



OPEN ACCESS

Edited by:

Manuel Corpas,
Cambridge Precision Medicine,
United Kingdom

Reviewed by:

Ranjit Das,
Yenepoya University, India
Naim Akhtar Khan,
Université de Bourgogne, France

***Correspondence:**

Pankaj Jha
panjha@gmail.com
Mitali Mukerji
mitali@iitj.ac.in

†Present address:

P. Prakrithi,
Indian Institute of Technology,
University of Queensland- IIT Delhi
Academy of Research (UQIDAR),
New Delhi, India
Pankaj Jha,
MedGenomes Labs Ltd., Bangalore,
India
Mitali Mukerji,
Department of Bioscience and
Bioengineering, Indian Institute of
Technology, Jodhpur, India

‡These authors have contributed
equally to this work and share first
authorship

Specialty section:

This article was submitted to
Human and Medical Genomics,
a section of the journal
Frontiers in Genetics

Received: 17 February 2022

Accepted: 31 May 2022

Published: 12 July 2022

Citation:

Prakrithi P, Jha P, Jaiswal J, Sharma D,
Bhoyar RC, Jain A, Imran M,
Senthilvel V, Divakar MK, Mishra A,
Scaria V, Sivasubbu S and Mukerji M
(2022) Landscape of Variability in
Chemosensory Genes Associated
With Dietary Preferences in Indian
Population: Analysis of 1029
Indian Genomes.
Front. Genet. 13:878134.
doi: 10.3389/fgene.2022.878134

Landscape of Variability in Chemosensory Genes Associated With Dietary Preferences in Indian Population: Analysis of 1029 Indian Genomes

P. Prakrithi^{1†‡}, Pankaj Jha^{1†‡*}, Jushta Jaiswal¹, Disha Sharma^{1,2}, Rahul C. Bhoyar¹, Abhinav Jain^{1,2}, Mohamed Imran^{1,2}, Vigneshwar Senthilvel^{1,2}, Mohit Kumar Divakar^{1,2}, Anushree Mishra¹, Vinod Scaria^{1,2}, Sridhar Sivasubbu^{1,2} and Mitali Mukerji^{1,3*†‡}

¹CSIR- Institute of Genomics and Integrative Biology, Delhi, India, ²Academy of Scientific and Innovative Research (AcSIR), Ghaziabad, India, ³Department of Bioscience and Bioengineering, Indian Institute of Technology, Jodhpur, India

Perception and preferences for food and beverages determine dietary behaviour and health outcomes. Inherent differences in chemosensory genes, ethnicity, geo-climatic conditions, and sociocultural practices are other determinants. We aimed to study the variation landscape of chemosensory genes involved in perception of taste, texture, odour, temperature and burning sensations through analysis of 1,029 genomes of the IndiGen project and diverse continental populations. SNPs from 80 chemosensory genes were studied in whole genomes of 1,029 IndiGen samples and 2054 from the 1000 Genomes project. Population genetics approaches were used to infer ancestry of IndiGen individuals, gene divergence and extent of differentiation among studied populations. 137,760 SNPs including common and rare variants were identified in IndiGenomes with 62,950 novel (46%) and 48% shared with the 1,000 Genomes. Genes associated with olfaction harbored most SNPs followed by those associated with differences in perception of salt and pungent tastes. Across species, receptors for bitter taste were the most diverse compared to others. Three predominant ancestry groups within IndiGen were identified based on population structure analysis. We also identified 1,184 variants that exhibit differences in frequency of derived alleles and high population differentiation ($F_{ST} \geq 0.3$) in Indian populations compared to European, East Asian and African populations. Examples include *ADCY10*, *TRPV1*, *RGS6*, *OR7D4*, *ITPR3*, *OPRM1*, *TCF7L2*, and *RUNX1*. This is a first of its kind of study on baseline variations in genes that could govern cuisine designs, dietary preferences and health outcomes. This would be of enormous utility in dietary recommendations for precision nutrition both at population and individual level.

Keywords: chemosensory perceptions, genetic diversity, population differentiation, nutrigenomics, food preferences, precision nutrition

INTRODUCTION

Indian cuisines are extremely diverse across the country and reflect inherent differences in ethnic, geo-climatic, and socio-cultural practices. These could have co-evolved in populations with differences in perception of taste, odour, and flavour that guide dietary preferences and behaviour. As dietary behaviour impacts chronic and complex diseases, it becomes more important to understand the variability in chemosensory perception for precision nutrition recommendations. The ability to perceive distinct tastes (sweet, sour, pungent, bitter, salty, umami), flavour and texture has reported genetic underpinnings (Dotson et al., 2012). Variability in these genes is associated with dietary preferences, food intake behaviour and risk for diseases (Newcomb et al., 2012; Diószegi et al., 2019). For instance, individuals classified as “Supertasters,” “Moderate-Tasters” and “Non-Tasters” differ in their ability to perceive bitter tasting foods with supertasters having extreme sensitivity to very low thresholds due to variations in a bitter taste receptor gene *TAS2R38* (Duffy et al., 2004). Because of this, supertasters avoid alcohol, caffeine containing beverages and also healthy cruciferous vegetables and have higher preference for sweet foods. Similarly, variability in olfactory genes *OR7D4*, *OR11H7P*, *OR6A2* govern differences in preference for pork meat, cheesy food, and cilantro respectively (Menashe et al., 2003; Eriksson et al., 2012; Lunde et al., 2012). Variations in genes, for example, *CD36* also govern the ability to perceive textures of fatty food that impart flavour in food preparations (Keller et al., 2012). A significant number of chemosensory SNPs have also been associated with different non-metabolic diseases. For instance, the human bitter taste receptors (T2Rs) are implicated in chronic rhinosinusitis, asthma, thyroid levels, cystic fibrosis, and risk for cancer (Clark et al., 2015; Shaik et al., 2016). Similarly, disrupting the function of sodium channels such as the transient receptor potential cation subfamily V member 1 channel (*TRPV1*) and ENaC (*SCNN1B* gene) and variants associated with these genes modulate the perception for salt taste (Dias et al., 2013). A recent study of 2,854 SNPs in 50 taste-related genes reported association of a taste 1 receptor member 2 (*TAS1R2*)- rs11261087 variant with pancreatic cancer risk (Gentiluomo et al., 2019). Understanding baseline variations in chemosensory genes is likely to form an integral basis for precision nutrition for management of health and disease conditions.

Whole genome sequencing across different world populations provides a huge amount of baseline variation data. India houses a major fraction of global diversity. The IndiGen initiative of sequencing of 1,029 genomes has recently filled the gap in the variation data from Indian populations that are relatively under-represented in the global databases (Jain et al., 2021). We utilized this data to study the extent of variability in chemosensory genes that are likely to have evolved differently due to varied ethnicity, geo-climatic conditions, socio-cultural and culinary practices across India. We report patterns of genetic variations in 80 chemosensory genes across the representative IndiGenomes derived from diverse genetic clusters of the Indian reference populations (Indian Genome Variation Consortium, 2008). Our study shows differences from the global populations of

the 1000 Genomes Project (The 1000 Genomes Project Consortium, 2015) and also highlights some key genes that could have evolved differently in Indian populations (Indian Genome Variation Consortium, 2008). This study provides important baseline genetic information on chemosensory genes that would be of great importance in precision nutrition studies.

MATERIALS AND METHODS

Chemosensory Gene Set

80 genes associated with variations in taste perceptions and food preferences were manually curated from literature (**Supplementary Table S1**). The genes included have been associated with variations in 1) perception of five primary tastes: bitter, sweet, umami, salty, and sour 2) perceptions as well as preferences for food of different textures and odours and 3) variations in temperature thresholds and burning sensations in the mouth. Nearly 230 variants (SNPs) in these genes have earlier reported associations (**Supplementary Table S2**).

Study Populations and Datasets

The study population included data of 1) 1,029 unrelated healthy Indian individuals from the IndiGen Genome Project (Jain et al., 2021) 2) 390 samples from 28 diverse populations of India included in the Indian Genome Variation (IGV) Consortium (Indian Genome Variation Consortium, 2008) 3) 2,504 samples from the 1000 Genomes project (The 1000 Genomes Project Consortium, 2015). The SAS populations of the 1,000 Genomes include four representative populations from South Asian ethnic groups and were also used to separately compare with the IndiGen genomes.

All the variants in the genes listed in **Supplementary Table S1** and the variations residing in their flanking regions of 10 kb were retrieved from 1,029 unrelated healthy Indian individuals from the IndiGen Genome Project in variant call format (VCF, GRCh38/hg38). The IndiGen variants were prioritized after performing genotype level missingness tests. Variants with more than 10% missing genotypes were not included in the analysis. A total of 137,760 single nucleotide polymorphisms (SNPs) were finally used for analysis.

SNP genotypes were also retrieved for the same set of genes from the 1000 Genomes Project Phase 3 data (reference build GRCh37/hg19). 235,931 variants within 80 genes were obtained from 1,000 Genomes. 63,600 SNPs (46%) in the IndiGen dataset were common with 1,000 Genomes datasets. PLINKv1.09 (Purcell et al., 2007) and VCFTOOLS (Danecek et al., 2011) were used for variant filtering. We used the Liftover tool of UCSC (Hinrichs, 2006) for conversion of genomic coordinates of 1,000 Genomes data from hg19 to hg38. There were 2,898 variants that were shared between IndiGen and IGV. These were studied to estimate the frequency spectrum of chemosensory genes in diverse Indian ethnic groups. The frequency of the variants was calculated using VCFTOOLS and the SNPs were classified into three groups as common variants, rare variants and private variants based on their Minor Allele Frequency (MAF) cut-off of

5% and above, 1%–5%, and below 1% respectively. The number of variants in each group were compared between the 1,000 Genomes and IndiGen data to identify overlaps and differences between the two datasets as well as novel variants in the IndiGen data.

We further affirmed the novelty of unreported variants in IndiGen in two other publically available sources, the Simon Genome diversity Project (SGDP); $n = 263$ (37 of South Asian ancestry) and the Human Genome Diversity Project (HGDP); $n = 929$ (188 of South Asian ancestry).

Variant Annotation

ANNOVAR (Wang et al., 2010) was used to annotate the variants using dbSNP v150 and RefGene annotations. The functional impact of exonic variations was estimated using CADD_phred scores annotated with the hg38_dbnsfp35a.txt database file. The variants that had a score above or equal to 20 were considered potentially deleterious. To assess differential enrichment of SNPs in genes and regulatory regions, the average percentage of SNPs in the genic and 10 kb flanking regions normalized to their lengths, were estimated across the different perception categories (bitter, sweet/fatty, salty/pungent, sour/astringent, and olfactory).

Population Genetic Analysis

To estimate the population substructure within the IndiGen dataset, we carried out a conjoint analysis with the genome-wide common variants data that were shared between Indian Genome Variation (IGV), IndiGenomes and 1,000 Genomes populations. Combined variants ($n = 100,800$ SNPs) were used for principal component analysis (PCA) using PLINK v1.9 and the results were visualized using R-packages. To classify the IndiGen individuals based on their genetic ancestry, ADMIXTURE and PCA analyses were performed on a common dataset between IndiGen, 1,000 Genomes populations (YRI, CEU, FIN, CHB, and JPT) and samples from the Indian Genome Variation Consortium. The analysis was repeated with different sets of SNPs after applying HWE filters (p -value = 0.001) and LD r^2 cut-off of 0.1, 0.2, 0.5. To assign the IndiGen samples to sub-populations of the IGV, PCA analysis followed by KNN calculations were performed. We further used the maximum likelihood methods implemented in ADMIXTURE software (Alexander et al., 2009) to estimate the individual-wise ancestry of IndiGen samples (Jain et al., 2021). The admixture analysis for each k was replicated 10 times and the k run which yielded the lowest CV error was selected to define the population structure. We also looked at signatures of selection in the chemosensory genes using different methods. The population differentiation (F_{ST}) was estimated based on the Weir and Cockerham method (1983) using VCFtools. Pairwise F_{ST} analysis was performed using the hierfstat of R package (Goudet, 2005). The ancestral allele information was obtained from the 1000 Genomes Project. Annotation of ancestral allele information was carried out using VCFTOOLS and the dataset for subsequent analysis was generated with PLINK. 48,824 sites with known ancestral alleles were used for derived allele frequency calculation.

Divergence of Chemosensory Genes During Evolution

Protein sequences of chemosensory genes from 24 species that include primates and non-primates were retrieved using BLAST (Altschul et al., 1990). A phylogenetic analysis was performed to study the divergence patterns. A similar number of a random set of non-chemosensory genes were also included.

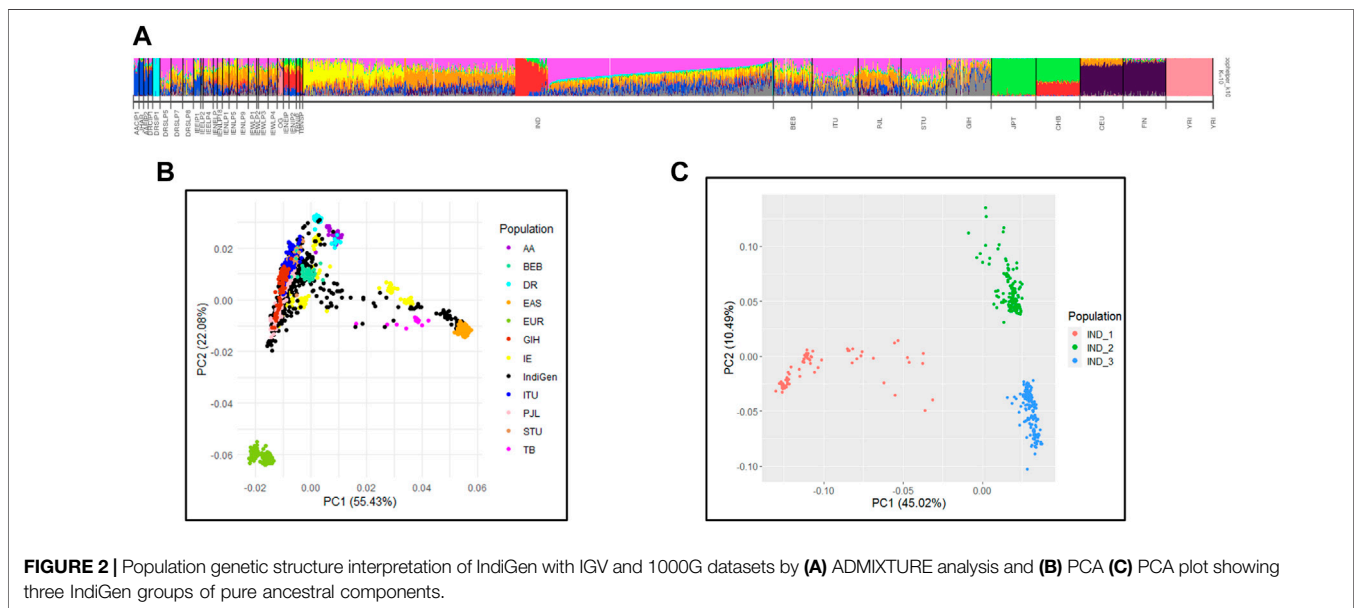
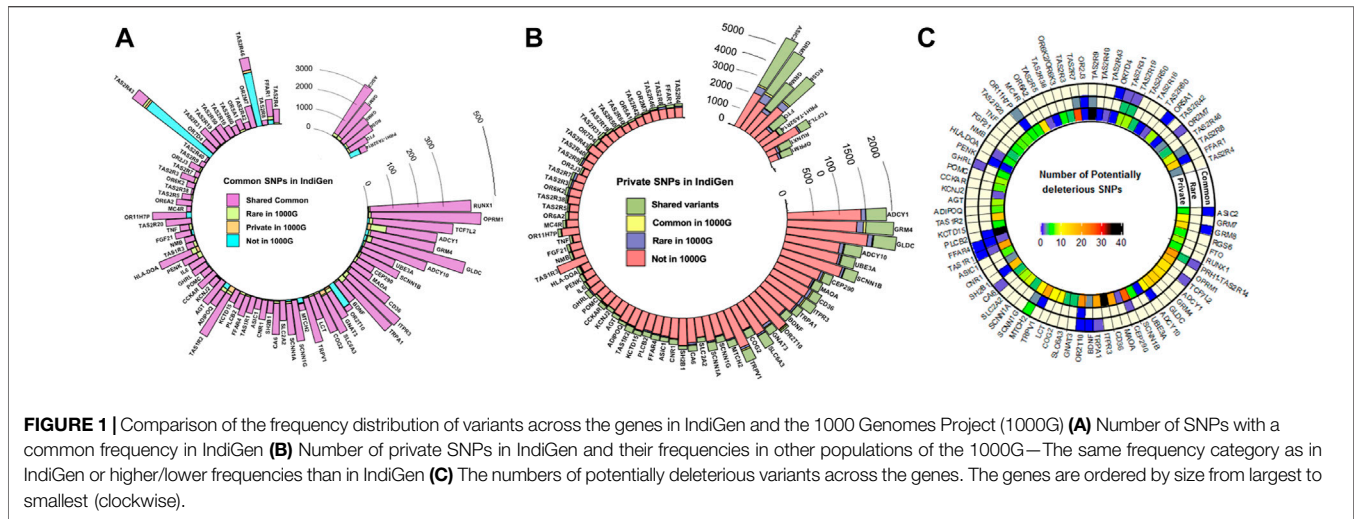
RESULTS

Pattern of Variations in Chemosensory Genes Among the Indian Genomes

A total of 137,760 SNPs in the 80 chemosensory genes including 10 kb upstream and downstream regions were retrieved from 1,000 IndiGen genomes. In these genes, the 1,000 Genomes Phase 3 data ($n = 2015$) and SAS population ($n = 489$) harbored a corresponding set of 203,560 and 75,555 SNPs respectively. There were ~33,000 insertion and deletion polymorphisms comprising 18% of the total variants. These were not included in the present analysis. 68% of the variants were intronic and 2% were exonic as anticipated. Nearly 75% of variants were C to T transitions. The details of annotation and distribution of variants in chemosensory genes in IndiGen data have been depicted in **Supplementary Figures S1A–D**. Surprisingly, out of the 2,528 exonic variants, the fraction of nonsynonymous (NS) variants was higher than synonymous. A smaller fraction of the variants in the exonic regions were also frameshift causing insertions and deletions. There were nearly 65 nonsense mutants in the NS class and few splice site variants. Ten top mutated genes had more than 35 variants and included *CEP290*, *ITPR3*, *TAS1R2*, *TRPV1*, and *CD36* (**Supplementary Figure S2**). There did not seem to be any correlation with the number of potentially deleterious exonic variants to the size of the genes. A large fraction of the variants was private and the density of variation was observed to differ across genes. *MTCH1* harbours a higher number of potentially deleterious common variants, while *CEP290*, *ITPR3*, *PLCB2*, *TAS2R7*, and *TAS2R9* have higher overall numbers (**Figure 1C**). The average percentage of SNPs in the genic and 10 kb flanking regions, normalized for the length across the different perception categories (bitter, sweet/fatty, salty/pungent, sour/astringent, and olfactory) were also calculated. It was observed that the genes associated with olfaction harbored more SNPs (1.61%) followed by salt taste and pungency associated genes (*SCNN1A*, *SCNN1B*, *SCNN1G*, *TRPV1*) (1.17%).

Chemosensory Gene Variants in IndiGenomes vs. 1,000 Genomes Populations: Similarities and Uniqueness

Nearly 70% (17,774) of the common variants all-together in the study are shared between IndiGen and 1,000 genomes (**Figure 1** and **Supplementary Figure S3A**), which is 13% of all the variants in IndiGen. Interestingly, 8.2% of the common variants are also unique to IndiGenomes and not present in any of the 1,000 Genomes populations. An additional 3% IndiGen variants



overlap with SAS. However, 10% of the common variants shared between SAS and non-SAS populations of the 1,000 Genomes are not observed in IndiGen. The overlap of rare variants between the three cohorts is 9.6%. With respect to rare variants, the overlap is much more between SAS and non-SAS populations of 1,000 genomes. Only 3% of the private variants are shared across IndiGen and SAS and non-SAS of 1,000 genomes datasets. IndiGen data has a significantly higher proportion of unique private variants (27%) (**Figure 1**). In all the three sets, the extent of overlap of private variants in IndiGene with SAS is higher (9%) (**Supplementary Figure S3A**).

Amongst all the genes and the considered flanking regions, the bitter taste receptors in the IndiGen data have a higher number of common and rare variants compared to SAS and 1000G. The olfactory receptor gene *OR2T10*, has a higher number of SNPs in

IndiGen in all three categories (**Supplementary Figure S3A**). The bitter-associated genes *PRH1-TAS2R14* (which encompasses *TAS2R43* and *TAS2R46*), *OR2T10*, *OR11HP* (olfaction), *MTCH2*, *FFAR1*, *ADIPOQ* (sweet and fatty foods) have a higher number of novel common variants. Most of the genes are found to have higher proportions of private variants in IndiGen not represented in the 1000 Genomes Project, while some genes like *FTO*, *RGS6*, *OPRM1* (sweet and fatty foods), *GRM7*, *GRM8*, *ASIC2*, *RUNX1* (sour/marmite) have more private variants shared between the two datasets (**Figures 1A,B**).

We also analyzed the frequency distribution of 231 variants that have been previously associated with taste perceptions. Out of 231 variants, 195 variants were observed in the IndiGen dataset that also includes 3 monomorphic variants (**Supplementary Table S2**). Seven variants had a frequency less than 1% and

TABLE 1 | High F_{ST} variants differentiating world populations either earlier reported to be associated with taste perception or nonsynonymous functional variations.

SNP ID	Location	Gene	Reported associations with the SNP	Taste perception associated with the SNP/gene
rs11064153	intergenic	Near <i>SCNN1A</i>	Yang et al., 2014 (33)	Salt taste
rs10256873	Intronic	<i>GRM8</i>	Roos et al., 2017 (34)	Marmite taste
rs11760281	Intronic	<i>GNAT3</i>	Farook et al., 2012	Sucrose intake
rs150908	Intronic	<i>TRPV1</i>	Chamoun et al., 2018	Salt taste
rs1524600	Intronic	<i>GNAT3</i>	Knaapila et al., 2012	Sweet taste
rs224549	Intronic	<i>TRPV1</i>	Allen et al., 2014 ^a	Pungent and bitter taste
rs224550	Intronic	<i>TRPV1</i>	Allen et al., 2014 ^a	Pungent and bitter taste
rs224551	Intronic	<i>TRPV1</i>	Allen et al., 2014 ^a	Pungent and bitter taste
rs2277675	Intronic	<i>TRPV1</i>	Pirastu et al., 2012	Salt taste
rs3211913	Intronic	<i>CD36</i>	(Love-Gregory et al., 2008)	Fat taste perception
rs3818521	Intronic	<i>ITPR3</i>	Pirastu et al., 2012	Lamb meat and sheep cheese liking
rs4908563	Intronic	<i>TAS1R1</i>	Chamoun et al., 2018	Sweet/Umami Taste
rs769148	Intronic	<i>RGS6</i>	Sibbel et al., 2011	Higher intake of fat/oil/sweets
rs236514	UTR3	<i>KCNJ2</i>	Chamoun et al., 2018	Sour taste
rs2274333	nonsynonymous	<i>CA6</i>	Melis et al., 2013	PROP Sensitivity (Bitter)
rs1799971	nonsynonymous	<i>OPRM1</i>	Deb et al., 2010	Alcohol taste
rs2233998	nonsynonymous	<i>TAS2R4</i>	Choi et al., 2017	Bitter taste
rs936212	nonsynonymous	<i>PLCB2</i>	None reported	Sweet
rs10845281	nonsynonymous	<i>TAS2R20</i>	None reported	Bitter taste
rs10845279	nonsynonymous	<i>TAS2R20</i>	None reported	Bitter taste
rs2234002	nonsynonymous	<i>TAS2R4</i>	None reported	Bitter taste
rs35969491	nonsynonymous	<i>TAS2R42</i>	None reported	Bitter taste
rs1376251	nonsynonymous	<i>TAS2R50</i>	None reported	Bitter taste
rs224534	nonsynonymous	<i>TRPV1</i>	None reported	Salt taste
rs9262	nonsynonymous SNV	<i>C12orf29</i> (Upstream of <i>CEP290</i>)	None reported	Marmite

^aIn LD with rs224547 in the SAS population which is reported to be significantly associated with burning/stinging and bitterness.

the rest 185 variants had a frequency greater than 1% in the IndiGen dataset. Amongst these, 61 variants were exonic variants with 49 nonsynonymous, 11 synonymous and a stop-gain variant. Out of the 49 nonsynonymous variants, 8 were not observed in IndiGen.

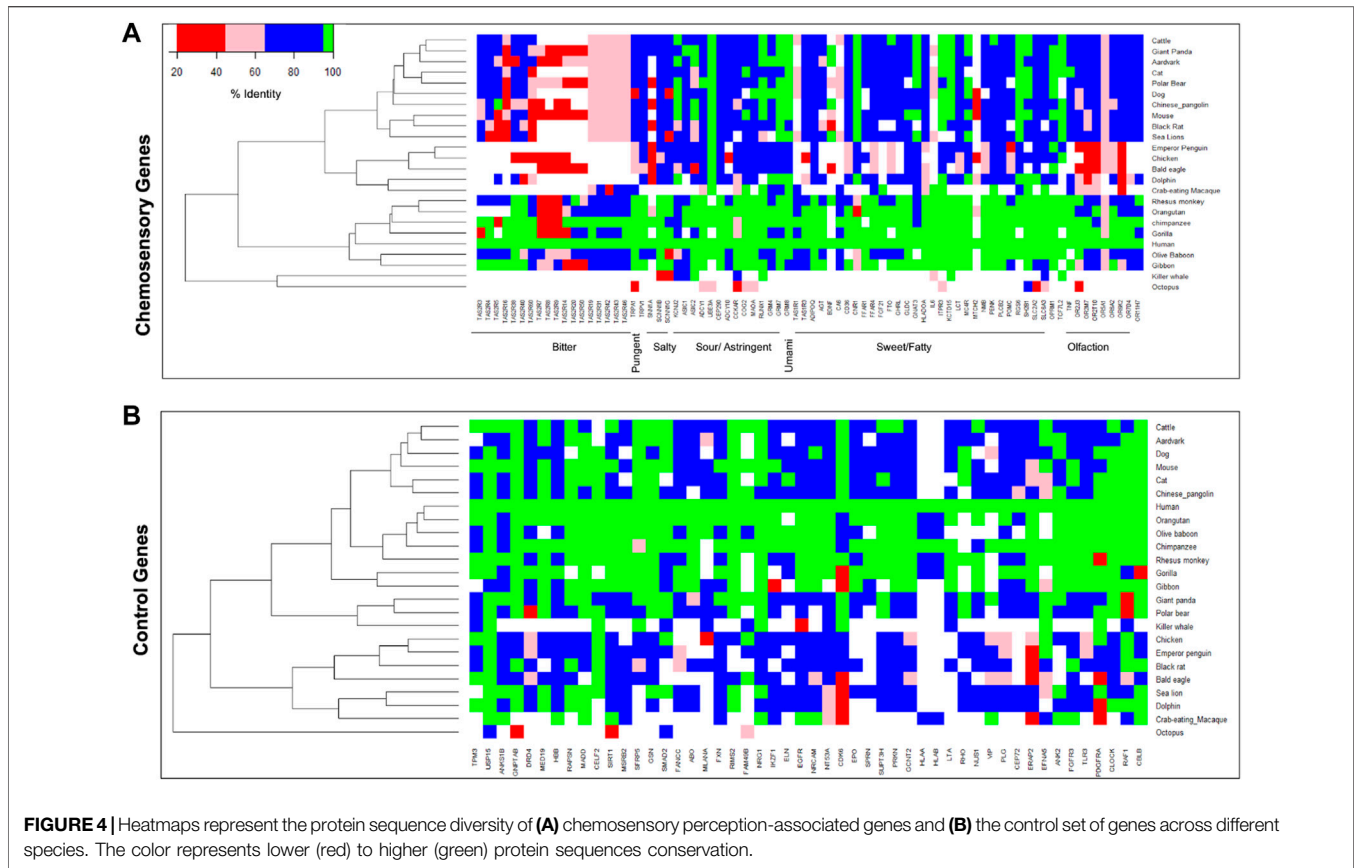
To check if the variants unique to IndiGen compared to the 1,000 Genomes dataset are novel and not reported elsewhere, we also compared the variants with HGDP and SGDP. 46% of the IndiGen variants were found to be novel (**Supplementary Figure S3C**), most of them being private (<1% MAF).

Mapping Variations in Chemosensory Perception to Genetic Ancestry: Relatedness and Admixture Analysis of IndiGenomes

In previous studies of the Indian Genome Variation Consortium (IGVC) we have reported that the Indian population maps to five major genetic clusters (Indian Genome Variation Consortium, 2008). We performed conjoint analysis of the 1,000 Genomes dataset with IGV and IndiGen to map ancestry and admixture proportions in the IndiGenomes. In the PCA analysis, PC1 separated African versus non-African populations while PC2 separated Europeans with East-Asian population (**Supplementary Figure S4**). None of the IndiGen individuals show closeness to African (YRI) or admixed African Indian (OG) populations. In the subsequent PCA analysis, a subset of IndiGen individuals show their closeness to East Asians and Tibeto-Burman populations of IGV in PC1, while second group of

IndiGen Individuals are closer to European of 1,000 Genomes and Indo-Europeans large populations of IGV in PC2 (**Figure 2B**). The third set of IndiGen samples are closer to IGV populations of Dravidian large population linguistic groups while the rest of the IndiGen individuals fall within the Indo-European and Dravidian linguistic cline. Individual-wise ancestry of IndiGen samples were further assessed using maximum likelihood-based methods implemented in ADMIXTURE. The lowest cross validation (CV) error was observed at $K = 10$ and used for structure interpretation (**Supplementary Figure S5**). A similar ancestry component was observed in admixture analysis where three predominant groups namely, Tibeto-Burman like (IND_1 group), Indo-European like (IND_2 group), and Dravidian like (IND_3 group) were classified (**Figure 2C**). The individual-wise ancestry estimates of these three predominant individuals are depicted in **Supplementary Figure S6**. The three predominant groups were used for further population comparisons analysis of chemosensory gene variants. Pairwise F_{ST} analysis shows that IndiGen individuals have maximal closeness with SAS populations of the 1,000 Genomes.

Derived allele frequency (DAF) of the variants within chemosensory genes were calculated which were further used to calculate the absolute delta DAF between pairwise populations. Variants were prioritized based on signatures of positive selection with both F_{ST} value ≥ 0.3 and delta DAF ≥ 0.5 between pairwise populations. Pairwise population comparisons provide 1,184 variants within chemosensory genes that show high F_{ST} and high derived allele frequency in IndiGen compared to other world populations (**Supplementary Table S3**). Genes with



these with other ethnic world populations. We observe a total of 137,760 SNPs within chemosensory genes in the IndiGen dataset including private variants. 9% of the common SNPs and 27% private variants in IndiGen seem to be Indian population specific as they are not represented in 1,000 Genomes further substantiating the diversity and uniqueness of IndiGen data even for chemosensory genes (Jain et al., 2021). Further, the bitter taste receptor genes display higher variation compared to other taste receptor gene sets. The protein sequence diversity estimates between human and other species also reveal higher divergence of bitter taste receptor genes compared to other chemosensory genes and control gene sets which reflects dietary habits of different species as well as provide potential variants that are involved in sensing of bitter taste to avoid ingestion of harmful food (Dong et al., 2009).

Conjoint population genetic analysis of IndiGen individuals with those from Indian Genome Variation Consortium and 1,000 Genomes identified sets of individuals that show proximity to three predominant ethnic groups namely, Tibeto-Burman group (IND_1), Indo-European large group (IND_2), and Dravidian large group (IND_3).

We speculated that differences in genetic variation and diversity of chemosensory genes between IndiGen data and other ethnic populations could also reflect differences in selective pressure. We thus looked at the extent of differentiation through pairwise F_{ST} and derived allele

frequency statistics. Based on higher F_{ST} value and higher delta DAF between pairwise population comparison we identified 1,181 variants within chemosensory genes that display differences in derived allele frequency between IndiGen populations and representative European (CEU), African (YRI) and East Asian (CHB) populations. The top genes were *ADCY10*, *TRPV1*, *RUNX1*, *GRM7*, *TCF7L2*, *OR7D4*. These could be potential targets of positive selection.

The baseline variability information in chemosensory genes could have utility in nutrigenomic and precision medicine studies especially the ones that show high F_{ST} values (Table 1). For example, a variant rs3818521 *ITPR3* has been studied for its association with lamb-meat and cheese-liking (umami). Its minor allele (C) reported to be associated with lamb meat liking (Pirastu et al., 2012) has lower frequency in the Indian population and highest in African. Another interesting SNP was rs769148 in *RGS6* wherein the carriers of C allele are prone to obesity development due to higher intake of fat/oil/sweets that are palatable energy-dense foods (Sibbel et al., 2011). The frequency of derived allele T was highest (81%) in Ind_1 of IndiGen set who are related to the Tibeto-Burman lineage. This is also highest in its most proximal East Asian (CHB; 93%) population from 1,000 Genomes. The T allele is associated with lower preferences for fat/oily food thereby contributing to reduced risk for obesity in these populations. A recent study has reported a greater prevalence of the G allele of rs1799971

(*OPRM1*) in both alcoholic and opioid addicts in India compared with those in the normal population (Deb et al., 2010) that also corroborates with similar findings from a study in Sweden (Bart et al., 2005). This allele has also been reported to be associated with a higher preference for sweet and fatty foods (Davis et al., 2011). The frequency spectrum of this variant displays differences amongst diverse Indian Populations in the IGVC (**Supplementary Table S5**). When looking at the variants with reported association in IGV populations, some of them seem to be more striking. Among them, rs2274333 of the *CA6* (Carbonic Anhydrase 6) gene is associated with PROP sensitivity where the AA genotype (homozygous ancestral allele associated with a fully-functional gustin protein) and the GG genotypes are more frequent in supertasters and nontasters respectively. This suggests that variation in gustin may be associated with differences in papillae densities and oral chemosensory abilities in individuals sensing PROP (Melis et al., 2013). The frequency of derived allele was higher in Asian populations compared to Europeans and Africans (**Figure 3A**) however, the frequency of derived allele was much higher in Indo-Europeans of North and North-Eastern populations and lower in Indo-Europeans of Western India. Variation in frequency spectrum suggests that populations from western part could have more supertasters compared to the north. Even within the Dravidian population, differences are observed. This polymorphism could also affect the intake of coffee and white cabbage (Mikołajczyk-Stecyna et al., 2017). The A allele of rs150908 (*TRPV1*) is associated with a higher salt sensitivity and hence a lower preference for salty foods (Chamoun et al., 2018). Another *TRPV1* variant rs2277675 is likely to be associated with sweet perception as its minor allele (C) is associated with less liking and preference for beet (Pirastu et al., 2012). It would be interesting to see whether some of these variations could contribute to altered perceptions between individuals and populations. Likewise, the C allele of rs4790522 (*TRPV1*) is known to be associated with high salt preference and a risk for cardiovascular diseases (Precone et al., 2019). Generally with the increase in altitude, taste perception for salt decreases. The frequency of this allele was higher, mostly in the North-Eastern parts of India (high altitude) (**Supplementary Figure S7**). We anticipate that conducting chemosensory/taste perception tests on a large scale and correlation with the underlying genotypes would shed light on the genetic basis of chemosensory perceptions and food preferences and would be useful in the field of precision nutrition.

As taste, flavours and texture forms the basis of determination of food selection and dietary habits, inter-individual variations in chemosensory perception can prove to be highly useful in the characterization of differences in nutrient intake. Since this is also directly linked to the health and general well-being of an individual, evaluation of these polymorphisms in taste receptor genes can assist in understanding the association between food selection and the occurrence of numerous metabolic disorders. Moreover, a comprehensive knowledge of genetic variations in taste receptor genes can aid the development of personalized

nutrition models that can help design dietary recommendations for preempting disorders due to dietary habits and enhance the general quality of life of individuals. A catalogue of curated variants could thus be enormously helpful in discerning differences in food preferences both within a population and between different global populations.

WEB RESOURCES

IndiGen database: <http://clingen.igib.res.in/indigen/>
 1,000 Genomes database: <ftp://ftp-trace.ncbi.nih.gov/1000genomes/ftp/release/20130502/>
 1,000 Genomes ancestral allele information: <ftp://ftp.1000genomes.ebi.ac.uk>
 IGV database: <http://igvdb.res.in/>
 PLINK software: <http://zzz.bwh.harvard.edu/plink/www.cog-genomics.org/plink/1.9/>
 VCFTOOLS: <http://vcftools.sourceforge.net/>
 UCSC Genome Browser: <https://genome.ucsc.edu/>
 ANNOVAR: <https://annovar.openbioinformatics.org/>
 GGV: <http://popgen.uchicago.edu/ggv/>

DATA AVAILABILITY STATEMENT

The SNP information from the IndiGen project (Jain et al., 2020) are available at <https://clingen.igib.res.in/indigen/> and the genotype data can be made available upon request. The codes used for the analyses can be accessed at https://github.com/Prakrithi-P/IndiGen_Chemosensory-Landscape and the results from the current study can be found in the **Supplementary Material**.

AUTHOR CONTRIBUTIONS

The authors' responsibilities were as follows—MM: conceived and designed the project; PP and PJ: carried out the data analyses; PP, JJ, PJ, and MM: wrote the manuscript; PP: generated the visualizations; VSC., SS, DS, RCB, AJ, MI, VSE, MD, and AM: were involved in sample collection, sequencing and data generation of the IndiGen project.

FUNDING

This project was supported by the Council of Scientific and Industrial Research, India (MLP1809/MLP201) and CSIR fellowship (GAP0206) (to PP).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.878134/full#supplementary-material>

REFERENCES

- Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast Model-Based Estimation of Ancestry in Unrelated Individuals. *Genome Res.* 19, 1655–1664. doi:10.1101/gr.094052.109
- Allen, A. L., McGeary, J. E., and Hayes, J. E. (2014). Polymorphisms in TRPV1 and TAS2Rs Associate with Sensations from Sampled Ethanol. *Alcohol. Clin. Exp. Res.* 38, 2550–2560. doi:10.1111/acer.12527
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic Local Alignment Search Tool. *J. Mol. Biol.* 215 (3), 403–410. doi:10.1016/S0022-2836(05)80360-2
- Bart, G., Kreek, M. J., Ott, J., LaForge, K. S., Proudnikov, D., Pollak, L., et al. (2005). Increased Attributable Risk Related to a Functional μ -Opioid Receptor Gene Polymorphism in Association with Alcohol Dependence in Central Sweden. *Neuropsychopharmacology* 30, 417–422. doi:10.1038/sj.npp.1300598
- Chamoun, E., Carroll, N., Duizer, L., Qi, W., Feng, Z., Darlington, G., et al. (2018). The Relationship between Single Nucleotide Polymorphisms in Taste Receptor Genes, Taste Function and Dietary Intake in Preschool-Aged Children and Adults in the Guelph Family Health Study. *Nutrients* 10, 990. doi:10.3390/nu10080990
- Choi, J.-H., Lee, J., Yang, S., and Kim, J. (2017). Genetic Variations in Taste Perception Modify Alcohol Drinking Behavior in Koreans. *Appetite* 113, 178–186. doi:10.1016/j.appet.2017.02.022
- Clark, A. A., Dotson, C. D., Elson, A. E. T., Voigt, A., Boehm, U., Meyerhof, W., et al. (2015). TAS2R Bitter Taste Receptors Regulate Thyroid Function. *FASEB J.* 29, 164–172. doi:10.1096/fj.14-262246
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., et al. (2011). The Variant Call Format and VCFtools. *Bioinformatics* 27, 2156–2158. doi:10.1093/bioinformatics/btr330
- Davis, C., Zai, C., Levitan, R. D., Kaplan, A. S., Carter, J. C., Reid-Westoby, C., et al. (2011). Opiates, Overeating and Obesity: a Psychogenetic Analysis. *Int. J. Obes.* 35 (10), 1347–1354. doi:10.1038/ijo.2010.276
- Deb, I., Chakraborty, J., Gangopadhyay, P. K., Choudhury, S. R., and Das, S. (2010). Single-nucleotide Polymorphism (A118G) in Exon 1 of OPRM1 Gene Causes Alteration in Downstream Signaling by μ -Opioid Receptor and May Contribute to the Genetic Risk for Addiction. *J. Neurochem.* 112, 486–496. doi:10.1111/j.1471-4159.2009.06472.x
- Dias, A. G., Rousseau, D., Duizer, L., Cockburn, M., Chiu, W., Nielsen, D., et al. (2013). Genetic Variation in Putative Salt Taste Receptors and Salt Taste Perception in Humans. *Chem. Senses* 38, 137–145. doi:10.1093/chemse/bjs090
- Diószegi, J., Llanaj, E., and Ádány, R. (2019). Genetic Background of Taste Perception, Taste Preferences, and its Nutritional Implications: A Systematic Review. *Front. Genet.* 10, 1272. doi:10.3389/fgene.2019.01272
- Dong, D., Jones, G., and Zhang, S. (2009). Dynamic Evolution of Bitter Taste Receptor Genes in Vertebrates. *BMC Evol. Biol.* 9, 12. doi:10.1186/1471-2148-9-12
- Dotson, C. D., Babich, J., and Steinle, N. I. (2012). Genetic Predisposition and Taste Preference: Impact on Food Intake and Risk of Chronic Disease. *Curr. Nutr. Rep.* 1, 175–183. doi:10.1007/s13668-012-0021-3
- Duffy, V. B., Davidson, A. C., Kidd, J. R., Kidd, K. K., Speed, W. C., Pakstis, A. J., et al. (2004). Bitter Receptor Gene (TAS2R38), 6-N-Propylthiouracil (PROP) Bitterness and Alcohol Intake: Alcohol. *Clin. Exp. Res.* 28, 1629–1637. doi:10.1097/01.ALC.0000145789.55183.D4
- Eriksson, N., Wu, S., Do, C. B., Kiefer, A. K., Tung, J. Y., Mountain, J. L., et al. (2012). A Genetic Variant Near Olfactory Receptor Genes Influences Cilantro Preference. *Flavour* 1, 22. doi:10.1186/2044-7248-1-22
- Farook, V., Puppala, S., Schneider, J., Fowler, S., Chittoor, G., Dyer, T., et al. (2012). Metabolic Syndrome Is Linked to Chromosome 7q21 and Associated with Genetic Variants in CD36 and GNAT3 in Mexican Americans. *Obes. Silver Spring Md* 20, 2083–2092. doi:10.1038/oby.2012.74
- Gentiluomo, M., Lu, Y., Canzian, F., and Campa, D. (2019). Genetic Variants in Taste-Related Genes and Risk of Pancreatic Cancer. *Mutagenesis* 34, 391–394. doi:10.1093/mutage/gez032
- Goudet, J. (2005). Hierfstat, a Package for R to Compute and Test Hierarchical F-Statistics. *Mol. Ecol. Notes* 5, 184–186. doi:10.1111/j.1471-8286.2004.00828.x
- Hinrichs, A. S. (2006). The UCSC Genome Browser Database: Update 2006. *Nucleic Acids Res.* 34, D590–D598. doi:10.1093/nar/gkj144
- Indian Genome Variation Consortium (2008). Genetic Landscape of the People of India: a Canvas for Disease Gene Exploration. *J. Genet.* 87, 3–20. doi:10.1007/s12041-008-0002-x
- Jain, A., Bhojar, R. C., Pandhare, K., Mishra, A., Sharma, D., Imran, M., et al. (2021). IndiGenomes: a Comprehensive Resource of Genetic Variants from over 1000 Indian Genomes. *Nucleic Acids Res.* 49, D1225–D1232. doi:10.1093/nar/gkaa923
- Keller, K. L., Liang, L. C. H., Sakimura, J., May, D., van Belle, C., Breen, C., et al. (2012). Common Variants in the CD36 Gene Are Associated with Oral Fat Perception, Fat Preferences, and Obesity in African Americans. *Obesity* 20, 1066–1073. doi:10.1038/oby.2011.374
- Knaapila, A., Hwang, L.-D., Lysenko, A., Duke, F. F., Fesi, B., Khoshnevisan, A., et al. (2012). Genetic Analysis of Chemosensory Traits in Human Twins. *Chem. Senses* 37, 869–881. doi:10.1093/chemse/bjs070
- Love-Gregory, L., Sherva, R., Sun, L., Wasson, J., Schappe, T., Doria, A., et al. (2008). Variants in the CD36 Gene Associate with the Metabolic Syndrome and High-Density Lipoprotein Cholesterol. *Hum. Mol. Genet.* 17 (11), 1695–1704. doi:10.1093/hmg/ddn060
- Lunde, K., Egeland, S., Skuterud, E., Mainland, J. D., Lea, T., Hersleth, M., et al. (2012). Genetic Variation of an Odorant Receptor OR7D4 and Sensory Perception of Cooked Meat Containing Androstenedione. *PLoS ONE* 7, 7. doi:10.1371/journal.pone.0035259
- Melis, M., Atzori, E., Cabras, S., Zonza, A., Calò, C., Muroli, P., et al. (2013). The Gustin (CA6) Gene Polymorphism, Rs2274333 (A/G), as a Mechanistic Link between PROP Tasting and Fungiform Taste Papilla Density and Maintenance. *PLoS ONE* 8, 15. doi:10.1371/journal.pone.0074151
- Menashe, I., Man, O., Lancet, D., and Gilad, Y. (2003). Different Noses for Different People. *Nat. Genet.* 34, 143–144. doi:10.1038/ng1160
- Mikołajczyk-Stecyna, J., Malinowska, A. M., and Chmurzynska, A. (2017). TAS2R38 and CA6 Genetic Polymorphisms, Frequency of Bitter Food Intake, and Blood Biomarkers Among Elderly Woman. *Appetite* 116, 57–64. doi:10.1016/j.appet.2017.04.029
- Newcomb, R. D., Xia, M. B., and Reed, D. R. (2012). Heritable Differences in Chemosensory Ability Among Humans. *Flavour* 1, 9. doi:10.1186/2044-7248-1-9
- Pirastu, N., Robino, A., Lanzara, C., Athanasakis, E., Esposito, L., Tepper, B. J., et al. (2012). Genetics of Food Preferences: A First View from Silk Road Populations. *J. Food Sci.* 77, S413–S418. doi:10.1111/j.1750-3841.2012.02852.x
- Precone, V., Beccari, T., Stuppia, L., Baglivo, M., Paolacci, S., Manara, E., et al. (2019). Taste, Olfactory and Texture Related Genes and Food Choices: Implications on Health Status. *Eur. Rev. Med. Pharmacol. Sci.* 23 (3), 1305–1321. doi:10.26355/eurrev_201902_17026
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., et al. (2007). PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* 81, 559–575. doi:10.1086/519795
- Roos, T. R., Kulemin, N. A., Ahmetov, I. I., Lasarow, A., and Grimaldi, K. (2017). Genome-Wide Association Studies Identify 15 Genetic Markers Associated with Marmite Taste Preference. *bioRxiv*, 185629. doi:10.1101/185629
- Shaik, F. A., Singh, N., Arakawa, M., Duan, K., Bhullar, R. P., and Chelikani, P. (2016). Bitter Taste Receptors: Extraoral Roles in Pathophysiology. *G Protein-Coupled Recept. GPCRs* 77, 197–204. doi:10.1016/j.biocel.2016.03.011
- Sibbel, S. P., Talbert, M. E., Bowden, D. W., Haffner, S. M., Taylor, K. D., Chen, Y.-D. I., et al. (2011). RGS6 Variants Are Associated with Dietary Fat Intake in Hispanics: The IRAS Family Study. *Obesity* 19, 1433–1438. doi:10.1038/oby.2010.333

- The 1000 Genomes Project Consortium (2015). A Global Reference for Human Genetic Variation. *Nature* 526, 68–74. doi:10.1038/nature15393
- Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: Functional Annotation of Genetic Variants from High-Throughput Sequencing Data. *Nucleic Acids Res.* 38, e164. doi:10.1093/nar/gkq603
- Yang, X., He, J., Gu, D., Hixson, J. E., Huang, J., Rao, D. C., et al. (2014). Associations of Epithelial Sodium Channel Genes with Blood Pressure Changes and Hypertension Incidence: The GenSalt Study. *Am. J. Hypertens.* 27, 1370–1376. doi:10.1093/ajh/hpu060

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

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

Copyright © 2022 Prakrithi, Jha, Jaiswal, Sharma, Bhoyar, Jain, Imran, Senthilvel, Divakar, Mishra, Scaria, Sivasubbu and Mukerji. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.