The core with both derived alleles is the only one showing a large signal of selection; all others have been combined and are represented by one line in the graph. To simplify the plots, we combined the populations with similar core haplotype frequencies (using Fisher's exact test) represented by the more intense lines. The region abbreviations are the same as those in Figure 2. Other individual populations showing evidence of selection have been omitted, but data are in Table 2.

linearized REHH. For 120 kb upstream, the curve is $y = 0.285 - 0.553 \ln x$ ($R^2 = 0.227$, $F = 17.872$, $P < 0.001$), and the transformation formula is $\text{IREHH} = (0.285 - \text{REHH})/0.553 \ln(\text{Freq})$. The IREHH values of haplotype AG increase with the haplotype frequency, while those of the other haplotypes remain constant (Figure 5B). It is reasonable to assume that the stronger the positive selection, the higher the frequency would be. The African populations that showed evidence of selection in the study of Soranzo et al. (2005) have quite low IREHH values except for Mbuti,

Table 2. Selection Test Values for the AG Core Haplotype Formed by rs846664 and rs702424

<i>Distance from Core</i>		<i>100kb Downstream</i>			<i>120kb Upstream</i>		
<i>Population</i>	<i>Frequency</i>	<i>EHH</i>	<i>REHH</i>	<i>IREHH</i>	<i>EHH</i>	<i>REHH</i>	<i>IREHH</i>
AAM	0.113	0.251	1.990	1.088	0.316	3.060	2.053
BIA	0.198	0.330	3.029	2.204	0.173	1.125	0.953
CGA	0.080	0.095	1.206	0.572	0.048	0.450	0.268
HAS	0.103	0.143	1.486	0.780	0.143	0.966	0.626
IBO	0.043	0.167	1.815	0.694	0.333	2.571	1.270
MAS	0.158	0.200	2.408	1.546	0.133	1.204	0.922
MBU	0.708	0.623	3.663	10.793^b	0.565	4.034	8.476^b
SND	0.287	0.265	1.635	1.524	0.344	3.577	3.668
YOR	0.089	0.218	1.808	0.894	0.321	2.028	1.250
ASH	0.325	0.716	3.946	4.061	0.595	2.423	2.673
DRU	0.327	0.627	5.228	5.408	0.401	2.597	2.876
ETJ	0.125	0.536	5.691	3.257	0.321	2.740	1.909
SAM	0.013	1.000	3.367	0.939	1.000	5.225	1.945
ADY	0.315	0.631	4.119	4.735	0.588	2.073	4.310
DAN	0.265	0.652	4.721	3.093	0.468	3.982	2.954
EAM	0.304	0.584	3.518	3.773	0.556	3.011	3.734
FIN	0.300	0.729	3.871	4.567	0.519	3.523	4.938
HGR	0.317	0.742	4.734	4.231	0.516	4.695	3.620
IRI	0.333	0.699	4.197	4.695	0.521	3.332	3.796
RUA	0.303	0.721	4.469	4.425	0.558	3.390	4.810
RUV	0.385	0.465	4.551	4.659	0.701	4.547	5.381
YMJ	0.279	0.699	3.886	3.760	0.428	4.374	2.092
CHV	0.300	0.438	2.842	2.742	0.507	2.346	2.467
JPN	0.480	0.730	5.528^a	8.440^b	0.286	2.098	3.037
KMZ	0.468	0.868	3.788	5.607	0.710	3.686	5.229
KOR	0.481	0.688	4.836	7.402^b	0.317	2.476	3.590
KTY	0.357	0.889	4.319	4.824	0.459	3.356	3.927
YAK	0.370	0.351	2.131	2.460	0.236	1.744	2.089
AMI	0.526	0.656	1.979	3.408	0.345	1.993	3.113
ATL	0.625	0.679	1.420	3.223	0.302	0.647	1.187
CBD	0.540	0.513	2.495	4.461	0.179	2.162	3.455
CHS	0.434	0.620	3.530	4.791	0.351	2.493	3.339
CHT	0.404	0.474	2.184	2.747	0.330	2.191	2.787
HKA	0.256	0.405	0.868	0.745	0.462	1.709	1.646
LAO	0.538	0.440	2.842	5.054	0.227	1.759	2.802
MCR	0.279	0.480	2.506	2.287	0.444	1.347	1.359
CHY	0.491	0.403	1.574	2.473	0.422	1.739	2.564
MAY	0.293	0.464	1.898	1.798	0.293	0.964	1.000
PMA	0.330	0.506	1.069	1.114	0.348	1.001	1.115
PMM	0.316	0.534	2.340	2.353	0.595	2.158	2.340
QUE	0.348	0.600	2.310	2.522	0.283	1.044	1.202
SUR	0.209	0.889	1.247	0.938	1.000	1.244	1.081
TIC	0.180	0.745	3.066	2.112	0.745	3.667	2.973

a. $P < 0.05$.

b. $P < 0.01$.

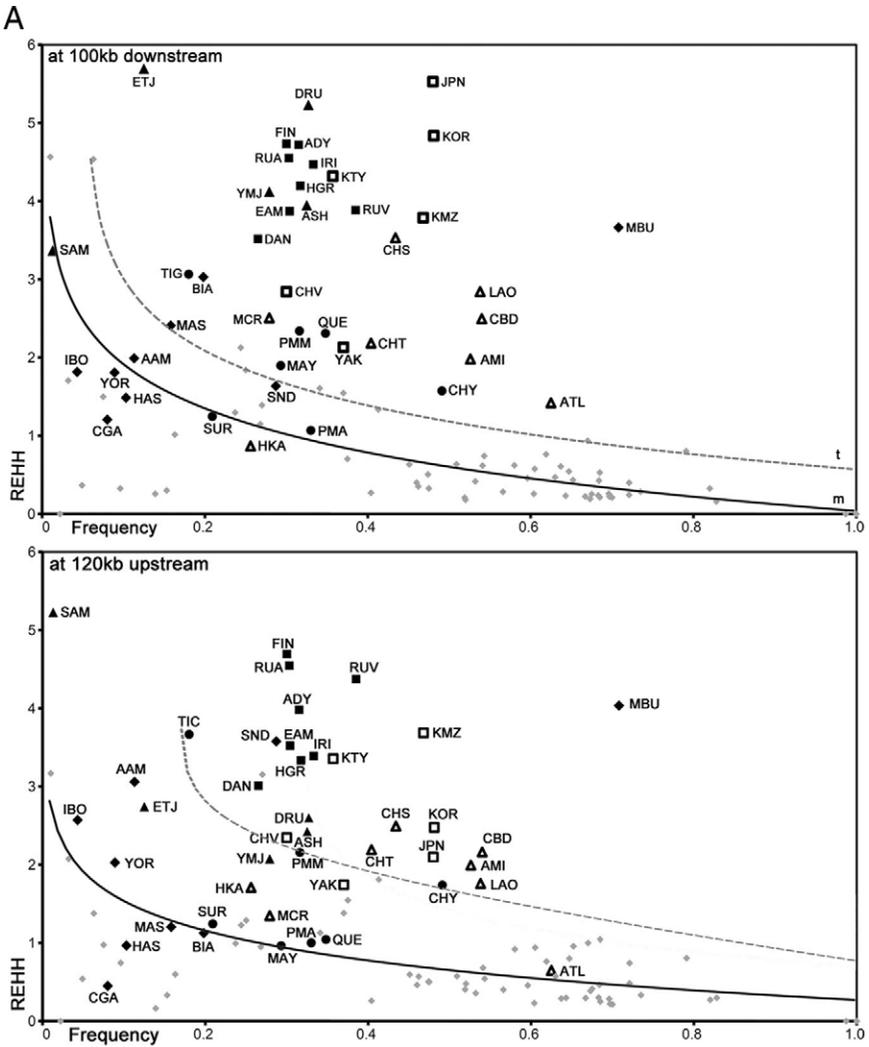


Figure 5. Reference frame of REHH-Haplotype frequency (A) and IREHH-Haplotype frequency (B). Curve *m* was estimated based on all of the REHH values of the core haplotypes other than AG, and curve *t* was estimated based on the highest REHH values of the same haplotypes.

while the North Asian populations have quite high IREHH values, indicating strong selection signals.

Geographic Distribution of IREHH. In Figure 5, the IREHH values for haplotype AG seem to be associated with geographic regions. Most high IREHH values appear in Eurasia except for Mbuti, a hunter-gatherer Pygmy population. Two areas in Eurasia, Mesopotamia and southeast China, display especially low IREHH. We also examined the correlation between IREHH and latitude. The

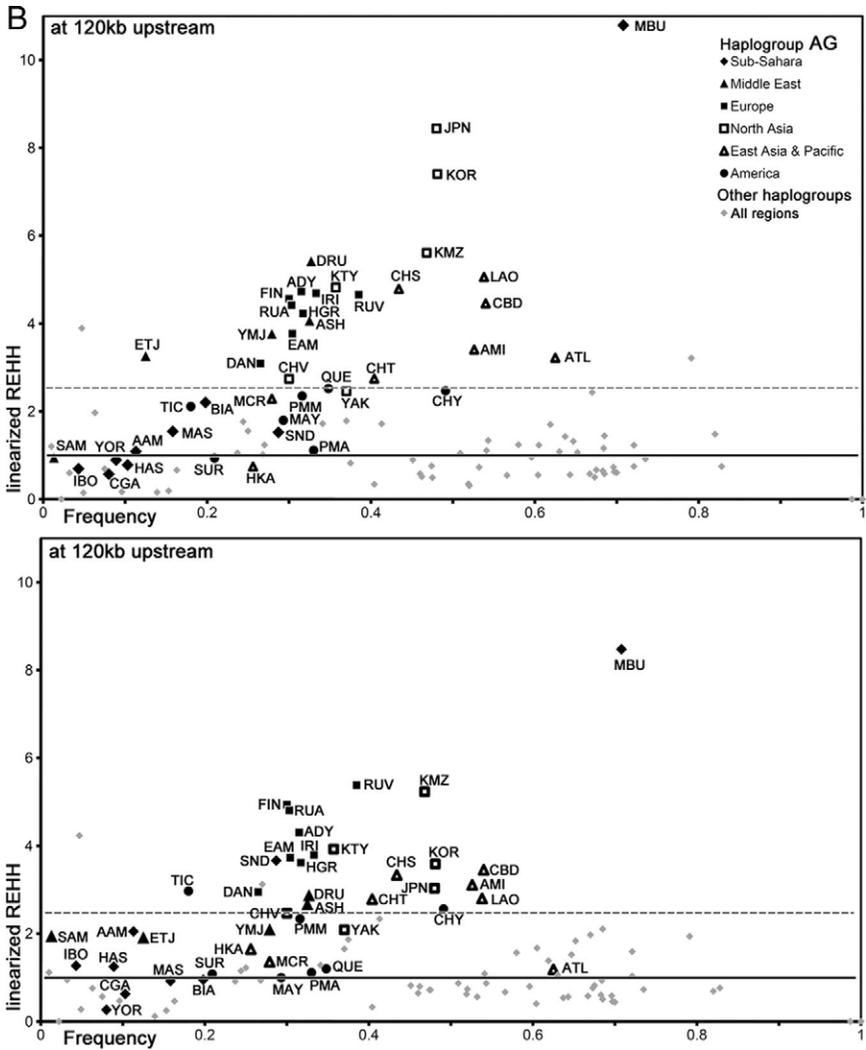


Figure 5. (continued)

correlation is marginally significant although not high (Pearson correlation $r = 0.317$, $P = 0.039$), suggesting a stronger signal for the populations that migrated to the north, especially in East Asia.

Discussion

Bitter taste is thought to provide an important protection against toxins in food (Rozin and Vollmecke 1986). Any positive selection on *TAS2R16* that generated the quite variable REHH we observe must have occurred during more recent human history, after human expansion around the world. Different

methods are optimal for detecting the signal of selection that occurred at different past times. The LRH test is limited to detecting signals of selection occurring during the most recent ten thousand years (Sabeti et al. 2002). This test did not find a selection signal in most of the African populations, whereas Fay and Wu's H test did (Soranzo et al. 2005). Modern humans evolved in Africa (Cann et al. 1987; Ingman et al. 2000; Gilbbons 1997; Jin and Su 2000) and, hence, modern African populations have a history much deeper than non-African populations. The selection in Africa for 172Asn was estimated to be old, predating the expansion of modern humans out of Africa. However, the 172Asn allele did not reach fixation. The near fixation of that allele outside of Africa is consistent with it having been selected for but could be attributed to the bottleneck that occurred with the expansion out of Africa. In either case, the near fixation precludes using the H or any other statistics to test for selection on that variant. Our finding argues that much more recent selection has occurred because one specific 172Asn haplotype acquired an additional variant providing an advantage to that haplotype relative to the others.

We hypothesize that by 10,000 yrs ago human populations in Africa would have already accumulated enough knowledge to distinguish edible plants and subsequently not have had to depend on bitter taste. That may be why we cannot find a recent selection signal in most African populations. The selection signal in the Mbuti is strange under this scenario, and more information is needed to explain it. One possibility is that this population's small effective population size, indicated by other studies (e.g., Kidd and Kidd 2008; Tishkoff and Kidd 2004), has resulted in a rapid increase in this haplotype simply because of random genetic drift.

The strong selection signals in Eurasia suggest quite recent selection on bitter taste across the entire region. We do not know the molecular nature or functional consequences of the new variation responsible for the evidence of selection, but we can speculate. When modern humans expanded out of Africa into Eurasia ~80 thousand years ago, they were challenged to recognize the safe foods in the new environment. Accumulating knowledge of edible plants may have extended over a long period both because humans continuously expanded into new ecosystems and because hunting may have been a more significant source of nutrition. As gathering became a more important source of food, the selective advantage of some aspect of bitter taste may have become even stronger. Our results suggest that a gathering lifestyle was maintained at least into the most recent 10,000 yrs, especially in East Asia and in northern latitudes, regions in which the ecological changes associated with the end of the last glacial maximum may have been great.

Several possible explanations exist for the absence of any signal of selection in Native Americans. Their distinctive history may not be long enough to accumulate enough selection influences for the LRH test, or there has never been any selective force, possibly because of different plants being consumed. We think one likely explanation is that the cryptic (to our study) variation under selection does not exist in the Native Americans. These populations experienced another bottleneck when the

ancestors of our sampled populations left central Asia, migrated across Beringia, and then expanded throughout the Americas. Also, in the northerly climes they occupied for many centuries, populations likely carried on a primarily hunting lifestyle. The population that then expanded to temperate and tropical regions experienced new plant sources of nutrition. Many Native American populations were among the very first to domesticate plants and developed a diverse assembly of domesticated plants in early Central American agriculture, involving also many latitudes and climates (Mazoyer 2006; Meillassoux 1983). There are also many bitter toxic plants chosen as food resources in the Americas such as Cassava. Therefore, selection on bitter taste should have lasted for some thousands of years, enough time to accumulate selection signals if there were relevant genetic variation. We see no evidence that selection has occurred through much of the pre-agricultural and agricultural history of Native Americans.

That taste may have been an important aspect of past human culture can be found in the legend of *Shennong*, the God of Medicine in both China and India. *Shennong* was believed to be an emperor of China around 5,000 yrs ago. Legend says at that time the Chinese population increased rapidly, and there were not enough wild animals for food any more. As people tried to gather plant food instead, many people died of toxic foods as they could not recognize them. *Emperor Shennong* saved the people by tasting and recording hundreds of herbs and crops. That was the traditional beginning of Chinese medicine and agriculture, and the basis for regarding *Shennong* as the God of Medicine and Agriculture (Hoizey and Hoizey 1993; Tsuei 1992). This legend could contain a social memory reflecting early history rather than being purely myth and could indicate that the gathering lifestyle in East Asia was not as ancient as we tend to believe and could have served as a relatively recent selective force on bitter taste. The earliest agriculture started in Mesopotamia with wheat and in East China with rice (Zong et al. 2007) and could have decreased the selective stress in these two areas. Yet, adoption of an agricultural economy and lifestyle was not uniform. For example, the strength of evidence of selection at *ADH1B* in East Asia varies according to how early the population adopted agriculture (Li et al. 2008).

Both the non-synonymous SNP rs846664 and the regulatory-region SNP rs702424 seem to be related to selection, as only the core haplotype with both derived alleles exhibits the selection signal. This region is 4.7 kb long, and it is possible that other variants in this small region are in (nearly) complete LD with these two SNPs. Some variants have already been identified (cf dbSNP) and the ongoing whole genome resequencing is underway in many labs, and perhaps targeted resequencing will potentially determine whether or not other relevant variants exist. In any case, future functional studies of *TAS2R16* should consider the regulatory region. For instance, the ancestral allele of rs846664 (C or 172Lys), which occurs primarily in Africans, was strongly associated with higher risk for alcoholism (Hinrichs et al. 2006; Wang et al. 2007) in African Americans. However, we now see that most African samples contain both ancestral alleles, C of rs846664 and A of rs702424; therefore, the effects on risk of alcoholism of these two SNPs cannot be separated in

existing data and population studies will not easily distinguish effects of the two because of the nearly complete LD. More studies of African populations for both SNPs and risk of alcoholism are needed to clarify the relationships of genotypes and risk because our study has shown that the evolution of the variation at *TAS2R16*, both at the molecular and the geographic levels, is much more complex than a single variant can show.

Acknowledgments We thank all the colleagues who helped us assemble the population samples, as well as the Coriell Institute for Medical Research (NIGMS Human Genetic Cell Repository) and the National Laboratory for the Genetics of Israeli Populations at Tel-Aviv University. Special thanks are due the many hundreds of individuals from these populations who volunteered to give blood samples for studies such as this. This research was supported in part by U.S. Public Health Service Grants AA009379 and GM057672 and by National Science Foundation Grant BCS0725180. None of the authors has any conflicts of interest related to the data presented here.

Received 23 November 2010; revision accepted for publication 2 February 2011.

Literature Cited

- Adler, E., M. A. Hoon, K. L. Mueller et al. 2000. A novel family of mammalian taste receptors. *Cell* 100:693–702.
- Bufe, B., T. Hofmann, D. Krautwurst et al. 2002. The human TAS2R16 receptor mediates bitter taste in response to beta-glucopyranosides. *Nat. Genet.* 32:397–401.
- Cann, H. M., C. de Toma, L. Cazes et al. 2002. A human genome diversity cell line panel. *Science* 296:261–262.
- Cann, R. L., M. Stoneking, and A. C. Wilson. 1987. Mitochondrial DNA and human evolution. *Nature*. 325:31–36.
- Chandrashekar, J., M. A. Hoon, N. J. Ryba et al. 2006. The receptors and cells for mammalian taste. *Nature*. 444:288–294.
- Endo, T., K. Ikeo, and T. Gojobori. 1996. Large-scale search for genes on which positive selection may operate. *Mol. Biol. Evol.* 13:685–690.
- Fay, J. C. and C.-I. Wu. 2000. Hitchhiking under positive Darwinian selection. *Genetics* 155:1405–1413.
- Gilbbons, A. Y. 1997. Y chromosome shows that Adam was African. *Science* 278:31–36.
- Go, Y., Y. Satta, O. Takenaka et al. 2005. Lineage-specific loss of function of bitter taste receptor genes in humans and non-human primates. *Genetics* 170:313–326.
- Hinrichs, A. L., J. C. Wang, B. Bufo et al. 2006. Functional variant in a bittertaste receptor (hTAS2R16) influences risk of alcohol dependence. *Am. J. Hum. Genet.* 78:103–111.
- Hoizey, D., and M. J. Hoizey. 1993. *A History of Chinese Medicine*. Edinburgh, U.K.: Edinburgh University Press Ltd., 47.
- Ingman, M., H. Kaessmann, S. Paabo et al. 2000. Mitochondrial genome variation and the origin of modern humans. *Nature*. 408:708–13.
- Jin, L., and B. Su. 2000. Natives or immigrants: Origin and migrations of modern humans in East Asia. *Nat. Rev. Genet.* 1:126–133.
- Kidd, K. K., and J. R. Kidd. 2008. Human genetic variation of medical significance. In *Evolution in Health and Disease*, 2nd Edition, S. C. Stearns and J. C. Koella., eds. New York, NY: Oxford University Press, 51–62.

- Kingsbury, J. 1964. *Poisonous Plants of the United States and Canada*. Englewood Cliffs, NJ: Prentice Hall.
- Mazoyer, M., and L. Roudart. 2006. *A History of World Agriculture: From Neolithic Age to the Current Crisis*. Earthscan, London/Sterling, VA: Monthly Review Press, 528.
- Meillassoux, C. 1983. *Maiden, Meal and Money: Capitalism and the Domestic Community*. New York, NY: Cambridge University Press, 196.
- Osier, M. V., K. H. Cheung, J. R. Kidd et al. 2002. ALFRED: An allele frequency database for Anthropology. *Am. J. Phys. Anthropol.* 119:77–83.
- Rozin, P., and T. A. Vollmecke. 1986. Food likes and dislikes. *Annu. Rev. Nutr.* 6:433–456.
- Sabeti, P. C., D. E. Reich, J. M. Higgins et al. 2002. Detecting recent positive selection in the human genome from haplotype structure. *Nature.* 419:832–837.
- Sabeti, P. C., P. Varilly, B. Fry et al. 2007. Genome-wide detection and characterization of positive selection in human populations. *Nature.* 449:913–918.
- Stephens, M., and P. Scheet. 2005. Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. *Am. J. Hum. Genet.* 76:449–462.
- Soranzo, N., B. Bufo, P. C. Sabeti et al. 2005. Positive selection on a high-sensitivity allele of the human bitter-taste receptor TAS2R16. *Curr. Biol.* 15:1257–1265.
- Steiner, J. E. 1994. Behavior manifestations indicative of hedonics and intensity in chemosensory experience. In *Olfaction and Taste XI*, K. Kurihara, N. Suzuki, and H. Ogawa, eds. Tokyo, Japan: Springer, 284–287.
- Tishkoff, S. A., and K. K. Kidd. 2004. Implications of biogeography of human populations for “race” and medicine. *Nat. Genet.* 36:S21–S27.
- Tsuei, W. 1992. *Roots of Chinese Culture and Medicine*. Selangor Darul Ehsan, Malaysia: Pelanduk Publications, 9.
- Wang, J. C., A. L. Hinrichs, S. Bertelsen et al. 2007. Functional variants in TAS2R38 and TAS2R16 influence alcohol consumption in high-risk families of African-American origin. *Alcohol Clin. Exp. Res.* 31:209–215.
- Zong, Y., Z. Chen, J. B. Innes et al. 2007. Fire and flood management of coastal swamp enabled first rice paddy cultivation in east China. *Nature.* 449:459–462.