Genes from the *TAS1R* and *TAS2R* Families of Taste Receptors: Looking for Signatures of Their Adaptive Role in Human Evolution

Cristina Valente[1,2,3], Luis Alvarez[1,2], Patrícia Isabel Marques[1,2], Leonor Gusmão[4], António Amorim[1,2,3], Susana Seixas[1,2], and Maria João Prata[1,2,3,*]

[1]I3S, Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Portugal

[2]IPATIMUP, Institute of Molecular Pathology and Immunology, University of Porto, Portugal

[3]Faculty of Sciences, University of Porto, Portugal

[4]DNA Diagnostic Laboratory (LDD), State University of Rio de Janeiro (UERJ), Brazil

*Corresponding author: E-mail: mprata@ipatimup.pt.

Accepted: March 27, 2018

Abstract

Taste perception is crucial in monitoring food intake and, hence, is thought to play a significant role in human evolution. To gain insights into possible adaptive signatures in genes encoding bitter, sweet, and *umami* taste receptors, we surveyed the available sequence variation data from the 1000 Genomes Project Phase 3 for *TAS1R* (*TAS1R1-3)* and *TAS2R* (*TAS2R16* and *TAS2R38)* families. Our study demonstrated that genes from these two families have experienced contrasting evolutionary histories: While *TAS1R1* and *TAS1R3* showed worldwide evidence of positive selection, probably correlated with improved *umami* and sweet perception, the patterns of variation displayed by *TAS2R16* and *TAS2R38* were more consistent with scenarios of balancing selection that possibly conferred a heterozygous advantage associated with better capacity to perceive a wide range of bitter compounds. In *TAS2R16*, such adaptive events appear to have occurred restrictively in mainland Africa, whereas the strongest evidence in *TAS2R38* was detected in Europe. Despite plausible associations between taste perception and the *TAS1R* and *TAS2R* selective signatures, we cannot discount other biological mechanisms as driving the evolutionary trajectories of those *TAS1R* and *TAS2R* members, especially given recent findings of taste receptors behaving as the products of pleiotropic genes involved in many functions outside the gustatory system.

Key words: TAS1R and TAS2R, taste receptors, signatures of selection.

Introduction

Several lines of evidence support the idea that taste perception plays a key role in food preference, dietary habits, and many other health issues. Humans are able to discriminate five tastes: sweet, *umami*, sour, salty, and bitter, which are commonly referred as the basic tastes. However, this apparently limited repertoire seems enough to accommodate the evolutionary demand for recognition of essential dietary elements, while avoiding potential dietary threats with negative impacts on nutritional and physiological status (Chandrashekar et al. 2006).

Sweet, bitter, and *umami* are considered the most important tastes for food acceptance in humans (Temussi 2009). They depend on the activation of different receptors of the seven-transmembrane G-protein-coupled receptors (GPCRs) superfamily, more specifically from the TAS1R and TAS2R classes, which are proteins coexpressed in distinct subpopulations of taste bud cells of the human gustatory system. So far, *umami* and sweet have been reported to be perceived uniquely by TAS1R receptors, a small family of GPCR proteins that in humans, as in many other mammals, includes three members—TAS1R1, TAS1R2, and TAS1R3. These proteins only function as heterodimers and, whereas the dimer TAS1R1 + TAS1R3 acts as the main receptor for *umami*, TAS1R2 + TAS1R3 responds to a broad variety of natural and artificial sweet ligands (Li et al. 2002; Nelson et al. 2002). In turn, bitter is mainly mediated by receptors of the TAS2R family, which in humans contains at least 25 members, all of them active as monomeric receptors. The most studied receptors of this GPCR family are TAS2R38, which determines

phenylthiocarbamide (PTC) and 6-n-propylthiouracil (PROP) sensitivity (Kim et al. 2003), and TAS2R16, which responds to β-glucopyranosides such as salicin found in willow bark (Bufe et al. 2002).

Large interindividual variability exists in human taste perception but, to date, its genetic basis is largely unclear (Bachmanov and Beauchamp 2007). Still, it is already known that a substantial proportion of the differences in PTC/PROP sensitivity is explained by the three *TAS2R38* nonsynonymous substitutions responsible for amino acid changes at residues 49,262, and 296 (rs713598, rs1726866, rs10246939, respectively). The haplotypes defined by the various combinations of variants in these positions are strongly correlated with bitter perception, allowing the *TAS2R38* haplotypes associated with bitter sensitivity or insensitivity to be classified as "tasters" and "nontasters," respectively (Wooding et al. 2004). In *TAS2R16*, the nonsynonymous variant G516T (rs846664) was identified as influencing phenotypic taste variation in response to salicin and other naturally occurring glycosides (Soranzo et al. 2005).

Likewise, *TAS1R1* and *TAS1R3* contain several coding variants that have been correlated in vitro with dose response to *umami*: in *TAS1R1*, the C329T (rs41278020) and G1114A (rs34160967), and in *TAS1R3* the G13A (rs76755863) and C2269T (rs307377) (Raliou et al. 2009; Shigemura et al. 2009). In addition, two promoter region variants in *TAS1R3* (-T1572C (rs307355) and -T1266C (rs35744813)) were recently shown to influence in vitro gene expression and to affect sensitivity to sucrose in humans, explaining almost 16% of sweet taste variability (Fushan et al. 2009). Lastly, a candidate variant in *TAS1R2* was reported to possibly affect responsiveness to sweet *stimuli*; this variant, G571A (rs35874116), is a nonsynonymous substitution found to be associated with habitual consumption of sugars (Eny et al. 2010).

Since diversity in *TAS1R/2R* genes is likely associated with a wide range of food preferences observed across distinct geographic regions and ethnic groups, it has been hypothesized that *TAS1R/2R* evolution was driven by diet-related selective pressures. A few studies have already shed some light on signatures of selection at genes encoding the bitter receptors *TAS2R38* (Wooding et al. 2004; Campbell et al. 2012; Risso et al. 2016) and *TAS2R16* (Soranzo et al. 2005; Li et al. 2011; Campbell et al. 2014), as well as the intriguing departures from neutrality at the three *TAS1R* genes (Kim et al. 2006).

Here, we use the human variation data released by the 1000 Genomes Project consortium to interrogate the site frequency spectrum of *TAS1R1*, *TAS1R2*, *TAS1R3, TAS2R16*, and *TAS2R38*, in order to gain further insight into the evolutionary scenarios of these genes across a wider spectrum of human populations. To our knowledge, this represents the first work to: 1) combine the investigation of *TAS1R* and *TAS2R* gene families; 2) apply common analytical approaches across several unlinked genes; 3) use the identical panel of human populations for all genomic regions investigated; and 4) analyze *TAS1R3* promoter variation in an evolutionary framework.

Overall, the findings of this study reinforce the evidence for adaptive evolution of *TAS1R* and *TAS2R* genes throughout more recent (100–600 ka) or ancient (2–7 Ma) human history, respectively, while raising a number of questions to address in the future. Which environmental changes triggered human preference for *umami* and sweet tastes (*TAS1R1* and *TAS1R3*)? Could this preference be connected with human dependence on cooked food? Which evolutionary forces are maintaining a wide spectrum of bitterness perceptions (*TAS2R16* and *TAS2R38*) over several million years? Why is the *TAS2R16* selective signature only observed in Africans?

Materials and Methods

Samples and Sequence Databases

The coordinates, based on the genomic built GRCh37: CM000663.1, for the sequences used in this study are: *TAS1R1*—chr1: 6,615,241–6,639,817; *TAS1R2*—chr1: 19,166,093–19,186,176; *TAS1R3*—chr1: 1,265,014–1,270,651; *TAS2R16*—chr7: 122,634,759–122,635,754; and *TAS2R38*—chr7: 141,672,501–141,673,397. All sequences were retrieved from the 1000 Genomes Project Phase 3 database (www.1000genomes.org). After excluding all potentially admixed populations, a panel of 20 populations representing major geographical regions was selected for this study: Eastern Africa—LWK (Luhya, Kenya); Western Africa—YRI (Yoruba, Nigeria), GWD (Gambian, Gambia), MSL (Mende, Sierra Leone), ESN (Esan, Nigeria); North/Western Europe—CEU (Utah residents with Northern and Western European ancestry), GBR (British, England, and Scotland), FIN (Finnish, Finland); South Europe—IBS (Iberian, Spain), TSI (Tuscan, Italy); South Asia—PJL (Punjabi, Pakistan), BEB (Bengali, Bangladesh), STU (Sri Lankan Tamil, from UK), ITU (Indian Telugu, from UK), GIH (Gujarati Indian, from Houston); and East Asia—CHB (Chinese, from Beijing, China), CHS (Southern Han Chinese, China), JPT (Japanese, Japan), CDX (Chinese Dai, China), KHV (Kinh, Vietnam). For each variant, ancestral allele information was obtained using BioMart tool (http://www.ensembl.org/biomart/martview).

Using the Neanderthal Genome browser (http://neanderthal.ensemblgenomes.org/index.html), we inspected the available Neanderthal genome, which is based on the sequencing of three individuals (Vi33.16; Vi33.25; and Vi33.26) at very limited coverage. On the other hand, the UCSC browser (http://genome.ucsc.edu/) was used to query the Denisova genome, which results from the sequencing of a single individual at high coverage (30×).

Statistical Analyses

Phased data were retrieved from the 1000 Genomes Project Phase 3 browser (http://phase3browser.1000genomes.org/index.html), and SPIDER software (Lischer and Excoffier 2012) was used to convert .vcf files into .xml files. To

identify signatures of selection, we calculated several statistics based on the allele frequency spectrum for each population. The DivSat program (Soares et al. 2015) was used to estimate the following statistics: number of segregating sites (S), nucleotide diversity (π), or average number of pairwise differences between sequences (Nei and Li 1979), Watterson's estimator of the population mutation rate parameter ($\theta_W = 4Ne\mu$) (Watterson 1975), and Tajima's *D* (Tajima 1989). The latter is the statistic most widely used for detecting departures from neutrality by evaluating discrepancies between π and θ_W values. Briefly, Tajima's *D* is expected to have a value close to 0 under neutrality. A negative value generally indicates an excess of rare variants, probably resulting from positive selection or population expansion, whereas a positive value suggests an excess of intermediate-frequency variants caused either by balancing selection or population structure.

To address the evolutionary significance of Tajima's *D* results, we ran 100,000 coalescent simulations using "ms" program (Hudson 2002) in all populations for the constant size model and additionally for the best-fit demographic models suited for specific population sets when Tajima's *D* values reached statistical significance under the constant model. The Gravel et al. (2011) models inferred from the 1000 Genomes Phase 1 data for YRI, CEU, and CHB+ JPT, were applied to the identical populations and to other samples located in the same geographical regions. Specifically, the YOR model, which was inferred for a population from Sub-Saharan Africa, was also applied to GWD, MSL, ESN, and LWK; the CEU model, which corresponds to a sample of Utah residents with Northern and Western European ancestry, was also applied to GBR and FIN; and the CHB + JPT model, which is expected to represent East Asian demography, was also applied to CHS, CDX, and KHV samples. In addition, the Voight et al. (2005) model inferred from 50 noncoding autosomal regions for an Italian population was applied to the two South European samples, TSI and IBS. To our knowledge, no proxy of best-fit model is available for South Asians, and thus no model other than the constant size model was applied to PLJ, ITU, STU, GIH, and BEB samples.

Estimates of the population recombination parameter Rho (ρ), which combines information on effective population size (N_e) and recombination rate (*r*) as summarized in the equation $\rho = 4N_e r$, were based on values of r for each gene obtained from HapMap Phase II (McVean et al. 2004), assuming an ancestral population size of N_e=7,300 individuals according to Gravel et al. (2011). Independent of the simulated demographic model, the null distributions of Tajima's *D*s were used to calculate either the 5th (for negative Tajima's *D* values) or 95th (for positive Tajima's *D* values) percentiles.

Moreover, to perform an empirical comparison for Tajima's *D*s using genome-wide data, tracks available at the POPHUMAN browser (Casillas et al. 2018) were downloaded. POPHUMAN contains Tajima's *D* statistics calculated in nonoverlapping sliding windows of 10 kb covering almost 90% of genome data from 1000 Genomes Project Phase 3. For each population tracks from chromosome 1 (where the *TAS1R* genes locate) and chromosome 7 (where the *TAS2R* genes locate) were used to build Tajima's *D* boxplots.

Phylogenetic relationships between haplotypes were assessed through the median-joining algorithm implemented in the NETWORK v.4.6.1.0 program (Bandelt et al. 1999). Time to Most Recent Common Ancestor (TMRCA) was estimated using the coalescent method implemented in GENETREE v.9.0 software (Griffiths and Tavare 1994), using the maximum likelihood of θ (theta) also given by GENETREE (Coop and Griffiths 2004). Given that the GENETREE coalescent method does not assume recombination, the rare and recombinant haplotypes producing incompatibilities were removed from the analysis. Mutation rate per generation per base pair, $\mu = (D_{xy}/2n)L$, was estimated considering: D_{xy} to be the differences in sequence between human and chimpanzee calculated with the DnaSP software (Rozas 2009); $2n$ to be the number of generations elapsed since the human/chimpanzee divergence (5.4 Ma) (Patterson et al. 2006) with a generation time of 25 years; and *L* as the length of genomic sequence. For TAS2R16, due to the unusual TMRCA value obtained with the chimpanzee as reference, the mutation rate per generation was also calculated, using human/orangutan divergence within a time frame of 15.2 Ma (http://www.timetree.org/). Considering that the theta parameter computed by GENETREE is the population mutation rate given by $\theta = 4N_e\mu$ (Watterson 1975), it was possible from the θ values to derive time scaled in $2N_e$ generations, converting coalescent units into years ($2N_e t$).

F_{ST} measures of genetic distances between African (LWK, YRI, GWD, MSL, ESN), Asian (CHB, CHS, JPT, CDX, KHV, PJL, BEB, STU, ITU, GIH), and European (CEU, GBR, FIN, IBS, TSI) populations were calculated in ARLEQUIN software ver. 3.5.1 (Excoffier and Lischer 2010).

Results

The *TAS1R* Family and the *Umami* and Sweet Tastes

Polymorphism Levels and Neutrality Tests

The *TAS1R* family comprises the genes *TAS1R1, TAS1R2*, and *TAS1R3*, which are distributed throughout chromosome 1 and share a common organization in five exon–introns. Analysis of the site frequency spectrum of *TAS1R* genes shows that all tend to exhibit substantially higher nucleotide diversity in Africa than in Eurasia (see π values in tables 1–3), as could be expected from the "Out-of-Africa" model of human migrations. As modern humans originated in Africa ~200,000 years ago and only a small group dispersed into Eurasia <50,000 years ago, European and Asian populations still represent a subset of the genetic variation found in African populations (reviewed in Gomez et al.2014).

Table 1
Summary Statistics for *TAS1R1*

Geographic Region	Population	*N*	*S*	π	Θ_W	Tajima's *D*[a]
Africa						
Eastern	LWK	198	294	13.785	50.145	−1.038
Western	YRI	216	241	12.671	40.503	−0.733
	GWD	226	234	13.130	39.029	−0.549
	MSL	170	224	12.741	39.229	−0.652
	ESN	198	211	12.075	35.989	−0.560
Europe						
North/Western	CEU	198	97	3.742	16.545	−1.377*
	GBR	182	92	4.553	15.921	−0.927
	FIN	198	86	3.682	14.668	−1.182
South	IBS	214	115	5.085	19.357	−1.100
	TSI	214	133	5.196	22.387	−1.340
Asia						
South	PJL	192	105	5.220	18.004	−0.898
	BEB	172	122	5.100	21.322	−1.304
	STU	204	126	5.290	21.382	−1.222
	ITU	204	88	5.070	14.933	−0.513
	GIH	206	107	5.110	18.127	−0.955
East	CHB	206	156	6.912	26.429	−1.124
	CHS	210	155	6.987	26.174	−1.080
	JPT	208	159	5.142	26.893	−1.665*
	CDX	186	148	8.085	25.516	−0.703
	KHV	198	161	7.972	27.460	−0.906

Note.—The total length of the *TAS1R1* region analyzed was 24,576 bp. *N*, number of chromosomes; *S*, number of segregating sites; π, nucleotide diversity per base pair ($\times 10^{-4}$); θ_W, population mutation rate parameter: Watterson's estimator of θ ($4Ne\mu$) (Watterson 1975) per base pair ($\times 10^{-4}$).
[a]Tajima's *D* statistic (Tajima 1989).
*$P < 0.05$ for constant model;
†$P < 0.05$ for best-fit model (Voight et al. 2005; Gravel et al. 2011).

Table 2
Summary Statistics for *TAS1R2*

Geographic Region	Population	*N*	*S*	π	Θ_W	Tajima's *D*[a]
Africa						
Eastern	LWK	198	226	18.337	38.547	−0.144
Western	YRI	216	214	19.121	35.965	0.211
	GWD	226	218	20.683	36.360	0.444
	MSL	170	211	20.177	36.952	0.307
	ESN	198	210	19.581	35.818	0.307
Europe						
North/Western	CEU	198	140	16.024	23.879	1.087
	GBR	182	125	16.164	21.632	1.571*
	FIN	198	132	15.366	22.514	1.157
South	IBS	214	135	16.347	22.724	1.381*
	TSI	214	155	15.304	26.091	0.554
Asia						
South	PJL	192	174	15.851	29.835	0.209
	BEB	172	164	16.200	28.662	0.432
	STU	204	184	16.400	31.224	0.170
	ITU	204	169	16.055	28.679	0.388
	GIH	206	175	16.000	29.648	0.242
East	CHB	206	137	13.510	23.210	0.523
	CHS	210	154	13.743	26.005	0.189
	JPT	208	112	12.619	18.943	1.044
	CDX	186	126	14.239	21.723	0.991
	KHV	198	137	13.794	23.367	0.579

Note.—The total length of the *TAS1R2* region analyzed was 20,083 bp. *N*, number of chromosomes; *S*, number of segregating sites; π, nucleotide diversity per base pair ($\times 10^{-4}$); θ_W, population mutation rate parameter: Watterson's estimator of θ ($4Ne\mu$) (Watterson 1975) per base pair ($\times 10^{-4}$).
[a]Tajima's *D* statistic (Tajima 1989).
*$P < 0.05$ for constant model;
†$P < 0.05$ for best-fit model (Voight et al. 2005; Gravel et al. 2011).

In this family, independent of the population analyzed, *TAS1R2* is the more diverse gene and *TAS1R3* (without the promoter, to be comparable to the other *TAS1R* genes) is less polymorphic (table 3). Considering that *TAS1R3* can dimerize with either *TAS1R1* or *TAS1R2*, the low variability observed in *TAS1R3* might indicate that this gene has experienced stronger evolutionary constraints due to its dual role as a functional unit of *umami* and of sweet receptors.

To search for potential signatures of selection among *TAS1R* genes, Tajima's *D* values were calculated (tables 1–3) and compared with those from null distributions assuming a neutral equilibrium model in a population of constant size. Since Tajima's *D* is highly influenced by population demography, its statistical significance was further evaluated by coalescent simulations with the various best-fit models available for 15 of the 20 populations being studied. In general, for *TAS1R1* and *TAS1R3* only negative Tajima's *D* values were obtained. Particularly for *TAS1R3* (promoter region not included), Tajima's *D* values were significantly lower: not only lower than the constant-size expectations in all samples (except the YRI) but even lower than the predictions based on best-fit demographic models in several European and African samples. Consistently, in the empirical comparison of Tajima's *D* values with those estimated for chromosome 1 in nonoverlapping sliding windows of 10 kb, *TAS1R3* also stood out, fitting in most populations the lower quartile of the distribution (supplementary fig. S1*A*, Supplementary Material online).

Conversely, Tajima's *D* values for *TAS1R2* tended to be slightly positive, although none was high enough to reach statistical significance under geographic specific models, including European samples in which the highest scores were obtained (GBR: 1.57 and IBS: 1.38).

Haplotype Structure and Gene Genealogies

The phylogenetic relationships between *TAS1R1* haplotypes revealed a double star-like network (fig. 1*A*), in which the central haplotypes, CA and CG, are defined by the two nonsynonymous variants (C329T and G1114A) in strong linkage disequilibrium (LD), known to influence glutamate sensitivity. These two haplotypes, both displaying a worldwide distribution, diverge by a single substitution at position G1114A that is assumed to shift *umami* perception from intermediate (CG haplotype) to high sensitivity (CA haplotype). The remaining *TAS1R1* haplotype identified exhibits the TG configuration, presumably leading to the lowest *umami* sensitivity

Table 3
Summary Statistics for *TAS1R3*

Geographic Region	Population	Genic Region	*N*	*S*	π	Θ_W	Tajima's *D*[a]
Africa							
Easthern	LWK	Promoter	198	27	16.000	4.605	−1.172
		Exon–intron		46	10.000	7.846	−1.523*
		Total		73	11.489	12.451	−1.464*
Western	YRI	Promoter	216	27	15.900	4.538	−1.144
		Exon–intron		42	10.000	7.059	−1.288
		Total		69	11.705	11.596	−1.301
	GWD	Promoter	226	27	14.300	4.503	−1.291
		Exon–intron		44	10.000	7.339	−1.384*,†
		Total		71	11.080	11.842	−1.424*,†
	MSL	Promoter	170	30	15.100	5.254	−1.494*,†
		Exon–intron		46	9.400	8.056	−1.626*,†
		Total		76	11.000	13.310	−1.653*,†
	ESN	Promoter	198	27	15.800	4.605	−1.1957
		Exon–intron		47	10.000	8.016	−1.467*,†
		Total		74	11.800	12.622	−1.437*,†
Europe							
North/Western	CEU	Promoter	198	13	5.000	2.2173	−1.537*
		Exon–intron		26	2.800	4.435	−2.101*,†
		Total		39	3.400	6.652	−2.069*
	GBR	Promoter	182	13	6.500	2.250	−1.290
		Exon–intron		25	3.300	4.326	−1.945*
		Total		38	4.260	6.577	−1.858*
	FIN	Promoter	198	10	5.900	1.706	−0.989
		Exon–intron		21	3.500	3.58	−1.664*
		Total		31	4.200	5.287	−1.575*
South	IBS	Promoter	214	15	5.600	2.525	−1.586*
		Exon–intron		30	2.800	5.050	−2.204*,†
		Total		45	3.603	7.575	−2.144*,†
	TSI	Promoter	214	18	4.800	3.030	−1.923*,†
		Exon–intron		42	2.800	7.070	−2.461*,†
		Total		60	3.360	10.100	−2.434*,†
Asia							
South	PJL	Promoter	192	19	8.600	3.258	−1.492*
		Exon–intron		30	4.000	5.144	−1.970*
		Total		49	5.341	8.402	−1.911*
	BEB	Promoter	172	16	9.800	2.796	−1.086
		Exon–intron		30	4.900	5.243	−1.819*
		Total		46	6.311	8.039	−1.668*
	STU	Promoter	204	14	9.200	2.376	−0.879
		Exon–intron		29	5.100	4.921	−1.668*
		Total		43	6.291	7.297	−1.506*
	ITU	Promoter	204	12	8.200	2.036	−0.793
		Exon–intron		31	4.300	5.261	−1.930*
		Total		43	5.410	7.297	−1.706*
	GIH	Promoter	206	14	8.600	2.372	−0.990
		Exon–intron		23	4.200	3.897	−1.577*
		Total		37	5.455	6.268	−1.470*
East	CHB	Promoter	206	19	6.300	3.219	−1.777*
		Exon–intron		25	3.500	4.235	−1.852*
		Total		44	4.343	7.454	−1.970*
	CHS	Promoter	210	11	6.500	1.857	−0.992
		Exon–intron		26	4.100	4.390	−1.762*

(continued)

Table 3 Continued

Geographic Region	Population	Genic Region	*N*	*S*	π	Θ_W	Tajima's *D*[a]
		Total		37	4.745	6.248	−1.648*
	JPT	Promoter	208	14	7.400	2.368	−1.200
		Exon–intron		23	4.600	3.890	−1.453*
		Total		37	5.402	6.258	−1.481*
	CDX	Promoter	186	8	5.100	1.3792	−0.857
		Exon–intron		21	3.700	3.621	−1.619*
		Total		29	4.099	4.999	−1.526*
	KHV	Promoter	198	18	6.300	3.070	−1.730*
		Exon–intron		27	3.600	4.605	−1.947*
		Total		45	4.349	7.675	−2.008*

NOTE.—The total length of *TAS1R3* analyzed was 5,637 bp, encompassing 1,692 bp from the promoter and 3,945 bp from the exon–intron region. *N*, number of chromosomes; *S*, number of segregating sites; π, nucleotide diversity per base pair ($\times 10^{-4}$); θ_W, population mutation rate parameter: Watterson's estimator of θ ($4Ne\mu$) (Watterson 1975) per base pair ($\times 10^{-4}$).

[a]Tajima's *D* statistic (Tajima 1989).

*$P < 0.05$ for constant model;

†$P < 0.05$ for best-fit model (Voight et al. 2005; Gravel et al. 2011).

(Raliou et al. 2009; Shigemura et al. 2009), and was detected only in non-African populations at very low frequencies.

The *TAS1R1* tree (fig. 2*A*) uncovered a recent TMRCA of 129 ± 11 ka, considerably less than the average estimates for other autosomal loci that, according to Blum and Jakobsson (2011), date to ~1.5 Ma (first quartile = 950,000 years; third quartile = 1,700,000 years). The CG haplotype, placed as the ancestral haplotype, was clearly associated with an excess of rare variants that was also observed to a lesser extent in the CA haplotype, which originated ~82 ± 10 ka through the G1114A mutation.

For *TAS1R3*, given that all populations had significant levels of LD across the promoter and the exon–intron region (supplementary fig. S2, Supplementary Material online), a single network was reconstructed for the whole gene sequence (fig. 1*C*, but see supplementary fig. S3, Supplementary Material online, for separated networks of promoter and exon–intron regions). Five haplotype classes were defined based on combinations of two promoter variants correlated with human sweet perception, -T1572C and -T1266C, and two exon–intron variants reported to influence *umami* perception, G13A and C2269T. As figure 1*C* illustrates, the CCGC haplotype conferring high sweet and *umami* sensitivity also displays a double star-like structure associated with -T1266C and with another promoter variant rs35946613 (-G1221A). Two less-common haplotypes (<0.022 frequency in the full panel) radiate from the CCGC network core: CCAC (high sweet and intermediate *umami* sensitivity) and CTGC (intermediate sweet and high *umami*). The other two detected haplotypes were both associated with low sweet sensitivity and bear either the TTGC configuration (high *umami*, global frequency 0.06), which is only a few mutations away from the major CCGC haplotype, or the TTGT configuration (intermediate *umami*, global frequency 0.16), which is separated by multiple positions from the central haplotype.

The gene genealogy for *TAS1R3* (fig. 2*B* and supplementary fig. S4, Supplementary Material online, for independent trees of the promoter and the exon–intron region) yielded a TMRCA age of ~592 ± 8.75 ka, also quite recent when compared with other autosomal loci estimates (Blum and Jakobsson 2011). In *TAS1R3*, a strong signal of rapid lineage diversification was detected in the CCGC haplotype; this harbors three derived alleles at the promoter segment (-T1572C, -T1266C, and -G1221A - rs35946613), while retaining the ancestral configuration at the exon–intron region. Consistent with the overall TMRCA estimated for *TAS1R3*, the three promotor mutations were also young (~589 ± 4.78 ka for -T1266C; ~434 ± 2.17 ka for -T1572C; and ~348 ± 1.3 ka for G1221A). All three predate the G13A (~80 ± 0.72 ka) and C2269T (~255 ± 2.62 ka) substitutions located in the exon–intron region.

For *TAS1R2*, the haplotype network (fig. 1*B*) revealed to be extremely reticulated, indicating the strong influence of recombination in the region and highlighting the genomic mechanism driving *TAS1R2* to high levels of nucleotide diversity. Analysis of LD patterns within *TAS1R2* confirmed the occurrence of a recombination hotspot between exons 3 and 4 (supplementary fig. S5, Supplementary Material online), which was further supported by deCODE recombination maps (Kong et al. 2002). Since the method used to build gene genealogies assumes an infinite-site model without recombination, it was impossible to produce a tree for *TAS1R2* due to the excessive number of haplotypes violating that assumption.

The *TAS2R* Family and the Bitter Receptors

Polymorphism Levels and Neutrality Tests

In the *TAS2R* family, we investigated *TAS2R16* and *TAS2R38*, two genes on chromosome 7 separated by ~19 Mb that, importantly, are intronless genes. According to previous

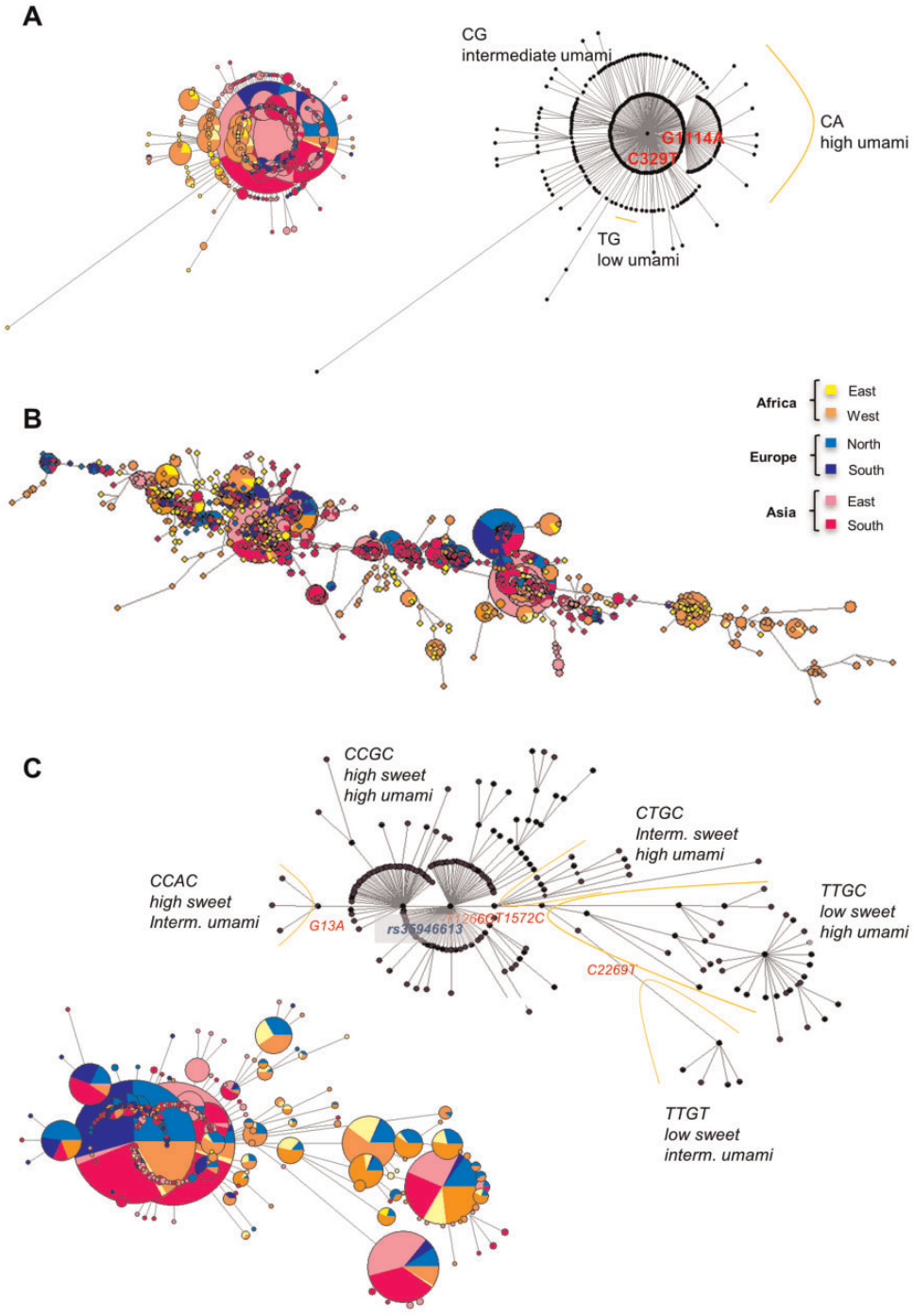

Fig. 1.—Networks from *TAS1R* genes. Each node represents a different haplotype, and the line length is proportional to the number of mutations along the branch. In the colored networks, circle sizes are proportional to frequencies, which were not considered in the black-and-white network versions. (*A*) *TAS1R1*. The haplotypes discriminated are defined by the nonsynonymous variations shown in red, C329T and G1114A. (*B*) *TAS1R2*. (*C*) *TAS1R3* encompassing the promoter and the exon–intron region. The haplotypes discriminated are defined by the variations shown in red -T1572C and -T1266C from the promoter, plus G13A and C2269T from the exon–intron region. Variation rs35946613 shown in blue is located in the promoter.

studies in different mammalian species, these types of genes may experience quite different evolutionary rates than genes with an exon–intron organization (Shabalina et al. 2010; Yu et al. 2014). Accordingly, a direct comparison of the summary statistics of nucleotide variation between genes from *TAS1R* and *TAS2R* families was hampered.

Centering, then, on the *TAS2R* family, *TAS2R16* (table 4) was nearly four times more diverse in Africans than in

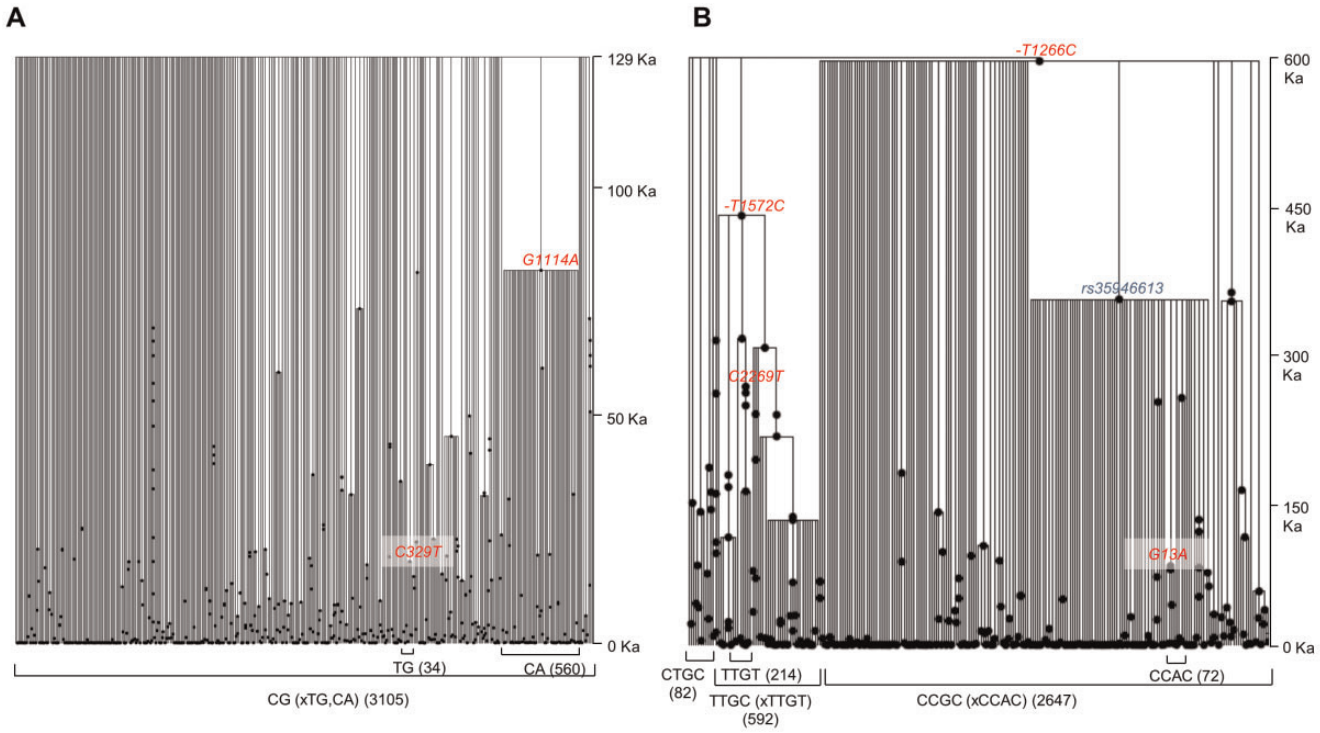

Fig. 2.—Coalescent-based genealogy, TMRCA of global variation, and ages of individual mutations at *TAS1R* genes. Mutations are represented by black dots, and figures on the bottom of branches correspond to the number of individuals with that haplotype (x means except). (*A*) *TAS1R1* and (*B*) *TAS1R3*. Variations in red and blue as in legend for figure 1.

Europeans or Asians, whereas *TAS2R38* (table 5) displayed similar levels of diversity across all studied populations. Regarding Tajima's *D*, for *TAS2R16* (table 4) almost every population exhibited positive values, but only in ESN did the estimated Tajima's *D* surpass the expectation under the best-fit demographic model. In the empirical comparison with Tajima's *D* values estimated for chromosome 7, *TAS2R16* in ESN but also YRI and GWD lied above the upper quartile, whereas in non-African populations in most cases it fell closer to the distribution mean (supplementary fig. S1*B*, Supplementary Material online). For *TAS2R38*, significant positive Tajima's *D* were found in Europe, and values in three of the five analyzed populations remained statistically significant after assuming the best-fit model (table 5). Consistently, in the comparison with chromosome 7 genome wide distribution of Tajima's *D* those three values were near or above the upper quartile of values (supplementary fig. S1*B*, Supplementary Material online).

Haplotype Structure and Gene Genealogies

The phylogenetic relationships of *TAS2R16* haplotypes depicted in figure 3*A* show that the G516T substitution influencing salicin recognition splits the network into two major haplotype groups: one found only in Africa, characterized by the "low-sensitivity" G allele, and the other, apparently fixed in Eurasian populations, defined by the "high-sensitivity" T allele. In the latter haplotype group, there is a cluster of lineages defined by a nonsynonymous substitution A665G (p.Arg222His, rs860170) that, according to Campbell et al. (2014), does not contribute to salicin sensitivity or to cell surface expression of receptors.

In contrast to the shorter TMRCA estimates of the *TAS1R* family, the genealogy obtained for *TAS2R16* (fig. 4*A*) coalesced into a much deeper TMRCA of ~7.02 ± 1.99 Ma (using human–chimpanzee divergence) or ~4.03 ± 1.99 Ma (using human–orangutan divergence, instead). These age estimates are several times older than the value of ~1.75 ± 0.75 Ma reported by Campbell et al. (2014), and much greater than the 3.0 Ma upper limit of TMRCA estimates obtained for other autosomal loci (Blum and Jakobsson 2011). The presence of two highly divergent branches corresponding to "low-sensitivity" and "high-sensitivity" haplotypes seems to explain the unusual TMRCA obtained for *TAS2R16*, as well as the ancient origin estimated for the G516T mutation (~6.54 ± 1.70 Ma or ~3.59 ± 1.70 Ma). Together, these findings place the shift toward an enhanced bitter perception very early in *Homo* evolution, and long before the emergence of anatomically modern humans ~200 ka.

The *TAS2R38* network (fig. 3*B*) uncovers four major haplotypes linked through a chain of three amino acid substitutions that explain most of the interindividual differences in

Table 4
Summary Statistics for *TAS2R16*

Geographic Region	Population	*N*	*S*	π	Θ_W	Tajima's *D*[a]
Africa						
Eastern	LWK	198	11	19.401	1.876	0.087
Western	YRI	216	8	23.843	1.344	1.425
	GWD	226	7	21.689	1.168	1.530
	MSL	170	9	22.407	1.576	0.754
	ESN	198	6	22.830	1.023	2.194*,†
Europe						
North/Western	CEU	198	3	4.919	0.512	−0.167
	GBR	182	2	5.260	0.346	0.577
	FIN	198	2	6.398	0.341	1.023
South	IBS	214	2	5.369	0.337	0.657
	TSI	214	3	5.191	0.505	−0.067
Asia						
South	PJL	192	10	6.082	1.71465	−1.570*
	BEB	172	4	4.976	0.69908	−0.612
	STU	204	3	5.861	0.50909	0.117
	ITU	204	5	5.314	0.84848	−0.796
	GIH	206	6	6.004	1.01649	−0.909
East	CHB	206	2	5.482	0.339	0.690
	CHS	210	3	5.577	0.50659	0.041
	JPT	208	2	5.410	0.33828	0.666
	CDX	186	2	5.623	0.34481	0.718
	KHV	198	8	6.148	1.36450	−1.280

Note.—The total length of the *TAS2R16* region analyzed was 995 bp. *N*, number of chromosomes; *S*, number of segregating sites; π, nucleotide diversity per base pair ($\times 10^{-4}$); θ_W, population mutation rate parameter: Watterson's estimator of θ ($4Ne\mu$) (Watterson 1975) per base pair ($\times 10^{-4}$).
[a]Tajima's *D* statistic (Tajima 1989).
*$P < 0.05$ for constant model;
†$P < 0.05$ for best-fit model (Voight et al. 2005; Gravel et al. 2011).

Table 5
Summary Statistics for *TAS2R38*

Geographic Region	Population	*N*	*S*	π	Θ_W	Tajima's *D*[a]
Africa						
Eastern	LWK	198	8	17.662	1.364	0.356
Western	YRI	216	9	16.959	1.513	0.013
	GWD	226	8	17.531	1.334	0.388
	MSL	170	6	17.537	1.051	1.032
	ESN	198	7	16.998	1.194	0.590
Europe						
North/Western	CEU	198	3	16.296	0.512	2.980*,†
	GBR	182	3	16.279	0.519	2.946*,†
	FIN	198	4	16.450	0.682	2.072*
South	IBS	214	5	16.950	0.842	1.533
	TSI	214	3	16.660	0.505	3.109*,†
Asia						
South	PJL	192	6	16.829	1.029	0.954
	BEB	172	5	15.076	0.874	1.072
	STU	204	6	14.526	1.018	0.566
	ITU	204	7	15.618	1.188	0.380
	GIH	206	5	16.539	0.847	1.435
East	CHB	206	5	15.351	0.847	1.195
	CHS	210	4	15.020	0.675	1.757*
	JPT	208	5	17.728	0.846	1.679*
	CDX	186	5	14.294	0.862	0.944
	KHV	198	5	13.137	0.853	0.733

Note.—The total length of the *TAS2R38* region analyzed was 896 bp. *N*, number of chromosomes; *S*, number of segregating sites; π, nucleotide diversity per base pair ($\times 10^{-4}$); θ_W, population mutation rate parameter: Watterson's estimator of θ ($4Ne\mu$) (Watterson 1975) per base pair ($\times 10^{-4}$).
[a]Tajima's *D* statistic (Tajima 1989).
*$P < 0.05$ for constant model;
†$P < 0.05$ for best-fit model (Voight et al. 2005; Gravel et al. 2011).

PTC/PROP taste perception. These haplotypes have previously been called as PAV, for the ancestral "taster" haplotype; AVI, for most derived and "nontaster" configuration; and AAI or AAV for intermediate haplotypes with midsensitivity to PTC/PROP (Wooding et al. 2004). The two opposite haplotypes, PAV and AVI, are present at high frequencies worldwide, whereas, among the intermediate configurations, AAI appears to be most prevalent in Africa.

The gene genealogy for *TAS2R38* (fig. 4*B*) yielded a coalescence time of $\sim 2.49 \pm 0.77$ Ma, which is slightly longer than the TMRCA of $\sim 2.1 \pm 0.45$ Ma calculated by Campbell et al. 2012. Nevertheless, both times exceed the 1.7 Ma third quartile of estimates for other autosomal loci (Blum and Jakobsson 2011). Likewise, the mean estimates for the three amino acid substitution variants associated with PTC/PROP lower-sensitivity phenotypes are ancient and range from $\sim 2.08 \pm 0.71$ Ma for P49A, and $\sim 1.06 \pm 0.41$ Ma for V266I, to $\sim 710 \pm 306$ ka for A262V. Therefore, the chronological order of appearance inferred for *TAS2R38* haplotypes is: PAV→AAV→AAI→AVI, all of them having originated by single mutational steps preceding the emergence of anatomically modern humans.

Population Differentiation of *TAS1R* and *TAS2R* Genes

Levels of genetic differentiation were assessed across worldwide populations by computing intercontinental F_{ST} values for *TAS1R* and *TAS2R* sequences. To view the results in a genome-wide context, we referred to the estimates reported by the 1000 Genomes Project (Genomes Project et al. 2010): the mean value of F_{ST} was 0.071 between CEU and YRI, 0.083 between YRI and CHB + JPT, and 0.052 between CEU and CHB + JPT. Akey et al. (2002) had previously calculated the mean as F_{ST}=0.12, using 25,549 autosomal SNPs genotyped in African–American, East Asian, and European–American populations, while Shriver et al. (2004) analyzed 8,525 autosomal SNPs in African–American, European–American, Chinese, and Japanese individuals, and pinpointed a mean of F_{ST}=0.13.

The estimates obtained here, grouping populations into African, Asian, and European, were F_{ST}=0.135 for *TAS1R1*, F_{ST}=0.075 for *TAS1R2*, F_{ST}=0.200 for *TAS1R3* total, F_{ST}=0.305 for *TAS1R3* promoter, F_{ST}=0.102 for *TAS1R3* exon–intron, F_{ST}=0.278 for *TAS2R16*, and F_{ST}=0.070 for *TAS2R38*.

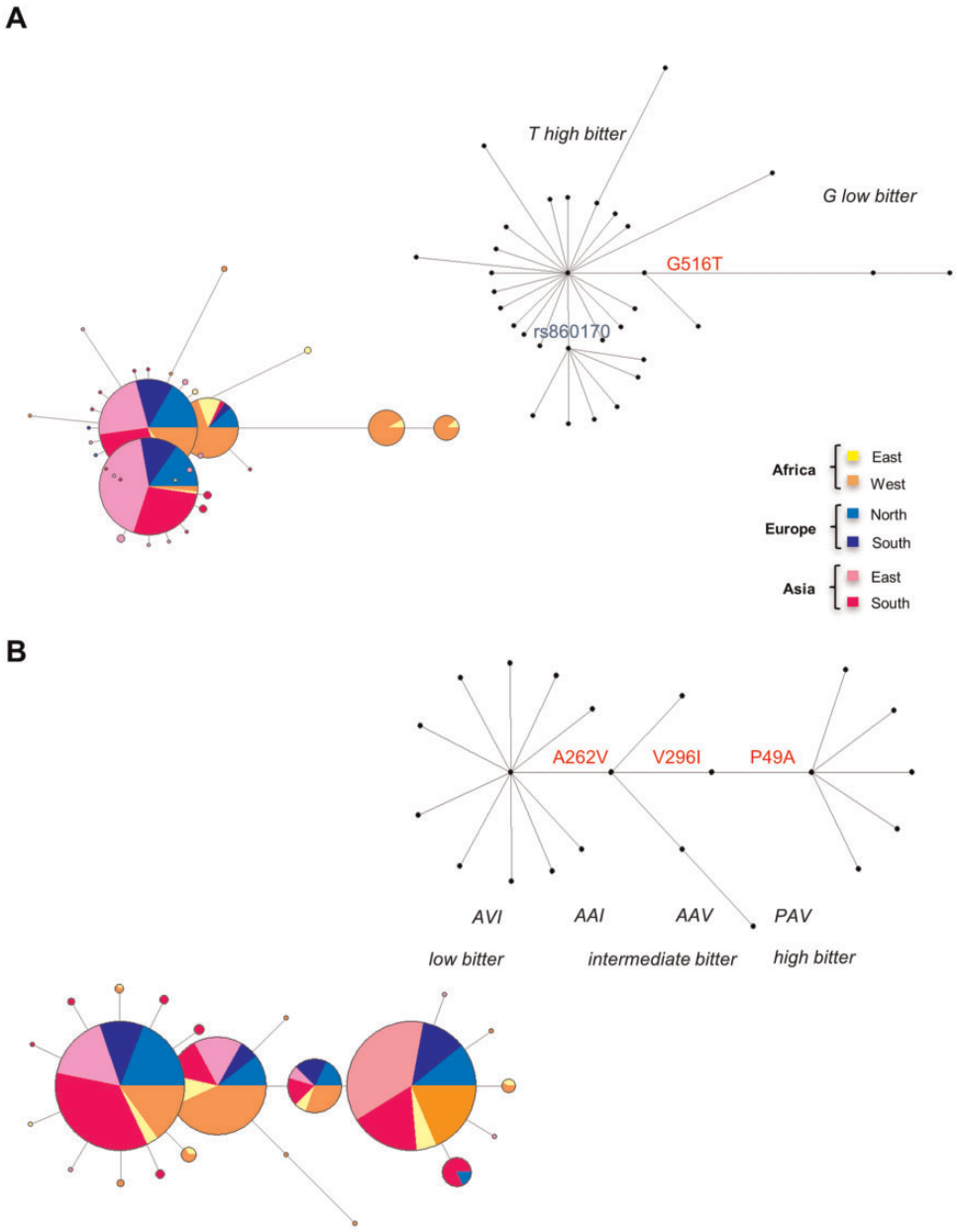

Fig. 3.—Networks from *TAS2R* genes. Each node represents a different haplotype and line length is proportional to the number of mutations along the branch. In the colored networks, circle sizes are proportional to frequencies, which were not considered in the black-and-white network versions. (*A*) *TAS2R16*. The alleles discriminated are from the variation shown in red G516T. rs860170, shown in blue, is a non-synonymous substitution. (*B*) *TAS2R38*. The haplotypes discriminated are defined by the variations shown in red P49A, A262V, and V296I (according to residue substitution).

In light of the above-mentioned reference F_{ST} values, our estimates support an atypically high level of population differentiation on a continental scale in two instances: One was observed in the promoter of *TAS1R3* (0.305) and the other in *TAS2R16* (0.278). Conversely, *TAS2R38* yielded the lowest F_{ST} value (0.070) of the genomic sequences investigated.

Discussion

The sense of taste is crucial for evaluating foods' toxicity and nutrient content, with consequential impact on nutritional status. Therefore, it is highly likely that taste perception has played a significant role in human evolutionary history, during which food preferences and aversions have changed

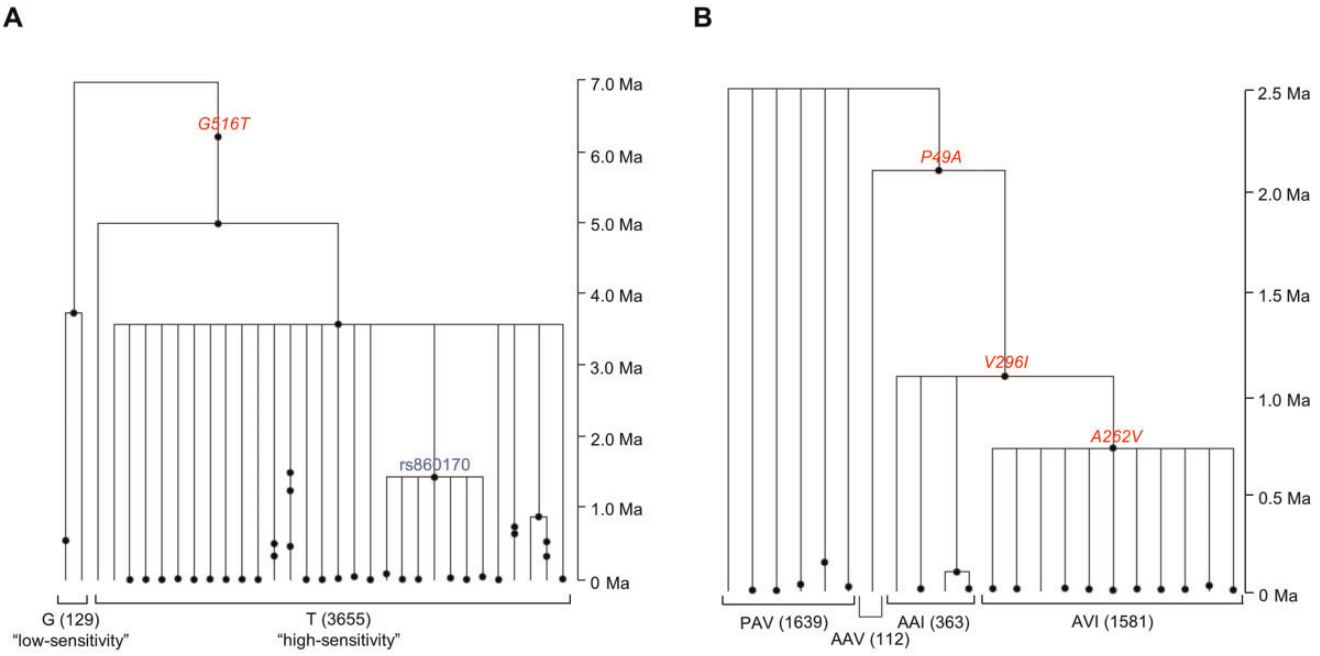

Fig. 4.—Coalescent-based genealogy, TMRCA of global variation, and ages of individual mutations at *TAS2R* genes. Mutations are represented by black dots, and figures on the bottom of branches correspond to the number of individuals with that haplotype. (*A*) *TAS2R16*. (*B*) *TAS2R38*. Variations in red and blue as in legend for figure 3.

(Breslin 2013; Doty 2015). Our in-depth study of the major forces driving the evolution of bitter, sweet, and *umami* receptors scrutinized *TAS1R1, TAS1R2, TAS1R3, TAS2R16*, and *TAS2R38* sequence variations across major human populations: Africans (Western and Eastern), Europeans (Northern/Western and Southern), and Asians (Southern and Eastern).

From a global perspective, we demonstrated that two genes from the *TAS1R* family, *TAS1R1* and *TAS1R3*, experienced an evolution unlike that of the two genes of the *TAS2R* family, *TAS2R16* and *TAS2R38*. Although some differences would be expected, due to their gene organization in exon–intron or intronless structures, respectively, the departures from neutrality that were detected are more likely attributable to distinct selective pressures that prompted independent adaptive events correlated with *umami* and sweet perception in *TAS1R* and bitter taste in *TAS2R* genes.

In the *TAS1R* family, the genes encoding TAS1R1 and TAS1R3, which directly interact as heterodimers of *umami* taste receptors, were found to display patterns of diversity compatible with signatures of positive selection. Our findings agree with those previously reported by Kim et al. (2006), although their conclusions were based on small population sample sizes, a shortcoming seemingly overcome in the current work. First, *TAS1R1* and *TAS1R3* networks were both associated with star-like structures in which common haplotypes were connected to multiple rare haplotypes differing by only one or two mutations. Second, the genomic sequences coalesced at recent times (< 600,000 years ago), about half as long ago as would be expected for autosomal genes. Third, Tajima's *D* neutrality tests were systematically associated with negative values in African, European, and Asian populations, though being considerably lower in *TAS1R3* than in *TAS1R1*.

For *TAS1R1*, the lack of strong negative Tajima's *D* could still be compatible with a model of selection based on standing variation, whereby the target variant is already segregating in a population at the time of the selective event and so can be swept to higher frequencies along with the pre-existing linked neutral variation (Przeworski et al. 2005; Fu and Akey 2013). Once selection on standing variation predictably leaves less striking footprints than a standard selective sweep, the possibility could explain not only our *TAS1R1* findings, where two major haplotypes (corresponding to the ancestral CG configuration and its one-step derived CA haplotype) display evidence for a rapid increase in frequency but also the *TAS1R3* results that provide similar signs at two CCGC subhaplotypes.

Interestingly, in *TAS1R1* and *TAS1R3*, the haplotypes associated with "star-like" structures, CG/CA and CCGC (with and without rs35946613), respectively, coalesce within very short timescales (130–600 ka), falling in the Middle Pleistocene (126–781 ka; Hedges and Kumar 2009), a geological period during which evolution from archaic hominins to anatomically modern humans took place (reviewed in Groucutt et al. 2015). This finding together with the worldwide distribution of negative Tajima's *D* values disclosed by both genes, suggest that a common selective pressure might have shaped *TAS1R1* and *TAS1R3* diversities, guiding a coordinated evolutionary process at the dawn of modern humans.

The implied phenotypic changes were probably related to the shared role of TAS1R1 plus TAS1R3 in the dimer functioning as an *umami* receptor, given the possibility that enhanced *umami* taste perception was instigated by the acquisition of new dietary regimes. Notably, it is currently being claimed that a major dietary shift associated with the introduction of cooked food by *Homo* occurred precisely in the Middle Pleistocene. According to the cooking hypothesis, humans have adapted to the obligatory usage of cooked food in their diets, without which we cannot survive. When this obligation developed has been the subject of controversy (reviewed in Wrangham 2017). However, the scarce archeological evidence for a controlled domain of fire before 400 ka, together with data supporting that the best-known human adaptations to cooked food date ~550–750 ka, are now relegating such dependence on cooking, with consequent effects on human biology and behavior, to the Middle Pleistocene period (Wrangham 2017).

Although we have argued for a selective advantage associated with an improved perception of *umami* as the main phenotypic effect, consequences linked to other tastes should not be dismissed. As a matter of fact, the strongest candidate variant of *TAS1R3* to have been under selection was -T1266C, located in the promoter sequence, a region that until now was only connected with sugar taste sensitivity, recalling that TAS1R3 + TAS1R2 forms the unique heterodimer known to work as a sugar receptor. To our knowledge, the influence of *TAS1R3* promoter variants in *umami* sensitivity has never been investigated, nor has the contribution of variability at the *TAS1R3* exon–intron region to sugar sensitivity, meaning that functional studies are yet needed to clarify this issue.

In the *TAS2R* family, due to the intronless nature of *TAS2R16* and *TAS2R38*, we anticipated a stronger effect of purifying selection leading to depletion of deleterious mutations along with some linked neutral variation. However, if purifying selection indeed exerted a significant impact on these genes, it was not sufficient to mask other stronger adaptive signatures, but perhaps might have contributed to further distort the shape of *TAS2R* genealogies. In this respect, the few low frequency variants associated with *TAS2R16* low-bitter haplotypes are impressive for their ancient origins, as is the reduced number of mutations associated with *TAS2R38* intermediate-bitter haplotypes (AAI and AAV). Although both genes have old TMRCAs, they also retain other signatures not reconciled with the expectations for sequences that evolve uniquely under purifying selection.

TAS2R16 revealed a particularly old TMRCA (~7.02–4.03 Ma), dating backward to a time that precedes the emergence of the earliest *Homo* specimens. The presence of two highly divergent branches corresponding to haplotypes conferring "low sensitivity" and "high sensitivity" to salicin certainly plays a part in the deep TMRCA associated with *TAS2R16*. Remarkably, the "low-sensitivity" haplotypes—those retaining the ancestral allele (G) at position G516T—are uniquely found in Africa, explaining the atypical high F_{ST} value (0.278) observed across continental populations. In Africa, the "low-sensitivity" and "high-sensitivity" haplotypes occur at global frequencies of ~33% and ~67%, respectively, and all populations generated positive Tajima's *D*s, suggesting that in this geographic region balancing selection has acted to maintain high phenotypic variability in perception of bitterness.

Before this study, a selective hypothesis for the evolution of *TAS2R16* had already been raised by at least three independent works, all agreeing on proposing the "high-sensitivity" haplotypes as targets of positive selection in Eurasia and Africa (Soranzo et al. 2005; Li et al. 2011; Campbell et al. 2014). Here, after inspecting the 1000 Genomes panel, the single population that could potentially support such hypothesis is PJL, which is associated with a strongly negative Tajima's *D* value. However, the "high-sensitivity" lineage is old enough (~6.25–3.59 Ma) to have drifted to higher frequencies in the absence of positive selection; and the Neanderthal and Denisovan genomes were consistently found to carry "high-sensitivity" alleles. Thus, the possible excess of low frequency variants is more likely to be associated with some specific feature of the PJL sample. Conversely, the occurrence of "low-sensitivity" and "high-sensitivity" haplotypes at intermediate frequencies in Africa, and the trend toward overtly positive Tajima's *D* values in Western Africa, suggest a scenario of long-standing balancing selection. We recognize that ancestral population substructure can reproduce patterns of genetic variation expected under balancing selection, and that could represent a possible cause for the architecture of *TAS2R16* diversity observed in Africa. But it seems a little odd that ancestral substructure could evenly affect all five African groups here examined, which despite representing only Bantu speaking populations, are dispersed across Africa, including the Western and Eastern regions. Taking into account the expectation that long-standing balancing selection will maintain beyond regular coalescent times (1–1.5 Ma) with at least two divergent lineages that are generally correlated with a heterozygous advantage, we hypothesize that balancing selection persisted until the present as a major pressure in modeling *TAS2R16* variation in Africa. Plausibly, improved ability to recognize a wide spectrum of bitter tastes provided better discrimination between nutritive raw foods and others containing toxic elements.

Concerning *TAS2R38*, our findings agree with previous studies that point to balancing selection as the best model to explain the worldwide distribution of "taster" and "nontaster" haplotype classes (Wooding et al. 2004; Campbell et al. 2012). Nevertheless, a recent study revisiting the current patterns of *TAS2R38* variation reported signs of ancient balancing selection followed more recently by a relaxation of selective forces, given that no significant departure from neutrality was detected (Risso et al. 2016). To the contrary, our results provide evidence for such a departure, at

least in current-day Europeans, in which Tajima's *D* values were often highly significant under best-fit demographic models, residing concordantly above the average Tajima's *D* across entire chromosome 7. Notably, the study of Risso et al. differs from the present work by the extension of the sequence analyzed: limited in our study to the coding region (896 bp), while Risso et al. also included additional flanking 5′ and 3′ sequences (1,143 bp). That difference might account for loss of the signal of balancing selection in the latter work, given that signatures of balancing selection are mainly expected to affect short genomic segments (Andrés et al. 2009). On the other hand, Risso et al. only provided averaged Tajima's *D* values for African and non-African populations, not showing individual values for each population. Thus, the amalgamation of populations might also have contributed to the loss of significant signals of selection, which we have detected here mainly in European populations.

Collectively, the data gathered in this study, and in Wooding et al. 2004 or in Campbell et al. 2012, are consistent with a scenario of long-standing balancing selection that, for an extended time frame, has maintained two divergent haplotypes of *TAS2R38*. Furthermore, the discovery of an overlap in the selection mode operating at *TAS2R38* and *TAS2R16* reinforces the assumption that the capacity of humans to perceive a multiplicity of bitter tastes is, or was until recently, essential to survival worldwide.

Although this study has relied on taste perception to ground the evolutionary hypotheses, we note that taste perception does not exhaust the functions of TAS proteins. Current research is showing that *TAS* genes are expressed in a plethora of nonoral tissues, and taste receptors are involved in many other biological processes—namely, being integrated into respiratory and gastrointestinal pathways (Behrens and Meyerhof 2011)—and that, for instance, bitter and sweet receptors are implicated in the regulation of human upper respiratory innate immunity (Lee et al. 2014; Gil et al. 2015). Taste receptors' emergence as the products of pleiotropic genes, which participate in a wide range of signaling pathways outside of taste perception, allows us to anticipate a broader role of *TAS* genes in human health and evolution (Campa et al. 2012).

In conclusion, our results indicate *TAS1R* (*TAS1R1* and *TAS1R3*) genes as being recently driven by positive selection in a coevolution process possibly correlated with a better capacity to perceive *umami* or sweet tastes, which probably represented an adaptation to cooked food. On the other hand, *TAS2R* (*TAS2R16* and *TAS2R38*) genes appear to have been under balancing selection for a long time before the emergence of modern humans in Africa. Their role was most likely to prevent consumption of dangerous raw foods, taking advantage of a wide spectrum of sensations of bitterness. In the near future, elucidation of the pleiotropic relationships in genes that code for taste receptors and their involvement in diverse biological functions promises to expand our understanding of *TAS1R* and *TAS2R* molecular evolution.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

IPATIMUP is included in the i3S Consortium, which is partially supported by the Portuguese Foundation for Science and Technology (FCT). This work is also funded by FEDER funds through the Operational Programme Competitiveness Factors (COMPETE) and National Funds through the FCT (projects PEst-C/SAU/LA0003/2013 and POCI-01-0145-FEDER-007274, fellowships SFRH/BD/63343/2009 to C.V., SFRH/BPD/65000/2009 to L.A., and SFRH/BPD/120777/2016 to P.I.M.), and by Programa Operacional Regional do Norte (Norte 2020), through FEDER funds under the Quadro de Referência Estratégico Nacional (QREN; NORTE-01-0145-FEDER-000029). We thank Jacquelyn Beals for the careful editing of the work. We also thank Sònia Casillas for the helpful instructions about the tracks from the PopHuman browser, and Miguel Arenas and Eduardo Conde Sousa for the aid in extracting data from those tracks.

Literature Cited

Akey JM, Zhang G, Zhang K, Jin L, Shriver MD. 2002. Interrogating a high-density SNP map for signatures of natural selection. Genome Res. 12(12):1805–1814.

Andrés AM, et al. 2009. Targets of balancing selection in the human genome. Mol Biol Evol. 26(12):2755–2764.

Bachmanov AA, Beauchamp GK. 2007. Taste receptor genes. Annu Rev Nutr. 27:389–414.

Bandelt HJ, Forster P, Rohl A. 1999. Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol. 16(1):37–48.

Behrens M, Meyerhof W. 2011. Gustatory and extragustatory functions of mammalian taste receptors. Physiol Behav. 105(1):4–13.

Blum MG, Jakobsson M. 2011. Deep divergences of human gene trees and models of human origins. Mol Biol Evol. 28(2):889–898.

Breslin PA. 2013. An evolutionary perspective on food and human taste. Curr Biol. 23(9):R409–R418.

Bufe B, Hofmann T, Krautwurst D, Raguse JD, Meyerhof W. 2002. The human TAS2R16 receptor mediates bitter taste in response to beta-glucopyranosides. Nat Genet. 32(3):397–401.

Campa D, et al. 2012. Bitter taste receptor polymorphisms and human aging. PLoS One 7(11):e45232.

Campbell MC, et al. 2012. Evolution of functionally diverse alleles associated with PTC bitter taste sensitivity in Africa. Mol Biol Evol. 29(4):1141–1153.

Campbell MC, et al. 2014. Origin and differential selection of allelic variation at TAS2R16 associated with salicin bitter taste sensitivity in Africa. Mol Biol Evol. 31(2):288–302.

Casillas S, et al. 2018. PopHuman: the human population genomics browser. Nucleic Acids Res. 46(D1):D1003–D1010.

Chandrashekar J, Hoon MA, Ryba NJ, Zuker CS. 2006. The receptors and cells for mammalian taste. Nature 444(7117):288–294.

Coop G, Griffiths RC. 2004. Ancestral inference on gene trees under selection. Theor Popul Biol. 66(3):219–232.
Doty RL. 2015. Handbook of olfaction and gustation. Hoboken, New Jersey: Wiley-Blackwell.
Eny KM, Wolever TM, Corey PN, El-Sohemy A. 2010. Genetic variation in TAS1R2 (Ile191Val) is associated with consumption of sugars in overweight and obese individuals in 2 distinct populations. Am J Clin Nutr. 92(6):1501–1510.
Excoffier L, Lischer HE. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol Ecol Resour. 10(3):564–567.
Fu W, Akey JM. 2013. Selection and adaptation in the human genome. Annu Rev Genomics Hum Genet. 14:467–489.
Fushan AA, Simons CT, Slack JP, Manichaikul A, Drayna D. 2009. Allelic polymorphism within the TAS1R3 promoter is associated with human taste sensitivity to sucrose. Curr Biol. 19(15):1288–1293.
Gil S, et al. 2015. Genotype-specific regulation of oral innate immunity by T2R38 taste receptor. Mol Immunol. 68(2 Pt C):663–670.
Gomez F, Hirbo J, Tishkoff SA. 2014. Genetic variation and adaptation in Africa: implications for human evolution and disease. Cold Spring Harb Perspect Biol. 6(7):a008524.
Gravel S, et al. 2011. Demographic history and rare allele sharing among human populations. Proc Natl Acad Sci U S A. 108:11983–11988.
Griffiths RC, Tavare S. 1994. Sampling theory for neutral alleles in a varying environment. Philos Trans R Soc Lond B Biol Sci. 344(1310):403–410.
Groucutt GHS, et al. 2015. Rethinking the dispersal of *Homo sapiens* out of Africa. Evol Anthropol. 24(4):149–164.
Hedges B and Kumar S, editors with foreword by Watson JD. 2009. The timetree of life. New York: Oxford University Press.
Hudson RR. 2002. Generating samples under a Wright-Fisher neutral model of genetic variation. Bioinformatics 18(2):337–338.
Kim UK, et al. 2003. Positional cloning of the human quantitative trait locus underlying taste sensitivity to phenylthiocarbamide. Science 299(5610):1221–1225.
Kim UK, Wooding S, Riaz N, Jorde LB, Drayna D. 2006. Variation in the human TAS1R taste receptor genes. Chem Senses 31(7):599–611.
Kong A, et al. 2002. A high-resolution recombination map of the human genome. Nat Genet. 31(3):241–247.
Lee RJ, et al. 2014. Bitter and sweet taste receptors regulate human upper respiratory innate immunity. J Clin Invest. 124(3):1393–1405.
Li H, Pakstis AJ, Kidd JR, Kidd KK. 2011. Selection on the human bitter taste gene, TAS2R16, in Eurasian populations. Hum Biol. 83(3):363–377.
Li X, et al. 2002. Human receptors for sweet and umami taste. Proc Natl Acad Sci U S A. 99(7):4692–4696.
Lischer HE, Excoffier L. 2012. PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. Bioinformatics 28(2):298–299.
McVean GA, et al. 2004. The fine-scale structure of recombination rate variation in the human genome. Science 304(5670):581–584.
Nei M, Li WH. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci U S A. 76(10):5269–5273.
Nelson G, et al. 2002. An amino-acid taste receptor. Nature 416(6877):199–202.
Patterson N, Richter DJ, Gnerre S, Lander ES, Reich D. 2006. Genetic evidence for complex speciation of humans and chimpanzees. Nature 441(7097):1103–1108.
Przeworski M, Coop G, Wall JD. 2005. The signature of positive selection on standing genetic variation. Evolution 59(11):2312–2323.
Raliou M, et al. 2009. Nonsynonymous single nucleotide polymorphisms in human tas1r1, tas1r3, and mGluR1 and individual taste sensitivity to glutamate. Am J Clin Nutr. 90(3):789S–799S.
Risso DS, et al. 2016. Global diversity in the TAS2R38 bitter taste receptor: revisiting a classic evolutionary PROPosal. Sci Rep. 6:25506.
Rozas J. 2009. DNA sequence polymorphism analysis using DnaSP. Methods Mol Biol. 537:337–350.
Shabalina SA, et al. 2010. Distinct patterns of expression and evolution of intronless and intron-containing mammalian genes. Mol Biol Evol. 27(8):1745–1749.
Shigemura N, Shirosaki S, Sanematsu K, Yoshida R, Ninomiya Y. 2009. Genetic and molecular basis of individual differences in human umami taste perception. PLoS One 4(8):e6717.
Shriver MD, et al. 2004. The genomic distribution of population substructure in four populations using 8,525 autosomal SNPs. Hum Genomics 1(4):274–286.
Soares I, Moleirinho A, Oliveira GN, Amorim A. 2015. DivStat: a user-friendly tool for single nucleotide polymorphism analysis of genomic diversity. PLoS One 10(3):e0119851.
Soranzo N, et al. 2005. Positive selection on a high-sensitivity allele of the human bitter-taste receptor TAS2R16. Curr Biol. 15(14):1257–1265.
Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123(3):585–595.
Temussi PA. 2009. Sweet, bitter and umami receptors: a complex relationship. Trends Biochem Sci. 34(6):296–302.
The 1000 Genomes Project Consortium. 2010. A map of human genome variation from population-scale sequencing. Nature 467(7319):1061–1073.
Voight BF, et al. 2005. Interrogating multiple aspects of variation in a full resequencing data set to infer human population size changes. Proc Natl Acad Sci U S A. 102(51):18508–18513.
Watterson GA. 1975. On the number of segregating sites in genetical models without recombination. Theor Popul Biol. 7(2):256–276.
Wooding S, et al. 2004. Natural selection and molecular evolution in PTC, a bitter-taste receptor gene. Am J Hum Genet. 74(4):637–646.
Wrangham R. 2017. Control of fire in the Paleolithic: evaluating the cooking hypothesis. Curr Anthropol. 58 (S16):S303–S313.
Yu C, et al. 2014. DFA7, a new method to distinguish between intron-containing and intronless genes. PLoS One 9(7):e101363.

Associate editor: Partha Majumder