

UC Merced

UC Merced Previously Published Works

Title

Global population genetics and diversity in the TAS2R bitter taste receptor family

Permalink

<https://escholarship.org/uc/item/4bk4d338>

Authors

Wooding, Stephen P

Ramirez, Vicente A

Publication Date

2022

DOI

10.3389/fgene.2022.952299

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed



## OPEN ACCESS

EDITED BY  
Daniel Liang-Dar Hwang,  
The University of Queensland, Australia

REVIEWED BY  
Masha Y. Niv,  
Hebrew University of Jerusalem, Israel  
Emma Louise Feeny,  
University College Dublin, Ireland

\*CORRESPONDENCE  
Stephen P. Wooding,  
swooding@ucmerced.edu

SPECIALTY SECTION  
This article was submitted to  
Nutrigenomics,  
a section of the journal  
Frontiers in Genetics

RECEIVED 24 May 2022  
ACCEPTED 25 August 2022  
PUBLISHED 11 October 2022

CITATION  
Wooding SP and Ramirez VA (2022),  
Global population genetics and diversity  
in the TAS2R bitter taste receptor family.  
*Front. Genet.* 13:952299.  
doi: 10.3389/fgene.2022.952299

COPYRIGHT  
© 2022 Wooding and Ramirez. This is an  
open-access article distributed under  
the terms of the [Creative Commons  
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,  
distribution or reproduction in other  
forums is permitted, provided the  
original author(s) and the copyright  
owner(s) are credited and that the  
original publication in this journal is  
cited, in accordance with accepted  
academic practice. No use, distribution  
or reproduction is permitted which does  
not comply with these terms.

# Global population genetics and diversity in the TAS2R bitter taste receptor family

Stephen P. Wooding<sup>1\*</sup> and Vicente A. Ramirez<sup>2</sup>

<sup>1</sup>Department of Anthropology, University of California, Merced, Merced, CA, United States,

<sup>2</sup>Department of Public Health, University of California, Merced, Merced, CA, United States

Bitter taste receptors (TAS2Rs) are noted for their role in perception, and mounting evidence suggests that they mediate responses to compounds entering airways, gut, and other tissues. The importance of these roles suggests that *TAS2Rs* have been under pressure from natural selection. To determine the extent of variation in *TAS2Rs* on a global scale and its implications for human evolution and behavior, we analyzed patterns of diversity in the complete 25 gene repertoire of human *TAS2Rs* in ~2,500 subjects representing worldwide populations. Across the *TAS2R* family as a whole, we observed 721 single nucleotide polymorphisms (SNPs) including 494 nonsynonymous SNPs along with 40 indels and gained and lost start and stop codons. In addition, computational predictions identified 169 variants particularly likely to affect receptor function, making them candidate sources of phenotypic variation. Diversity levels ranged widely among loci, with the number of segregating sites ranging from 17 to 41 with a mean of 32 among genes and per nucleotide heterozygosity ( $\pi$ ) ranging from 0.02% to 0.36% with a mean of 0.12%.  $F_{ST}$  ranged from 0.01 to 0.26 with a mean of 0.13, pointing to modest differentiation among populations. Comparisons of observed  $\pi$  and  $F_{ST}$  values with their genome wide distributions revealed that most fell between the 5th and 95th percentiles and were thus consistent with expectations. Further, tests for natural selection using Tajima's  $D$  statistic revealed only two loci departing from expectations given  $D$ 's genome wide distribution. These patterns are consistent with an overall relaxation of selective pressure on *TAS2Rs* in the course of recent human evolution.

## KEYWORDS

taste, genetics, perception, selection, evolution

## Introduction

Bitter taste perception plays key roles in vertebrate biology and fitness. By enabling the detection of toxic defense compounds found in plants, it allows herbivores to monitor intake, preventing overexposure (Crozier et al., 2006; Reed and Knaapila, 2010). The protective role of bitter taste appears to be reduced in modern humans, whose diets consist mainly of domestic crops bred to be nontoxic. Nonetheless, bitter perception continues to shape behavior and health. For instance, many substances perceived as bitter

are nontoxic, and avoidance is unnecessary (Nissim et al., 2017; Glendinning, 2021). Such is the case with cruciferous vegetables. They are healthy but often contain harmless bitter compounds that discourage consumption. Complexity is added by the fact that the toxicity of bitter substances is not binary but continuous, ranging from harmless to highly dangerous (Nissim et al., 2017). Bitter perception also plays less obvious roles. Notably, the molecular mechanisms underlying bitter taste are found in chemosensory cells in the gut and airways, where they mediate cellular responses to ingested and inhaled substances (Deshpande et al., 2010; Calvo and Egan, 2015). Further, variation in bitter taste sensitivity is associated with ingestion related health measures such as tobacco use and body mass index (BMI), as well as extraoral traits including glucose homeostasis and resistance to respiratory infection (Dotson et al., 2008; Lee et al., 2012).

Bitter taste responses are initiated by receptors in the TAS2R family, a series of G protein-coupled receptors (GPCRs) allied to other sensory receptors including opsins (OPNs) and olfactory receptors (ORs) (Chaudhari and Roper, 2010). TAS2Rs are expressed on the surface of taste receptor cells (TRCs) in taste buds, where they are exposed to ingested substances. When exposed to agonists they trigger a transduction cascade that depolarizes the cell and generates a neural signal. Humans express 25 TAS2R isoforms, each of which is receptive to a different range of compounds. Scores of TAS2R agonists have been identified experimentally, and it is now known that most TAS2Rs are capable of responding to numerous compounds, indicating that the full range of compounds capable of eliciting bitter tastes is vast (Meyerhof et al., 2010; Dagan-Wiener et al., 2019). The extent to which TAS2R expression and response patterns in extraoral tissues mirror those in TRCs is not well investigated. However a number of TAS2Rs are expressed in chemosensory cells the airway and gut, suggesting those tissues too are responsive to diverse agonists (Martens et al., 2021).

The genes encoding TAS2R receptors harbor extensive polymorphism, including numerous nonsynonymous variants (Kim et al., 2005). Nonsynonymous variation is a common source of phenotypic variation, and its presence in TAS2Rs results in variation in perception. The most thoroughly documented example of the relationship is between variation at *TAS2R38* and perception of isothiocyanates (ITCs), a class of compounds that imparts bitterness to many vegetables (Kim et al., 2003; Wooding et al., 2010). *TAS2R38* harbors nonsynonymous variants that alter its responses to ITCs, and associations between *TAS2R38* genotypes and taste sensitivity to goitrin, a bitter yet harmless ITC found in many vegetables, are strong (Bufe et al., 2005; Wooding et al., 2010). Variation in *TAS2R38* also associates with preferences for downstream health variables relating to vegetable consumption such as adiposity (Tepper et al., 2009; Duffy et al., 2010). Thus, *TAS2R38* mediates a path from perception to health.

Findings similar to those surrounding *TAS2R38* continue to emerge at other loci. For instance, *TAS2R44* harbors numerous

nonsynonymous variants shaping its responses to the bitter off-tastes of artificial sweeteners such as saccharin, which associate with perception (Roudnitzky et al., 2011). Likewise, while some variants of *TAS2R44* associate with sensitivity to the bitterness of saccharin, others do not, and whether they associate with behaviors such as consumption of low calorie drinks is not known. Additional associations between variation in TAS2Rs, functional variation, and phenotypic variation have also emerged (Pronin et al., 2007; Intelmann et al., 2009; Roudnitzky et al., 2011; Hayes et al., 2015; Nolden et al., 2016). Given the large number of TAS2Rs in humans and their responses to a broad constellation of compounds, it is likely many more such relationships exist. An essential step forward in defining them will be to fully characterize of the extent of polymorphism in TAS2Rs, its diversity and population structure, and its evolutionary origins.

In this study, we sought to determine the extent of genetic diversity in TAS2Rs on a global scale and its implications for the evolution of bitter perception, functional variation, and similarities and differences among populations. To obtain a full portrait of variation in TAS2Rs, we cataloged variation across the complete TAS2R family in humans by analyzing whole genome sequences in ~2,500 worldwide subjects. We then used computational methods to predict the functional consequences of discovered variants and tested for signatures of natural selection to determine their influences during recent human evolution. Finally, we determined the distribution of TAS2R variation across human populations to shed light on the potential for phenotypic similarities and differences between them.

## Materials and methods

We examined genetic variation across the TAS2R family in data available from Phase 3 of the 1,000 Genomes Project (1,000GP) (The 1000 Genomes Project Consortium, 2015) (Table 1). The 1,000GP sample comprises 2,504 healthy, randomly selected subjects from 26 worldwide populations in five continental super populations. The complete TAS2R repertoire in humans includes 25 annotated functional loci and seven pseudogenes. However, because we focused on functional variation, we excluded the pseudogenes from analysis. Two TAS2Rs known to have high frequency whole gene deletion alleles, *TAS2R43* and *TAS2R45*, were also excluded, because the presence of deletion alleles causes errors in genotype calls estimates of allele frequency (Roudnitzky et al., 2011). Thus, our final sample was composed of complete polymorphism data from 23 TAS2Rs in 2,405 individuals.

We obtained the genomic coordinates of TAS2Rs from the Ensembl hg19/GRCh37 human genome assembly, the reference genome for the 1,000GP (Table 2). Ensembl Gene ID and Transcript IDs were obtained using the

TABLE 1 Sampled populations and super populations.

Super population	Population
Africa ( <i>N</i> = 661)	African Caribbeans in Barbados ( <i>N</i> = 96)
	Americans of African Ancestry in SW United States ( <i>N</i> = 61)
	Esan in Nigeria ( <i>N</i> = 99)
	Gambian in Western Divisions in the Gambia ( <i>N</i> = 113)
	Luhya in Webuye, Kenya ( <i>N</i> = 99)
	Mende in Sierra Leone ( <i>N</i> = 85)
	Yoruba in Ibadan, Nigeria ( <i>N</i> = 108)
Americas ( <i>N</i> = 347)	Colombians from Medellin, Colombia ( <i>N</i> = 94)
	Mexican Ancestry from Los Angeles, United States ( <i>N</i> = 64)
	Peruvian from Lima, Peru ( <i>N</i> = 85)
	Puerto Rican in Puerto Rico ( <i>N</i> = 104)
East Asia ( <i>N</i> = 504)	Chinese Dai in Xishuangbanna, China ( <i>N</i> = 93)
	Han Chinese in Beijing, China ( <i>N</i> = 103)
	Japanese in Tokyo, Japan ( <i>N</i> = 104)
	Kinh in Ho Chi Minh City, Vietnam ( <i>N</i> = 99)
	Southern Han Chinese ( <i>N</i> = 105)
Europe ( <i>N</i> = 503)	British in England and Scotland ( <i>N</i> = 91)
	Finnish in Finland ( <i>N</i> = 99)
	Iberian population in Spain ( <i>N</i> = 107)
	Toscani in Italia ( <i>N</i> = 107)
South Asia ( <i>N</i> = 489)	Utah residents (CEPH) with European ancestry ( <i>N</i> = 99)
	Bengali from Bangladesh ( <i>N</i> = 86)
	Gujarati Indian from Houston, TX ( <i>N</i> = 103)
	Indian Telegu from the United Kingdom ( <i>N</i> = 102)
	Punjabi from Lahore, Pakistan ( <i>N</i> = 96)
Sri Lankan Tamil from the United Kingdom ( <i>N</i> = 102)	

Ensembl browser. The transcript IDs were used to extract variation data from the 1,000GP database in variant call format (vcf) using the Tabix software package (Li, 2011). Variants were then categorized as synonymous, nonsynonymous, premature stop, lost stop, lost start, and frameshift using annotations obtained from the Ensembl database using the Variant Effect Predictor (VEP) (McLaren et al., 2016).

Genetic diversity was assessed at each locus using three measures, the number of segregating sites (*S*), nucleotide diversity ( $\pi$ ) and population substructure ( $F_{ST}$ ). *S* was defined as the number of variable nucleotide positions in a sequence.  $\pi$  was defined as mean pairwise nucleotide difference between sequences normalized to sequence length.  $F_{ST}$  is the fraction of diversity accounted for by between-population differences. Here, we calculated  $F_{ST}$  using the weighted method of Weir and

Cockerham (1984) These calculations were performed using VCFtools and the R packages PopGenome, heifstat, pegas, adegenet, gdsfmt, and SNPRelate packages (Goulet, 2004; Paradis, 2010; Danecek et al., 2011; Jombart and Ahmed, 2011; Pfeifer et al., 2014).

We used two algorithms to predict the functional impact of non-synonymous variants, PolyPhen-2 and SIFT (Kumar et al., 2009; Adzhubei et al., 2010). PolyPhen-2 predicts the impact of amino acid substitutions based on the position of the changes in the three dimensional structure of the protein, conservation relative to homologous genes, and the biochemical characteristics of the substituted amino acids. PolyPhen-2 scores range from 0.0 to 1.0, and represent the probability that the substitution is damaging. PolyPhen-2 also provides qualitative ratings based on the probabilities, including benign (probability 0.000–0.446), possibly damaging (0.446–0.908), and probably damaging (0.908). SIFT predicts the impact of amino acid substitutions based on evolutionary conservation relative to homologous genes, and physiochemical similarity between the amino acids substituted. SIFT scores range from 0.0 to 1.0 and estimate the probability that the variant is deleterious or tolerated. SIFT scores < 0.05 are categorized as deleterious and scores > 0.05 are categorized as tolerated. To further examine the distribution of variation with respect to TAS2Rs' molecular structure, which is also potentially indicative of variants' functional effects, we aligned the amino acid sequences of the 23 putatively functional receptors using Clustal W and determined the abundance of variable sites across secondary structures and individual residues across the length of the molecule (Sievers et al., 2011).

Tests for natural selection were performed using Tajima's *D* statistic, which compares the number of nucleotide differences between alleles with the number of variable sites, which are differentially affected by selective processes (Tajima, 1989). *D* tests are designed to detect selective effects in nonrecombining genomic regions but are applicable with elevated conservativeness to regions with recombination. Because TAS2Rs are small (~1 kb–2 kb), recombination rates in them are nonzero but likely low. Therefore, the results of *D* tests were slightly conservative in our sample. These analyses were performed using the PopGenome R package.

A key assumption of selection tests using Tajima's *D* is that population sizes have been constant. This assumption is violated in humans, whose populations have grown explosively over the last 60,000 years (Henn et al., 2012). Various strategies can be used control for this (Williamson et al., 2005). We took an empirical approach, using noncoding regions of the genome as a basis for comparisons. Roughly 98% of the human genome is thought to be noncoding. These regions do contain

TABLE 2 *TAS2R* genes, coordinates, and sizes.

Gene	GRCh 37 coordinates	Transcript length	# Codons
<i>TAS2R1</i>	5:9629109–9630463	1,355	300
<i>TAS2R3</i>	7:141463897–141464997	1,101	317
<i>TAS2R4</i>	7:141478242–141479235	994	300
<i>TAS2R5</i>	7:141490017–141491166	1,150	300
<i>TAS2R7</i>	12:10954131–10955226	1,096	319
<i>TAS2R8</i>	12:10958650–10959892	1,243	310
<i>TAS2R9</i>	12:10961693–10962767	1,075	313
<i>TAS2R10</i>	12:10977916–10978957	1,042	308
<i>TAS2R13</i>	12:11060525–11062161	1,637	304
<i>TAS2R14</i>	12:11090005–11091862	1,858	318
<i>TAS2R16</i>	7:122634759–122635754	996	292
<i>TAS2R19</i>	12:11174218–11175219	1,002	300
<i>TAS2R20</i>	12:11149094–11150474	1,381	310
<i>TAS2R30</i>	12:11285557–11287243	1,687	320
<i>TAS2R31</i>	12:11182986–11184006	1,021	310
<i>TAS2R38</i>	7:141672431–141673573	1,143	334
<i>TAS2R39</i>	7:142880512–142881528	1,017	339
<i>TAS2R40</i>	7:142919130–142920162	1,033	324
<i>TAS2R41</i>	7:143174966–143175889	924	308
<i>TAS2R42</i>	12:11338599–11339543	945	315
<i>TAS2R43</i>	na:na-na	na	na
<i>TAS2R45</i>	na:na-na	na	na
<i>TAS2R46</i>	12:11213964–11214893	930	310
<i>TAS2R50</i>	12:11138512–11139511	1,000	300
<i>TAS2R60</i>	7:143140546–143141502	957	319

substantial functional sequence (Encode Project Consortium, 2012); however, the maximum estimated functional content in noncoding regions is < 15% and is likely less than half that (Graur, 2017). Therefore, to approximate the neutral distribution of  $D$  in our sample, we iteratively calculated it in adjacent 1 kb windows spanning the length of the 1,000GP genomes, excluding annotated functional elements and known unstable and repetitive regions such as telomeres. We then defined the probability of observed  $D$  values as their percentile position in the genome-wide distribution. We denoted  $p$ -values from these empirical tests  $P_E$ -values to distinguish them from the  $p$ -values obtained in standard parametric tests. We used the same methodology to determine the  $P_E$  values of observed  $\pi$  and  $F_{ST}$  values. These analyses were performed using VCFTools and R (Danecek et al., 2011).

In addition to using  $F_{ST}$ , we examined population structure using principal components analyses (PCA). To gain insight into the potential for population differences in functional variation we performed these for all sites, non-synonymous sites, and synonymous sites. These analyses were performed using the *gdsfnt* and *SNPrelate* R packages and visualized using the *ggplot2* library (Zheng et al., 2012; Wickham, 2016).

## Results

Across our sample as a whole, we observed 721 single nucleotide polymorphisms, 9 short insertion-deletion polymorphisms, and 8 large structural variants (Tables 3, 4). The number of coding region variants ranged from 15 in *TAS2R13* to 40 in *TAS2R31*. The minor allele frequency (MAF) of variants across ranged from 0.0002 to 0.943 with an average of 0.033. However, their distribution was skewed strongly downward. The majority of sites ( $n = 447$ ) were observed just once in the sample, and thus were singletons. Eighty percent of sites had allele frequencies below 0.004 and 95% had allele frequencies below 0.34. Alleles with intermediate frequencies, between 0.25 and 0.75, accounted for 4.5% of sites.

Nucleotide diversity ranged from 0.005% in *TAS2R39* to 0.358% in *TAS2R20*, with a mean of 0.12%. Comparing the observed  $\pi$  values in *TAS2Rs* with the genome wide empirical distribution revealed that 20 of the 23 of observed values fell between the 5th and 95th percentiles, making them consistent with neutral expectations (Table 3). Two genes, *TAS2R20* and *TAS2R42*, had  $\pi$  values above the 95th percentile (percentiles 98.6 and 98.2 respectively). One, *TAS2R39*, had a  $\pi$  value below the 5th percentile, 0.005%. The mean  $\pi$ , 0.0012 fell at the 72nd percentile of the genomic empirical distribution, again well within expectations.

Of the 721 SNPs, 525 were nonsynonymous substitutions and 196 were synonymous. Among these, PolyPhen-2 and SIFT detected 239 variants predicted to alter *TAS2R* receptor function. Overall, 182 SNPs had PolyPhen-2 scores of Possibly or Probably Damaging and 188 had SIFT scores of Deleterious. The prediction tools were largely in agreement, with 131 SNPs predicted to be Possibly or Probably Damaging by PolyPhen-2 and Deleterious by SIFT. SIFT and PolyPhen-2 scores were also in agreement for 258 nonsynonymous SNPs predicted to be Benign by PolyPhen-2 and Tolerated by SIFT. Among sites where classification disagreed between the two methods, 51 were scored by PolyPhen-2 to be Possibly or Probably Damaging but not Deleterious by SIFT, and 57 were scored by were SIFT to be Deleterious but not Possibly or Probably damaging by PolyPhen-2. Sites likely to have functional effects but not scored by SIFT or PolyPhen-2 because they are not amino acid changes were also present. They included 27 SNPs resulting in premature stop codons, one resulting in the loss of a stop codon, and three resulting in a lost start codon. There were also nine small indels, seven of which resulted in frameshifts.

Because their characteristics predict that they will have functional effects, we denoted variants scored as possibly or probably damaging by PolyPhen and Deleterious by SIFT, indels, and changes altering start or stop codons as putatively high impact (PHI) sites, and their derived alleles PHI alleles. Under these criteria, we categorized 131 nonsynonymous SNPs along with the 9 indels as (PHI) sites. Notably, every *TAS2R* harbored at least one PHI site. Like

**TABLE 3** Summary of variation in *TAS2Rs*. The four main columns are segregating sites (*S*), nucleotide diversity ( $\pi$ ), Tajima's *D* statistic (*D*), and population differentiation ( $F_{ST}$ ). The percentile (% ile) of each value relative to the genome wide empirical distribution is given for  $\pi$ , *D*, and  $F_{ST}$ . Separate values for nonsynonymous (*n*) and synonymous (*s*) sites are given for *S* and  $F_{ST}$ .

Gene	<i>S</i>	( <i>n</i> , <i>s</i> )	$\pi$ (%)	(%ile)	<i>D</i>	(%ile)	$F_{ST}$	(%ile)	( <i>n</i> , <i>s</i> )
<i>TAS2R1</i>	27	(21, 6)	0.062	(42.1)	-2.02	(22.3)	0.18	(92.7)	(0.16, 0.01)
<i>TAS2R3</i>	26	(17, 9)	0.104	(65.2)	-1.49	(68.3)	0.10	(68.9)	(0.00, 0.14)
<i>TAS2R4</i>	30	(23, 7)	0.160	(84.1)	-1.22	(82.3)	0.14	(86.5)	(0.15, 0.02)
<i>TAS2R5</i>	33	(25, 8)	0.108	(67.3)	-1.90	(32.9)	0.13	(81.5)	(0.13, 0.04)
<i>TAS2R7</i>	40	(31, 9)	0.033	(22.5)	-2.20	(7.5)	0.05	(30.4)	(0.04, 0.01)
<i>TAS2R8</i>	32	(26, 6)	0.072	(48.4)	-1.98	(26.1)	0.21	(96.0)	(0.26, 0.20)
<i>TAS2R9</i>	31	(27, 4)	0.058	(39.1)	-1.91	(32.1)	0.17	(91.4)	(0.17, 0.00)
<i>TAS2R10</i>	32	(27, 5)	0.030	(20.8)	-2.17	(9.5)	0.14	(85.9)	(0.11, 0.20)
<i>TAS2R13</i>	17	(16, 1)	0.065	(44.0)	-1.66	(55.1)	0.23	(96.8)	(0.23, 0.00)
<i>TAS2R14</i>	30	(21, 9)	0.077	(51.6)	-1.82	(41.0)	0.14	(85.9)	(0.16, 0.15)
<i>TAS2R16</i>	36	(20, 16)	0.115	(70.5)	-1.73	(49.1)	0.22	(96.1)	(0.18, 0.25)
<i>TAS2R19</i>	31	(25, 6)	0.160	(84.1)	-1.25	(81.0)	0.10	(71.5)	(0.07, 0.13)
<i>TAS2R20</i>	36	(27, 9)	0.358	(98.6)	-0.25	(98.0)	0.26	(98.1)	(0.25, 0.30)
<i>TAS2R30</i>	33	(22, 11)	0.144	(80.0)	-1.25	(80.8)	0.07	(48.2)	(0.07, 0.07)
<i>TAS2R31</i>	41	(36, 5)	0.232	(93.8)	-1.17	(84.2)	0.06	(45.9)	(0.08, 0.00)
<i>TAS2R38</i>	32	(22, 10)	0.136	(77.7)	-1.32	(77.9)	0.08	(57.9)	(0.08, 0.00)
<i>TAS2R39</i>	32	(23, 9)	0.005	(3.5)	-2.26	(5.5)	0.01	(2.6)	(0.01, 0.00)
<i>TAS2R40</i>	28	(16, 12)	0.020	(14.9)	-2.13	(12.2)	0.03	(14.0)	(0.03, 0.00)
<i>TAS2R41</i>	31	(18, 13)	0.078	(51.7)	-1.81	(42.2)	0.09	(65.6)	(0.09, 0.09)
<i>TAS2R42</i>	36	(24, 12)	0.335	(98.2)	-0.47	(96.8)	0.22	(96.5)	(0.20, 0.30)
<i>TAS2R43</i>	na	(na, na)	na	(na)	na	(na)	na	(na)	(na, na)
<i>TAS2R45</i>	na	(na, na)	na	(na)	na	(na)	na	(na)	(na, na)
<i>TAS2R46</i>	32	(25, 7)	0.119	(71.8)	-1.58	(61.9)	0.05	(38.6)	(0.04, 0.10)
<i>TAS2R50</i>	28	(16, 12)	0.123	(73.4)	-1.46	(70.0)	0.23	(97.1)	(0.29, 0.19)
<i>TAS2R60</i>	34	(23, 11)	0.051	(33.3)	-2.02	(22.1)	0.10	(72.9)	(0.03, 0.12)

the overall distribution of allele frequencies, the distribution of allele frequencies at PHI sites was dominated by rare variants. Only two PHI alleles had frequencies above 0.20 and eight had frequencies above 0.01. The mean derived allele frequency among PHI sites was 0.004, with a minimum of 0.0002 and a maximum of 0.22.

Comparisons of the aligned *TAS2R* sequences revealed that polymorphism was present at 307 of 374 sites (82%; [Figure 1A](#)). Variation was found in all regions of the molecule and PHI sites were found in all but two regions, EL1 and EL3, which are both short (6 and 5 residues, respectively). The number of compared *TAS2Rs* harboring polymorphism at a given position also varied ([Figure 1B](#)). Sixty-seven positions were invariant in all *TAS2Rs*. However, the majority of positions (205 of 374) were polymorphic in two or more *TAS2Rs* and one, position 307, was polymorphic in nine of the 23 *TAS2Rs*.

Tajima's *D*, like the other diversity measures, varied substantially across loci. The highest and lowest values were observed at *TAS2R39* ( $D = -2.26$ ;  $P_E = 0.055$ ) and *TAS2R20* ( $D = -0.25$ ;  $P_E = 0.98$ ) and exceeded expectation under neutrality

for *TAS2R20* and *TAS2R42* ( $D = -0.47$ ;  $P_E = 0.97$ ) ([Figure 2](#)). The mean *D* across loci was  $-1.61$  ( $P_E = 0.60$ ), which is consistent with the overall excess of low frequency variants across bitter receptor genes.

$F_{ST}$  values ranged from 0.01 at *TAS2R39* ( $P_E = 0.026$ ) to 0.258 at *TAS2R20* ( $P_E = 0.98$ ). Differentiation exceeded expectation ( $P_E > 0.95$ ) for *TAS2R8*, *13*, *16*, *20*, *42*, and *50* when compared to the empirical distribution, and fell below expectation for one, *TAS2R39* ( $P_E = 0.026$ ) ([Figure 2](#)). The mean  $F_{ST}$  across loci was 0.13, which was consistent with previously published genome wide averages and lay at the 82nd percentile of the genomic distribution in the 1000GP ([Bamshad and Wooding, 2003](#)).  $F_{ST}$ s for nonsynonymous sites ranged from 0.0009 at *TAS2R3* to 0.29 at *TAS2R50*. Among synonymous sites  $F_{ST}$  ranged from 0.00 in *TAS2R13*, which had a single rare substitution, to 0.30 in *TAS2R20*.

Principal components analyses revealed overlaps in diversity among super populations ([Figure 3](#)). In the combined analysis of synonymous and nonsynonymous sites the first and second principal components together accounted for 17.4% of observed variance. Analyses in synonymous and



TABLE 4 PHI sites with minor allele frequency (MAF) > 0.001. Variant type abbreviations are nonsynonymous (Ns), stop gained (Sg), inframe change (If), frameshift change (Fs). Ref. and Alt. Indicate reference and alternate codons and amino acids at the given position.

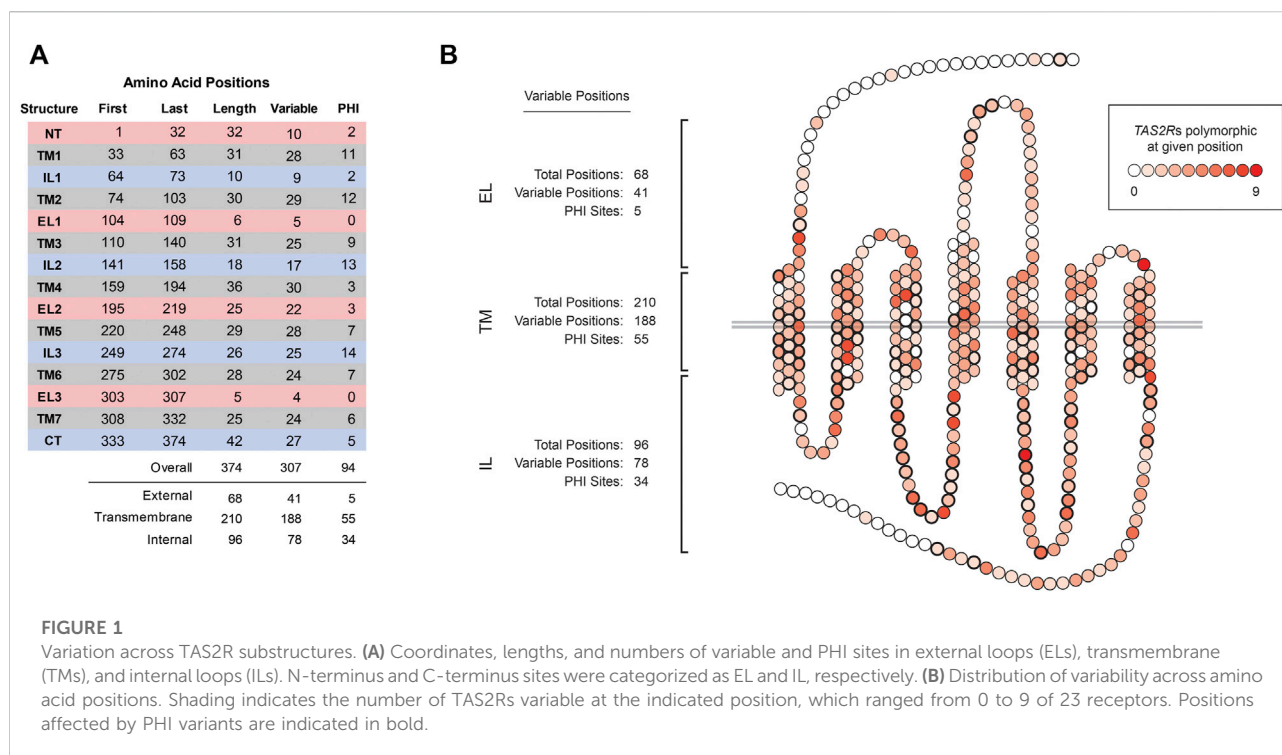
Gene	rsid	Var Type	Ref.Codon	Alt.Codon	Ref. AA	Alt. AA	MAF
TAS2R42	rs1669412	Ns	cGa	cAa	R	Q	0.2200
TAS2R46	rs2708381	Sg	tGg	tAg	W	Stop	0.2119
TAS2R60	rs35195910	If	gTCTtc	gtc	VF	V	0.0321
TAS2R7	rs77050900	Ns	aTt	aCt	I	T	0.0194
TAS2R31	rs139069360	Ns	tgG	tgT	W	C	0.0144
TAS2R46	rs201847607	Fs	Tgg	gg	W	X	0.0136
TAS2R9	rs113883583	Ns	gGg	gAg	G	E	0.0122
TAS2R31	rs116926686	Ns	ttA	ttT	L	F	0.0102
TAS2R10	rs117936881	Ns	Tgg	Cgg	W	R	0.0082
TAS2R8	rs200711805	Fs	ttc	ttTc	F	FX	0.0080
TAS2R5	rs2234014	Ns	cCg	cTg	P	L	0.0080
TAS2R8	rs41324347	Sg	Gaa	Taa	E	Stop	0.0076
TAS2R1	rs2234232	Ns	tGt	tAt	C	Y	0.0076
TAS2R10	rs201689842	Fs	Att	tt	I	X	0.0062
TAS2R14	rs35804287	Ns	Ctc	Ttc	L	F	0.0060
TAS2R7	rs150192473	Sg	Cga	Tga	R	Stop	0.0058
TAS2R39	rs184819681	Ns	aGc	aAc	S	N	0.0058
TAS2R20	rs116400924	Sg	tGg	tAg	W	Stop	0.0054
TAS2R8	rs142540719	Ns	aGa	aTa	R	I	0.0054
TAS2R42	rs139960283	Ns	tTg	tGg	L	W	0.0052
TAS2R7	rs139604652	Ns	gTg	gAg	V	E	0.0046
TAS2R46	rs150894148	Sg	Cag	Tag	Q	Stop	0.0042
TAS2R13	rs34885344	Ns	cAc	cGc	H	R	0.0042
TAS2R41	rs75955374	Ns	gTt	gAt	V	D	0.0040
TAS2R19	rs77837442	Sg	tgG	tgA	W	Stop	0.0038
TAS2R10	rs142507813	Ns	atG	atA	M	I	0.0038
TAS2R10	rs114006371	Ns	aTc	aAc	I	N	0.0038
TAS2R19	rs115193179	Ns	aAg	aCg	K	T	0.0038
TAS2R40	rs565742335	If	gAGAag	gag	EK	E	0.0036
TAS2R7	rs202246571	Ns	cTc	cAc	L	H	0.0034
TAS2R31	rs140958087	Ns	gGa	gTa	G	V	0.0034
TAS2R5	rs2234013	Ns	Ggt	Agt	G	S	0.0030
TAS2R8	rs61737282	Ns	aTa	aCa	I	T	0.0026
TAS2R9	rs148917754	Ns	tCa	tTa	S	L	0.0022
TAS2R8	rs114977408	Ns	Tac	Cac	Y	H	0.0020
TAS2R31	rs202165116	Ns	tTa	tCa	L	S	0.0020
TAS2R38	rs115966953	Ns	cGg	cAg	R	Q	0.0020

nonsynonymous sites separately yielded similar results, with the first two components accounting for 23.6% and 17.5% respectively. The results illustrated the extent of similarity and difference among individuals specifically with respect to variation in TAS2Rs, with coordinates near each other reflecting greater similarity and distant points reflecting less similarity. In all three data subsets (all sites, synonymous sites, and nonsynonymous sites), subjects from Asia, Europe, and the Americas showed extensive overlap, reflecting overlap in the

variability of TAS2R loci. The African subjects were somewhat distinct, falling in an partially overlapping but somewhat different cluster in all analyses.

## Discussion

The broad importance of bitter perception to vertebrate biology raises questions about its specific importance to



humans, who pose a conundrum. On the one hand humans have sophisticated cognitive abilities and might use cues other than bitter taste, such as learned recognition or social behavior, to gauge food safety. This could reduce selective pressures on TAS2Rs. On the other hand, human history over last 65,000 years has been characterized by explosive population growth and rapid diffusion, which brought humans into contact with myriad novel environments (Bergström et al., 2021). This could have enhanced TAS2Rs' value for evaluating the toxicity of unfamiliar plants, increasing selective pressures on them. Moreover, these alternatives are not mutually exclusive, and might vary across loci. Our findings clarify global patterns of diversity in TAS2Rs and evidence for selective pressures on them.

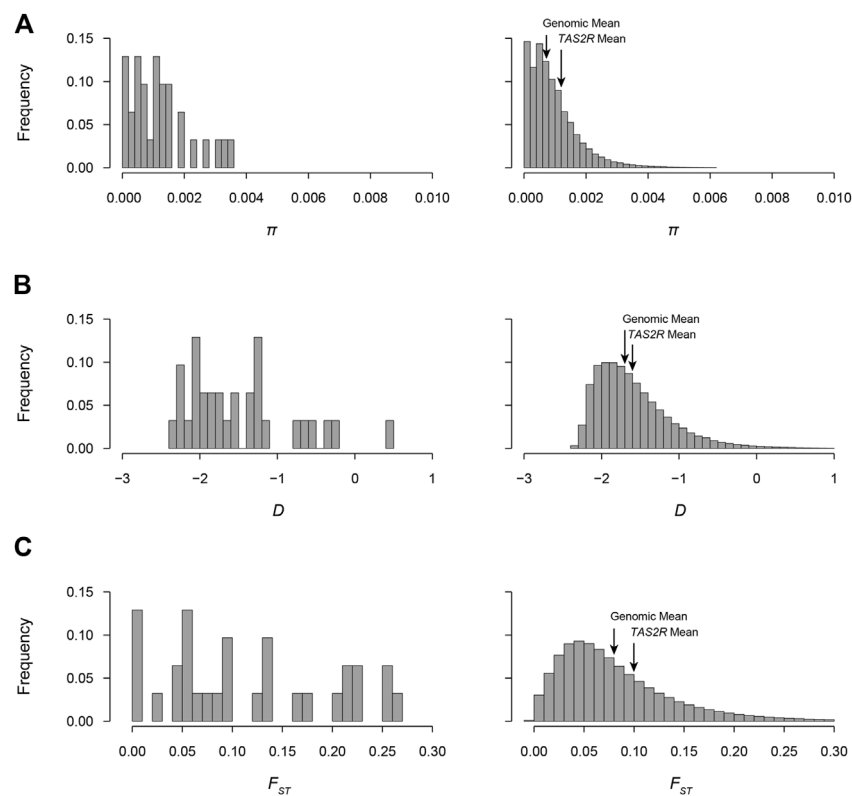
Consistent with previous studies, the variation we found across TAS2Rs in the 1,000GP was extensive with respect to both synonymous and nonsynonymous polymorphism (Tables 3, 4). The mean number of segregating sites per locus in our sample (31.7 per locus), as well as the aggregate number across loci (721 SNPs) far exceeded all previous reports in family wide studies of TAS2Rs. For instance, in a study across the TAS2R family in 22 globally distributed subjects Wang et al. (2004) observed 105 segregating sites (4.2 per locus), and a study in 55 subjects by Kim et al. (2005) yielded 151 sites (5.6 per locus). We observed a similar trend in single locus studies. While we cataloged 32 segregating sites at TAS2R38 in the 1,000GP, Wooding et al. (2004) observed 5 in a sample of 155 global subjects and Campbell et al. (2012) observed 21 in a sample of 611 Africans. Similarly, we found 36 segregating sites at TAS2R16

whereas Campbell et al. (2014) found 15 across 689 worldwide subjects and Soranzo et al. (2005) found 17 across 997. Therefore, from an absolute standpoint our findings provide the largest and most comprehensive inventory of TAS2R variation to date.

A second feature of our data, an abundance of nonsynonymous variation, was also consistent with previous studies. We found that the mean number of segregating sites per locus in our sample, 31.7, was accounted for by a mean of 23.1 nonsynonymous and 8.6 synonymous sites, such that 73% of substitutions were nonsynonymous changes. This places our estimate intermediate between those reported in previous multi-locus studies such as the 69% rate reported Wang et al. (2004) and the 75% rate reported by Kim et al. (2005). Nonsynonymous rates were also similar for individual loci. For instance we observed a rate of 69% at TAS2R38 while Campbell et al. (2012) found a 90% rate and Wooding et al. (2004) found a 100% rate. Likewise, we found a 56% rate at TAS2R16 while Campbell et al. (2014) found a rate of 47% and Soranzo et al. (2005) found a rate of 64%. Across all loci in our study we found a substantial range of rates, from a minimum of 56% (at TAS2R16) to 94% (at TAS2R13), with a mean of 73%. Because amino acid substitutions can alter the properties of proteins, and hence their function, the abundant nonsynonymous variation in TAS2Rs suggests that they could produce substantial variation in taste phenotypes.

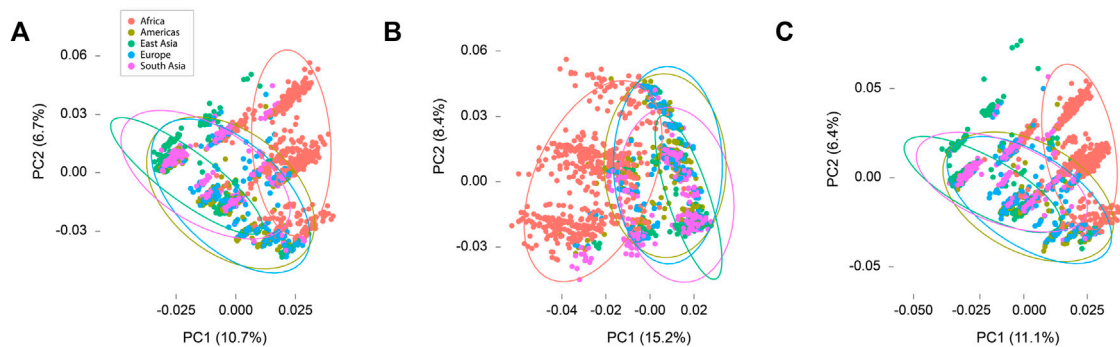
While the presence of nonsynonymous variation points to the presence of functional variation, it does not demonstrate it. Some amino acid substitutions are more impactful than others as the result of chemical similarities between residues, their





**FIGURE 2**

Distributions of  $\pi$ , Tajima's  $D$ , and  $F_{ST}$  in the 1000GP. Values observed in *TAS2Rs* and genome-wide sliding windows are shown in the left and right of each panel, respectively. **(A)**  $\pi$  distributions. Mean values in *TAS2Rs* and genome-wide were 0.0012 and 0.0009, respectively. **(B)**  $D$  distributions. Mean values in *TAS2Rs* and genome-wide were  $-1.61$  and  $-1.59$  and. **(C)**  $F_{ST}$  distributions. Mean values in *TAS2Rs* and genome-wide were 0.13 and 0.08.



**FIGURE 3**

Principal components plots. **(A)** All sites. **(B)** Synonymous sites only. **(C)** Nonsynonymous sites only.

positions within the protein structure, and other factors. As a result, some substitutions have no discernable effects, while others have intermediate or strong ones. This has implications

for the effects of substitutions on taste phenotypes. In particular, is likely that some of the substitutions we observed affect taste responses while others do not. These alternatives are exemplified

in TAS2Rs by the findings of Roudnitzky et al. (2011). In a combined *in vitro* and psychophysical analysis of variation in TAS2R31, Roudnitzky et al. identified nine amino acid substitutions. Among these one, TAS2R31-R35W had strong effects on receptor function and was associated with taste responses to the bitter off-tastes of saccharin and acesulfame potassium. However, other substitutions had more complex effects. For instance, a TAS2R31-D45H substitution altered receptor function only when combined with the W35 allele. Amino acid replacements in other TAS2Rs exhibit graded effects. For instance, Roudnitzky et al. (2015) found that two amino acid replacements in TAS2R46 combine to form three alleles with high, medium, and low responses *in vitro*. Similarly, Wooding et al. found that amino acid substitutions in TAS2R38 differentially affect its functional responses and taste responses to phenylthiocarbamide (PTC) and a bitter compound found in vegetables, goitrin.

The computational predictions we obtained using SIFT and PolyPhen 2 were consistent with the suggestion that TAS2Rs harbor extensive functional variation. In total, 131 SNPs had both high SIFT and high PolyPhen-2 scores, and were thus PHI sites, with an additional 35 PHI sites being altered stop codons, start codons and indels. Taken together, these findings point to the presence of numerous TAS2R variants capable of affecting taste phenotypes. Nonetheless, the frequency distribution of PHI alleles suggested that the contributions of most sites to phenotypic variance are small (Table 4). The distribution was strongly skewed downward such that although there were 166 PHI sites, just 21 had frequencies above 0.005 and eight had frequencies above 0.01. Only two sites were found at moderate frequencies, a premature stop in TAS2R46 (rs2708381) with a frequency of 0.21 and an arginine-glutamine substitution in TAS2R42 (rs1669412) with a frequency of 0.22. In addition, while most individual sites and all subdomains of TAS2Rs harbored variation, it was not obviously concentrated in any region of the molecule. Thus, while scores of variants in TAS2Rs were computationally predicted to result in functional variation, most of them were rare so their effects must be limited to a small number of subjects and contribute little to phenotypic variance more broadly.

Cross referencing PHI sites against functional sites detected in previous studies revealed potential weaknesses in the ability of SIFT and PolyPhen-2 to identify variants with phenotypic effects. In particular, the PHI sites identified with SIFT and PolyPhen-2 showed no overlap with sites previously demonstrated to have functional and phenotypic consequences. These included three sites in TAS2R38 known to mediate taste responses to PTC and PROP (Bufe et al., 2005), two sites in TAS2R43 and -44 mediating responses to aloin, saccharin, and acesulfame potassium (Pronin et al., 2007; Roudnitzky et al., 2011), and two sites in TAS2R16 mediating taste responses to glucopyranosides (Soranzo et al., 2005). Importantly, the lack of overlap between PHI sites and

known functional sites does not imply a failure on the part of SIFT and PolyPhen-2. First, the number of confirmed functional, phenotypically important polymorphisms in human TAS2Rs is currently small, limiting the potential for overlap. Second, both algorithms have high true positive rates and hence also have high false negative rates. This might have resulted in a high type-2 error rate in our analyses. Further, these algorithms are aimed at detecting the effects of changes at individual sites, not the effects of site-site interactions, which are known to occur in TAS2Rs, such as the aforementioned interactions at TAS2R31. Nonetheless, given that functional variation in TAS2Rs has been demonstrated unambiguously *in vitro*, we hypothesize that TAS2Rs do harbor substantial functional variation. They may be elucidated more effectively recent and sophisticated methods such as computational inferences, which are increasingly powerful (Bahia et al., 2018; Dagan-Wiener et al., 2019; Margulis et al., 2021). The emergence of new techniques for determining proteins' three dimensional structures, such as AlphaFold may be a particularly effective resource (Jumper et al., 2021).

Natural selection has long been recognized as a probable influence on bitter taste perception (Fisher et al., 1939). Because it is directly involved in feeding behaviors bitter perception seems likely to affect reproductive success, placing it under selective pressure. However, previous studies of natural selection in human TAS2Rs have reached contrasting conclusions. For instance, while Wooding et al. (2004) found patterns of diversity consistent with the effects of natural selection on TAS2Rs, Wang et al. (2004) concluded that selective pressures on TAS2Rs are relaxed. Our findings in the 1,000GP are more consistent with the latter perspective. Among the 23 loci we examined, only two had *D* values above the 95th percentile in the empirical distribution, which was assumed to represent neutral expectations (TAS2R20,  $P_E = 0.98$ ; TAS2R42,  $P_E = 0.97$ ) (Table 3). These two measures reflect significantly high *D* values, which typically result from diversifying pressures such as balancing natural selection or local adaptation {Enard, 2021 #1976}. Thus, while some loci with evidence of selection were present, more than 90% had *D* values consistent with selective neutrality. Further, the mean *D* value in TAS2Rs was also close to the mean of the genome wide empirical distribution (-1.61 vs. -1.59,  $P_E = 0.57$ ), the result expected under neutrality (Figure 2). Overall, this is evidence for a general decrease or absence of selective pressure on TAS2Rs in modern human populations, but there are important exceptions that merit further investigation.

Evidence that selective pressures have been relaxed on TAS2Rs raises questions about the timing of relaxation. One possibility is that selective pressures on TAS2Rs shifted with the advent of domestication, which enabled humans to produce large quantities of food using nontoxic crops such as wheat, rice, and corn. However, plants were first domesticated just ~10,000 years ago (Gross and Olsen, 2010). It is conceivable that the equilibration of diversity in TAS2Rs from expectations under

selection to neutral expectations occurred in that short time frame, but it may have taken place over a more extended period. We speculate that the relaxation of selective pressures on *TAS2Rs* is better explained by longer term factors in human evolution and behavior, particularly social behavior, learned foraging strategies, and cognitive developments. These have likely enhanced humans' ability to forage safely and efficiently without cues from taste for hundreds of thousands of years. If so, *TAS2Rs* probably began evolving neutrally prior to humans' dispersal from Africa. This would predict that other genes implicated in coping with phytotoxins, such as cytochrome P40 genes, experienced shifts in selective pressures over the time period. It would also predict that most or all *TAS2Rs* experienced reductions in selective pressure simultaneously. These suggestions are potentially testable using phylogenetic methods to compare *TAS2Rs* in humans with those in other primates or with genes predicted to be under different selective pressures over the same time period.

The proposal that human populations differ with respect to taste sensitivity emerged as soon as phenotypic variation in taste responses was discovered (Fox, 1932). Our findings suggest populations do differ with respect to functional variation at some *TAS2R* loci. However, the differences overall were no greater than expected given the genome wide neutral distribution. First,  $F_{ST}$  under panmixia, in which there is no differentiation among populations, is 0. We found that only one gene, *TAS2R39*, had an  $F_{ST}$  with  $P_E < 0.05$ , and was thus distributed more homogeneously than expected by chance. The remaining 22 *TAS2Rs* had  $F_{ST}$ s with  $P_E > 0.05$ , which is consistent with the presence of at least some genetic differentiation. In addition, six loci (*TAS2R8*, -13, -16, -20, -42, and -50) had  $P_E$  values  $> 0.95$ , and thus showed strong evidence of excess genetic differentiation that could result in phenotypic differences. Moreover, the  $F_{ST}$  at *TAS2R13*, 0.23 ( $P_E = 0.97$ ), was accounted for entirely by nonsynonymous variation, suggesting that it is particularly likely to underlie population differences. However, while a number of individual loci showed evidence of departing from expectations, the distribution of  $F_{ST}$  across the *TAS2R* family as a whole did not reveal differences from the genome wide empirical distribution. When compared across loci, the mean  $F_{ST}$  observed across *TAS2Rs*, 0.13, was somewhat higher than the genome wide mean but not extremely so, falling in the 81.5th percentile of the overall distribution (Figure 2). This suggests that populations are differentiated with respect to variation in *TAS2Rs*, but most loci are no more differentiated than expected under neutrality.

The extent of overlap among populations with respect to variation in *TAS2Rs* was emphasized by the results of principal components analyses (Figure 3). These revealed that the African super population was somewhat distinct from the others in every analysis, falling in an overlapping but slightly separate cluster from the other four super populations. This is a common pattern in humans (Bamshad et al., 2004). Nonetheless, the pattern in

nonsynonymous variants was noteworthy. While SIFT and PolyPhen-2 did not detect high frequency functional variants, their high false negative rates suggest that many functional variants could be present but undetected. Therefore, while the results of principal components analyses point to overall similarities in population differentiation with respect to synonymous and nonsynonymous variation, the presence of differentiation with respect to nonsynonymous sites suggests that there could be systematic differences among super populations with respect to those sites and, hence, with respect to their respective bitter taste phenotypes. Differences between African and non-African populations are particularly likely.

Variation in bitter taste sensitivity has long been recognized as a potential health factor because it is tied to ingestive behaviors, many of which have potential consequences for health. For instance, connections between taste sensitivity to PTC and thyroid health were proposed as early as 1949 (Harris et al., 1949). Later studies proposed connections between bitter taste and food preferences more generally, and between bitter taste and tobacco use (Fox, 1954; Fischer et al., 1963). The emergence of molecular genetic technologies transformed perspectives on connections between bitter taste perception and health by uncovering many specific compounds eliciting bitter responses as well as the *TAS2Rs* controlling them. The most complete portrait of these relationships is exemplified by *TAS2R38*. *TAS2R38* is responsive to a range of isothiocyanates (ITCs) impairing thyroid activity and harbors variants altering *TAS2R38*'s responses to them. Further, *TAS2R38* genotypes associate with taste responses to ITCs and preferences for foods containing them, as well as with downstream health traits shaped by ingestion behaviors, such as body mass index (Ortega et al., 2016).

Relationships between genetic variation, taste responses and health involving *TAS2R* loci other than *TAS2R38* are not as well defined, but they show evidence of similar effects. For instance the association of taste sensitivity to the bitter off-taste of noncaloric sweeteners with variants in *TAS2R31* suggests they may influence preferences for diet foods, with possible consequences for weight loss and long term BMI (Roudnitzky et al., 2011). Variants in *TAS2R43* affect the receptor's responses to aristolochic acid, a carcinogenic contaminant of wheat supplies in Eastern Europe, and predict kidney damage in exposed populations (Grollman et al., 2007; Pronin et al., 2007; Wooding et al., 2012). Other potentially important associations include relationships between variants in *TAS2R16*, receptor response, and the perception of salicin, and between variants in *TAS2R3*, -4, and -5 and perception of ethanol, suggesting they may shape consumption and its health effects (Campbell et al., 2014; Nolden et al., 2016). Our finding that *TAS2Rs* harbor hundreds of nonsynonymous variants suggests that the responses they mediate may be similarly extensive and can be anticipated to emerge in future studies.

It is important to note that while much focus in research on bitter taste and health has been on individual *TAS2Rs*, it is well recognized that other factors are important, particularly when

considering complex phenotypes such as food and drink likings (Dioszegi et al., 2019). These include the potential for contributions from multiple TAS2Rs simultaneously, contributions from individual TAS2Rs to multiple traits, contributions from other molecules participating in taste such as salivary proteins, contributions from neurological systems, and environmental factors. An important development in dissecting these has been the emergence of extremely large population samples, some of which now exceed 100,000 subjects. These provide a statistically powerful means of detecting both predicted and unpredicted influences. For instance, a recent study of the UK-Biobank (> 160,000 subjects) by May-Wilson et al. (2022) detected more than a thousand associations with food and drink likings, which implicated not just TAS2Rs but myriad other genes as well, such as *FGF21* (fibroblast growth factor), *ADH1B* (alcohol dehydrogenase), and *MHC* (major histocompatibility complex) loci.

While *TAS2Rs* are best known for their roles mediating bitter taste responses and ingestive behaviors, mounting evidence indicates that they play important roles outside perception, as well. For instance, they are highly expressed in cells in gut and bronchial smooth muscle responsive to ingested and inhaled compounds (Calvo and Egan, 2015; Cohen, 2017; An and Liggett, 2018; Tran et al., 2021). For instance, stimulation of gut expressed *TAS2Rs* triggers endocrine responses and gastric emptying (Dotson et al., 2008; Egan and Margolskee, 2008). Evidence that *TAS2Rs* can monitor quorum sensing in intestinal flora raises the possibility that they also mediate responses to changing conditions in the gut (Cohen, 2017; Fan and Pedersen, 2021). Recent findings point to another function of *TAS2Rs* in the gut, the detection of parasites emitting *TAS2R* agonists (Luo et al., 2019). Similar patterns are found in the airways, where *TAS2Rs* trigger responses to both endogenous and exogenous compounds entering the lungs, and likely mediate immune responses (Cohen, 2017; Tran et al., 2021). Investigations of the impact of genetic variation in *TAS2Rs* on these processes are so far scant, but those available point to important effects. For example, consistent with evidence that ingestion of bitter compounds can produce hormonal responses in gut cells, Dotson et al. (Dotson et al., 2008) found that variation in *TAS2R9* associates with glucose dysregulation and Clark et al. (Clark et al., 2015) found that variation in *TAS2R42* associates with thyroid hormone levels. Similarly, in an investigation of variation of associations between *TAS2R38* variants and respiratory health, Lee et al. (Lee et al., 2012) found that variation in *TAS2R38* associates with susceptibility to infection.

## Conclusion

Our findings highlight the complexity of relationships between bitter taste receptors, perception, evolution, and

health. From a broad perspective, the patterns we observe across the *TAS2R* family point to a reduction in selective pressure in the course of human evolution, possibly as the result of evolved changes in other aspects of human biology and behavior. However, some loci do show evidence of departures from neutrality that may represent recent or ongoing influences from selection. Further, while computational analyses in our sample did not identify high frequency sites predicted to have functional effects, they did not exclude them and the presence of hundreds of nonsynonymous variants, including some with high frequencies and frequency differences among populations, raises the possibility that they shape *TAS2R* mediated phenotypes. Ongoing developments in molecular genetics and computational biology make dissecting their effects a compelling prospect.

## Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: 1000 Genomes Project (<https://www.internationalgenome.org/>). Data files used for analysis are provided in the article [Supplementary Material](#).

## Author contributions

SW conceived project and authored manuscript. VR conducted analyses and co-authored manuscript.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.952299/full#supplementary-material>

## References

- Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., et al. (2010). A method and server for predicting damaging missense mutations. *Nat. Methods* 7, 248–249. doi:10.1038/nmeth0410-248
- An, S. S., and Liggett, S. B. (2018). Taste and smell GPCRs in the lung: Evidence for a previously unrecognized widespread chemosensory system. *Cell. Signal.* 41, 82–88. doi:10.1016/j.cellsig.2017.02.002
- Bahia, M. S., Nissim, I., and Niv, M. Y. (2018). Bitterness prediction *in-silico*: A step towards better drugs. *Int. J. Pharm.* 536, 526–529. doi:10.1016/j.ijpharm.2017.03.076
- Bamshad, M., and Wooding, S. P. (2003). Signatures of natural selection in the human genome. *Nat. Rev. Genet.* 4, 99–111. doi:10.1038/nrg999
- Bamshad, M., Wooding, S., Salisbury, B. A., and Stephens, J. C. (2004). Deconstructing the relationship between genetics and race. *Nat. Rev. Genet.* 5, 598–609. doi:10.1038/nrg1401
- Bergström, A., Stringer, C., Hajdinjak, M., Scerri, E. M. L., and Skoglund, P. (2021). Origins of modern human ancestry. *Nature* 590, 229–237. doi:10.1038/s41586-021-03244-5
- Bufe, B., Breslin, P. A., Kuhn, C., Reed, D. R., Tharp, C. D., Slack, J. P., et al. (2005). The molecular basis of individual differences in phenylthiocarbamide and propylthiouracil bitterness perception. *Curr. Biol.* 15, 322–327. doi:10.1016/j.cub.2005.01.047
- Calvo, S. S., and Egan, J. M. (2015). The endocrinology of taste receptors. *Nat. Rev. Endocrinol.* 11, 213–227. doi:10.1038/nrendo.2015.7
- Campbell, M. C., Ranciaro, A., Froment, A., Hirbo, J., Omar, S., Bodo, J. M., et al. (2012). Evolution of functionally diverse alleles associated with PTC bitter taste sensitivity in Africa. *Mol. Biol. Evol.* 29, 1141–1153. doi:10.1093/molbev/msr293
- Campbell, M. C., Ranciaro, A., Zinshteyn, D., Rawlings-Goss, R., Hirbo, J., Thompson, S., et al. (2014). Origin and differential selection of allelic variation at *TAS2R16* associated with salicin bitter taste sensitivity in Africa. *Mol. Biol. Evol.* 31, 288–302. doi:10.1093/molbev/mst211
- Chaudhari, N., and Roper, S. D. (2010). The cell biology of taste. *J. Cell Biol.* 190, 285–296. doi:10.1083/jcb.201003144
- Clark, A. A., Dotson, C. D., Elson, A. E., Voigt, A., Boehm, U., Meyerhof, W., et al. (2015). *TAS2R* bitter taste receptors regulate thyroid function. *FASEB J.* 29, 164–172. doi:10.1096/fj.14-262246
- Cohen, N. A. (2017). The genetics of the bitter taste receptor *T2R38* in upper airway innate immunity and implications for chronic rhinosinusitis. *Laryngoscope* 127, 44–51. doi:10.1002/lary.26198
- Crozier, A., Clifford, M. N., and Asihara, H. (2006). *Plant secondary metabolites: Occurrence, structure, and role in the human diet*. Oxford, UK: Blackwell Publishing, Ltd.
- Dagan-Wiener, A., Di Pizio, A., Nissim, I., Bahia, M. S., Dubovski, N., Margulis, E., et al. (2019). BitterDB: Taste ligands and receptors database in 2019. *Nucleic Acids Res.* 47, D1179–D1185. doi:10.1093/nar/gky974
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., et al. (2011). Genomes Project AnalysisThe variant call format and VCFtools. *Bioinformatics* 27, 2156–2158. doi:10.1093/bioinformatics/btr330
- Deshpande, D. A., Wang, W. C., McIlmoyle, E. L., Robinett, K. S., Schillinger, R. M., An, S. S., et al. (2010). Bitter taste receptors on airway smooth muscle bronchodilate by localized calcium signaling and reverse obstruction. *Nat. Med.* 16, 1299–1304. doi:10.1038/nm.2237
- Dioszegi, J., Llanaj, E., and Adany, R. (2019). Genetic background of taste perception, taste preferences, and its nutritional implications: A systematic review. *Front. Genet.* 10, 1272. doi:10.3389/fgene.2019.01272
- Dotson, C. D., Zhang, L., Xu, H., Shin, Y. K., Vignes, S., Ott, S. H., et al. (2008). Bitter taste receptors influence glucose homeostasis. *PLoS One* 3, e3974. doi:10.1371/journal.pone.0003974
- Duffy, V. B., Hayes, J. E., Davidson, A. C., Kidd, J. R., Kidd, K. K., and Bartoshuk, L. M. (2010). Vegetable intake in college-aged adults is explained by oral sensory phenotypes and *TAS2R38* genotype. *Chemosens. Percept.* 3, 137–148. doi:10.1007/s12078-010-9079-8
- Egan, J. M., and Margolskee, R. F. (2008). Taste cells of the gut and gastrointestinal chemosensation. *Mol. Interv.* 8, 78–81. doi:10.1124/mi.8.2.5
- Encode Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57–74. doi:10.1038/nature11247
- Fan, Y., and Pedersen, O. (2021). Gut microbiota in human metabolic health and disease. *Nat. Rev. Microbiol.* 19, 55–71. doi:10.1038/s41579-020-0433-9
- Fischer, R., Griffin, F., and Kaplan, A. R. (1963). Taste thresholds, cigarette smoking, and food dislikes. *Med. Exp. Int. J. Exp. Med.* 210, 151–167. doi:10.1159/000135346
- Fisher, R. A., Ford, E. B., and Huxley, J. (1939). Taste-testing the anthropoid apes. *Nature* 144, 750. doi:10.1038/144750a0
- Fox, A. L. (1954). “A new approach to explaining food preferences,” in *Abstracts of papers of the American chemical society’s 126th meeting, division of agriculture and food chemistry* 14.
- Fox, A. L. (1932). The relationship between chemical constitution and taste. *Proc. Natl. Acad. Sci. U. S. A.* 18, 115–120. doi:10.1073/pnas.18.1.115
- Glendinning, J. I. (2021). What does the taste system tell us about the nutritional composition and toxicity of foods? *Handb. Exp. Pharmacol.* 275, 321–351. doi:10.1007/164\_2021\_451
- Goudet, J. (2004). *hierfstat*, a package for R to compute and test hierarchical *F*-statistics. *Mol. Ecol. Notes* 5, 184–186. doi:10.1111/j.1471-8286.2004.00828.x
- Graur, D. (2017). An upper limit on the functional fraction of the human genome. *Genome Biol. Evol.* 9, 1880–1885. doi:10.1093/gbe/evx121
- Grollman, A., Shibutani, S., Moriya, M., Miller, F., Wu, L., Moll, U., et al. (2007). Aristolochic acid and the etiology of endemic (Balkan) nephropathy. *Proc. Natl. Acad. Sci. U. S. A.* 104, 12129–12134. doi:10.1073/pnas.0701248104
- Gross, B. L., and Olsen, K. M. (2010). Genetic perspectives on crop domestication. *Trends Plant Sci.* 15, 529–537. doi:10.1016/j.tplants.2010.05.008
- Harris, H., Kalmus, H., and Trotter, W. R. (1949). Taste sensitivity to phenylthiourea in goitre and diabetes. *Lancet* 254, 1038. doi:10.1016/s0140-6736(49)91605-0
- Hayes, J. E., Feeney, E. L., Nolden, A. A., and Mcgeary, J. E. (2015). Quinine bitterness and grapefruit liking associate with allelic variants in *TAS2R31*. *Chem. Senses* 40, 437–443. doi:10.1093/chemse/bjv027
- Henn, B. M., Cavalli-Sforza, L. L., and Feldman, M. W. (2012). The great human expansion. *Proc. Natl. Acad. Sci. U. S. A.* 109, 17758–17764. doi:10.1073/pnas.1212380109
- Intelmann, D., Batram, C., Kuhn, C., Haseleu, G., Meyerhof, W., and Hofmann, T. (2009). Three *TAS2R* bitter taste receptors mediate the psychophysical responses to bitter compounds of Hops (*Humulus lupulus* L.) and Beer. *Chemosens. Percept.* 2, 118–132.
- Jombart, T., and Ahmed, I. (2011). *Adegenet* 1.3-1: new tools for the analysis of genome-wide SNP data. *Bioinformatics* 27, 3070–3071. doi:10.1093/bioinformatics/btr521
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589. doi:10.1038/s41586-021-03819-2
- Kim, U. K., Jorgenson, E., Coon, H., Leppert, M., Risch, N., and Drayna, D. (2003). Positional cloning of the human quantitative trait locus underlying taste sensitivity to phenylthiocarbamide. *Science* 299, 1221–1225. doi:10.1126/science.1080190
- Kim, U., Wooding, S., Ricci, D., Jorde, L. B., and Drayna, D. (2005). Worldwide haplotype diversity and coding sequence variation at human bitter taste receptor loci. *Hum. Mutat.* 26, 199–204. doi:10.1002/humu.20203
- Kumar, P., Henikoff, S., and Ng, P. C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.* 4, 1073–1081. doi:10.1038/nprot.2009.86
- Lee, R. J., Xiong, G., Kofonow, J. M., Chen, B., Lysenko, A., Jiang, P., et al. (2012). *T2R38* taste receptor polymorphisms underlie susceptibility to upper respiratory infection. *J. Clin. Invest.* 122, 4145–4159. doi:10.1172/JCI64240
- Li, H. (2011). Tabix: Fast retrieval of sequence features from generic tab-delimited files. *Bioinformatics* 27, 718–719. doi:10.1093/bioinformatics/btq671
- Luo, X. C., Chen, Z. H., Xue, J. B., Zhao, D. X., Lu, C., Li, Y. H., et al. (2019). Infection by the parasitic helminth *Trichinella spiralis* activates a *Tas2r*-mediated signaling pathway in intestinal tuft cells. *Proc. Natl. Acad. Sci. U. S. A.* 116, 5564–5569. doi:10.1073/pnas.1812901116
- Margulis, E., Dagan-Wiener, A., Ives, R. S., Jaffari, S., Siems, K., and Niv, M. Y. (2021). Intense bitterness of molecules: Machine learning for expediting drug discovery. *Comput. Struct. Biotechnol. J.* 19, 568–576. doi:10.1016/j.csbj.2020.12.030
- Martens, K., Steelant, B., and Bullens, D. M. A. (2021). Taste receptors: The gatekeepers of the airway epithelium. *Cells* 10.
- May-Wilson, S., Matoba, N., Wade, K. H., Hottenga, J. J., Concas, M. P., Mangino, M., et al. (2022). Large-scale GWAS of food liking reveals genetic determinants and genetic correlations with distinct neurophysiological traits. *Nat. Commun.* 13, 2743. doi:10.1038/s41467-022-30187-w



- Mclaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R., Thormann, A., et al. (2016). The Ensembl variant effect predictor. *Genome Biol.* 17, 122. doi:10.1186/s13059-016-0974-4
- Meyerhof, W., Batram, C., Kuhn, C., Brockhoff, A., Chudoba, E., Bufo, B., et al. (2010). The molecular receptive ranges of human TAS2R bitter taste receptors. *Chem. Senses* 35, 157–170. doi:10.1093/chemse/bjp092
- Nissim, I., Dagan-Wiener, A., and Niv, M. Y. (2017). The taste of toxicity: A quantitative analysis of bitter and toxic molecules. *IUBMB Life* 69, 938–946. doi:10.1002/iub.1694
- Nolden, A. A., Mcgeary, J. E., and Hayes, J. E. (2016). Differential bitterness in capsaicin, piperine, and ethanol associates with polymorphisms in multiple bitter taste receptor genes. *Physiol. Behav.* 156, 117–127. doi:10.1016/j.physbeh.2016.01.017
- Ortega, F. J., Aguera, Z., Sabater, M., Moreno-Navarrete, J. M., Alonso-Ledesma, I., Xifra, G., et al. (2016). Genetic variations of the bitter taste receptor TAS2R38 are associated with obesity and impact on single immune traits. *Mol. Nutr. Food Res.* 60, 1673–1683. doi:10.1002/mnfr.201500804
- Paradis, E. (2010). *pegas*: an R package for population genetics with an integrated-modular approach. *Bioinformatics* 26, 419–420. doi:10.1093/bioinformatics/btp696
- Pfeifer, B., Wittelsburger, U., Ramos-Onsins, S. E., and Lercher, M. J. (2014). PopGenome: An efficient Swiss army knife for population genomic analyses in R. *Mol. Biol. Evol.* 31, 1929–1936. doi:10.1093/molbev/msu136
- Pronin, A. N., Xu, H., Tang, H., Zhang, L., Li, Q., and Li, X. (2007). Specific alleles of bitter receptor genes influence human sensitivity to the bitterness of aloin and saccharin. *Curr. Biol.* 17, 1403–1408. doi:10.1016/j.cub.2007.07.046
- Reed, D. R., and Knaapila, A. (2010). Genetics of taste and smell: Poisons and pleasures. *Prog. Mol. Biol. Transl. Sci.* 94, 213–240. doi:10.1016/B978-0-12-375003-7.00008-X
- Roudnitsky, N., Behrens, M., Engel, A., Kohl, S., Thalmann, S., Hubner, S., et al. (2015). Receptor polymorphism and genomic structure interact to shape bitter taste perception. *PLoS Genet.* 11, e1005530. doi:10.1371/journal.pgen.1005530
- Roudnitsky, N., Bufo, B., Thalmann, S., Kuhn, C., Gunn, H. C., Xing, C., et al. (2011). Genomic, genetic, and functional dissection of bitter taste responses to artificial sweeteners. *Hum. Mol. Genet.* 20, 3437–3449. doi:10.1093/hmg/ddr252
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7, 539. doi:10.1038/msb.2011.75
- Soranzo, N., Bufo, B., Sabeti, P. C., Wilson, J. F., Weale, M. E., Marguerie, R., et al. (2005). Positive selection on a high-sensitivity allele of the human bitter-taste receptor TAS2R16. *Curr. Biol.* 15, 1257–1265. doi:10.1016/j.cub.2005.06.042
- Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123, 585–595. doi:10.1093/genetics/123.3.585
- Tepper, B. J., White, E. A., Koelliker, Y., Lanzara, C., D'adamo, P., and Gasparini, P. (2009). Genetic variation in taste sensitivity to 6-n-propylthiouracil and its relationship to taste perception and food selection. *Ann. N. Y. Acad. Sci.* 1170, 126–139. doi:10.1111/j.1749-6632.2009.03916.x
- The 1000 Genomes Project Consortium (2015). A global reference for human genetic variation. *Nature* 526, 68–74.
- Tran, H. T. T., Stetter, R., Herz, C., Spottel, J., Krell, M., Hanschen, F. S., et al. (2021). Allyl isothiocyanate: A TAS2R38 receptor-dependent immune modulator at the interface between personalized medicine and nutrition. *Front. Immunol.* 12, 669005. doi:10.3389/fimmu.2021.669005
- Wang, X., Thomas, S. D., and Zhang, J. (2004). Relaxation of selective constraint and loss of function in the evolution of human bitter taste receptor genes. *Hum. Mol. Genet.* 13, 2671–2678. doi:10.1093/hmg/ddh289
- Weir, B. S., and Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution* 38, 1358–1370. doi:10.1111/j.1558-5646.1984.tb05657.x
- Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. New York: Springer-Verlag.
- Williamson, S. H., Hernandez, R., Fledel-Alon, A., Zhu, L., Nielsen, R., and Bustamante, C. D. (2005). Simultaneous inference of selection and population growth from patterns of variation in the human genome. *Proc. Natl. Acad. Sci. U. S. A.* 102, 7882–7887. doi:10.1073/pnas.0502300102
- Wooding, S., Gunn, H., Ramos, P., Thalmann, S., and Meyerhof, W. (2010). Genetics and bitter taste responses to goitrin, a plant toxin found in vegetables. *Chem. Senses* 35, 685–692. doi:10.1093/chemse/bjq061
- Wooding, S., Kim, U. K., Bamshad, M. J., Larsen, J., Jorde, L. B., and Drayna, D. (2004). Natural selection and molecular evolution in *PTC*, a bitter-taste receptor gene. *Am. J. Hum. Genet.* 74, 637–646. doi:10.1086/383092
- Wooding, S. P., Atanasova, S., Gunn, H. C., Staneva, R., Dimova, I., and Toncheva, D. (2012). Association of a bitter taste receptor mutation with Balkan Endemic Nephropathy (BEN). *BMC Med. Genet.* 13, 96–97. doi:10.1186/1471-2350-13-96
- Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., and Weir, B. S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* 28, 3326–3328. doi:10.1093/bioinformatics/bts606