



OPEN

Adaptive selection drives TRPP3 loss-of-function in an Ethiopian population

Sandra Walsh¹, Mercè Izquierdo-Serra², Sandra Acosta¹, Albert Edo²,
 María Lloret², Roser Moret¹, Elena Bosch^{1,3}, Baldo Oliva⁴, Jaume Bertranpetit¹✉ &
 José Manuel Fernández-Fernández²✉

TRPP3 (also called PKD2L1) is a nonselective, cation-permeable channel activated by multiple stimuli, including extracellular pH changes. TRPP3 had been considered a candidate for sour sensor in humans, due to its high expression in a subset of tongue receptor cells detecting sour, along with its membership to the TRP channel family known to function as sensory receptors. Here, we describe the functional consequences of two non-synonymous genetic variants (R278Q and R378W) found to be under strong positive selection in an Ethiopian population, the Gumuz. Electrophysiological studies and 3D modelling reveal TRPP3 loss-of-functions produced by both substitutions. R278Q impairs TRPP3 activation after alkalisation by mislocation of H⁺ binding residues at the extracellular polycystin mucolipin domain. R378W dramatically reduces channel activity by altering conformation of the voltage sensor domain and hampering channel transition from closed to open state. Sour sensitivity tests in R278Q/R378W carriers argue against both any involvement of TRPP3 in sour detection and the role of such physiological process in the reported evolutionary positive selection past event.

The action of positive (also called adaptive) selection can be identified in our genome by exploring the genome-wide landscape of the footprints of the so-called classical sweeps¹. Recognizing where in the genome adaptation has happened is a fully blind process in relation to the information of the genome, and many methods have been described to have full genome scans of adaptive selection (see² for a review). However, the biological interpretation of such footprints is not direct. Indeed, besides identifying where in the genome the sweep occurred, the obvious next step is to recognize the genetic bases of the selected phenotype. The understanding of the relationship between the selection signal and the selected trait is, in general, much more complex, with many cases not solved. So far, the most powerful and interesting method to associate phenotypic adaptations is population specific adaptation^{3–5}. There are some well-known cases for which the adaptation has been fully interpreted, from the detection of the signal to the specific genetic change and into the adaptive phenotype, like the lactase persistence^{6,7}, the adaptation to high altitudes^{8–11} and diet¹². In the recent years, the explosion of the production of genomic data has allowed to undertake whole genome scans of positive selection for several human populations. So far, the most reliable functional adaptive selection are associated to protein-coding genes, but for many cases, the detection of a selection signal is far from being functionally interpreted, being the main reason the lack of precision of the selection footprints, which usually comprise several kilobases with many different (and often pleiotropic) genes and several functional genetic variants, many still not well known or annotated in the genome.

As in the case of the identification of adaptive selection in fat metabolism in Inuit¹² and lactase persistence mainly in Europe^{6,7}, food resources are main drivers of adaptation. Therefore natural selection may have played an important role in shaping the genetic variation of the human taste system. Food resources and potential dangers vary across different environments and thus the human taste system has potentially evolved in a population specific way. These potential adaptations are likely to be very population specific as a survey on selection in the main taste-related genes in three divergent human populations (Yoruba from Africa, Han from China and Central Europeans) did not find main adaptive differences in taste, while specific adaptations were found for biotransformation¹³.

¹Institut de Biologia Evolutiva (UPF-CSIC), Universitat Pompeu Fabra, Dr. Aiguader, 88, 08003 Barcelona, Catalonia, Spain. ²Laboratory of Molecular Physiology, Department of Experimental and Health Sciences, Universitat Pompeu Fabra, 08003 Barcelona, Spain. ³Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM), 43206 Reus, Spain. ⁴Structural Bioinformatics Lab, Department of Experimental and Health Sciences, Universitat Pompeu Fabra, 08003 Barcelona, Spain. ✉email: jaume.bertranpetit@upf.edu; jmanuel.fernandez@upf.edu

The mammalian tongue contains gustatory receptors regulating basic taste types, providing an evolutionary old hedonic compass for what to ingest and what not to ingest¹⁴. One of the main hypotheses of the evolutionary sour tasting function is that it could warn against the acidic ingestion of rotten, unripe or fermented food¹⁵. In the recent years, transient receptor potential (TRP) and particularly TRPP3 had been considered a candidate for acid receptor since the ablation of taste receptor cells expressing TRPP3 in mice deplete the acid-induced responses in gustatory nerve recordings^{16,17}. TRPP3 (Transient Receptor Potential Polycystin-3, also called Polycystic Kidney Disease 2-Like 1 (PKD2L1) or polycystin-L) belongs to the TRPP subfamily of ion channels that are characterized by large extracellular domains^{18,19}. Human TRPP3 forms a nonselective, cation-permeable channel in a tetrameric structure. Each TRPP3 monomer has a voltage sensor domain (VSD) formed by the S1 to S4 transmembrane α -helices, a pore domain formed by the S5 to S6 helices, and a polycystin mucolipin domain (PMD) in the extracellular loop between transmembrane segments S1 and S2, which rests on top of the VSD of the same monomer and the adjacent pore domain of the tetrameric channel¹⁹. TRPP3 responds to multiple stimuli that include voltage membrane changes, mechanical stress, hypo-osmotic shock, temperature changes and alterations in extracellular pH^{20–23}. Hence, TRPP3 is co-expressed and interacts with the Polycystic Kidney Disease 1-Like 3 (PKD1L3) protein in a subset of receptor cells at the tongue taste buds of mammals²⁴, where the TRPP3/PKD1L3 channel complex was suggested to partly contribute to sour taste responses due to its activation after washout of an acid stimulus (acid-evoked off-response)^{25–28}. Accordingly, in humans, two patients with sour taste ageusia have been reported and neither had detectable TRPP3 transcripts²⁹. However, whereas it is well accepted that ablation of taste receptor cells expressing TRPP3 eliminates sour taste responses³⁰, receptors other than TRPP3 and PKD1L3 proteins have been proposed to be the main responsible for acid/sour detection. Indeed, genetic or functional knockout of TRPP3 and/or PKD1L3 in mice only produces a mild reduction (by 25–45%) in acid taste responses²⁷ or has no effect at all³¹. Interestingly, Otopetrin 1 (OTOP1) has been identified in TRPP3-expressing, acid-detecting taste receptor cells of mice as the ion channel required for conduction of H⁺ currents in response to extracellular acidification³². There, OTOP1 is essential for sour sensing in the taste system^{33,34}. OTOP1 is evolutionarily conserved from nematodes to humans^{32,35,36}, and human OTOP1 also works as an H⁺ channel of similar properties to murine OTOP1³². Overall, these results point out to OTOP1 as the main sour receptor in humans.

In this study we depart from a signal of selection in the genome, recognize the gene *TRPP3* and two non-synonymous variants as the selection target, and use molecular physiology techniques measuring channel activity and 3D molecular models to characterize the functional alterations produced by the two putative adaptive substitutions. Finally, in order to establish a link between the adaptive genotype and a potential adaptive external phenotype we look for differences in sour taste recognition (as the minimum citric acid concentration where a subject can recognize sour taste). Our results show a total decrease of the double mutant channel activity as seen in the electrophysiological experiments, which is in agreement with the 3D structure modelling of the two substitutions. On the phenotype, our results give further support to previous doubts on the involvement of TRPP3 in sour taste recognition.

Results

Identification of adaptive selection in the TRPP3 gene in the Gumuz population. The allele frequency increase of functional genetic variants is the major driver of human population adaptation to the wide variety of environments by adaptive (positive) selection. In a recent study of positive selection in several populations from Ethiopia from our lab (Walsh et al., in press) several interesting signals were uncovered using, by one hand, the SFselect software³⁷ based in detecting changes in the site frequency spectrum, which uncovers mainly ancient signals and by the other hand, the iHS test³⁸, based in the extension on linkage disequilibrium, which detects recent events that are mainly population-specific. One of the top signals of positive selection of the iHS analysis (showing also a high population differentiation as measured with the F_{st} index) was a region in the long arm of chromosome 10 that contains the *TRPP3* gene (Fig. 1A). The signal is not found in other populations such as the European (CEU) or the West African (YRI) populations from 1000 Genomes Project. We also checked the PopHuman selection browser³⁹ for signals of selection in a set of 26 worldwide populations from the Phase 3 1000 Genomes Project and did not find evidence of positive selection in any of them. Thus, this signal is found exclusively in a Nilo-Saharan-speaking population from Ethiopia, the Gumuz. This population is not affected by the extensive admixture from the Levant that is found in other Ethiopian populations such as the Amhara, Oromo, Somali and Wolayta⁴⁰. In the *TRPP3* gene, there are two non-synonymous variants with high iHS values, and a very strong population differentiation. In the Gumuz population these exonic variants, rs17112895 and rs7909153, are present in its derived form (T and A, respectively) both at a frequency of 0.71, much higher than in other African populations, with allele frequencies decreasing towards the west (Fig. 1B and Table 1A–C); in all non-African populations, the allele frequency of the derived allele is zero. These two variants encode for non-synonymous genetic variants (R278Q and R378W) located in the most external α -helix of the extracellular polycystin mucolipin domain (PMD) and in the intracellular linker between transmembrane segments S2 and S3 at the voltage sensor domain (VSD) of the TRPP3 channel, respectively (Fig. 2). There is a strong linkage disequilibrium in the region, even considering that these are sub-Saharan populations, characterized with general low levels of linkage disequilibrium^{41–43}. In fact, pairwise linkage disequilibrium measurements between both variants indicate that they are in strong linkage disequilibrium ($r^2 = 1$ and $D' = 1$). It is interesting to note that the two variants have an ancient age within the spectrum of the age of human SNPs. According to the Genealogical Estimation of Variant Age⁴⁴, rs7909153 and rs17112895 would have an age of 290,000 and 683,000 years respectively (with a generation time of 25 years). The interesting points here are that they have different ages and that both are much older than the Out of Africa expansion of modern humans. Together with the fact that both

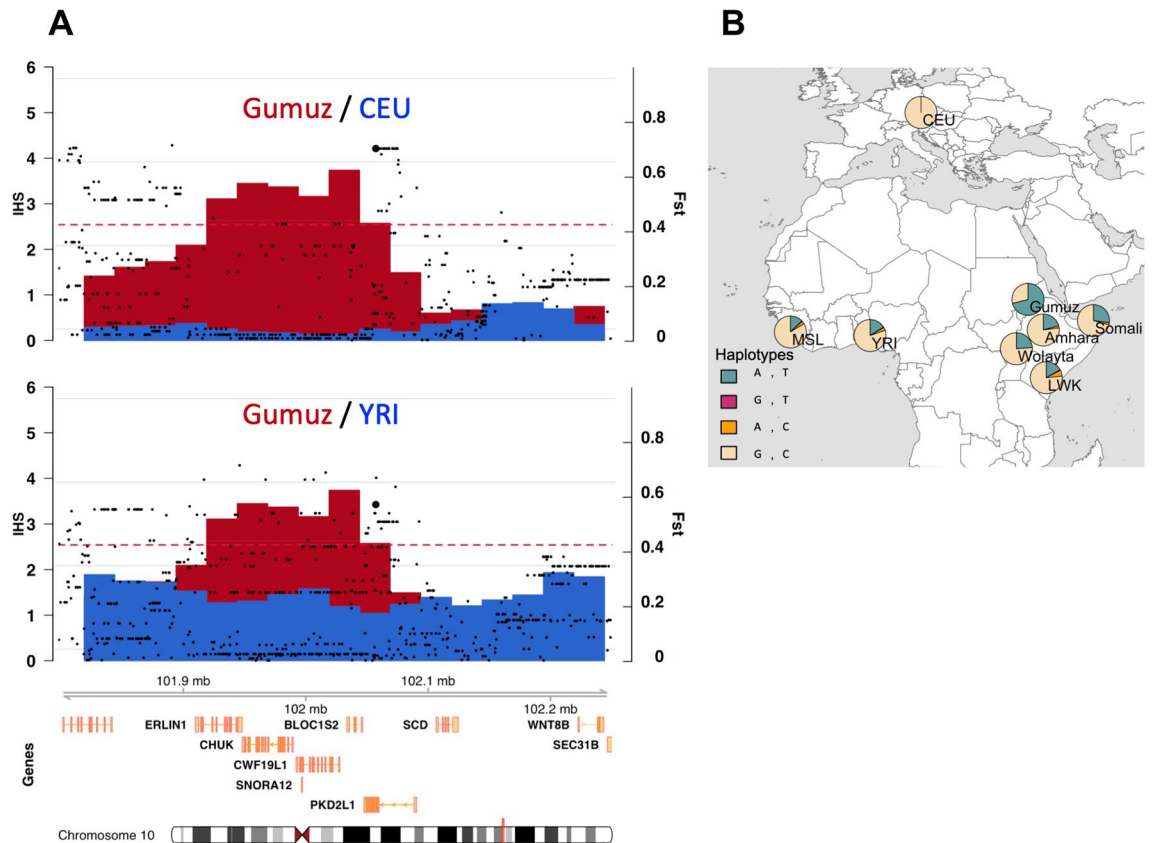


Figure 1. Positive selection signal in the *TRPP3* region of the Gumuz population. **(A)** Positive selection analysis (iHS and F_{st}) of the Gumuz population compared to CEU (upper panel) and YRI (lower panel) populations. Blue and red bars represent the absolute mean iHS of a 30 kilobase (kb) window (left y-axis). Red dotted lines indicate the 99.99th percentile thresholds of the absolute mean iHS per 30 kb window reported in Walsh et al., in press that were obtained from extensive population genetics neutral simulations. Black dots correspond to the F_{st} values per site (as a measure of population differentiation) between the Gumuz and CEU (upper panel) and YRI (lower panel) population; the right y-axis indicates the scale of F_{st} values. Variants rs17112895 and rs7909153 are indicated in bigger black dots. **(B)** Spatial distribution of haplotypes per population. Pie plots indicate the population frequencies of each combination of haplotypes containing both substitutions in the Gumuz, Amhara, Somali, Wolayta from Ethiopia, YRI: Yoruba from Nigeria, LWK: Luhya from Kenya, MSL: Mende from Sierra Leone, CEU: Europeans. Map and pie plots were generated with R (v3.5.3)⁷² using the packages rworldmap (v1.3-6, http://journal.r-project.org/archive/2011-1/RJournal_2011-1_South.pdf), rworldxtra (v1.01, <https://CRAN.R-project.org/package=rworldxtra>) and mapplots (v1.5.1, <https://CRAN.R-project.org/package=mapplots>).

variants are found at lower frequencies in other African populations suggests that selection acted in standing variation that was previously neutrally segregating in Africa.

Electrophysiological analysis reveals impaired function of double mutant TRPP3. Next, we aimed to characterize the functional consequences of the two *TRPP3* non-synonymous genetic variants (R278Q and R378W) by comparing channel activity of the wild-type (WT) *TRPP3* channel with that of channels carrying either the double substitution (R278Q/R378W) or the single substitutions (R278Q or R378W). For that purpose, we performed whole-cell patch-clamp recordings on GFP positive HEK293 cells co-transfected with PKD1L3 and GFP-fused *TRPP3* channels (*TRPP3*-GFP). We measured robust channel activity in response to voltage changes in cells expressing WT *TRPP3*. As previously reported²⁰, they exhibited outwardly rectifying currents with large tail currents after repolarization to -100 mV at extracellular pH 7.4, whereas no such currents were recorded from untransfected HEK293 cells (Fig. 3A,E,G and Supplementary Fig. S1A,B online). Besides, recordings at extracellular pH 9.0 revealed largest *TRPP3* steady-state and tail currents (Fig. 3B), as reported before for HEK293 cells expressing only *TRPP3* channels in the absence of PKD1L3²². The average current–voltage relationships for WT steady-state currents and tail currents at each test pulse and pH condition are shown in Fig. 3E–H. Furthermore, we also found that at extracellular pH 9, the deactivation time course of WT *TRPP3* tail currents was accelerated when compared to the extracellular pH 7.4 condition (see Supplementary Fig. S2A,C online), another known effect of alkalinisation on *TRPP3* tail currents²².

The presence of the double substitution R278Q/R378W reduced the amplitude of both *TRPP3* steady-state and tail currents recorded at pH 7.4 when compared to the WT channel (Fig. 3C,E,G). Moreover, unlike the

rsID	Genomic position (hg19)	Ancestral nucleotide	Derived nucleotide	Amino acid site	Ancestral amino acid	Derived amino acid	Gumuz	Amhara	LWK	YRI	MSL	CEU
(A)												
rs17112895	102057262	C	T	278	R	Q	0.71	0.25	0.22	0.18	0.13	0
rs7909153	102056790	G	A	378	R	W	0.71	0.27	0.25	0.21	0.15	0
rsID	Genomic position (hg19)	Ancestral nucleotide	Derived nucleotide	Gumuz fraction of ancestral homozygotes	Gumuz fraction of heterozygotes	Gumuz fraction of derived homozygotes						
(B)												
rs17112895	102057262	C	T	0.04	0.5	0.46						
rs7909153	102056790	G	A	0.04	0.5	0.46						
Haplotype	Amino Acid	Gumuz	Amhara	YRI	LWK	MSL	CEU					
(C)												
A,T	Q278, W378	0.71	0.21	0.16	0.04	0.13	0					
G,T	R278, W378	0	0	0.01	0.13	0.01	0					
A,C	Q278, R378	0	0.02	0.03	0.19	0.03	0					
G,C	R278, R378	0.29	0.77	0.8	0.64	0.84	1					

Table 1. Summary of substitutions rs17112895 and rs7909153. **(A)** Genomic location and allele frequencies in several populations (Gumuz, Amhara, Somali, Wolayta from Ethiopia, YRI: Yoruba from Nigeria, LWK: Luhya from Kenya, MSL: Mende from Sierra Leone, CEU: Europeans). **(B)** Fraction of Gumuz individuals homozygous and heterozygous for both substitutions. **(C)** Haplotype frequencies of both substitutions in worldwide context.

response of the WT channel to alkalinisation, currents through the double variant TRPP3 channel were not enhanced at the extracellular pH 9 condition (Fig. 3D,F,H).

When recorded at pH 7.4, TRPP3 steady-state and tail currents through the R278Q mutant channel were of similar magnitude to those found for the WT channel (Fig. 3E,G and Supplementary Fig. S1D online). However, no significant increase in R278Q TRPP3 current amplitude neither faster deactivation kinetic was observed at pH 9 (Fig. 3F,H and Supplementary Fig. S1E, S2B,D online).

The presence of the single substitution R378W also impaired TRPP3 channel activity at pH 7.4, as observed for the double substitution R278Q/R378W (Fig. 3E,G and Supplementary Fig. S1F online). The magnitude of the whole-cell currents recorded from cells expressing these mutant channels was indistinguishable from that obtained in untransfected cells (Fig. 3C,E,G and Supplementary Fig. S1A,F online). Yet, R378W TRPP3 steady-state and tail currents were slightly but significantly largest at pH 9 (Fig. 3F,H, and Supplementary Fig. S1G online).

Because of the low open probability of TRPP3 at hyperpolarized potentials and the large amplitude of the single channel currents it is possible to record spontaneous single-channel activity in the whole-cell configuration^{20,22}, which was never observed in untransfected HEK293 cells (see Fig. 4 and Supplementary Fig. S3 online). Under these experimental conditions, we observed that the alkaline condition (pH 9) only increased the open probability (NP_o) of the WT channel (Fig. 4A–C). Besides, as observed for macroscopic TRPP3 steady-state and tail currents, WT and the R278Q channels showed similar activity only at extracellular pH 7.4, whereas the open probability of R378W and R278Q/R378W channels is lower at both neutral (pH 7.4) and alkaline (pH 9) conditions when compared to the WT channel (Fig. 4A–C and Supplementary Fig. S3 online). As reported before²², WT TRPP3 showed similar single-channel conductance at extracellular pH 7.4 and 9, and it was not altered in the TRPP3 channels with the substitutions (Fig. 4D,E).

To verify that the decrease in channel activity produced by the double substitution R278Q/R378W was due to a functional alteration of the channel and not to lower presence of the protein in the cell membrane, we evaluated the level of co-localization of both, the ectopically-expressed WT and mutated TRPP3-GFP channels, with the plasma membrane marker concanavalin A. Confocal imaging analysis indicates that both the WT and the TRPP3 proteins with the substitutions are expressed similarly in the plasma membrane (Fig. 5).

Altogether, in agreement with the extreme Combined Annotation Dependent Depletion (CADD) scores, rs17112895 for R278Q and rs7909153 for R378W, have CADD scores of 23 and 24.8 respectively, values that suggest potentially very strong functional effects at protein level. In fact, our data indicate that both genetic variants alone and in combination, exert loss-of-function effects on TRPP3 channels.

After coexpression of WT and double mutant R278Q/R378W channels we did not observe a dominant negative effect of the double mutation (see Supplementary Fig. S1B,C,H,I and S3 online). However, given the high proportion of homozygotes in the Gumuz population carrying the two mutations (Table 1B) our functional characterization of the mutant channel remains relevant.

Further evidence of TRPP3 mutant impaired function through protein 3D modelling. We next explored by 3D structural modelling the structural modifications driven by the R378W substitution, that changes the charged side-chain of arginine by the aromatic chain of tryptophan. This stabilizes the helix formed in the linker between S2 and S3 in the closed-channel conformation by locating the side-chain upwards towards

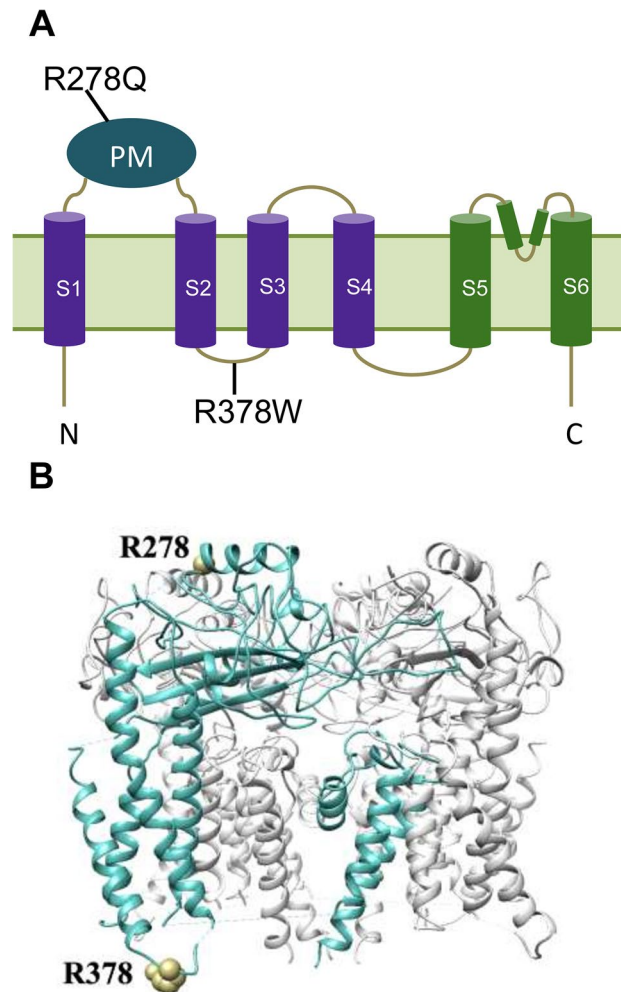


Figure 2. Schematic representation of TRPP3 topology in the cellular membrane and location of R278Q and R378W variants. **(A)** Cartoon illustrating the membrane topology of a TRPP3 monomer. S1–S4 transmembrane segments constitute the voltage sensor domain (VSD) and S5–S6 contribute to the pore region. The polycystin mucolipin domain (PMD) locates in the extracellular loop between S1 and S2 transmembrane segments. The variants are located in the PMD (R278Q) and in the intracellular linker between the S2 and S3 α -helices (R378W). **(B)** Precise location of residues R278 and R378 according to the human TRPP3 cryo-EM structure of the tetrameric channel (Worldwide Protein Data Bank (6DU8) and Electron Microscopy Data Bank (8912))¹⁹. Affected residues are only shown in the TRPP3 monomer coloured in cyan.

the inner membrane (see model in Supplementary Fig. S4 online). Besides, the comparison of the structures of the wild type and mutant form in open conformation (see Supplementary Fig. S5 online) shows that the connecting helix between S2 and S3 is unfolded, with a movement of helix S4 that orients K455 towards D390 in the wild type, according to⁴⁵. The structural analysis of the models shows a similar distribution of contacts at hydrogen bonding distance (H-bond, considering the maximum acceptable distance to form the hydrogen bond at 3.5 Å), between D390 and K452 and between D390 and K455, for both mutant and wild type forms (see Supplementary Fig. S6 online). Also, the distribution of distances between D390 and K455 when they are not at H-bond distance has an average around 8.7 Å, but the deviation is larger for R378W (2.3 Å) than for the wild type form (1.9 Å), which produces a strong significant difference (Mann–Whitney *U*-test P-value of $1.1e^{-106}$). Furthermore, wild type and R378W forms have a significant different distribution of the contacts between D390 and K452 when they are not at H-bond distance (with a Mann–Whitney *U*-test P-value of $7.3e^{-54}$). The capacity to form a hydrogen bond between D390 and K452 is more often lost in the wild type form than in the R378W form. This is reflected by the ratio of H-bonds D390–K452 preserved (49% for the wild type versus 57% for the mutant form R378W). These trends confirm that while in the wild type form the hydrogen bond swaps from D390–K452 to D390–K455 when opening the channel, in the R378W mutant the hydrogen bond is more restrained and preserves the closed conformation, which implies the loss-of-function.

Phenotypic analysis in human taste sense. After demonstrating that both substitutions impair the TRPP3 activity, we next examined the putative phenotypic differences for the main function initially associated to the gene, sour tasting, between individuals carrying the derived or the ancestral variants. For that, we

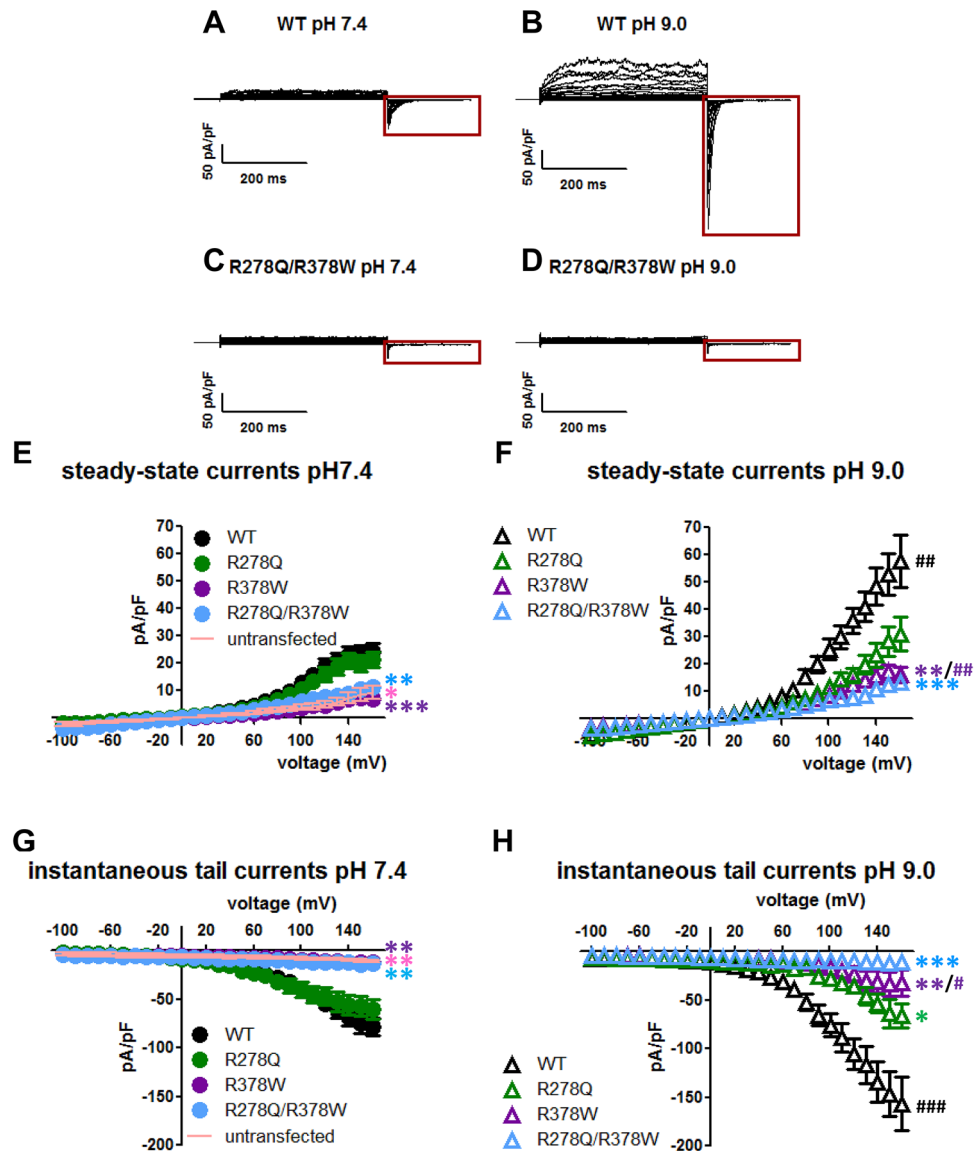


Figure 3. Loss of TRPP3 function induced by mutations R278Q and R378W. (A–D) Representative whole-cell currents from HEK293 cells co-transfected with cDNAs encoding PKD1L3 and wild-type (WT) or double mutant (R278Q/R378W) TRPP3-GFP and exposed to either extracellular pH 7.4 or pH 9.0, as indicated. Currents were elicited by step pulses from -100 to $+160$ mV in 10 mV increments with a postpulse to -100 mV, and normalized by cell size (membrane capacitance). Red boxes show tail currents after membrane repolarization to -100 mV. (E–H) Average current–voltage (I–V) relationships for steady-state currents (E,F) and instantaneous tail currents (G,H) at each test pulse in untransfected HEK293 cells and HEK293 cells expressing PKD1L3 and wild-type (WT) or mutants (R278Q, R378W, R278Q/R378W) TRPP3-GFP and exposed to either extracellular pH 7.4 or pH 9.0, as indicated. Area under the I–V curves (AUCs) were calculated for statistical analysis (only values at positive voltages were taken under consideration for steady-state I–V curves). Obtained AUC values for steady-state currents were: control untransfected cells at pH 7.4 (pink line, $n=6$) 533.15 ± 145.3 ; WT at pH 7.4 (black circle, $n=19$) 1662.6 ± 233.9 ; WT at pH 9 (black triangle, $n=17$) 3213.9 ± 444.2 ; R278Q at pH 7.4 (green circle, $n=9$) 1375.6 ± 184.3 ; R278Q at pH 9 (green triangle, $n=10$) 1663.9 ± 291.7 ; R378W at pH 7.4 (purple circle, $n=5$) 492 ± 72.5 ; R378W at pH 9 (purple triangle, $n=6$) 1102.3 ± 176 ; R278Q/R378W at pH 7.4 (blue circle, $n=10$) 620.3 ± 100.4 ; R278Q/R378W at pH 9 (blue triangle, $n=6$) 856.1 ± 132.5 . Obtained AUC values for instantaneous tail currents were: control untransfected cells at pH 7.4 (pink line, $n=6$) -1689 ± 359.4 ; WT at pH 7.4 (black circle, $n=19$) -6034.2 ± 649.3 ; WT at pH 9 (black triangle, $n=17$) $-10,882.7 \pm 1437.6$; R278Q at pH 7.4 (green circle, $n=9$) -5563.5 ± 1010.1 ; R278Q at pH 9 (green triangle, $n=10$) -5034 ± 920.5 ; R378W at pH 7.4 (purple circle, $n=5$) -1860 ± 218.1 ; R378W at pH 9 (purple triangle, $n=6$) -2992 ± 538.2 ; R278Q/R378W at pH 7.4 (blue circle, $n=10$) -2005.8 ± 229.2 ; R278Q/R378W at pH 9 (blue triangle, $n=6$) -1979.2 ± 117.5 . * $P < 0.05$, ** $P < 0.01$ and *** $P < 0.001$ (when compared to the corresponding TRPP3 channel at pH 7.4; one-tail Student's *t*-test or one-tail Mann–Whitney *U*-test, as appropriate); * $P < 0.05$, ** $P < 0.01$ and *** $P < 0.001$ (when compared to the WT channel at the corresponding extracellular pH condition; Kruskal–Wallis, followed by Dunn post hoc test).

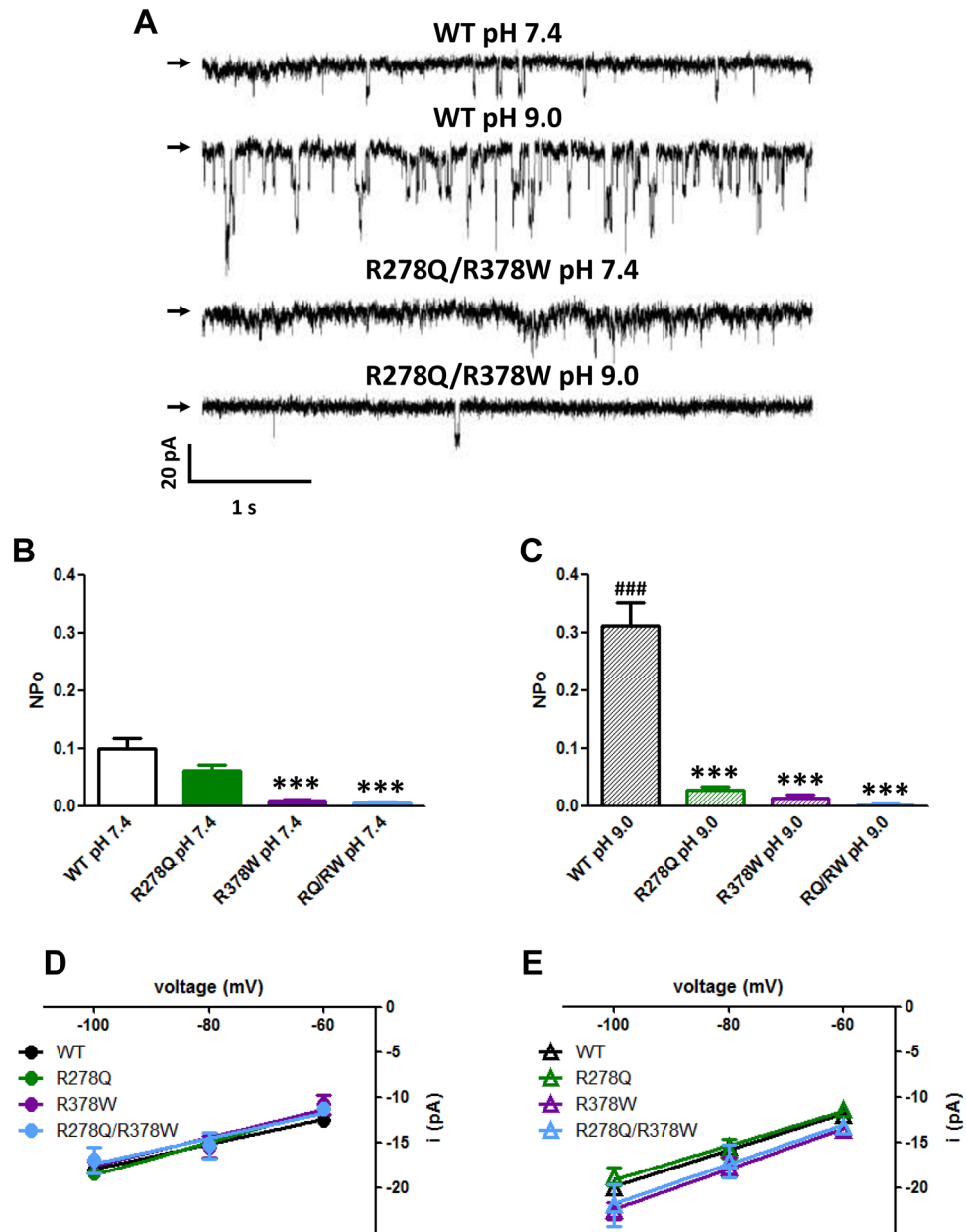


Figure 4. Effect of mutations R278Q and R378W on spontaneous TRPP3 single-channel activity. (A) Representative single-channel activity of wild-type (WT) or double mutant (R278Q/R378W) TRPP3 channels in whole-cell recordings obtained at negative membrane potential (-80 mV) from HEK293 cells co-transfected with cDNAs encoding PKD1L3 and wild-type (WT) or double mutant R278Q/R378W TRPP3-GFP, and exposed to either extracellular pH 7.4 or pH 9.0, as indicated. Arrows indicate the zero current level. (B,C). Average channel activity (NP_o) in whole-cell recordings obtained at negative membrane potentials (-60 , -80 and -100 mV) from HEK293 cells co-transfected with cDNAs encoding PKD1L3 and wild-type (WT) or mutants (R278Q, R378W, R278Q/R378W) TRPP3-GFP, and exposed to either extracellular pH 7.4 or pH 9.0, as indicated. $###P < 0.0001$ (when compared to WT channels at pH 7.4; one-tail Mann-Whitney *U*-test); $***P < 0.001$ (when compared to the WT channel at the corresponding extracellular pH condition; Kruskal-Wallis, followed by Dunn post hoc test). (D,E) Average single-channel current amplitudes and slope conductances at negative potentials (between -60 and -100 mV) obtained in whole-cell recordings from HEK293 cells co-transfected with cDNAs encoding PKD1L3 and wild-type (WT) or mutants (R278Q, R378W, R278Q/R378W) TRPP3-GFP, and exposed to either extracellular pH 7.4 or pH 9.0, as indicated. No significant differences were found between WT and mutant TRPP3 channels or between pH conditions (one-way ANOVA, $P = 0.1974$). Single-channel conductance values (in pS) for the different TRPP3 channels were: WT at pH 7.4 (black circle, $n = 16$) 137.8 ± 16.3 ; WT at pH 9 (black triangle, $n = 14$) 201.7 ± 20.1 ; R278Q at pH 7.4 (green circle, $n = 9$) 180 ± 13.7 ; R278Q at pH 9 (green triangle, $n = 7$) 189.4 ± 35.3 ; R378W at pH 7.4 (purple circle, $n = 3$) 155.9 ± 35.6 ; R378W at pH 9 (purple triangle, $n = 7$) 222.7 ± 30.4 ; R278Q/R378W at pH 7.4 (blue circle, $n = 4$) 139.5 ± 46.3 ; R278Q/R378W at pH 9 (blue triangle, $n = 4$) 217.9 ± 73.4 .

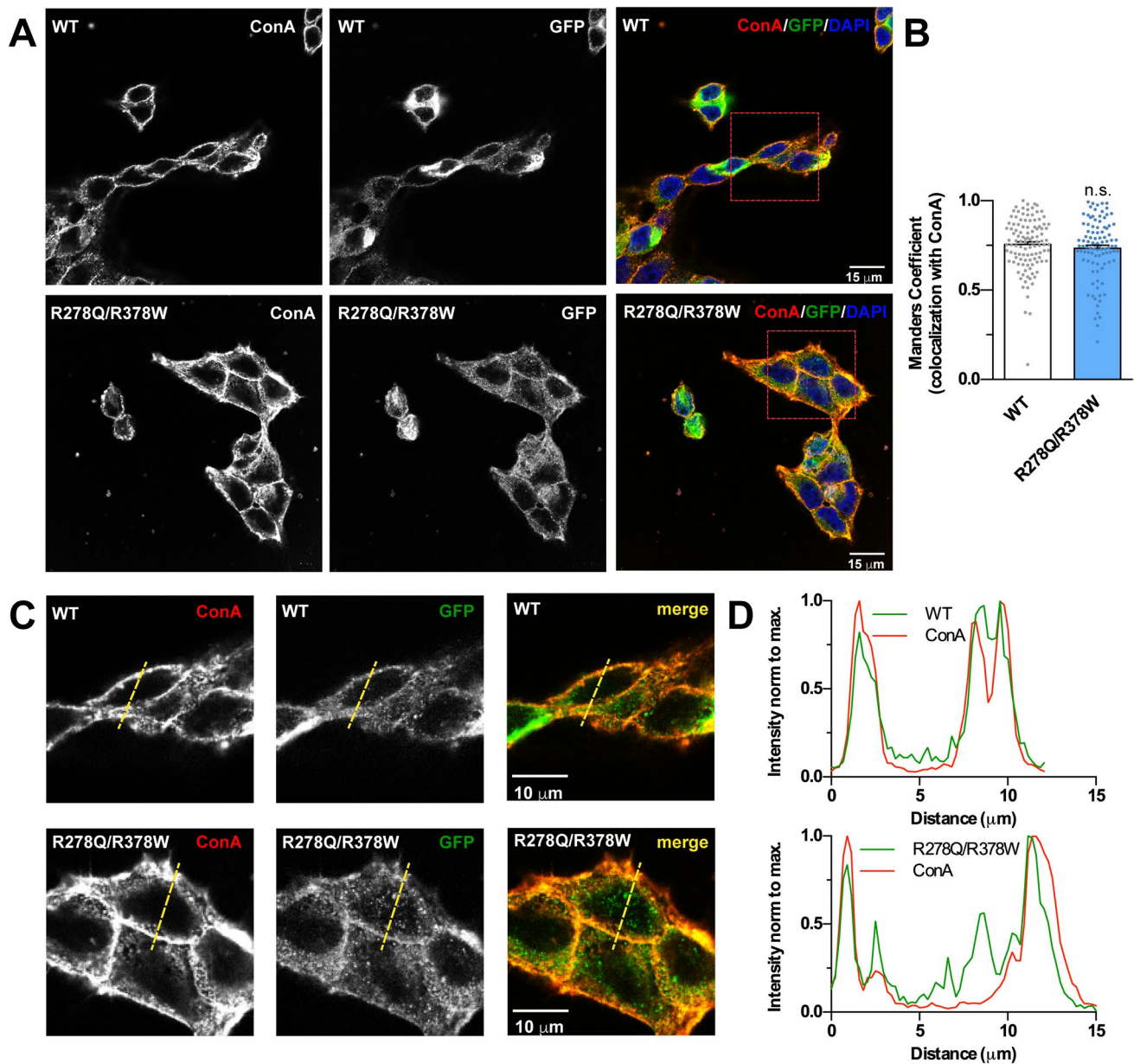


Figure 5. Substitutions R278Q and R378W do not modify the plasma membrane expression level of heterologously expressed TRPP3 channels. **(A)** Confocal images (single z plane) of HEK293 cells co-transfected with cDNAs encoding PKD1L3 and wild-type (WT, top) or double mutant (R278Q/R378W, bottom) TRPP3-GFP channels (middle panels). Cell membranes were stained with ConA-Rhodamine (left) and the overlaid image is shown on the right. Red-dashed boxes indicate the zoomed areas in **(C)**. Scale bar 15 μm . **(B)** Manders colocalization coefficient indicating the fraction of TRPP3-channel WT and R278Q/R378W colocalizing with plasma membrane ConA-Rhodamine. Double mutant ($n=97$) does not show significant differences when compared to the wild-type ($n=112$) TRPP3 channel (Mann–Whitney test, $P=0.5158$). **(C)** Zoomed images of the area indicated in **A** showing two cells with similar expression levels of wild-type (WT, top images) and double mutant (R278Q/R378W, bottom images) TRPP3-GFP channels. **(D)** Normalized pixel intensity profiles of crossed sections indicated by the yellow dashed line in **C** for WT (top) and double mutant (bottom) TRPP3 channels.

recruited a total of 44 individuals of Western African (13), Ethiopian (13) and European (18) ancestry, and genotyped them for both *TRPP3* variants (not all individuals could be genotyped, see Table 2A,B and Supplementary Table S1 online) from a saliva sample. The Ethiopian sampled populations were of Amhara and Oromo ancestry, none of the individuals had a Gumuz ancestry (a population we cannot access for the test). For the phenotypic analysis, we determined their sour taste recognition threshold following the described method⁴⁶. Briefly, the individuals were asked to blindly taste solutions of citric acid at different concentrations and fill a questionnaire

Population	Chromosomes	Males/ females	Genotyped chromosomes	Q278 frequency	W378 frequency	278 ^{R/R} 378 ^{R/R}	278 ^{R/Q} 378 ^{R/W}	278 ^{Q/Q} 378 ^{W/W}	Mean RT (mM)	Mean pH
(A)										
WA	26	7/6	26	0.31	0.31	9	4	0	0.240	–
ET	26	9/4	24	0.25	0.25	10	1	1	0.225	6.9
EU	36	6/12	30	0	0	15	0	0	0.250	6.6
	Samples	Males/ females	Mean RT (mM)	Mean pH						
(B)										
278 ^{R/R} 378 ^{R/R}	34	15/19	0.24	6.8						
278 ^{R/Q} 378 ^{R/W}	5	4/1	0.24	6.25						
278 ^{Q/Q} 378 ^{W/W}	1	1/0	0.20	7						

Table 2. Summary of the sampled individuals for the sour taste phenotypic analysis. **(A)** Number of chromosomes per ancestry, Ethiopian (ET), West African (WA) or European (EU), number of males and females, frequency of the substitutions, number of individuals carrying the three genotypes and mean recognition threshold (RT, citric acid in mM) and mean saliva pH. **(B)** Number of samples, number of males and females and mean recognition threshold and mean saliva pH of individuals carrying the three possible genotypes.

related to qualitative sour tasting foods. Due to differential TRPP3 channel activity through alkalinisation²², we also determined the saliva pH of some of the individuals.

Among the genotyped individuals, we found 34 homozygous for the ancestral alleles (278^{R/R} 378^{R/R}), 1 double homozygous for the derived alleles (278^{Q/Q} 378^{W/W}) and 5 double heterozygous (278^{R/Q} 378^{R/W}) indicating that both loci are in strong linkage disequilibrium (Table 2B). All individuals (6) carrying a derived state of an allele are of African origin, as expected by the general population allele frequencies that have a zero allele frequency outside Africa. The limited number of individuals carrying the derived alleles mirrors the low population allele frequencies of the variants within African (non-Gumuz Ethiopians and West Africans) populations.

The common saliva pH ranges from 6.2 to 7.6⁴⁷. For the analysed cohort, we correlated the pH and the recognition threshold for sour taste. We observed a significant positive correlation between saliva pH and the sour recognition threshold (Fig. 6E, Pearson's $r = 0.45$, p -value = 0.018). We also report significant differences between males and females of sour recognition thresholds (Fig. 6A,C). Nonetheless, no significant sour recognition differences were observed between the three populations after an analysis of covariance controlling for sex and saliva pH (Fig. 6B, $F = 1.43$, $df = 2$, p -value = 0.26) or between the three different observed genotypes (Fig. 6D, $F = 1.36$, $df = 2$, p -value = 0.28). The limited sample size of individuals carrying the derived alleles does not allow a clear result, but at this level we cannot claim any threshold difference related to the presence or not of the two TRPP3 substitutions.

Discussion

We report the functional consequences on TRPP3 channel activity of two non-synonymous genetic variants (R278Q and R378W) that we identified as positively selected amino acid substitutions in the Gumuz population from Ethiopia. Our results show that both genetic variants lead to TRPP3 loss-of-function without alteration of channel trafficking to the plasma membrane, which is similar for both wild-type (WT) and double mutant R278Q/R378W TRPP3 channels. Substitution R278Q, at the most external α -helix of the extracellular polycystin mucolipin domain (PMD), impairs TRPP3 activation by alkaline pH. And substitution R378W, at the intracellular linker between transmembrane domains S2 and S3 of the voltage sensor domain (VSD), strongly reduces channel activity at both neutral (pH 7.4) and alkaline (pH 9) conditions but still show some degree of activation in response to alkalinisation. Double substitution R278Q/R378W TRPP3 channels show both nearly no activity at neutral pH and impaired modulation by alkaline pH.

Concerning TRPP3 structure–function relationship, our results reveal an important role of the PMD in channel sensitivity to pH. Because of its structural organization, PMD seems to facilitate the assembly of polycystin domains in the tetrameric channel and maintain their stability^{18,19}. In this process it seems to be involved the PMD residue E290 (located at the same external α -helix containing the R278Q substitution), which forms a hydrogen bond with residue N176 of the adjacent PMD in the tetramer (see Supplementary Fig. S7 online). Besides, PMD interacts with the pore domain and it is in close structural relationship with the VSD (in particular the S3 transmembrane segment and the extracellular S3-S4 helical region)^{18,19}. With regards of TRPP3 modulation by pH, from the comparison of structural data obtained for TRPP3 and another pH sensitive channel (TRPML3) at neutral and acid pH conditions, it has been proposed that the TRPP3 acid-evoked off-response may be due to the release of an inhibition produced by the binding of H⁺ to certain negatively charged residues of the PMD (E135, D143 and D156) after acid washout^{18,48}. Thus, TRPP3 activation in response to alkalinisation would be the consequence of conformational changes in the PMD region interacting with H⁺ that in turn transduce a signal affecting both accessibility to the external pore region and VSD movement. Accordingly, we can speculate that the neutralization of the positive charge produced by the R278Q substitution in the upper α -helix of the PMD might modify its global structure organization and affect PMD assembly and steadiness, and/or the relative position of the negatively charged amino acids interacting with H⁺ (for instance, residue D156

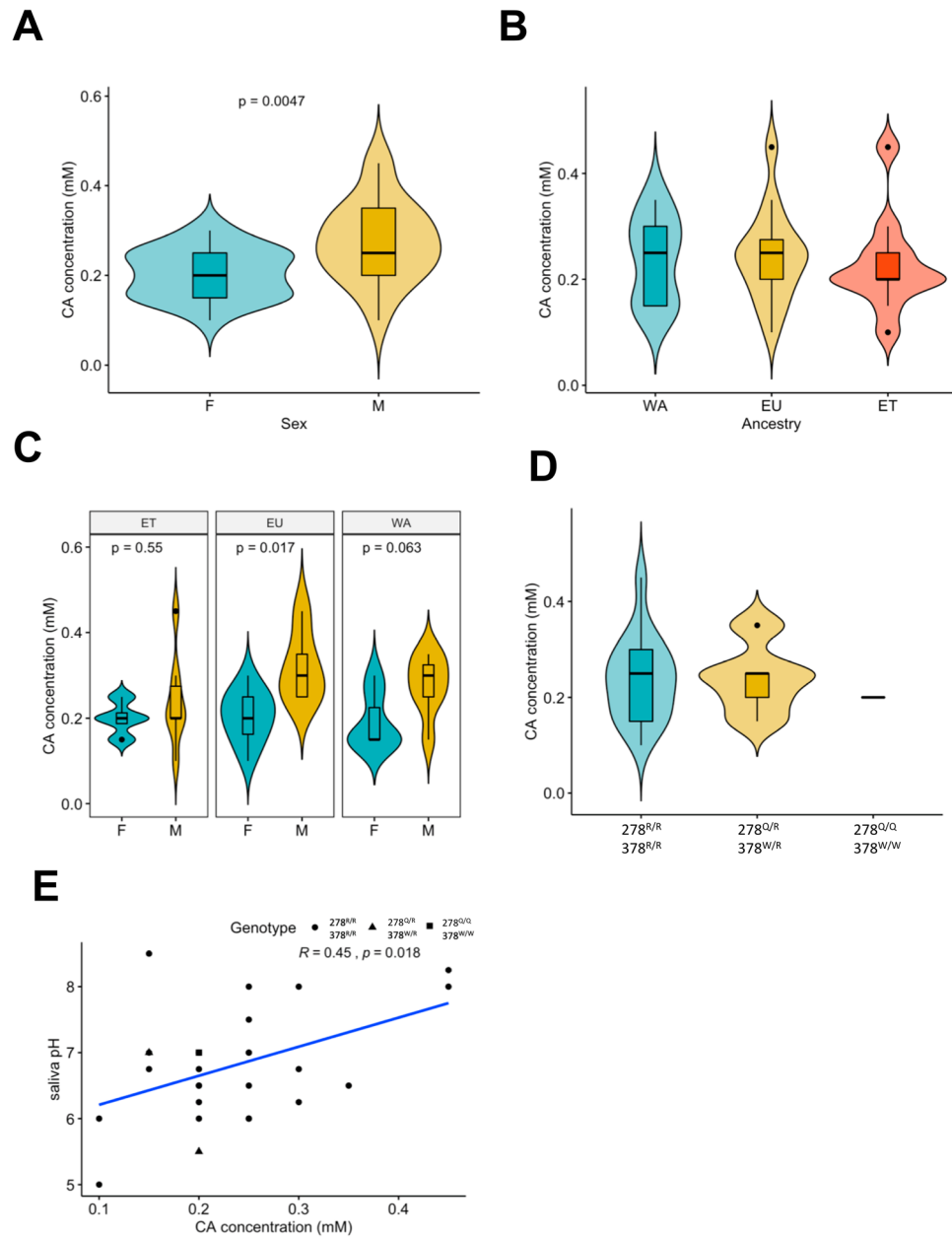


Figure 6. Genotype–phenotype analysis of the sour recognition thresholds. **(A)** Significant differences between males and females of sour recognition threshold (Wilcoxon test $W = 116.5$ p -value = 0,005). **(B)** Sour recognition threshold of the three sampled populations, Ethiopia (ET), West Africa (WA), Europe (EU). No significant recognition threshold differences are observed between populations after applying an analysis of covariance of ranked data controlling for sex and saliva pH. **(C)** Sour recognition thresholds of the three studied populations stratified by sex. Wilcoxon test was performed to test for differences between sexes within a same population and p -values were adjusted with holm correction. **(D)** Genotype analysis of individuals and sour recognition thresholds. We found individuals with both substitutions at an ancestral homozygous state ($278^{R/R} 378^{R/R}$), heterozygous state ($278^{Q/R} 378^{W/R}$) or derived homozygous state ($278^{Q/Q} 378^{W/W}$). No significant differences are found between the different genotypes after applying an analysis of covariance of ranked data controlling for sex and saliva pH. **(E)** Positive correlation between saliva pH and sour recognition threshold.

located in the α -helix immediately below the one containing the R278Q substitution)^{18,19}. In this respect, it is worth to notice that in this second α -helix, residue D167 is close to residue R278¹⁹ and the RQ change might alter a possible electrostatic interaction among them, modify the relative position of these two α -helices and therefore affect the location of residue D156 involved in H^+ binding (see Supplementary Fig. S7 online). Such conformational alteration might explain the impaired activation in response to alkaline pH 9 of TRPP3 channels carrying the R278Q variant.

The polycystin domains of TRPP3 share similar structural features and high degree of homology to PKD2 (also termed TRPP2)^{18,19,49}, that together with PKD1 play crucial roles in autosomal dominant polycystic kidney disease (ADPKD), a common and potentially lethal disorder resulting in major renal disorders. Loss-of-function mutations of either PKD1 or PKD2 account for the majority of ADPKD cases⁵⁰. Multiple ADPKD missense mutations affect the polycystin domain of PKD2 [The Autosomal Dominant Polycystic Kidney Disease Database (PKDB)]. None of them are located at the equivalent TRPP3 external α -helix containing the R278 residue (which, along with D167, is conserved in PKD2, corresponding to R399 and D289). However, there are three substitutions of conserved amino acids nearby, according to the 3D channel structure: W280R, Y292C and W414G (W158, Y170 and W293 in TRPP3, respectively)^{51–53}. The first two changes are located at the α -helix close to the most external one containing the R399/R278 residue (in PKD2 and TRPP3, respectively). The third change affects the first amino acid of the loop that follows at this external α -helix. The three PKD2 mutations may have an important effect on the correct folding of this region of the polycystin domain, resulting in a hampered protein function that leads to polycystic kidney disease. These findings altogether suggest a relevant role of this polycystin domain area around the most external α -helix in the correct function of both PKD2 and TRPP3 channels.

TRPP3 channel gating in response to voltage or other physiological stimuli seems to involve transitions from closed to open triggered by pore coupling with VSD motion. Such coupling would be mediated by interactions between residues at the S5 and S4 α -helices of adjacent monomers, leading to the widening of both the lower and the upper gates¹⁸. Two lysines at the cytoplasmic end of S4 (K452 and K455), which are potentially stabilized by amino acids from S1 (Y107), S2 (Y366), and S3 (D390), appear to explain TRPP3 voltage-dependent properties¹⁸. In agreement with this view, charge-neutralizing mutations K452Q and K455Q strongly weaken TRPP3 voltage dependence, and these mutant channels do not reach maximal activation even upon very strong depolarisation levels (up to +300 mV)⁵⁴. By structural modelling it has been suggested that the functional impairment after neutralization of these two positively charged amino acids is due to a marked restriction of the voltage sensor movement⁵⁴. Accordingly, more recent structural modelling shows that these two gating charges (K452 and K455) can generate S4 transitional and lateral movements in response to membrane depolarization that in turn will be coupled to pore opening. During this VSD motion, the negatively charged D390 of S3 segment would stabilize K452 in the closed state and K455 in the open state. Furthermore, Y366 would interact sequentially first with K452 (in the deactivated state) and then with K455 (in the activated state) to provide the required energy for channel opening⁴⁵. This coupling between VSD movement and pore opening seems to be a general mechanism for TRPP3 activation applying also for other physiological stimuli. Thus, neutralization of one of these two gating charges strongly impairs channel gating regardless of the stimulus employed (voltage, hypoosmotic shock or temperature)⁴⁵.

It is remarkable that all VSD residues mentioned above are also conserved in PKD2 (K572 and K575 at S4, Y227 at S1, Y487 at S2 and D511 at S3)^{18,49}. Moreover, loss-of-function mutation D511V is a frequent pathogenic missense variant linked to ADPKD^{55,56}. Structural analysis of both PKD2 and TRPP3 channels suggest that the neutralization of this negatively charged residue at S3 by valine substitution most probably alter the stabilization of S4 lysine residues and perturb both VSD structure and function, and thus subsequent pore gating in response to voltage changes^{18,49}. The R378W substitution we found at high frequency in the Gumuz population, by affecting the S2–S3 intracellular linker may well have a similar effect by altering VSD conformation, in particular the correct position of the S3 segment, and thus impair the stabilization of S4 lysines. The structural analysis of R378W variant shows the preference of the closed conformation over the open. Although the models are not able to capture the full motion of S4 described by⁴⁵, they can be used to unveil notable differences between the wild type form and the R378W mutation, suggesting the larger stability of the D390–K542 hydrogen bond and the helix conformation of the linker S2–S3 in the mutant form, which may hamper the motion towards the open state. This might be the underlying cause of the strong reduction in channel activity we observed for TRPP3 channels carrying the R378W variant.

Given the strong linkage disequilibrium of the two ancestral and derived alleles, if PKD1L3/TRPP3 channel complex were somehow involved in sour taste detection^{25–28} it would be expected that individuals carrying the loss-of-function R278Q/R378W genetic variants had impaired the recognition of acidic stimuli across the tongue. However, we did not find significant sour recognition differences between genotypes (R278/R378 double homozygous, Q278/W378 double homozygous, or double heterozygous). This conclusion is limited by the low number of R278Q/R378W carriers (one in homozygosis and five in heterozygosis) in the group of Africans participating in the sour taste sensitivity test (as none of them had a Gumuz ancestry). Yet, the fact that all six R278Q/R378W carriers could detect acidic stimulation without apparent loss, argues against both any involvement of PKD1L3/TRPP3 in sour detection and the role of such physiological process in the evolutionary positive selection we report here. Altogether, these results agree with the current idea that considers the proton channel OTO1 as the real sour sensor in TRPP3-expressing, acid-detecting taste receptor cells^{32–36}.

Several functional roles have been proposed for TRPP3. Thus, this channel interacts and co-localizes with β_2 -adrenergic receptor (β_2 AR) on the primary cilia of neuronal cells in mice brain, where TRPP3 deficiency leads to the loss of β_2 AR on neuronal primary cilia, reduction of brain cAMP levels, increased neuronal excitability and higher susceptibility to seizure induced by the chemoconvulsant pentylenetetrazol⁵⁷. Besides, in zebrafish, *Trpp3* is expressed in cerebrospinal fluid-contacting neurons and required for the responses to pressure and maintenance of spine curvature. Indeed, zebrafish that do not express *Trpp3* present an exaggerated spine curvature similar to human kyphosis²¹. Lastly, it has been shown that although TRPP3 knockout mice had a decreased cytosolic and mitochondrial calcium entry in cardiomyocytes when they were fed with a normal-salt diet, TRPP3 absence results in a mitochondrial calcium overload in muscle cardiac cells of animals under high-salt diet. Such mitochondrial dysregulation in murine cardiomyocytes exacerbates high-salt diet-induced cardiac hypertrophy and dysfunction, revealing the protective effect of TRPP3 on cardiac hypertrophy in mice exposed to high-salt loading⁵⁸. Notably, rs17112895 and rs7909153 are marginally associated to heart rate recovery on the GWAS

Atlas database⁵⁹ with p-values of 0.00073 and 0.003, respectively, even if their association signals did not pass the genome-wide multiple testing on the original study (describing 23 loci associated with heart rate increase or recovery based on 58,818 UK Biobank individuals)⁶⁰.

Further research on physiological and pathological traits of the Gumuz population at both neurological and cardiovascular levels is therefore needed in order to unveil the reasons behind the strong positive selection of the loss-of-function R278Q/R378W genetic variants at the TRPP3 channel in Gumuz Ethiopians.

Methods

Genomic data. The Gumuz individuals analyzed in this study are comprised within the European Genome-phenome Archive (EGA) repository with accession number EGAS00001000238. The dataset consists of five East African populations (Amhara, Oromo, Somali, Wolayta and Gumuz) with 24 individuals each from Pagani et al. 2015. Additional samples from the 1000 Genomes Project (Gibbs et al. 2015) also from Pagani et al. were included to calculate positive selection statistics and allele frequencies. The genome assembly of the data is GRCh37 (hg19).

Genomic analysis of positive selection. To identify that positive selection has acted on the PKD2L1 gene, this region was investigated after the results presented in Walsh et al. 2020. The linkage disequilibrium-based test iHS was used to detect recent events of selection in the Gumuz population and other populations for comparison (CEU and YRI). The software rehh 2.0 (Gautier et al. 2017) was used to calculate iHS for all the genetic variants with a minor allele frequency higher than 0.05 and excluded a variant from the calculation if a 20 kb gap was found when calculating the extension of the haplotype homozygosity. The mean iHS score (average of iHS scores across SNPs) in each 30 kb window was calculated. To assess the statistical significance of the scores, an iHS critical value was defined after performing extensive genetic neutral simulations. The demographic model implemented in the simulations was a typical three-population model representing Africans, Europeans and Asians (Jouganous et al. 2017). The 99.99th percentile of the neutral distribution of iHS scores was used as critical value, being 2.54.

cDNA constructs. cDNA of the human TRPP3 wild-type (WT) channel, cloned into the plasmid pCMV6-AC-GFP was obtained from OriGene, Maryland, United States. TRPP3 mutant channels (R278Q, R378W and R278Q/R378W) were generated by site-directed mutagenesis of the human WT TRPP3 cDNA using the QuikChange XL kit (Stratagene), forward (5'-CTTCCAGGATCCCAACAGGGTAGTGCA-3') and a reverse (5'-TGCACTACCCTGTTGGGATCCTGGAAAG-3') primers for R278Q substitution, and forward (5'-AGCTCCACATTCCTGGCTTCGCTACCT-3') and a reverse (5'-AGGTAGCGAAGCCAGTGAATGTGGAGCT primers for R378W substitution. cDNA of mouse PKD1L3, cloned into pDisplay vector was a gift from Dr. Yong Yu (Department of Biological Sciences, St. John's University, 8000 Utopia Parkway, Queens, New York, 11439, USA). All cDNA clones were sequenced previously to the study in order to confirm their integrity.

Heterologous expression. HEK293 cells were transfected using a linear polyethylenimine (PEI) derivative, the polycation ExGen500 (Fermentas Inc., Hanover, MD, USA) as previously reported (eight equivalents PEI/3.3 µg DNA/dish)⁶¹. Co-transfection was performed using the ratio for GFP-fused TRPP3 (WT or either mutant R278Q, R378W, or R278Q/R378W), and PKD1L3 cDNA constructs of 1:1. For testing possible dominant negative action of the double mutant, cells were co-transfected with WT, R278Q/R378W TRPP3-GFP and PKD1L3 cDNA constructs with a ratio of 1:1:2. Transfected cells were incubated for 24 h and seeded on 35 mm diameter dishes Poly-D-Lysine-coated (Sigma). Experiments were performed 48 h post-transfection at room temperature (22–24 °C).

Whole-cell patch-clamp recordings and electrophysiological analysis. Patch electrodes had a resistance of 2 MΩ when filled with a pipette solution containing (in mM): 80 CsMES, 20 NaCl, 2 MgCl₂, 10 HEPES and 5 BAPTA (pH 7.4, adjusted using TrisBase). Osmolarity was adjusted to 300 mOsm/l adding D-mannitol. Recordings were performed at two different extracellular pH values, 7.4 or 9. In the first case, the external bath solution contained (in mM): 150 NaCl, 2 CaCl₂, 10 HEPES (310 mOsm/l and pH 7.4, adjusted using Tris-base). The basic external bath solution contained (in mM) 150 NaCl, 2 CaCl₂ and 10 TAPS (310 mOsm/l and pH 9, adjusted using Tris-base). All chemicals were obtained from Sigma-Aldrich.

TRPP3 activity was assessed using the whole-cell configuration of the patch-clamp technique on GFP-positive HEK293 cells 48 h after co-transfection of PKD1L3 with cDNA encoding wild-type (WT), R278Q, R378W, R278Q/R378W or both WT and R278Q/R378W TRPP3 channels. Whole-cell cationic currents were evoked by the application of 400 ms step pulses from – 100 to + 160 mV in 10 mV increments followed by a 200 ms postpulse to – 100 mV. Besides, spontaneous single-channel currents were also recorded in the whole-cell configuration by using a 40 ms gap-free protocol at several negative holding-potentials (– 60 mV, – 80 mV and – 100 mV). All recordings were obtained with a D-6100 Darmstadt amplifier (List Medical, Germany), sampled at 10 kHz and filtered at 1 kHz. The pClamp8 or pClamp10.5 software (Molecular Devices, Sunnyvale, CA, USA) was used for pulse generation, data acquisition and subsequent analysis. By employing Fetchan and pStat softwares, the amplitude of single-channel currents was measured as the peak-to-peak distance in Gaussian fits of amplitude histograms. Channel activity (NP_o , where N is the number of channels and P_o is the open probability of a single channel) at negative voltages was calculated by dividing the mean current amplitude of recordings lasting 40 s by the single-channel amplitude obtained from the same traces.

Immunofluorescence. Transfected HEK293 cells were incubated with Concanavalin A (ConA)-Rhodamine (Molecular Probes Ref. C860) for 20 min, rinsed with PBS and fixed with PFA 4%. For immunofluorescence staining cells were fixed 24 or 48 h post-transfection with paraformaldehyde 4% for 30 min. Upon thorough rinsing, cells were treated with a blocking solution of 5% bovine serum albumin (BSA) and 3% foetal bovine serum with or without triton 0.1% (for permeabilization). Anti-GFP antibody (Life Technologies, Cat # A-11122) was incubated overnight and Alexa-488 secondary antibody was used for detection. DAPI was used for nuclear staining.

Confocal imaging was performed to detect membrane expression of the GFP in the membrane (shown as ConA-Rhodamine staining) using the Leica SP5. Manders colocalization coefficient was calculated using home-made pipeline of CellProfiler (v.3.1.9)⁶². Only cells that expressed TRPP3-GFP above a threshold of intensity per unit area were considered. Pixel intensity profiles were obtained using ImageJ.

Statistical analysis. Data are presented as the means \pm S.E.M. Statistical tests included one-tail Student's t-test, one-tail Mann–Whitney U-test, Kruskal–Wallis test followed by Dunn post hoc test, or one-way ANOVA followed by Bonferroni's post hoc test, as appropriate. Differences were considered significant if $P < 0.05$.

Analysis of the 3D structure. The structures of PDB⁶³ with code 5Z1W¹⁸, 5T4D⁴⁹ and 5MKE⁶⁴ were used to model the close conformation of TRPP3, and the structures with code 6DU8¹⁹, 6D1W⁶⁵ and 5Z1W¹⁸ for the open state. Sequences were compared with ClustalOmega⁶⁶, structural models were obtained with MODELLER⁶⁷ and optimized with Rosetta. Superposition and visualization of structures were obtained with Matchmaker and Chimera⁶⁸. Form with R278Q substitution was analysed by visual inspection of the neighbourhood of R278 after superimposing the structures 5Z1W and 6DU8. Conformations of the closed-state of wild type and R378W substitution forms were optimized with FIXBB and RELAX, from Rosetta package⁶⁹. Several conformations of the linker between S2 and S3 were modelled for the open-state conformation (a total of 2500 structures). The loop between amino-acid 378 and 384 was remodelled without template to allow for the flexible changes caused upon substitution. The side-chains of the structures were further relaxed with 10 iteration cycles of FIXBB. We calculated the distribution of distances between the amino chemical groups of K452 and K455 and the oxygens O δ 1 and O δ 2 of the carboxyl group of D390, grouping them in hydrogen-bonding (at less than 3.5Å) and non hydrogen-bonding (at more than 3.5Å), where we use an extra margin of 0.5Å to accept the hydrogen bond. We used Chimera for the visual inspection and analysis of selected structures with and without hydrogen bonds D390-K452 and D390-K455 of the open state conformations. We have to note that this is a study on static models, different than a molecular dynamics simulation of the protein embedded in a solvated environment. The study characterizes the main structural features of several models obtained by simulated annealing and optimization of the protein conformation.

Sensory data and sour taste sensitivity test. Sensory data was obtained from a total of 44 participants (22 males and 22 females ranging from 15 to 57 years old). 13 participants were self-reported as carrying West African ancestry, 18 European ancestry and 13 Afroasiatic Ethiopian ancestry. In this study it was not possible to obtain Gumuz participants among the Ethiopians.

A modified Harris-Kalmus test⁴⁶ was designed to determine the recognition threshold of sour taste of individuals. Citric acid (CA) or E-330 was used as stimuli since it is a harmless food additive which maximum doses are catalogued as “quantum satis” after food and health agencies. The citric acid used for this study was bought from Sigma-Aldrich with reference C0759.

The stimuli used were solutions of CA in Millipore-filtered, deionized water. They are prepared 24–48 h in advance of use, stored under refrigeration in amber glass bottles and warmed to room temperature before use. Water blanks are Millipore-filtered deionized water stored and handled the same way. The CA concentrations ranged from 5×10^{-2} to 1 mM in a 5×10^{-2} mM step.

As previously reported⁷⁰, during each trial, subjects received a 10 mL sample in a medicine cup. After holding the sample in their mouth for at least 5 s the subjects attempted to identify the quality of taste (sweet, sour, bitter, salty, water). Subjects rinsed at least twice with deionized water. The threshold run begins with the lowest concentration. Subjects sampled each concentration once in ascending order until they identify the target quality “sour”. We next ensure that the subjects experienced a reliable taste sensation with a sorting task. Subjects receive 3 samples at the concentration at which they identified the target quality with 3 more blank samples. All 6 samples are presented at the same time in a random position. The subjects are asked to sort 3 cups into “tastes” and 3 cups into “waters”. If the subject can correctly sort the 6 samples in 2 consecutive trials, the threshold run ends. If a subject fails to sort correctly the sorting task is repeated at the next higher concentration. The concentration that first allows successive correct sorts is the taste quality recognition threshold.

Saliva sample collection, DNA extraction, genotyping and allele calling. After having written informed consent from the participants and/or their parents, we obtained a saliva sample from which DNA was extracted using NaCl 6 M solution and subsequently precipitating the DNA on isopropanol⁷¹. Genotyping was performed through Sanger sequencing of PCR fragments. The primers used for the PCR are the following:

Target	Forward primer	Reverse primer
rs17112895 and rs7909153	5'-AAGATGACCACCAGGTCCAG-3'	5'-TTGATACGTGGAGCAGCAAG-3'
rs7909153	5'-GTCAGAAGGAGGGCTTAGGG-3'	5'-CTGATCCGCTATGTCAGCAA-3'

Target	Forward primer	Reverse primer
rs17112895	5'- GCACCTCCTGTAGCTGGAAA-3'	5'- CATTGCCTTGGGAGATGAAC-3'

The primers used for sequencing are:

	Sequencing Primers
Seq1	Reverse 5'-ACTGAACTCAAATGTAAATCTG-3'
Seq2	Forward 5'-CATGCACAAGACGGGAAC-3'
Seq3	Reverse 5'-CATTGCCTTGGGAGATGAAC-3'

The PCR reactions were sent to Eurofins to perform Sanger sequencing. The program 4Peaks was used to visualize the electropherograms and manually validate that the variants were appropriately called for each sample.

Ethical approval. The study of sour test sensitivity, with reference number 2018/8294/I, was approved by the ethical committee CEIm-Parc de Salut MAR and written informed consent to participate in the study and for the saliva donation was obtained from all participants and/or their parents. All methods to obtain samples and data were performed in accordance with the relevant guidelines and regulations, including the Helsinki Declaration of 1964, as revised in October 2013 (Fortaleza, Brazil).

Data availability

The authors declare that the data supporting the findings of this study are available within the article and its Supplementary Information file, or are available from the authors upon request.

Received: 6 May 2020; Accepted: 20 November 2020

Published online: 02 December 2020

References

- Sabeti, P. C. *et al.* Positive natural selection in the human lineage. *Science* **312**, 1614–1620 (2006).
- Vitti, J. J., Grossman, S. R. & Sabeti, P. C. Detecting natural selection in genomic data. *Annu. Rev. Genet.* **47**(1), 97–120 (2013).
- Scheinfeldt, L. B. & Tishkoff, S. A. Recent human adaptation: Genomic approaches, interpretation and insights. *Nat. Rev. Genet.* **14**(10), 692–702 (2013).
- Fan, S., Hansen, M. E. B., Lo, Y. & Tishkoff, S. A. Going global by adapting local: a review of recent human adaptation. *Science* **354**, 54–59 (2016).
- Booker, T. R., Jackson, B. C. & Keightley, P. D. Detecting positive selection in the genome. *BMC Biol.* **15**(1), 98 (2017).
- Tishkoff, S. A. *et al.* Convergent adaptation of human lactase persistence in Africa and Europe. *Nat. Genet.* **39**(1), 31–40 (2007).
- Enattah, N. S. *et al.* Identification of a variant associated with adult-type hypolactasia. *Nat. Genet.* **30**(2), 233–237 (2002).
- Simonson, T. S. *et al.* Genetic evidence for high-altitude adaptation in Tibet. *Science* **329**, 72–75 (2010).
- Beall, C. M. *et al.* Natural selection on EPAS1 (HIF2alpha) associated with low hemoglobin concentration in Tibetan highlanders. *Proc. Natl. Acad. Sci. U.S.A.* **107**(25), 11459–11464 (2010).
- Huerta-Sánchez, E. *et al.* Genetic signatures reveal high-altitude adaptation in a set of ethiopian populations. *Mol. Biol. Evol.* **30**(8), 1877–1888 (2013).
- Yi, X. *et al.* Sequencing of 50 human exomes reveals adaptation to high altitude. *Science* **329**, 75–78 (2010).
- Fumagalli, M. *et al.* Greenlandic Inuit show genetic signatures of diet and climate adaptation. *Science* **349**, 1343–1347 (2015).
- Dobon, B., Rossell, C., Walsh, S. & Bertranpetit, J. Is there adaptation in the human genome for taste perception and phase I biotransformation? *BMC Evol. Biol.* **19**(1), 39 (2019).
- Chikazoe, J., Lee, D. H., Kriegeskorte, N. & Anderson, A. K. Distinct representations of basic taste qualities in human gustatory cortex. *Nat. Commun.* **10**(1), 1–8 (2019).
- DeSimone, J. A., Heck, G. L. & DeSimone, S. K. Active ion transport in dog tongue: a possible role in taste. *Science* **214**, 1039–1041 (1981).
- Chandrashekar, J. *et al.* The taste of carbonation. *Science* **326**, 443–445 (2009).
- Zocchi, D., Wennemuth, G. & Oka, Y. The cellular mechanism for water detection in the mammalian taste system. *Nat. Neurosci.* **20**(7), 927–933 (2017).
- Su, Q. *et al.* Cryo-EM structure of the polycystic kidney disease-like channel PKD2L1. *Nat. Commun.* **9**(1), 1–12 (2018).
- Hulse, R.E., Li, Z., Huang, R.K., Zhang, J. & Clapham, D.E. Cryo-EM structure of the polycystin 2-1L ion channel. *ELife* **7**, pii: e36931 (2018).
- Shimizu, T., Janssens, A., Voets, T. & Nilius, B. Regulation of the murine TRPP3 channel by voltage, pH, and changes in cell volume. *Pflügers Arch. Eur. J. Physiol.* **457**(4), 795–807 (2009).
- Sternberg, J. R. *et al.* Pkd2l1 is required for mechanoreception in cerebrospinal fluid-contacting neurons and maintenance of spine curvature. *Nat. Commun.* **9**(1), 3804 (2018).
- Shimizu, T., Higuchi, T., Fujii, T., Nilius, B. & Sakai, H. Bimodal effect of alkalization on the polycystin transient receptor potential channel, PKD2L1. *Pflügers Arch. Eur. J. Physiol.* **461**(5), 507–513 (2011).
- Higuchi, T., Shimizu, T., Fujii, T., Nilius, B. & Sakai, H. Gating modulation by heat of the polycystin transient receptor potential channel PKD2L1 (TRPP3). *Pflügers Arch. Eur. J. Physiol.* **466**(10), 1933–1940 (2014).
- Ishimaru, Y. *et al.* Interaction between PKD1L3 and PKD2L1 through their transmembrane domains is required for localization of PKD2L1 at taste pores in taste cells of circumvallate and foliate papillae. *FASEB J.* **24**(10), 4058–4067 (2010).
- Ishimaru, Y. *et al.* Transient receptor potential family members PKD1L3 and PKD2L1 form a candidate sour taste receptor. *Proc. Natl. Acad. Sci. U.S.A.* **103**(33), 12569–12574 (2006).
- Inada, H. *et al.* Off-response property of an acid-activated cation channel complex PKD1L3-PKD2L1. *EMBO Rep.* **9**(7), 690–697 (2008).
- Kawaguchi, H. *et al.* Activation of polycystic kidney disease-2-like 1 (PKD2L1)-PKD1L3 complex by acid in mouse taste cells. *J. Biol. Chem.* **285**(23), 17277–17281 (2010).
- Horio, N. *et al.* Sour taste responses in mice lacking pkd channels. Meyerhof W, editor. *PLoS One* **6**(5), e20007 (2011).

29. Huque, T. *et al.* Sour ageusia in two individuals implicates ion channels of the ASIC and PKD families in human sour taste perception at the anterior tongue. *PLoS ONE* **4**(10), e7347 (2009).
30. Huang, A. L. *et al.* The cells and logic for mammalian sour taste detection. *Nature* **442**(7105), 934–938 (2006).
31. Nelson, T. M. *et al.* Taste function in mice with a targeted mutation of the Pkd1l3 gene. *Chem. Senses* **35**(7), 565–577 (2010).
32. Tu, Y. H. *et al.* An evolutionarily conserved gene family encodes proton-selective ion channels. *Science* **359**, 1047–1050 (2018).
33. Zhang, J. *et al.* Sour sensing from the tongue to the brain. *Cell* **179**(2), 392–402 (2019).
34. Teng, B. *et al.* Cellular and neural responses to sour stimuli require the proton channel Otop1. *Curr. Biol.* **29**(21), 3647–3656 (2019).
35. Hurler, B. *et al.* Non-syndromic vestibular disorder with otoconial agenesis in tilted/mergulhador mice caused by mutations in otopetrin 1. *Hum. Mol. Genet.* **12**(7), 777–789 (2003).
36. Hughes, I. *et al.* Identification of the Otopetrin Domain, a conserved domain in vertebrate otopetrins and invertebrate otopetrin-like family members. *BMC Evol. Biol.* **8**(1), 41 (2008).
37. Ronen, R., Udpa, N., Halperin, E. & Bafna, V. Learning natural selection from the site frequency spectrum. *Genetics* **195**, 181–193 (2013).
38. Sabeti, P. C. *et al.* Detecting recent positive selection in the human genome from haplotype structure. *Nature* **419**(6909), 832–837 (2002).
39. Casillas, S. *et al.* PopHuman: the human population genomics browser. *Nucleic Acids Res.* **46**(D1), D1003–D1010 (2018).
40. Pagani, L. *et al.* Ethiopian genetic diversity reveals linguistic stratification and complex influences on the Ethiopian gene pool. *Am. J. Hum. Genet.* **91**, 83–96 (2012).
41. Jakobsson, M. *et al.* Genotype, haplotype and copy-number variation in worldwide human populations. *Nature* **451**(7181), 998–1003 (2008).
42. Schlebusch, C. M. *et al.* Genomic variation in seven Khoe-San groups reveals adaptation and complex African history. *Science* **338**, 374–379 (2012).
43. Bosch, E. *et al.* Decay of linkage disequilibrium within genes across HGDP-CEPH human samples: most population isolates do not show increased LD. *BMC Genomics* **10**, 338 (2009).
44. Albers, P. K. & McVean, G. Dating genomic variants and shared ancestry in population-scale sequencing data. *PLoS Biol.* **18**(1), e3000586 (2020).
45. Ng, L. C. T., Vien, T. N., Yarov-Yaroyov, V. & DeCaen, P. G. Opening TRPP2 (PKD2L1) requires the transfer of gating charges. *Proc. Natl. Acad. Sci. U.S.A.* **116**(31), 15540–15549 (2019).
46. Harris, H. & Kalmus, H. The measurement of taste sensitivity to phenylthiourea. *Ann. Eugen.* **15**(1), 24–31 (1949).
47. Baliga, S., Muglikar, S. & Kale, R. Salivary pH: A diagnostic biomarker. *J. Indian Soc. Periodontol.* **17**(4), 461–465.
48. Zhou, X. *et al.* Cryo-EM structures of the human endolysosomal TRPML3 channel in three distinct states. *Nat. Struct. Mol. Biol.* **24**(12), 1146–1154 (2017).
49. Shen, P. S. *et al.* The structure of the polycystic kidney disease channel PKD2 in lipid nanodiscs. *Cell* **167**(3), 763–773 (2016).
50. Bergmann, C. *et al.* Polycystic kidney disease. *Nat. Rev. Dis. Primers.* **4**(1), 50 (2018).
51. Tan, Y. C. *et al.* Aberrant PKD2 splicing due to a presumed novel missense mutation in autosomal-dominant polycystic kidney disease. *Clin. Genet.* **80**(3), 287–292 (2011).
52. Veldhuisen, B. *et al.* A spectrum of mutations in the second gene for autosomal dominant polycystic kidney disease (PKD2). *Am. J. Hum. Genet.* **61**(3), 547–555 (1997).
53. The Autosomal Dominant Polycystic Kidney Disease (ADPKD) Database (PKDB).
54. Numata, T. *et al.* Integrative approach with electrophysiological and theoretical methods reveals a new role of S4 positively charged residues in PKD2L1 channel voltage-sensing. *Sci. Rep.* **7**(1), 9760 (2017).
55. Reynolds, D. M. *et al.* Aberrant splicing in the PKD2 gene as a cause of polycystic kidney disease. *J. Am. Soc. Nephrol.* **10**(11), 2342–2351 (1999).
56. Ma, R. *et al.* PKD2 functions as an epidermal growth factor-activated plasma membrane channel. *Mol. Cell. Biol.* **25**(18), 8285–8298 (2005).
57. Yao, G. *et al.* Disruption of polycystin-L causes hippocampal and thalamocortical hyperexcitability. *Hum. Mol. Genet.* **25**(3), 448–458 (2016).
58. Lu, Z. *et al.* Deficiency of PKD2L1 (TRPP3) exacerbates pathological cardiac hypertrophy by augmenting NCX1-mediated mitochondrial calcium overload. *Cell Rep.* **24**, 1639–1652 (2018).
59. Watanabe, K. *et al.* A global overview of pleiotropy and genetic architecture in complex traits. *Nat. Genet.* **51**(9), 1339–1348 (2019).
60. Verweij, N., van de Vegte, Y. J. & van der Harst, P. Genetic study links components of the autonomous nervous system to heart-rate profile during exercise. *Nat. Commun.* **9**(1), 898 (2018).
61. Izquierdo-Serra, M. *et al.* Stroke-like episodes and cerebellar syndrome in phosphomannomutase deficiency (PMM2-CDG): Evidence for hypoglycosylation-driven channelopathy. *Int. J. Mol. Sci.* **19**(2), E619 (2018).
62. McQuin, C. *et al.* CellProfiler 3.0: Next-generation image processing for biology. *PLoS Biol.* **16**(7), e2005970 (2018).
63. Burley, S. K. *et al.* RCSB Protein Data Bank: Biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. *Nucleic Acids Res.* **47**(D1), D464–D474 (2019).
64. Wilkes, M. *et al.* Molecular insights into lipid-assisted Ca²⁺ regulation of the TRP channel Polycystin-2. *Nat. Struct. Mol. Biol.* **24**(2), 123–130 (2017).
65. Zheng, W. *et al.* Hydrophobic pore gates regulate ion permeation in polycystic kidney disease 2 and 2L1 channels. *Nat. Commun.* **9**(1), 2302 (2018).
66. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
67. Webb, B. & Sali, A. Protein structure modeling with MODELLER. In: *Methods in Molecular Biology*. Humana Press Inc. pp. 39–54 (2017).
68. Yang, Z. *et al.* UCSF Chimera, MODELLER, and IMP: an integrated modeling system. *J. Struct. Biol.* **179**(3), 269–278 (2012).
69. Conway, P., Tyka, M. D., DiMaio, F., Kondering, D. E. & Baker, D. Relaxation of backbone bond geometry improves protein energy landscape modeling. *Protein Sci.* **23**(1), 47–55 (2014).
70. Wise, P. M. & Breslin, P. A. S. Individual differences in sour and salt sensitivity: detection and quality recognition thresholds for citric acid and sodium chloride. *Chem. Senses* **38**, 333–342 (2013).
71. Solé-Morata, N., Bertranpetit, J., Comas, D. & Calafell, F. Y-chromosome diversity in Catalan suRNAme samples: insights into suRNAme origin and frequency. *Eur. J. Hum. Genet.* **23**(11), 1549–1557 (2015).
72. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/> (2019).

Acknowledgements

We thank Dr. F. Rubio-Moscardó and Ms. C. Plata for excellent technical assistance. This work was funded by the Spanish Ministry of Science and Innovation, the State Research Agency (AEI, Agencia Estatal de Investigación) and FEDER Funds (Fondo Europeo de Desarrollo Regional): Grants BFU2016-77961-P to J.B. and E.B., RTI2018-094809-B-I00 to J.M.F.F., PID2019-110933GB-I00/AEI/10.13039/501100011033 to E.B. and J.B., and

CEX2018-000792-M through the “María de Maeztu” Programme for Units of Excellence in R&D to “Departament de Ciències Experimentals i de la Salut”. With the support of Secretaria d’Universitats i Recerca del Departament d’Economia i Coneixement de la Generalitat de Catalunya (GRC 2017 SGR 702). S.W. has had a F.P.I. grant (BES-2014-068994) and M.I.-S. a “Juan de la Cierva-Incorporación” Fellowship funded by the Spanish Ministry of Science and Innovation.

Author contributions

J.B. and J.M.F.-F. conceived and designed the project. S.W., M.I.-S., S.A., A.E., M.LL., R.M., E.B., B.O., and J.M.F.-F. contributed to data acquisition and analysis. All authors contributed to interpretation of the data and manuscript edition. S.W., J.B. and J.M.F.F. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-78081-z>.

Correspondence and requests for materials should be addressed to J.B. or J.M.F.-F.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020