2-1-2022

A Distinctive Pattern of Diversity for the TAS2R38 Gene in North Africa

Soufia Mourali-Chebil
University of Tunis El Manar

Sarra Elkamel
University of Tunis El Manar

Sami Boussetta
University of Tunis El Manar

Andrew J. Pakstis
Yale University

Kenneth K. Kidd
Yale University

See next page for additional authors

A Distinctive Pattern of Diversity for the *TAS2R38* Gene in North Africa

Soufia Mourali-Chebil,[1]* Sarra Elkamel,[1] Sami Boussetta,[1] Andrew J. Pakstis,[2] Kenneth K. Kidd,[2] Amel Benammar-Elgaaied,[1] and Lotfi Cherni[1,3]

[1]Laboratory of Genetics, Immunology, and Human Pathologies, Faculty of Sciences of Tunis, University of Tunis El Manar, Tunis, Tunisia.

[2]Department of Genetics, School of Medicine, Yale University, New Haven, Connecticut, USA.

[3]High Institute of Biotechnology, University of Monastir, Monastir, Tunisia.

*Correspondence to: Soufia Mourali-Chebil, Laboratory of Genetics, Immunology, and Human Pathologies, Faculty of Sciences of Tunis, University of Tunis El Manar, Tunis 2092, Tunisia. E-mail: mcsoufia@gmail.com.

Short Title: Diversity of *TAS2R38* in North Africa

KEY WORDS: *TAS2R38*, BITTER TASTE, AUTOSOMAL SNP, DIVERSITY, NORTH AFRICA.

Abstract

The TAS2R38 gene is involved in bitter taste perception. This study documents the distinctive diversity patterns in Northern Africa of functional SNPs rs713598 and rs1726866 at the TAS2R38 locus and places those patterns in the context of global TAS2R38 diversity. We analyzed data previously genotyped with Taqman Applied Biosystem for rs713598 and rs1726866 for 375 unrelated subjects (305 Tunisians from seven locations: Mahdia, Sousse, Kesra, Nebeur, Kairouan, Smar and Kerkennah plus 70 Libyans). Data were analyzed to present haplotypes and genotypes and were compared to the data from worldwide populations. We provide information about TAS2R38 diversity in a part of the world that is relatively under-studied. Considering respectively the two SNPs rs713598 and rs1726866, the (C-A) nucleotide haplotype leading to PV amino acid haplotype is extremely rare almost everywhere, but it is relatively frequent (between 6% and 10%) in Northern Africa where it does coexist with the globally common haplotypes (PA, AA and AV). Given its higher frequency in Northern Africa, we propose the (C-A) haplotype as a biogeographic marker for forensic purposes.

The present study focuses on the SNP variation of the *TAS2R38* (TASTE 2 Receptor member 38) locus in Northern Africa. This locus has 1002 base pairs coding for bitter taste receptors binding the thiourea "–N–C=S" group (Drayna et al. 2003, Kim *et al.* 2003, Bufe et al. 2005) present in compounds such as PTC (Phenylthiocarbamide) and the less toxic PROP (6-n-propylthio-uracil) often used to test the sensitivity to bitterness leading to the classification of subjects into Tasters, Non-Tasters and Super-Tasters (Fox 1932, Bartoshuk et al.1994).

A GWAS (Genome Wide Association Study) conducted by Genick et al. (2011), suggests that *TAS2R38* may be the specific gene responsible for the variation underlying PROP taste sensitivity. *TAS2R38* is a member of the 43 human TAS2R genes involved in bitter tasting and was mapped by Gross (2011) to chromosome 7q34.

Go et al. 2005 have shown that the genes implicated in bitter tasting had duplicated in non-human primates and more intensively in humans, leading to pseudo-genes and to a gradual loss of the function of tasting uncommon bitter substances. Lalueza Fox et al. (2009) have sequenced the *TAS2R38* gene in a male Neanderthal individual (labelled SD 1253) from Spain (El Sidron) and found he was a heterozygote, implying that diversification at this locus was prior to the divergence between modern humans and Neanderthals. De Sousa Cardoso e Valente dos Santos (2015) studied the differences in *TAS2R38* genotype frequencies between different lifestyles (herders, farmers and hunters-gatherers) and found that there were significant differences in Eurasians but not in Africans; while Sjöstrand et al. (2019) found, between hunter-gatherers and farmers in Africa, significant differences in taste sensitivity to quinine as well as a great divergence of genes involved in tongue morphogenesis and in taste signal transduction.

Dozens of SNPs are known in *TAS2R38*, including SNP polymorphisms altering bitter taste sensitivity to isothiocyanates (see *1000 Genomes project* website), but essentially three non-synonymous mutations affecting the receptor's activity were described on *TAS2R38*

(Behrens et al. 2013, Robino et al. 2014). These SNPs (rs713598, rs1726866 and rs10246939) are separated by a short physical distance and form haplotypic combinations with strong linkage disequilibrium (LD) as was confirmed by *1000 Genomes* analyses on *TAS2R38* (R^2= 0.80−0.9, D′ = 0.997−0.998) in worldwide populations (Risso et al. 2016). In the case of our genetically diverse Northern African populations, LD between rs713598 and rs1726866 is very low.

The first SNP, rs713598 lying in position 145, either encodes for proline "P" or alanine "A" amino acid in position 49 on the coded receptor. According to Drayna (2005), a "P" in position 49 always leads to a Taster phenotype. The second SNP (rs1726866) lying in position 785 either encodes for alanine "A" or valine "V" amino acid in position 262 at the protein level; whereas the third (rs10246939) in position 886 is either associated to isoleucine "I" or valine "V" in position 296. The meta-analysis conducted by Sjöstrand et al. (2019) showed that the three SNPs of *TAS2R38* were all significantly associated to PROP perception. It also found that PAV haplotype carriers were more often PROP Super Tasters than non-PAV. According to Genick et al. (2011), a fourth SNP (rs4726481) located very close to *TAS2R38* is also involved in PROP taste sensitivity. However, individual SNPs are less strongly associated to phenotypes than are haplotypes (Kim et al. 2003). The PAV haplotype, associated with the Taster phenotype, occurs at common frequencies >40% in most populations studied and is considered the ancestral one as it is found in the great apes (Kim et al. 2003; Wooding et al. 2004). The AVI haplotype, which is associated with the NonTaster phenotype, is also common, whereas PAI and PVI haplotypes occur at lower frequencies and PVV is almost never encountered (Kim & Drayna 2005, Kim et al. 2005, Risso et al. 2016). As for AAI, although it is rare globally, it is present with a moderate frequency in Africa (Risso et al. 2016, Sjöstrand et al. 2019) and apparently it is not

associated with PTC thresholds (Behrens et al. 2013). AAV is found at low frequencies in Africans (Campbell et al. 2011) and in Tajiks (Sjöstrand et al. 2019).

The distribution of *TAS2R38* phenotypes varies according to gender and age, among and within groups of populations (Harris & Kalmus 1949, Guo & Reed 2001).

Selection studies of the distribution of genes of bitter taste receptors have detected signals of recent selection (Li et al. 2011, Sjöstrand et al. 2019) while other studies detected ancient selective forces acting before the "Out of Africa" event (Risso et al. 2016). Campbell et al. 2012 stated that ancient balancing selection likely affected patterns of variation at the TAS2R38 gene in global populations and that recent selection pressures may also have acted in Africa raising the frequencies of some rare variants. Another hypothesis claims that AV genotypes (who are normally non-Tasters) are possibly able to detect an unknown bitter substance (Wooding S. et al. 2004). Other authors (Rozengurt and Sternini 2007, Dotson et al. 2008, Shah et al. 2009, Deshpande et al. 2010, Lee et al. 2012, Gil et al. 2015) report pleiotropic effects of *TAS2R38* alleles influencing various physiological functions which then makes other selective factors than diet, good candidates for the explanation of the genetic diversity at this locus. Indeed, the pleiotropic aspects of *TAS2R38* manifested by physiological and immune functions (Tran et al. 2021) may interact with pathogens to produce selective effects. These functions could have an antagonistic effect as in balancing selection or, on the contrary, a combined action as in directional selection.

Given the documented (Duffy et al. 2004, Mangold et al. 2008, Dioszegi et al. 2019) relationship between bitter taste perception and dietary behavior and the implication of the latter in public health issues, there is an increasing interest in the *TAS2R38* polymorphism studies. So, it is increasingly important to understand the underlying patterns of food preferences and consequently of food intake which may lead to an increased morbidity in human populations. The multiplication of genotype-phenotype association studies is

potentially helpful in predicting phenotypes related to dietary behavior such as the *TAS2R38* polymorphisms. *TAS2R38* bitter taste receptors reportedly can act either as risk or as protective factors in some digestive cancers (Carrai et al. 2011, Yamaki et al. 2017) as well as contribute to prevention of dental caries (Wendell et al.2010). Taste receptors are not only expressed on the tongue but also by ciliated epithelial cells in airways (Shah et al.2009, Lee et al. 2012). It has also recently been reported that *TAS2R38* polymorphisms play a role in the severity of cystic fibrosis (Castaldo et al. 2020) and that *TAS2R38* is also expressed in adipocyte differentiation and leads to lipid accumulation in obese subjects (Cancello et al. 2020).

In the context of this short review (for more details see Pasquet & Hladick 1999, Diószegi et al. 2019), our purpose in the present study is to analyze the pattern of diversity of *TAS2R38* gene in the North African populations and discuss the impact of ancient admixture as well as selection, in the specific profile observed in such populations. Indeed, North African populations we have studied show a higher average heterozygosity than most of the world populations (Cherni et al. 2016) studied to date. Our sampled Northern African populations are also Mediterranean populations and are located in a geographic area of ancient human interactions within successive invading civilizations (Phoenician, Roman, Byzantine, Arab, Andalusians, Ottomans, French and South Europeans) over a Berber substrate (Despois 1964) deriving from the Neolithic Capsians and the Paleolithic IberoMaurusians. The present work was also motivated by the scarceness of autosomal SNPs variation studies reported for Northern Africa. Identifying the *TAS2R38* diversity in these Northern African populations and positioning those patterns of diversity in a global context figure among the first required steps to achieve this purpose.

Materials and Methods

Two of the three functional SNPs (rs713598 and rs1726866) at the *TAS2R38* gene have genotypes available in our population samples listed in the Supplementary Table S1. These SNPs have been genotyped by Taqman Applied Biosystem (Cherni *et al.* 2016) on unrelated individuals: 305 Tunisians coming from 7 populations spread from North West to South East (Nebeur: 12, Kesra: 42, Kairouan: 46, Sousse: 49, Mahdia: 47, Kerkennah: 47 and Smar: 62) as well as 70 Libyans coming from 6 Northern Libyan urban areas (Ghdames, Zwara, Tripoli, Masrata, Sabrata and Benghazi). The geographical locations of the North African populations are shown in Figure 1. The samples are comprised of a majority of men, since they have been used in another study involving the Y chromosome, without distinction between age cohorts.

Data analysis, relying on Phase.2.1.1 software (Stephens *et al.* 2001; Stephens & Scheet, 2005) which is based on Bayesian statistics to resolve the haplotypic phase on the chromosome, allowed us to impute haplotypes in all these populations for which genetic structures were studied using ARLEQUIN (ver 3.5.2.2) software (Excoffier & Lischer 2010). The haplotype and genotype distributions in the studied samples reveal the expected heterozygosity. The Hardy-Weinberg equilibrium was tested for each SNP separately in each population sample relying on an exact test using a Markov chain with forecasted chain length: 1000000, de-memorization steps: 100000 and steps done: 1001000. (Levene 1949, Guo & Thompson 1992). An exact test of differentiation has also been carried out with ARLEQUIN (ver 3.5.2.2) basing on a Markov chain length of 100000 steps (Raymond & Rousset 1995).

The last step of our study was to compare the results with worldwide populations (Supplementary Table S1). We compared the allele frequencies with population samples studied at the Kidd Lab. as well as from the 1000 Genomes Consortium (https://www.internationalgenome.org/), presenting the comparison in a bar plot image.

Results

1. The Following Results Were Obtained Using ARLEQUIN (ver 3.5.2.2) Software

Allele Frequencies. Figure 2 shows the SNP allele frequency profiles in our Northern African samples as well as in populations from around the world.

Expected Heterozygosity. The mean expected heterozygosity for our Tunisian and Libyan populations was around 40% for rs713598 and close to 50% for rs1726866 (Table 1). The mean expected heterozygosity for both loci was similar for our populations (~0.45).

Hardy-Weinberg Equilibrium. For rs1726866: all populations were in Hardy-Weinberg equilibrium. For rs713598: all populations were in Hardy-Weinberg equilibrium except Nebeur (Supplementary Table S3). However, the nominal deviation for the Nebeur population is not meaningful (p-value = 0.030 for a 5% threshold). Given the total number of Hardy-Weinberg tests done, the one result below the 5% level is not very different from what one would expect to occur by chance.

Exact Test of Differentiation. Supplementary Table S4 shows the result of the exact test of differentiation between populations., The population samples considered by pairs (Table S4), show two populations (Mahdia and Nebeur) significantly differentiated (Non differentiation: exact P value = 0.0422±0.0049) from each other but not to the other six. The tests have been applied to all of the 28 possible pairwise comparisons. We performed a Bonferroni correction to account for the multiple testing. The adjusted α was equal to 0.05/28 = 0.00179. Then, since the P value of 0.0422±0.0049 is greater than 0.00179, we are not rejecting the null hypothesis of Non differentiation between Mahdia and Nebeur.

2. The Following Results Were Obtained Using Phase.2.1.1

Relying on Phase2.1.1 software, we reconstructed the genotypes for the 2 SNPs of *TAS2R38* in our Northern African samples and we present the frequencies for each of the four possible haplotypes in Table 2. This table also shows the correspondence between SNP haplotypes and amino acids at the receptor level: the CG nucleotide haplotype corresponds to the PA amino acid haplotype and the GA nucleotide haplotype corresponds to the AV amino acid haplotype. Similarly, GG corresponds to AA and finally CA corresponds to PV. We will use the amino acid symbols for the alleles and haplotypes. The haplotypes are ordered from left to right in Table 2 by most to least frequent worldwide. The most frequent *TAS2R38* amino acid haplotype in the world is PA, while it is not the most frequent in Northern Africa. Indeed, the most frequent *TAS2R38* amino acid haplotype in Northern Africa appears (from Table 2) to be AV. The least frequent *TAS2R38* amino acid haplotype elsewhere in the world (PV) is present in all our Northern African samples at frequencies ranging from 6% to 14%.

We compared our Northern African samples to other studied populations from all over the world. The extent of the variation is displayed by the Figure 3 with the bar plots of the haplotype frequencies showing the populations clustered by geographical region.

The PV haplotype is extremely rare in the worldwide studied populations (Figure 3, Cherni et al. 2016), and was especially observed in Northern Africa and at lower frequencies in some other populations (Mexicans, Southern and Central Asians, Sub-Saharan Africans). In fact, PV haplotype has been observed 82 times (out of 2164 total observations), of which 68 observations have occurred in North African individuals (supplementary data Table S2). Both haplotypes PA and AV are present in every continent, while AA was practically absent in Asia.

We observed great haplotype diversity in all the North African studied populations (Figure 3). All four of the possible haplotypes occur at common frequencies in North Africa.

In other world regions only two or three of the haplotypes typically occur at common frequencies. Indeed, we found a relatively high proportion of AA and AV haplotypes (around respectively 30% and 40%), while PA and PV haplotypes are found in lower proportions. Libya and Tunisia showed almost the same frequency for each haplotype. Similar frequencies are also found in the case of Sousse and Mahdia on the one hand and Kesra and Nebeur on the other. These populations are geographically close to each other (Sousse and Mahdia are distant by 58.6 kms in the flat region of Sahel; Kesra and Nebeur are 90.9 kms apart in the High Tell mountainous region). We also noticed the presence of PV haplotype (mainly in Smar and Kairouan).

Discussion

We find a distinctive pattern of diversity of the *TAS2R38* haplotypes based on SNPs rs713598 and rs1726866 across world regions. All four of the possible *TAS2R38* haplotypes occur in our population samples from locales in Tunisia and Libya at common frequencies >6% (Figure 3) and this North African region is the most genetically diverse of the major world regions studied at this locus. In African populations that have been studied south of the Sahara, we see the next most diverse region such that three of the haplotype alleles (PA, AA and AV) are usually present at common frequencies. Beyond Africa, the populations in the regions for which we have frequencies typically have no more than one or two of the three present and occurring at common frequencies. Especially interesting and distinctive compared to other world regions is the relatively common occurrence in North Africa of the PV haplotype and the relatively low frequency of the ancestral PA haplotype. In a previous report analyzing a few hundred loci spread across all of the human autosomes, Northern Africa showed a distinct and ancient genetic background compared to other world regions

(Cherni *et al.* 2016). The *TAS2R38* locus is clearly one of the especially noteworthy gene regions contributing to North Africa's distinctiveness among the major world regions.

Genetic Structure of Northern African Populations

Non-Differentiation among North African Populations. After applying the Bonferroni correction over the results of the test of differentiation between all pairs of samples it appears that there are no populations significantly different from each other although Cherni et al. (2016) invoked a strong differentiation among localities, in relation with their complex population histories with differential genetic contributions from diverse peoples. For instance, in Mahdia, ancient capital located in Central Eastern Tunisia, Ottomans (who were themselves genetically diversified) have greatly contributed to its initially Berber-Arab genetic pool (Delarosbil 2006); while in the case of Nebeur (small locality in Northern Western Tunisia) the initial Berber genetic pool has only known admixture with Arab tribes (Monchicourt 1913).

Acting Processes. We described the genetic structures of Northern African samples in order to try to understand the reason why the PV haplotype frequency is higher. We have unfortunately not yet the adequate tools to test the effect of evolutionary pressures; thus, we will consider genetic drift as the most likely to raise the frequency of derived alleles over generations. Demographic events (such as bottlenecks or migrations) may have affected the frequency of some unusual haplotypes like PV. In the same way, the well documented ancient admixture in Northern Africa (Cherni et al. 2016) may have generated, given the huge number of generations available since Prehistoric times, new haplotype combinations explaining the presence of regionally different locus frequencies as well as the low Linkage Disequilibrium between non synonymous mutations. Theoretically, selection pressures could

have played a role in raising PV to common frequencies (in the 6 to 15% range) over time but assessing such a possibility must await collection of the appropriate evidence.

Comparisons of the North African Samples with the Rest of the World

As we have shown above, the comparison of our samples to those collected worldwide from the Kidd Lab and from the 1000 Genomes Consortium identifies North Africa as a distinct world region which is characterized by its higher haplotype diversity, especially the presence of the PV haplotype at common frequencies. Certainly, *TAS2R38* haplotype frequencies (combined with other gene region data) would contribute to differentiating Tunisia/Libyan populations from other world regions in an ancestry inference analysis like the STRUCTURE or ADMIXTURE computer programs.

Association with Tolerance to Capsaicin?

The PV haplotype is also found, at very low frequencies (less than 2%), in the Mexican population of Los Angeles as well as in Malaysia, Mongolia, Bengladesh, Pakistan and in Telugu speakers (from Southern Central Asia) as well as in Afro-Americans mostly of Western African ancestry (Figure 3). Interestingly, in all these populations, as well as in Tunisia, traditional meals are often very "spicy" and based on the consumption of hot chili pepper, for instance in 2000 the annual mean per habitant in Tunisia was reported by Khaldi et al. (2009) to be of 10.6 kg.

Is the PV haplotype causally related to the tolerance to hot spicy taste? Might the non-zero frequency of PV haplotype have been culturally influenced by the culinary preferences of the populations? If so, then this finding would support the relationship of the variation at *TAS2R38* locus with dietary behavior.

A reasonable speculation to explore in a future work would be that if much larger sample sizes had been collected and studied then more populations might be found to carry PV at rare to very low frequencies. Nonetheless, this exciting possibility should be explored using for example the Scoville scale (Scoville 1912) of spiciness perception (which depends on the content in capsaicin) and trying to detect a correlation between mean rates of tolerated capsaicin in populations and the frequencies of PV. This despite the fact that Nolden et al. (2020) have demonstrated that burn from capsaicin was associated to PROP bitterness but not to *TAS2R38* genotypes. Yet, they also showed that this correlation to PROP bitterness was increased when the subjects were grouped according to their *TAS2R38* genotypes.

Limitations

The third SNP (rs10246939) has not been investigated in the present study because it is only available in the study of Karmous et al. (2018) and over Tunisia as a whole without distinguishing between populations. They found a relationship between fat and bitter taste perceptions playing together a crucial role in obesity and have also shown that both *TAS2R38* SNPs, rs1726866 and rs10246939, are associated with obesity. However, the third SNP (rs10246939) and even the fourth (rs4726481) should be included in a future study to see if the diversity of *TAS2R38* haplotypes in Northern Africa is even more interesting than what we have reported here. What could be the effect on the Northern African pattern of *TAS2R38* diversity? Would it be further amplified? Essentially, this effect would tell us which of PVV or PVI is more frequent in Northern Africa. The strongest inferences will also be possible if taste perception phenotypes (and/or other phenotypic data) and genetic marker data are being assessed in the same study in order to determine if the PV haplotype is correlated with phenotypic variation.

Then sequencing rs10246939, as well as some of the dozens other *TAS2R38* known SNPs, on the same samples of populations may be of great help in disentangling the specific picture of Northern African genetic diversity. Assessing the role of multiple demographic factors properly will require larger sample sizes. Most of our sampled groups are from small communities where we may not be able to enlist enough random, unrelated individuals.

Conclusions

The present paper contributes to our understanding of the genetic relationships of autosomal diversity among North African populations and the extent of their resemblance to populations in other world regions. Overall, the population samples from Libya and Tunisia, are very similar. However, when we consider the populations individually, the specificity of each sample becomes obvious (Figure 2 and 3) although we finally did not find significant differences between our sampled populations. Nonetheless, in expectation of a genome-wide data analysis to better assess population structure, it is worth repeating the conclusion of Cherni et al. (2016) that Tunisians should be sampled according to their locality of origin given the potential for strong differentiation among localities, due in part to the differences in complexity of their known population histories with many contributions of Berber, Arab, Andalusian, Phoenician, etc ancestries.

Perspectives

The relatively high frequency of PV may be explained so far by genetic drift. However, it would be worthwhile testing other hypotheses in future studies. In the case of selection, for example, we need high density SNP data extending 100 to 200 kb sequence information on either side of the *TAS2R38* gene in North African population samples and from populations in other nearby geographical regions. Of course, separate studies looking simultaneously at

various genes affecting bitter-taste perception along with taste and olfactory phenotypes would also be very helpful in evaluating the possible explanations for the *TAS2R38* haplotype differences.

TAS2R38 SNP variation for rs713598 and rs1726866 shows high differentiation in Northern Africa. The (C-A) haplotype coding for PV, given its higher frequency in Northern Africa, would make a good biogeographic marker serving *forensic* purposes as it is, indeed, extremely rare elsewhere; even though (C-A) is not a microhaplotype, since its size is larger than 300bp (Oldoni *et al.* 2018). Then instead of detecting mixture it would be interesting at least in helping to infer ancestries.

It would be interesting to increase the number of sampled Northern African populations in order to better understand the evolutionary process which has been acting on the considered locus in this area. Arabian Peninsula samples should also be included in future studies, in order to disentangle which of the Berber or the Arab genetic contributions was responsible for the higher frequencies of PV.

Acknowledgments

This work was partially supported by the Tunisian Ministry of Higher Education and Scientific Research as well as by the University of Tunis El Manar. Special thanks go to the thousands of individuals around the world who volunteered to give blood and saliva samples to make this study possible.

Received 21 October 2020; accepted for publication 18 September 2021.

Literature Cited

Adhikari, K., J. C. Chacon-Duque, J. Mendoza-Revilla et al. 2017. The genetic diversity of the Americas. *Annu. Rev. Genomics Hum. Genet.* 18:277–296.

Allison, A. C. 1951. A note on taste-blindness in Kenya Africans and Arabs. *Man* 51:119–120.

Bartoshuk, L. M., V. B. Duffy, and I. J. Miller. 1995. PTC/PROP tasting: Anatomy, psychophysics, and sex effects. *Physiol. Behav.* 58:203.

Behrens, M., H. C. Gunn, and P. C. M. Ramos. 2013. Genetic, functional, and phenotypic diversity in *TAS2R38*-mediated bitter taste perception. *Chem. Senses* 38:475–484.

Birlouez, E. 2020. *Petite et Grande Histoire des Légumes.* Versailles, FR: Editions QUAE.

Campbell, M. C., A. Ranciaro, A. Froment et al. 2012. Evolution of functionally diverse alleles associated with PTC bitter taste sensitivity in Africa. *Mol. Biol. Evol.* 29:1,141–1,153.

Cancello, R., G. Micheletto, D. Meta et al. 2020. Expanding the role of bitter taste receptor in extra oral tissues: *TAS2R38* is expressed in human adipocytes. *Adipocyte* 9:7–15.

Carrai, M., V. Steinke, P. Vodicka et al. 2011. Association between TAS2R38 gene polymorphisms and colorectal cancer risk: A case-control study in two independent populations of Caucasian origin. *PloS One* 6:e20464.

Castaldo, A., G. Cernera, P. Iacotucci et al. 2020. TAS2R38 is a novel modifier gene in patients with cystic fibrosis. *Sci. Rep.* 10:5806.

Cherni, L., A. J. Pakstis, S. Boussetta et al. 2016. Genetic variation in Tunisia in the context of human diversity worldwide. *Am. J. Phys. Anthropol.* 161:62–71.

Delarosbil, D. 2006. *Mahdia: Histoire et Société.* Montréal, CA: UNESCO and the University of Montréal.

Deshpande, D. A., W. C. Wang, E. L. McIlmoyle et al. 2010. Bitter taste receptors on airway smooth muscle bronchodilate by localized calcium signaling and reverse obstruction. *Nat. Med.* 16:1,299–1,304.

Diószegi, J., E. Llanaj, and R. Ádány. 2019. Genetic background of taste perception, taste preferences, and its nutritional implications: A systematic review. *Front. Genet.* 10:1272.

Dotson, C. D., L. Zhang, H. Xu et al. 2008. Bitter taste receptors influence glucose homeostasis. *PLoS One* 3:e3974.

Drayna, D. 2005. Human taste genetics. *Annu. Rev. Genomics Hum. Genet*. 6:217–235.

Drayna, D., H. Coon, U. K. Kim et al. 2003. Genetic analysis of a complex trait in the Utah Genetic Reference Project: A major locus for PTC taste ability on chromosome 7q and a secondary locus on chromosome 16p. *Hum. Genet.* 112:567–572.

Duffy, V. B., A. C. Davidson, J. R. Kidd et al. 2004. Bitter receptor gene (*TAS2R38*), 6-n-propylthiouracil (PROP) bitterness and alcohol intake. *Alcohol. Clin. Exp. Res.* 28:1,629–1,637.

Earl, D. A., and B. M. vonHoldt. 2012. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* 4:359–361.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol.* 14:2,611–2,620.

Excoffier, L., and H. E. L. Lischer. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10:564–567.

Fox, A. L. 1932. The relationship between chemical constitution and taste. *Proc. Natl. Acad. Sci. U. S. A.* 18:115–120.

Genick, U. K., Z. Kutalik, M. Ledda et al. 2011. Sensitivity of genome-wide-association signals to phenotyping strategy: The PROP-*TAS2R38* taste association as a benchmark. *PLoS One* 6:e27745.

Gil, S., S. Coldwell, J. L. Drury et al. 2015. Genotype-specific regulation of oral innate immunity by T2R38 taste receptor. *Mol. Immunol.* 68:663–670.

Go, Y., Y. Satta, O. Takenaka et al. 2005. Lineage-specific loss of function of bitter taste receptor genes in humans and nonhuman primates. *Genetics* 170:313–326.

Gross, M. B. 5/12/2011. Personal Communication. Baltimore, MD.

Guo, S. W., and D. R. Reed. 2001. The genetics of phenylthiocarbamide perception. *Ann. Hum. Biol.* 28:111–142.

Guo, S. W., and E. A. Thompson. 1992. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics* 48:361–372.

Harris, H., and H. Kalmus. 1949. The measurement of taste sensitivity to phenylthiourea. *Ann. Eugen.* 15:24–31.

Karmous, I., J. Plesník, A. S. Khan et al. 2018. Orosensory detection of bitter in fat-taster healthy and obese participants: Genetic polymorphism of CD36 and *TAS2R38*. *Clin. Nutr*. 37:313–320.

Karrer, T., and L. Bartoshuk. 1991. Capsaicin desensitization and recovery on the human tongue. *Physiol. Behav.* 49:757–764.

Khaldi, R., B. Dhehibi, Z. Rached et al. 2009. La consommation de la pomme de terre et des légumes frais en Tunisie: Un diagnostic contrasté et des effets prix et dépenses alimentaires différenciés par milieu d'habitat. *New Medit* 8:14–22.

Kim, U. K., and D. Drayna. 2005. Genetics of individuals differences in bitter taste perception: Lessons from the PTC gene. *Clin. Genet.* 67:275–280.

Kim, U. K., E. Jorgenson, H. Coon et al. 2003. Positional cloning of the human quantitative trait locus underlying taste sensitivity to phenylthiocarbamide. *Science* 299:1,221–1,225.

Kim, U. K., S. Wooding, D. Ricci et al. 2005. Worldwide haplotype diversity and coding sequence variation at human bitter taste receptor loci. *Hum. Mutat.* 26:199–204.

Lalueza-Fox, C., E. Gigli, M. de la Rasilla et al. 2009. Bitter taste perception in Neanderthals through the analysis of the *TAS2R38* gene. *Biol. Lett.* 5:809–811.

Lee, R. J., G. Xiong, J. M. Kofonow et al. 2012. T2R38 taste receptor polymorphisms underlie susceptibility to upper respiratory infection. *J. Clin. Invest.* 122:4,145–4,159.

Levene, H. 1949. A new measure of sexual isolation. *Evolution* 3:315–321.

Li, H., A. J. Pakstis, J. R. Kidd et al. 2011. Selection on the human bitter taste gene, TAS2R16, in Eurasian populations. *Hum. Biol.* 83:363–377.

Mangold, J. E., T. J. Payne, J. Z. Ma et al. 2008. Bitter taste receptor gene polymorphisms are an important factor in the development of nicotine dependence in African Americans. *J. Med. Genet.* 45:578–582.

Monchicourt, C. 1913. La région du haut tell en Tunisie (le Kef, Teboursouk, Mactar, Thala): Essai de monographie géographique. PhD diss., University of Paris.

Nolden, A., J. E. McGeary, and J. E. Hayes. 2020. Predominant qualities evoked by quinine, sucrose, and capsaicin associate with PROP bitterness, but not TAS2R38 genotype. *Chem. Senses* 45:383–390.

Oldoni, F., K. K. Kidd, and D. Podini. 2019. Microhaplotypes in forensic genetics. *Forensic Sci. Int. Genet.* 38:54–69.

Pasquet, P., and C. M. Hladik. 1999. Évolution des comportements alimentaires: Adaptations morphologiques et sensorielles. *Bull. Mém. Soc. Anthropol. Paris* 11:307–332.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945–959.

Raymond, M., and F. Rousset. 1995. An exact for population differentiation. *Evolution* 49:1,280–1,283.

Reddy, F. E., L. M. Bartoshuk, I. J. Miller et al. 1993. Relationship among papillae, taste pores and 6-n-propylthiouracile (PROP) suprathreshold taste sensitivity. *Chem. Senses* 18:618–619.

Risso, D. S., M. Mezzavilla, L. Pagani et al. 2016. Global diversity in the *TAS2R38* bitter taste receptor: Revisiting a classic evolutionary proposal. *Sci. Rep.* 6:25506.

Robino, A., M. Mezzavilla, N. Pirastu et al. 2014. A population-based approach to study the impact of PROP perception on food liking in populations along the Silk Road. *PLoS One* 9:e91716.

Rozengurt, E., and C. Sternini. 2007. Taste receptor signaling in the mammalian gut. *Curr. Opin. Pharmacol*. 7:557–562.

Scoville, W. 1912. Note on Capsicums. *J. Am. Pharm. Assoc.* 1:453–454.

Shah, A. S., Y. Ben-Shahar, T. O. Moninger et al. 2009. Motile cilia of human airway epithelia are chemosensory. *Science* 325:1,131–1,134.

Sjöstrand, A. E., P. Sjödin, T. Hegay et al. 2019. Taste perception and lifestyle: Insights from phenotype and genome data among Africans and Asians. *Eur. J. Hum. Genet.* 29:325–337.

Stephens, M., and P. Scheet. 2005. Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. *Am. J. Hum. Genet.* 76:449–462.

Stephens, M., N. J. Smith, and P. Donnelly. 2001. A new statistical method for haplotype reconstruction from population data. *Am. J. Hum. Genet.* 68:978–989.

Tepper, B. J., E. A. White, Y. Koelliker et al. 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.

Tran, H. T. T., R. Stetter, C. Herz et al. 2021. Allyl isothiocyanate: A TAS2R38 receptor-dependent immune modulator at the interface between personalized medicine and nutrition. *Front. Immunol.* 12:669005.

V dos Santos, C. S. 2015. Lifestyle and genetic diversity: A study on African populations. PhD diss., University of Porto.

Wang, J. C., A. L. Hinrichs, S. Bertelsen et al. 2007. Functional variants in *TAS2R38* and *TAS2R16* influence alcohol consumption in high-risk families of African-American origin. *Alcohol. Clin. Exp. Res.* 31:209–215.

Wendell, S., X. Wang, M. Brown et al. 2010. Taste genes associated with dental caries. *J. Dent. Res.* 89:1,198–1,202

Wooding, S. 2006. Phenylthiocarbamide: A 75-year adventure in genetics and natural selection. *Genetics* 172:2015.

Wooding, S., U. Kim, M. J. Bamshad et al. 2004. Natural selection and molecular evolution in PTC, a bitter-taste receptor gene. *Am. J. Hum. Genet.* 74:637–646.

Yamaki, M., H. Saito, K. Isono et al. 2017. Genotyping analysis of bitter-taste receptor genes TAS2R38 and TAS2R46 in Japanese patients with gastrointestinal cancers. *J. Nutr. Sci. Vitaminol. (Tokyo)* 63:148–154.

Table 1. Expected Heterozygosity of rs713598 and rs1726866 in North African Samples

Locus	Libya	Mahdia	Kerkennah	Smar	Kairouan	Sousse	Kesra	Nebeur	Mean	s.d.
rs713598	0.449	0.379	0.436	0.460	0.433	0.378	0.425	0.344	0.411	0.041
rs1726866	0.503	0.496	0.502	0.502	0.505	0.499	0.503	0.518	0.504	0.006
Mean	0.476	0.496	0.436	0.481	0.469	0.439	0.453	0.431	0.460	0.024
s.d.	0.194	0.000	0.000	0.172	0.225	0.292	0.267	0.351	0.187	0.128

(s.d. for standard deviation)

Table 2. Frequencies of Amino Acid Haplotypes in North African Samples

Haplotypes ordered left to right by most to least frequent worldwide.

Haplotypes (alleles)		**CG**	**GA**	**GG**	**CA**
Amino acid labeling for haplotypes		**PA**	**AV**	**AA**	**PV**
Sample Size	**2N**				
Nebeur	**24**	0.125	0.375	0.417	0.083
Kerkennah	**94**	0.234	0.447	0.245	0.074
Kairouan	**92**	0.228	0.424	0.261	0.087
Kesra	**84**	0.190	0.381	0.345	0.083
Sousse	**98**	0.204	0.490	0.245	0.061
Mahdia	**94**	0.202	0.521	0.213	0.064
Smar	**124**	0.210	0.387	0.258	0.145
Lybia	**140**	0.250	0.407	0.243	0.100

Amino acid haplotypes (49-785): (PA: Proline-Alanine; AV: Alanine-Valine; AA: Alanine-Alanine; PV: Proline-Valine)

Supplementary Table S1. List of 113 Population Samples Listed below Including the 26 Populations from 1000 Genomes

Abbreviations for two of the 1000 Genomes groups adjusted here (Finns=FN1 instead of FIN, Southern Han Chinese=HCS instead of CHS) in order to avoid conflict with pre-existing Kidd Lab pop abbreviations.

OutofAfr Pop Rank #	World Region assignment	Population sample Description	Abbrev	Location,Country where individuals were collected	Project Source 1000 Genomes or Kidd Lab	Comments/Notes
1	Africa, Central	Biaka	BIA	Central Afr Rep.	Kidd Lab	
2	Africa, Central	Mbuti	MBU	Dem.Rep.Congo	Kidd Lab	
3	Africa, Central	Lisongo	LIS	Central Afr Rep.	Kidd Lab	
4	Africa, West	Gambians	GWD	Gambia	1000 Genomes	
5	Africa, West	Mende, Sierra Leone	MSL	Sierra Leone	1000 Genomes	
6	Africa, West	Ibo	IBO	Nigeria	Kidd Lab	
7	Africa, West	Hausa	HSA	Nigeria	Kidd Lab	
8	Africa, West	Esan	ESN	Nigeria	1000 Genomes	
9	Africa, West	Yoruba, Ibadan	YRI	Nigeria	1000 Genomes	
10	Africa, West	Yoruba, Benin City	YOR	Nigeria	Kidd Lab	
11	Africa, East	Chagga	CGA	Tanzania	Kidd Lab	
12	Africa, East	Sandawe	SND	Tanzania	Kidd Lab	
13	Africa, East	Masai	MAS	Tanzania	Kidd Lab	
14	Africa, East	Zaramo	ZRM	Tanzania	Kidd Lab	
15	Africa, East	Luhya, Kenya	LWK	Kenya	1000 Genomes	
16	Africa, East	Ethiopian Jews	ETJ	Israel	Kidd Lab	
17	Africa, East	Somali	SOM	Pakistan	Kidd Lab	
18	Admixed: W.Afr./Europe/Amer	AfrAmerSoWest	ASW	USA	1000 Genomes	Mostly W.Afr. ancestry
19	Admixed: W.Afr./Europe/Amer	African Americans	AAM	USA	Kidd Lab	Mostly W.Afr. ancestry
20	Admixed: W.Afr./Europe/Amer	Afro-Caribbeans, Barbados	ACB	Barbados,Caribbean island	1000 Genomes	Mostly W.Afr. ancestry
21	Admixed: E.Afr/S.C.Asia	Makrani	MKR	Pakistan	Kidd Lab	E.Afr. and SoCenAsian ancestry
22	Africa, North	**Nebeur**	NBR	Tunisia	Kidd Lab	
23	Africa, North	**Kerkennah**	KRK	Tunisia	Kidd Lab	
24	Africa, North	**Kairouan**	KRN	Tunisia	Kidd Lab	
25	Africa, North	**Kesra**	KSR	Tunisia	Kidd Lab	
26	Africa, North	**Sousse**	SOU	Tunisia	Kidd Lab	
27	Africa, North	**Mahdia**	MHD	Tunisia	Kidd Lab	
28	Africa, North	**Smar**	SMR	Tunisia	Kidd Lab	
29	Africa, North	**Tunisians Southern**	TNS	Tunisia	Kidd Lab	
30	Africa, North	**Libyans**	LYB	Libya	Kidd Lab	
31	SouthWest Asia	Yemenite Jews	YMJ	Israel	Kidd Lab	
32	SouthWest Asia	Kuwaiti	KWT	Kuwait	Kidd Lab	
33	SouthWest Asia	Saudi	SAU	Saudi Arabia	Kidd Lab	
34	SouthWest Asia	Druze	DRU	Israel	Kidd Lab	
35	SouthWest Asia	Palestinian Arabs	PLA	Israel	Kidd Lab	

36	SouthWest Asia	Samaritans	SAM	Israel	Kidd Lab	
37	SouthWest Asia/Europe	Ashkenazi Jews	ASH	Israel	Kidd Lab	
38	SouthWest Asia	Chaldeans	CHL	N. Iraq	Kidd Lab	Collected by: S.Dogan, M.Dogan, C.Gurkan
39	SouthWest Asia	Kurds	KRD	N. Iraq	Kidd Lab	Collected by: S.Dogan, M.Dogan, C.Gurkan
40	SouthWest Asia	Arabs_N_Iraq	NIA	N. Iraq	Kidd Lab	Collected by: S.Dogan, M.Dogan, C.Gurkan
41	SouthWest Asia	Shabaks	SHB	N. Iraq	Kidd Lab	Collected by: S.Dogan, M.Dogan, C.Gurkan
42	SouthWest Asia	Syriacs	SYR	N. Iraq	Kidd Lab	Collected by: S.Dogan, M.Dogan, C.Gurkan
43	SouthWest Asia	Turkmen	TKM	N. Iraq	Kidd Lab	Collected by: S.Dogan, M.Dogan, C.Gurkan
44	SouthWest Asia	Yazidis	YZD	N. Iraq	Kidd Lab	Collected by: S.Dogan, M.Dogan, C.Gurkan
45	SouthWest Asia	Turkish	TRK	Istanbul, Turkey	Kidd Lab	
46	SouthWest Asia	Iranians	IRN	Iran	Kidd Lab	
47	Europe, South	Adygei	ADY	Caucasus Mtns, Russia	Kidd Lab	
48	Europe, South	Greek_Cypriot	GCP	Greek Cyprus	Kidd Lab	Collected by: Pavlos Neophytou
49	Europe, South	Turkish_Cypriots	TCP	Turkish Cyprus	Kidd Lab	Collected by: Ozlem Bulbul
50	Europe, South	Greeks	GRK	Greece	Kidd Lab	
51	Europe, South	Roman Jews	RMJ	Italy	Kidd Lab	
52	Europe, South	Sardinians	SRD	Italy	Kidd Lab	
53	Europe, South	Iberians,Spain	IBS	Spain	1000 Genomes	
54	Europe, South	Toscani	TSI	Italy	1000 Genomes	
55	Europe, Mixed	EuroAmericans	EAM	USA	Kidd Lab	
56	Europe, North	Hungarians	HGR	Hungary	Kidd Lab	
57	Europe, North	N. & W. European ancestry	CEU	Utah, USA	1000 Genomes	
58	Europe, North	Danes	DAN	Denmark	Kidd Lab	
59	Europe, North	Irish	IRI	Ireland	Kidd Lab	
60	Europe, North	British from England,Scotland	GBR	United Kingdom	1000 Genomes	
61	Europe, North	Finns	FIN	Finland	Kidd Lab	
62	Europe, North	Finns	FN1	Finland	1000 Genomes	
63	Europe, North	Russians, Archangelsk	RUA	Russia	Kidd Lab	
64	Europe, North	Russians, Vologda	RUV	Russia	Kidd Lab	
65	Europe, North	Chuvash	CHV	Russia	Kidd Lab	
66	SouthCentral Asia	Mohanna	MHN	Pakistan	Kidd Lab	
67	SouthCentral Asia	Punjabi, Lahore	PJL	Lahore,Pakistan	1000 Genomes	
68	SouthCentral Asia	Gujarati	GIH	Houston, USA	1000 Genomes	
69	SouthCentral Asia	Telugu	ITU	United Kingdom	1000 Genomes	
70	SouthCentral Asia	Keralites,S.India	KER	Connecticut, USA	Kidd Lab	
71	SouthCentral Asia	Sri Lankan Tamil	STU	United Kingdom	1000 Genomes	
72	SouthCentral Asia	Thoti	THT	India	Kidd Lab	
73	SouthCentral Asia	Kachari from Assam	KCH	India	Kidd Lab	
74	SouthCentral Asia	Bengali from Bangladesh	BEB	Bangladesh	1000 Genomes	

75	Admixed: S.C.Asia/N.Asia	Hazara	HZR	Pakistan	Kidd Lab	
76	North Asia / Siberia West	Komi Zyriane	KMZ	W.Siberia,Russia	Kidd Lab	
77	North Asia / Siberia West	Khanty	KTY	W.Siberia,Russia	Kidd Lab	
78	North Asia / Siberia East	AltaiKazakh	AKZ	Mongolia	Kidd Lab	Collected by J.Brissenden and colleagues
79	North Asia / Siberia East	Yakut	YAK	E.Siberia, Russia	Kidd Lab	
80	North Asia / Siberia East	Tsaatan	TSA	Mongolia	Kidd Lab	
81	North Asia / Siberia East	Outer Mongolians	OMG	Mongolia	Kidd Lab	
82	East Asia	Koreans	KOR	South Korea	Kidd Lab	
83	East Asia	Japanese	JPN	Japan	Kidd Lab	
84	East Asia	Japanese, Tokyo	JPT	Japan	1000 Genomes	
85	East Asia	Han Chinese Beijing	CHB	China	1000 Genomes	
86	East Asia	Chinese Southern	HCS	China	1000 Genomes	samples collected from Hunan and Fujian provinces, China
87	East Asia	Chinese, SanFrancisco	CHS	USA	Kidd Lab	
88	East Asia	Chinese, Taiwan	CHT	Taiwan,China	Kidd Lab	
89	East Asia	Hakka	HKA	Taiwan,China	Kidd Lab	
90	East Asia	Ami	AMI	Taiwan,China	Kidd Lab	
91	East Asia	Atayal	ATL	Taiwan,China	Kidd Lab	
92	SouthEast Asia	Dai	CDX	Xishuangbanna,China	1000 Genomes	
93	SouthEast Asia	Laotians	LAO	Laos	Kidd Lab	
94	SouthEast Asia	Cambodians	CBD	USA	Kidd Lab	
95	SouthEast Asia	Vietnamese	KHV	Vietnam	1000 Genomes	
96	SouthEast Asia	Malaysians	MLY	Malaysia	Kidd Lab	
97	Pacific	Papuans--NewGuinea	PNG	New Guinea	Kidd Lab	
98	Pacific	Nasioi Melanesians	NAS	Solomon Islands	Kidd Lab	
99	Pacific	Micronesians	MCR	Micronesia, USA	Kidd Lab	
100	Pacific	Samoans	SMO	Samoa	Kidd Lab	
101	America, North	Plains AmerIndians	NPA	USA	Kidd Lab	
102	America, North	Southwest AmerIndians	SWA	USA	Kidd Lab	
103	America, North	Pima, Mexico	PMM	USA	Kidd Lab	
104	America, North	Maya	MAY	Yucatan,Mexico	Kidd Lab	
105	America, South	Guihiba speakers	GHB	Colombia	Kidd Lab	
106	America, South	Quechua	QUE	Peru	Kidd Lab	
107	America, South	Peruvians	PEL	Peru	1000 Genomes	European and Native American ancestry
108	America, South	Ticuna	TIC	Amazon,Brazil	Kidd Lab	
109	America, South	Rondonian Surui	SUR	Amazon,Brazil	Kidd Lab	
110	America, South	Karitiana	KAR	Amazon,Brazil	Kidd Lab	
111	Admixed: Amer/Europe	MexicanAmericans,LosAngeles	MXL	USA	1000 Genomes	NativeAmerican ancestry is low
112	Admixed: Europe/W.Afr./Amer	Colombians from Medellin	CLM	Colombia	1000 Genomes	NativeAmerican ancestry is low
113	Admixed: Europe/W.Afr./Amer	Puerto Ricans	PUR	USA	1000 Genomes	NativeAmerican ancestry is low

Supplementary Table S2. Allele Counts for the Populations Where PV Haplotype Is Observed

	PA	AV	AA	PV	
Ibo	53	24	18	1	96
Esan	94	61	42	1	198
Yoruba, BeninCity	81	31	43	1	156
Somali	17	13	1	1	32
African Americans	81	76	20	1	178
Nebeur	3	9	10	**2**	24
Kerkennah	22	42	23	**7**	94
Kairouan	21	39	24	**8**	92
Kesra	16	32	29	**7**	84
Sousse	20	48	24	**6**	98
Mahdia	19	49	20	**6**	94
Smar	26	48	32	**18**	124
Lybia	35	57	34	**14**	140
Keralites	24	35	0	1	60
Hazara	92	92	13	3	200
OuterMongolians	80	60	1	1	142
Pathans	90	121	10	3	224
MexicanAmericans,LosAngeles	85	39	3	1	128
N of haplotypes in populations where PV was observed	859	876	347	**82**	
N of world sampled haplotypes	6161	4924	911	82	

A=Alanine, P=Proline, V= Valine

Supplementary Table S3. Test of H-W Equilibrium of rs713598 and rs1726866 in North African Samples

Population	Locus	Observed Heterozygotes	Expected Heterozygotes	P-value	Observed Common Homozygotes	Observed rare Homozygotes	Expected Common Homozygotes	Expected rare Homozygotes	Observed Genotypes	N
Nebeur	**rs713598**	1	4.1304	0.03079	9 (CC)	2 (GG)	7.521 (CC)	0.521 (GG)	12	12
	rs1726866	5	5.958	0.59401	4 (GG)	3 (AA)	3.521 (GG)	2.521 (AA)	12	
Kesra	**rs713598**	15	16.90374	0.46163	23 (CC)	4 (GG)	22.149 (CC)	3.149 (GG)	42	42
	rs1726866	19	21.14448	0.54703	13 (GG)	10 (AA)	12.054 (GG)	9.054 (AA)	42	
Sousse	**rs713598**	16	18.1896	0.44810	28 (CC)	4 (GG)	27 (CC)	3 (GG)	48	49
	rs1726866	26	24.49461	0.77387	14 (AA)	9 (GG)	14.878 (AA)	9.878 (GG)	49	
Mahdia	**rs713598**	20	16.68964	0.24462	23 (CC)	1 (GG)	24.75 (CC)	2.75 (GG)	44	47
	rs1726866	27	22.3484	0.22506	12 (AA)	6 (GG)	8.45 (AA)	14.45 (GG)	45	
Kairouan	**rs713598**	22	19.5057	0.49441	20 (CC)	3 (GG)	21.356 (CC)	4.356 (GG)	45	46
	rs1726866	23	22.24156	1.00000	11 (GG)	10 (AA)	11.506 (GG)	10.506 (AA)	44	
Kerkennah	**rs713598**	19	20.0767	0.74095	22 (CC)	5 (GG)	21.571 (CC)	4.356 (GG)	46	47
	rs1726866	20	22.06908	0.55947	14 (AA)	10 (GG)	13.091(AA)	9.091 (GG)	44	
Smar	**rs713598**	27	28.07464	0.78479	26 (CC)	8 (GG)	25.578 (CC)	7.578 (GG)	61	62
	rs1726866	34	31.12214	0.60964	16 (AA)	12 (GG)	17.565 (AA)	13.565 (GG)	62	
Lybia	**rs713598**	25	30.11248	0.17899	32 (CC)	10 (GG)	29.556 (CC)	7.556 (GG)	67	70
	rs1726866	37	34.18496	0.62589	17 (AA)	14 (GG)	18.533 (AA)	15.533 (GG)	68	

A=Adénine, C=Cytosine, G=Guanine

rs713598: G>C

rs1726866: G>A

Supplementary Table S4. Differentiation Test between All Pairs of Samples

List of labels for population samples used below:

Label	Population name
1	Lybia
2	Mahdia
3	Kerkennah
4	Smar
5	Kairouan
6	Sousse
7	Kesra
8	Nebeur

Markov chain length : 100000 steps)

Non-differentiation exact P values:

	1	2	3	4	5	6	7
2	0.57182+-0.0125						
3	0.95249+-0.0042	0.58554+-0.0088					
4	0.91737+-0.0026	0.22725+-0.0107	0.49254+-0.0094				
5	0.67537+-0.0071	0.76652+-0.0093	0.75589+-0.0083	0.66265+-0.0103			
6	0.87465+-0.0046	0.33841+-0.0111	0.86506+-0.0069	0.35906+-0.0120	0.48726+-0.0055		
7	0.82073+-0.0057	0.12865+-0.0072	0.83526+-0.0070	0.55892+-0.0096	0.83640+-0.0050	0.66394+-0.0077	
8	0.59448+-0.0118	0.04224+-0.0049	0.44561+-0.0077	0.21802+-0.0079	0.21642+-0.0066	0.30313+-0.0081	0.76749+-0.0053

Table of significant differences (significance level=0.0500):

	1	2	3	4	5	6	7	8
1		-	-	-	-	-	-	-
2	-		-	-	-	-	-	+
3	-	-		-	-	-	-	-
4	-	-	-		-	-	-	-
5	-	-	-	-		-	-	-
6	-	-	-	-	-		-	-
7	-	-	-	-	-	-		-
8	-	+	-	-	-	-	-	

Histogram of the number of significant different populations (significance level=0.0500):

1	2	3	4	5	6	7	8
0	1	0	0	0	0	0	1

Figure Captions

Figure 1. Geographical location of North African samples used in this work. (From d-maps.com, https://d-maps.com/carte.php?num_car=25529&lang=en)

Figure 2. TAS2R38 SNPs allele frequencies profiles in worldwide populations. See abbreviations in Supporting Information Table S1.

Figure 3. Amino acid haplotype distribution in worldwide populations. See abbreviations in Supporting Information Table S1.

Figure 1.

Figure 2.

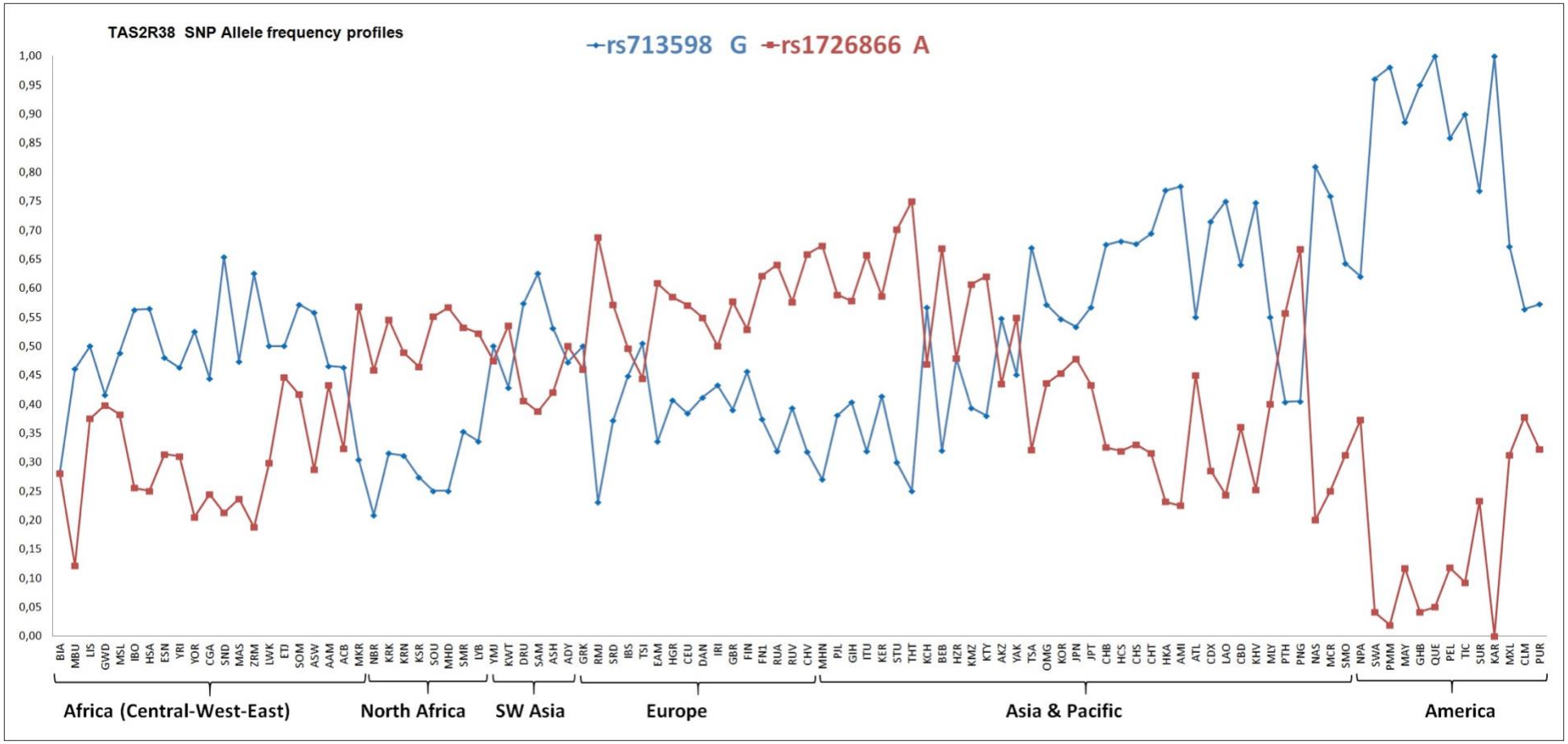
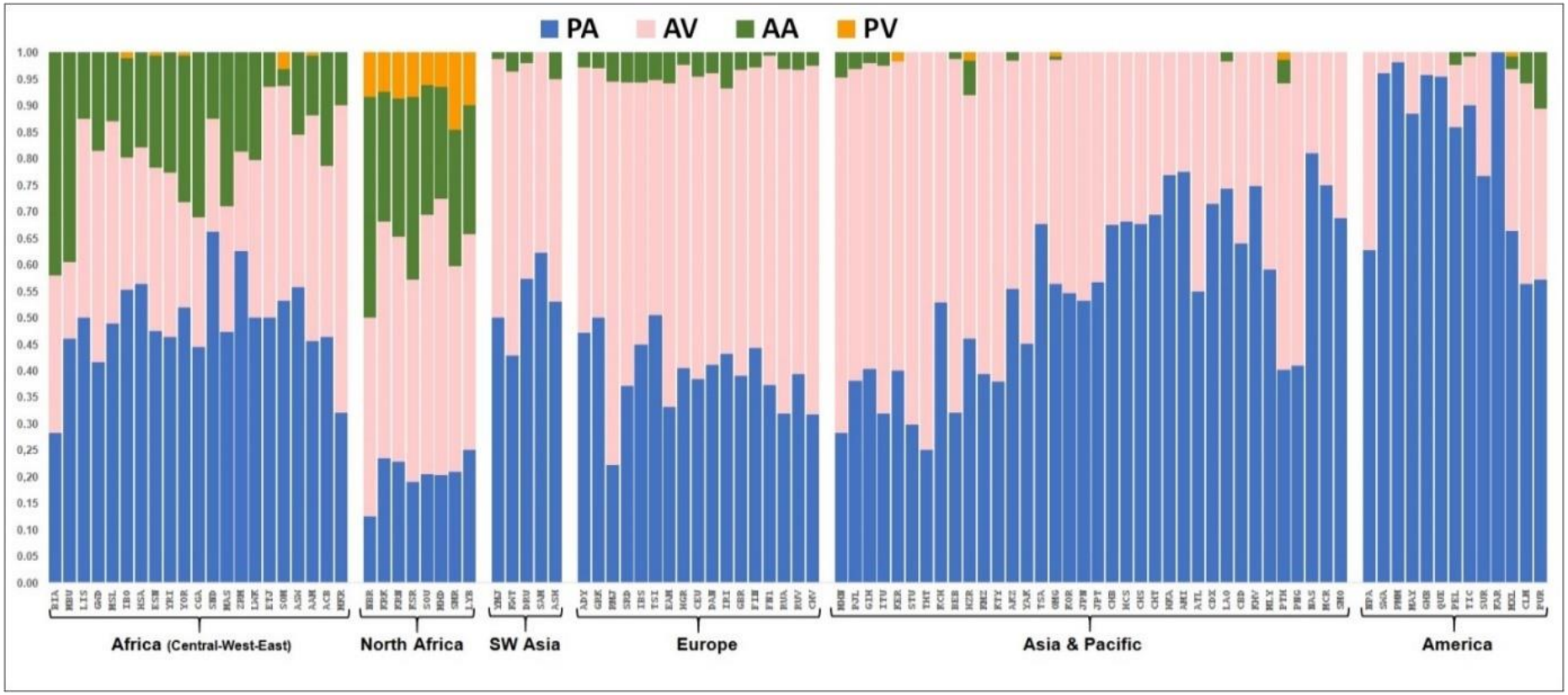


Figure 3.


PA
AV
AA
PV
1.00
0.95
0.90
0.85
0.80
0.75
0.70
0.65
0.60
0.55
0.50
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
Africa (Central-West-East)
North Africa
SW Asia
Europe
Asia & Pacific
America