Title: Hydrogen Sulfide (H_2S)-Producing Oral Bacteria May Protect Against COVID-19

Authors and Affiliation

Meghalbahen Vaishnani[1], Anupama Modi[1]*, Kshipra Chauhan[1] , Bhavin Parekh[2]*,

[1]School of Applied Sciences and Technology, Gujarat Technological University, Gujarat, India.

[2]Department of Validation Indic Knowledge through Advanced Research, Gujarat University

***Corresponding author**

Dr. Bhavin Parekh*****

Research fellow,

Department of Validation Indic Knowledge through Advanced Research

Gujarat University

Gujarat -380009

Email: bp2580@googlemail.com

Phone: +9192656 52730

Dr. Anupama Modi*

Assistant Professor

School of Applied Sciences and Technology

Gujarat Technological University

Gujarat -382424

Email: anupamarmodi@gmail.com

Phone:+91 97995 03950

Abstract

COVID-19 mortality rates have varied dramatically across the globe. Yet the reasons behind these disparities remain poorly understood. While recent research has linked gut microbes to these variations, the role of oral bacteria, a main port of entry for the coronavirus, remains unexplored.We investigated the relationship between oral microbiota and COVID-19 mortality rates across eightcountries. Raw sequencing data of 16S rRNA regions from oral microbiota in 244 healthy subjects from eight countries were obtained from public databases. We employed a generalized linear model (GLM) to predict COVID-19 mortality rates using oral microbiota composition. GLM revealed that high abundances of hydrogen sulfide (H_2S)-producing bacteria, particularly Treponema, predicted low COVID-19 mortality rates with a markedly low p-value. Unsupervised clustering using a combination of LIGER and t-SNE yielded four oral microbiome "orotypes." Orotypes enriched in H_2S-producing bacteria coincided with lower mortality rates, while orotypes harboring Haemophilus or Rothia were associated with increased vulnerability. To validate our findings, we analyzed influenza mortality data from the same countries, observing similar protective trends. Our findings suggest that oral bacteria-produced H_2S may serve as a critical initial defense against SARS-CoV-2 infection.H_2S from oral bacteria may prevent infection through antiviral activity, blocking ACE2 receptors, suppressing cytokines, and boosting antioxidants. This highlights the oral microbiome's role in COVID-19 outcomes and suggests new preventive and therapeutic strategies.

Keywords : COVID-19, Oral bacteria, COVID-19 mortality rates, Treponema, H_2S

Abbreviation

COVID-19: Coronavirus disease 2019

SARS-CoV-2: Severe Acute Respiratory Syndrome Coronavirus 2

GLM: Generalized linear model

LIGER: Linked Inference of Genomic Experimental Relationships

H_2S: Hydrogen sulfide

t-SNE: t-distributed Stochastic Neighbor Embedding

NCBI: National Center for Biotechnology Information

RNA: Ribonucleic Acid

DNA: Deoxyribonucleic acid

RSV: Respiratory syncytial virus

Nrf2: nuclear factor erythroid 2–related factor 2

Introduction

Why did COVID-19 hit some countries harder than others?[1]. This remains one of the pandemic's most perplexing questions—a critical piece of the puzzle for future pandemic preparedness [1] [2]. Conventional risk factors like age and comorbidities like diabetes and obesity fail to fully explain this aspect [2] [3]. Emerging research implicates the human microbiome as a potential arbiter of COVID-19 outcomes [3]. Researchers have shown that low levels of certain gut bacteria, such as Collinsella, predict high COVID-19 mortality rates. Collinsella produces ursodeoxycholate, which may prevent infection and mitigate acute respiratory distress syndrome by suppressing cytokine storm syndrome[3].

What about the oral microbiota—the second largest microbial community in the body [4], SARS-CoV-2's first contact point[5] and our frontline defense against respiratory viruses[6]?Exploring this overlooked aspect could further illuminate why some individuals succumb to the virus while others remain resilient[7]. Intrigued by this possibility, we hypothesize that dynamic interaction between the virus and specific oral microbiome may confer varying degrees of susceptibility or resilience to COVID-19 across populations.

To test this hypothesis, we analyzed the relationship between oral bacterial composition in 244 healthy subjects from eight countries and their respective COVID-19 mortality rates. Our findings reveal a negative correlation between four H_2S-prodcing genera and COVID-19 mortality, suggesting a potential protective role for this oral commensal. To validate these findings, we further analyzed influenza mortality rates in the same countries, confirming similar protective trends and strengthening the generalizability of our results to respiratory viral infections.

Materials and Methods

Dataset

To investigate the potential association between oral microbiota composition and COVID-19 mortality rates, this study utilized publicly available healthy oral microbiome datasets from the National Center for Biotechnology Information (NCBI). The analysis encompassed data from eight countries: the United States, Australia, Brazil, Italy, Sweden, Saudi Arabia, Egypt, and Russia. These datasets were selected based on their accessibility and comprehensive coverage, corresponding to accession numbers PRJNA606501, PRJNA384402, PRJNA256234, PRJNA267483, PRJNA598825, PRJNA227796, PRJNA292800, and PRJDB5153. COVID-19 mortality data, quantified as cumulative deaths per million population as of February 9, 2021, were obtained from 'https://ourworldindata.org/'. This timeframe was chosen to precede the widespread distribution of vaccines, thereby minimizing potential confounding effects of vaccination campaigns on mortality rates. To maintain the integrity and breadth of the analysis, no filtering criteria based on age or sex were applied to the datasets. This approach preserved a diverse and representative subject range, allowing for a more comprehensive examination of the potential relationship between oral microbiota and COVID-19 mortality across different populations.

Taxonomic Analysis of oral microbiota

To classify the oral microbiota, OmicsBox (version 3.0.29) developed by BioBam Bioinformatics, utilizing the Kraken2 tool, was employed. Kraken2 performs taxonomic classification by analyzing k-mers in DNA short reads and querying a comprehensive database of species-specific k-mer information. This approach allowed for thorough and detailed

taxonomic profiling, covering both metabarcoding (16S rRNA gene) and metagenomic sequencing reads, thus ensuring robust and accurate taxonomic identification.

Generalized Linear Model (GLM) Analysis

To investigate the relationship between oral microbiota composition and COVID-19 mortality, Generalized Linear Models (GLMs) were employed. The 15 most abundant genera served as predictors, with COVID-19 mortality as the outcome variable(Fig 1). Model selection was performed using Akaike Information Criterion (AIC), comparing Gaussian, Gamma, and Inverse Gaussian distributions. The optimal model (Gamma distribution) was used to identify significant associations between specific genera and mortality rates.

LIGER-tSNE Analysis

A combination of Linked Inference of Genomic Experimental Relationships (LIGER) and t-distributed stochastic neighbor embedding (t-SNE) was employed in our analysis to explore and visualize the complex oral microbiome structures across different populations (Fig 1). This LIGER-tSNE combination was found to be a powerful tool for dimensionality reduction and clustering, allowing subtle patterns in the oral microbiome data that might have been overlooked by traditional analysis methods to be discerned. This innovative approach, originally developed for single-cell RNA sequencing data, proved highly effective in identifying distinct oral microbiome orotypes and their relationship to COVID-19 mortality rates.

Results

Generalized Linear Model (GLM) Analysis

16S rRNA sequencing data from 244 healthy subjects across eight countries (Table 1) were analyzed to examine the effects of oral microbiota on COVID-19 mortality rates. Relative abundance analysis at the genus level (Fig 2) revealed *Streptococcus*, *Prevotella*, *Veillonella*, *Fusobacterium*, and *Haemophilus* as the most prevalent genera.

Using GLM, COVID-19 mortality rates were predicted with the 15 most abundant genera, with the Gamma distribution yielding the lowest AIC. GLM results (Fig 3) highlighted a marked negative predictive value of *Treponema* abundance for COVID-19 mortality rates ($p < 0.001$). *Campylobacter, Gemella, and Selenomonas* also showed negative correlations, albeit less significant.

Linked Inference of Genomic Experimental Relationships (LIGER) Analysis

Non-negative matrix factorization via LIGER identified four distinct oral microbiome clusters, termed "orotypes" (Fig 4A). For every orotype, the mean relative abundances of the top 15 genera are calculated in (S3 Table). The most abundant 5 genera in each orotype are given in Supplementary Table 1(S2 Table).. orotype distribution varied across countries in relation to COVID-19 mortality rates (Fig 4B). High mortality countries (e.g., Italy, USA) predominantly exhibited orotypes 2 and 3, while low mortality countries (e.g., Australia, Russia) showed a higher prevalence of orotype 4 (Fig 4B).Countries with intermediate mortality rates displayed a mixed distribution.Analysis of COVID-19 mortality rates across orotypes (Fig 4C) revealed a decreasing trend from orotypes 1 to 4, although not statistically significant ($p = 0.203$,

Jonckheere-Terpstra trend test, Fig 4D). Notably, *Treponema* relative abundance significantly increased from orotypes 1 to 4 (Fig 4E; p = 0.033, Jonckheere-Terpstra trend test, Fig 4F).

Other Genera of Interest and Combined Analysis

Analysis of Campylobacter, Gemella, and Selenomonas individually revealed no significant trends across orotypes (Figs S1a-f). However, combined analysis of Treponema, Gemella, Campylobacter, and Selenomonas demonstrated a highly significant increasing trend from orotypes 1 to 4 (p = 7.198e-05, Jonckheere-Terpstra trend test, Figs S1g-h), suggesting a potential synergistic effect of these bacteria.

Validation with Influenza Mortality Data

To validate these findings, influenza mortality rates in these eight countries (taken from 'https://ourworldindata.org/' (2019)) were analyzed, given the similarities between COVID-19 and influenza as respiratory viral illnesses. The influenza mortality chart (Fig S2) demonstrates that orotype 4 is linked to lower mortality rates in influenza, mirroring the trends observed with COVID-19. This further substantiates the potential protective role of orotype 4-associated microbial composition in respiratory viral infections.

Discussion

This study uncovers a fascinating link between specific oral bacteria—Treponema, Gemella, Campylobacter, and Selenomonas—and lower COVID-19 mortality rates (Fig 3). The unifying characteristic of these bacteria is their production of hydrogen sulfide (H_2S), a molecule now implicated as a pivotal protective agent against respiratory viral infections. Our data reveal a striking pattern in oral microbial communities. Those dominated by Haemophilus (orotype 1) or Rothia (orotype2) appear to be associated with increased vulnerability to SARS-CoV-2 (S2 Table). In contrast, communities enriched with H_2S-producing bacteria (orotypes 3 and 4) exhibit a protective effect.

This pattern echoes recent findings by Ren et al., who observed a similar association between Gemella and Rothia abundance in the upper respiratory tract of healthy individuals and COVID-19 patients, respectively. Wu et al. (2021) also reported significant microbiome alterations in COVID-19 patients, including elevated Rothia mucilaginosa levels in both oral and gut samples. Interestingly, this potential protective role of H_2S aligns with previous research on gut microbiota, where Collinsella, another H_2S-producing bacterium, was identified as a protective factor against COVID-19 [3] [8].

The H_2S-producing capacity of the identified bacteria is notable. Treponema species, particularly Treponema denticola, are notable for their high H_2S production capacity due to multiple cysteine-degrading enzymes, potentially explaining their strong negative correlation with COVID-19 mortality[9] [10].Gemella species such as Gemella morbillorum, while less studied in this context, has been identified as a H_2S producer, albeit at lower rates compared to some

other oral bacteria[9]. Campylobacter species, particularly C. rectus, and Selenomonas thrive in anaerobic conditions and are known H_2S producers[9].

But how exactly does H_2S exert its protective effects?Its antiviral properties, previously demonstrated against other RNA viruses including influenza and respiratory syncytial virus (RSV) infections [11] [12], likely extend to SARS-CoV-2. These include direct antiviral activity[12], modulation of viral entry via ACE2 and TMPRSS2 receptors[13], (Fig 5) anti-inflammatory effects[12], prevention of cytokine storm[14],and enhancement of antioxidant defenses through the Nrf2 pathway [15] (Fig-5). Our findings are consistent with recent clinical observations showing higher serum H_2S levels in COVID-19 pneumonia survivors [16], reinforcing its potential as both a prognostic marker and therapeutic target. Furthermore, the validation of our results using influenza mortality data strengthens the broader implications of our findings, suggesting that the protective effect of H_2S producing bacteria extends beyond COVID-19 to other respiratory viral infections.

Despite our promising findings, our study has several limitations. The observational nature of this study precludes establishing causality. Reliance on existing datasets may introduce biases, and the precise mechanisms of H_2S protection warrant further investigation. Additionally, individual variations in oral hygiene and diet could not be controlled. These limitations highlight the need for longitudinal studies, controlled experiments, and clinical trials.

In conclusion, our study identifies a novel association between H_2S-producing oral bacteria and reduced COVID-19 mortality, underscoring the oral microbiome's potential role in modulating respiratory viral infection outcomes. This work paves the way for innovative approaches in the fight against respiratory viral diseases, including the exploration of H_2S-based therapeutics and probiotics.

Acknowledgement: None

Disclosure of interest: No potential conflict was reported by the anthor(s)

Funding: None

Authour Contribution

BP conceptualized the study, performed data analysis, interpreted the data, and wrote the manuscript. MV conducted data analysis, participated in the creation of graphics, tables, and charts, and managed the submission process. AM reviewed, edited, and refined the manuscript, and coordinated the submission process. KC assisted with initial data collection efforts and offered input on the taxonomic analysis approach.

References

[1] A. Karlinsky and D. Kobak, "Tracking excess mortality across countries during the COVID-19 pandemic with the World Mortality Dataset," *elife*, vol. 10, p. e69336, 2021.

[2] D. Chang, X. Chang, Y. He, and K. J. K. Tan, "The determinants of COVID-19 morbidity and mortality across countries," *Scientific reports*, vol. 12, no. 1, p. 5888, 2022.

[3] M. Hirayama *et al.*, "Intestinal Collinsella may mitigate infection and exacerbation of COVID-19 by producing ursodeoxycholate," *PLoS One*, vol. 16, no. 11, p. e0260451, 2021.

[4] M. Kilian *et al.*, "The oral microbiome–an update for oral healthcare professionals," *British dental journal*, vol. 221, no. 10, pp. 657–666, 2016.

[5] A. Kusiak *et al.*, "COVID-19 manifestation in the oral cavity–a narrative literature review," *Acta Otorhinolaryngologica Italica*, vol. 41, no. 5, p. 395, 2021.

[6] M. Ptasiewicz *et al.*, "Armed to the teeth—the oral mucosa immunity system and microbiota," *International Journal of Molecular Sciences*, vol. 23, no. 2, p. 882, 2022.

[7] L. Tan *et al.*, "Potential interaction between the oral microbiota and COVID-19: a meta-analysis and bioinformatics prediction," *Frontiers in Cellular and Infection Microbiology*, vol. 13, p. 1193340, 2023.

[8] K. E. Murros, "Hydrogen sulfide produced by gut bacteria may induce Parkinson's disease," *Cells*, vol. 11, no. 6, p. 978, 2022.

[9] S. Persson, M. Edlund, R. Claesson, and J. Carlsson, "The formation of hydrogen sulfide and methyl mercaptan by oral bacteria," *Oral Microbiology and Immunology*, vol. 5, no. 4, pp. 195–201, Aug. 1990, doi: 10.1111/j.1399-302X.1990.tb00645.x.

[10] L. Phillips, L. Chu, and D. Kolodrubetz, "Multiple enzymes can make hydrogen sulfide from cysteine in Treponema denticola," *Anaerobe*, vol. 64, p. 102231, 2020.

[11] N. Bazhanov, O. Escaffre, A. N. Freiberg, R. P. Garofalo, and A. Casola, "Broad-range antiviral activity of hydrogen sulfide against highly pathogenic RNA viruses," *Scientific reports*, vol. 7, no. 1, p. 41029, 2017.

[12] V. Citi *et al.*, "Anti-inflammatory and antiviral roles of hydrogen sulfide: Rationale for considering H_2S donors in COVID-19 therapy," *British J Pharmacology*, vol. 177, no. 21, pp. 4931–4941, Nov. 2020, doi: 10.1111/bph.15230.

[13] G. Yang, "H_2S as a potential defense against COVID-19?," *American Journal of Physiology-Cell Physiology*, vol. 319, no. 2, pp. C244–C249, Aug. 2020, doi: 10.1152/ajpcell.00187.2020.

[14] E. Magli *et al.*, "H2S donors and their use in medicinal chemistry," *Biomolecules*, vol. 11, no. 12, p. 1899, 2021.

[15] A. R. Bourgonje *et al.*, "N-Acetylcysteine and Hydrogen Sulfide in Coronavirus Disease 2019," *Antioxidants & Redox Signaling*, vol. 35, no. 14, pp. 1207–1225, Nov. 2021, doi: 10.1089/ars.2020.8247.

[16] Z. Ren *et al.*, "Alterations in the human oral and gut microbiomes and lipidomics in COVID-19," *Gut*, vol. 70, no. 7, pp. 1253–1265, 2021.

List of legend

Table legend

Table 1: Eight 16S rRNA-seq datasets from eight countries and the mortality rates of COVID-19.

Figure legend

Fig.:1 The methodology for this bioinformatics pipeline (a) Healthy oral microbiome data was obtained from the National Center for Biotechnology Information (NCBI). (b) This data was processed and analyzed using OmicsBox software. (c) The most abundant bacterial genera were identified. (d) COVID-19 mortality rate data from Our World in Data was incorporated (e) A statistical model (Generalized Linear Model) was used to analyze the relationship between the 15 most abundant genera and COVID-19 mortality. (f) A specialized technique called LIGER, typically used for analyzing single-cell RNA sequencing data, was applied to classify the oral microbiota of 244 healthy individuals into four distinct groups, termed "orotypes".

Fig 2: Genera level abundance of health oral metagenomic

Fig 3: Plot of p-values of 15 genera in a generalized linear model (GLM) to predict the COVID-19 mortality rates.

Fig 4 (a) Unsupervised clustering of oral microbiota in 244 healthy subjects in ten countries by LIGER generated 4 orotypes. Each subject is plotted with t-SNE and is color-coded by its orotype. (b) Fractions of orotypes 1 to 4 in eight countries. eight countries are sorted in descending order of the COVID-19 mortality rates per million, which are indicated in parentheses. (c) The t-SNE plot is color-coded by the COVID-19 mortality rates in eight

countries. (d) Mean and standard error of the COVID-19 mortality rates in orotypes 1 to 4. $p = 0.203$ by Jonckheere-Terpstra trend test. (e) The t-SNE plot is color-coded by the relative abundance of genus *Treponema*. (f) Mean and standard error of the relative abundance of genus *Treponema* in orotypes 1 to 4. $P = 0.033$by Jonckheere-Terpstra trend test.

Fig 5: Proposed mechanisms of oral bacteria-produced hydrogen sulfide (H_2S) in protecting against SARS-CoV-2 infection. H_2S, produced by oral bacteria including Treponema, Gemella, Campylobacter, and Selenomonas, exhibits multifaceted protective effects against SARS-CoV-2. These include: (a) Direct antiviral activity, (b) Modulation of viral entry by altering ACE2 and TMPRSS2 receptors, (c) Anti-inflammatory effects and prevention of cytokine storm, and (d) Enhancement of antioxidant defenses through the Nrf2 pathway. These mechanisms, previously observed in other RNA viral infections, likely contribute to the potential protective role of H_2S-producing oral bacteria against severe COVID-19 outcomes.

Tables:

Table 1: Eight 16S rRNA-seq datasets from Eight countries and the mortality rates of COVID-19.

Dataset	City, Country	Accession number	Mortality rate per million[a]	The number of samples	Sequencing
1	Australia	PRJDB5153	35.26	18	16S rRNA (V3-V4)
2	Oklahoma, USA	PRJNA292800	1429.6	57	16S rRNA
3	Trento, Italy	PRJNA227796	1551.22	14	16S rRNA

4	Sweden	PRJNA598825	1188.41	12	16S rRNA (V3-V4)
5	Jazan, Saudi arabia	PRJNA267483	175.84	12	16S rRNA
6	Brazil, Sao Paolo, Piracicaba	PRJNA606501	1075.33	104	16S rRNA (V1-3 and V4-5)
7	Egypt	PRJNA384402	86.95	17	16S rRNA
8	Russia, Moscow	PRJNA256234	536.22	10	16S rRNA

286

287

Figures

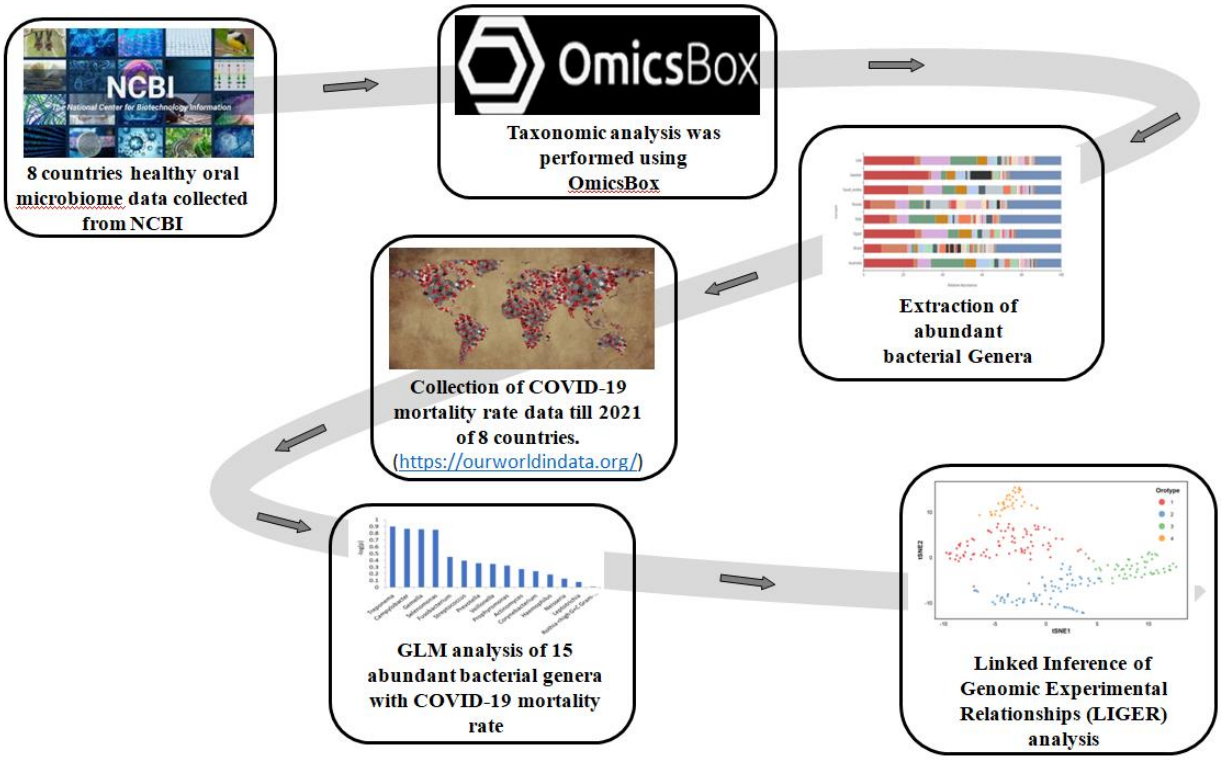
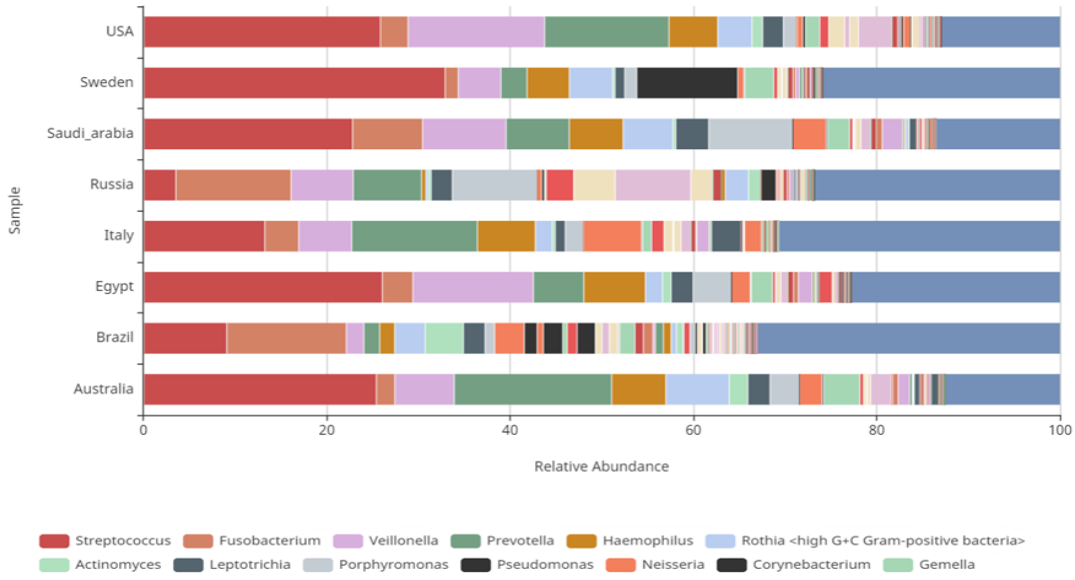
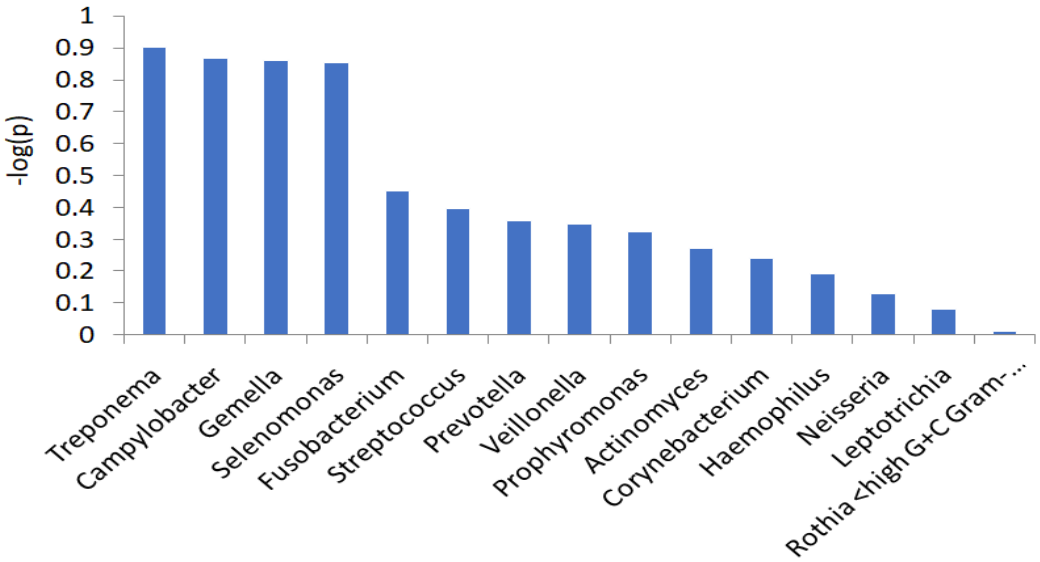


Fig 1: The methodology for this bioinformatics pipeline (a) Healthy oral microbiome data was obtained from the National Center for Biotechnology Information (NCBI). (b) This data was processed and analyzed using OmicsBox software. (c) The most abundant bacterial genera were identified. (d) COVID-19 mortality rate data from Our World in Data was incorporated (e) A statistical model (Generalized Linear Model) was used to analyze the relationship between the 15 most abundant genera and COVID-19 mortality. (f) A specialized technique called LIGER, typically used for analyzing single-cell RNA sequencing data, was applied to classify the oral microbiota of 244 healthy individuals into four distinct groups, termed "orotypes".



299

300 **Fig 2: Genera level abundance of health oral metagenomic**

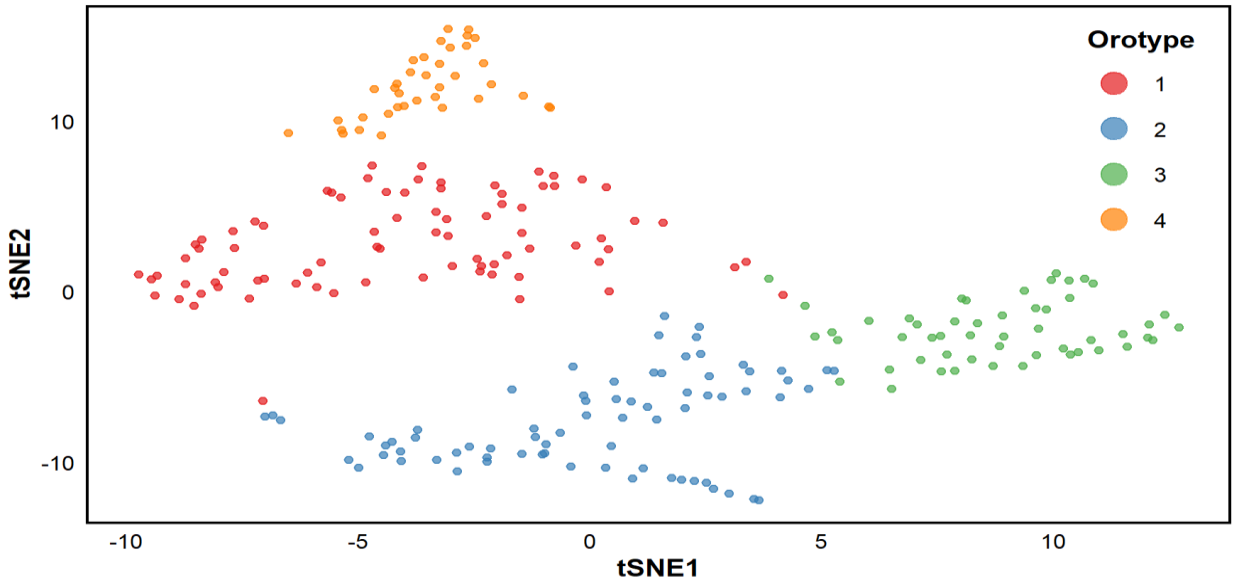


301

302

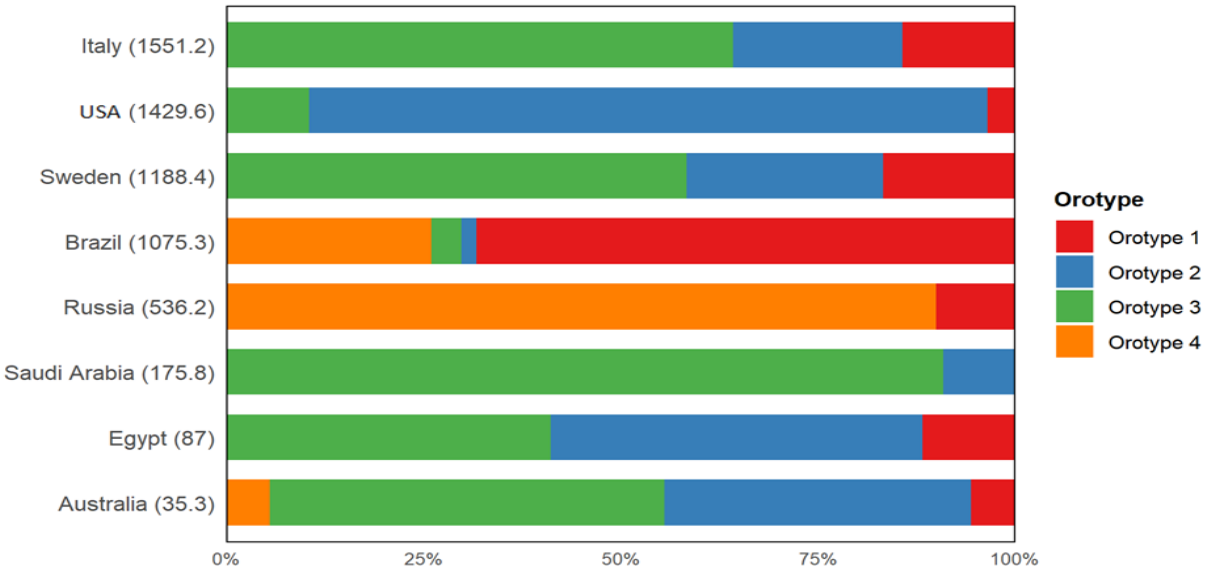
303 **Fig 3: Plot of p-values of 15 genera in a generalized linear model (GLM) to predict the**
304 **COVID-19 mortality rates.**

305



306

307 **Fig 4 (a):Unsupervised clustering of oral microbiota in 244 healthy subjects in ten countries**
308 **by LIGER generated 4 orotypes. Each subject is plotted with t-SNE and is color-coded by**
309 **its orotype.**



310

311 **Fig 4 (b): Fractions of orotypes 1 to 4 in eight countries. eight countries are sorted in**
312 **descending order of the COVID-19 mortality rates per million, which are indicated in**
313 **parentheses.**

314

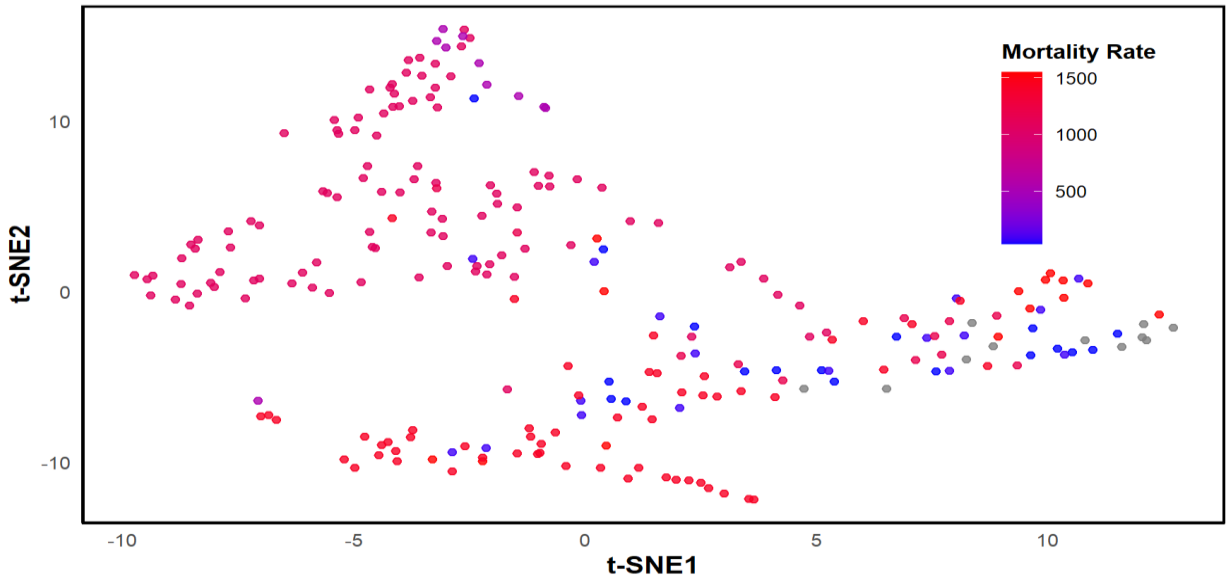


Fig 4 (c): The t-SNE plot is color-coded by the COVID-19 mortality rates in eight countries.

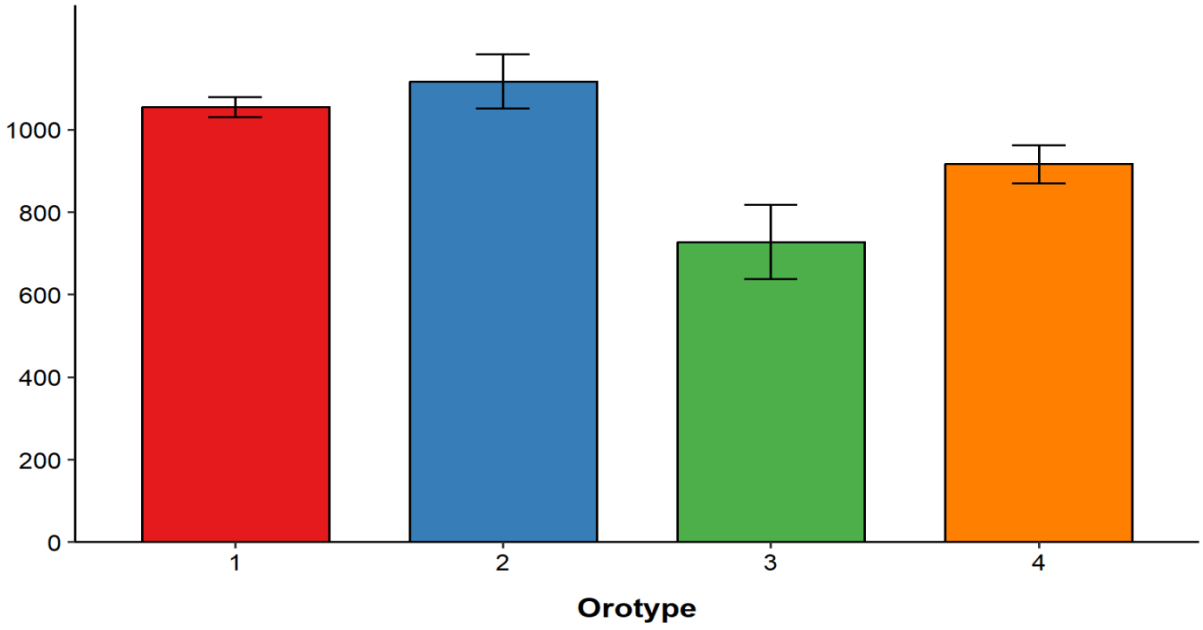


Fig 4 (d): Mean and standard error of the COVID-19 mortality rates in orotypes 1 to 4. *p* = *0.203* by Jonckheere-Terpstra trend test.

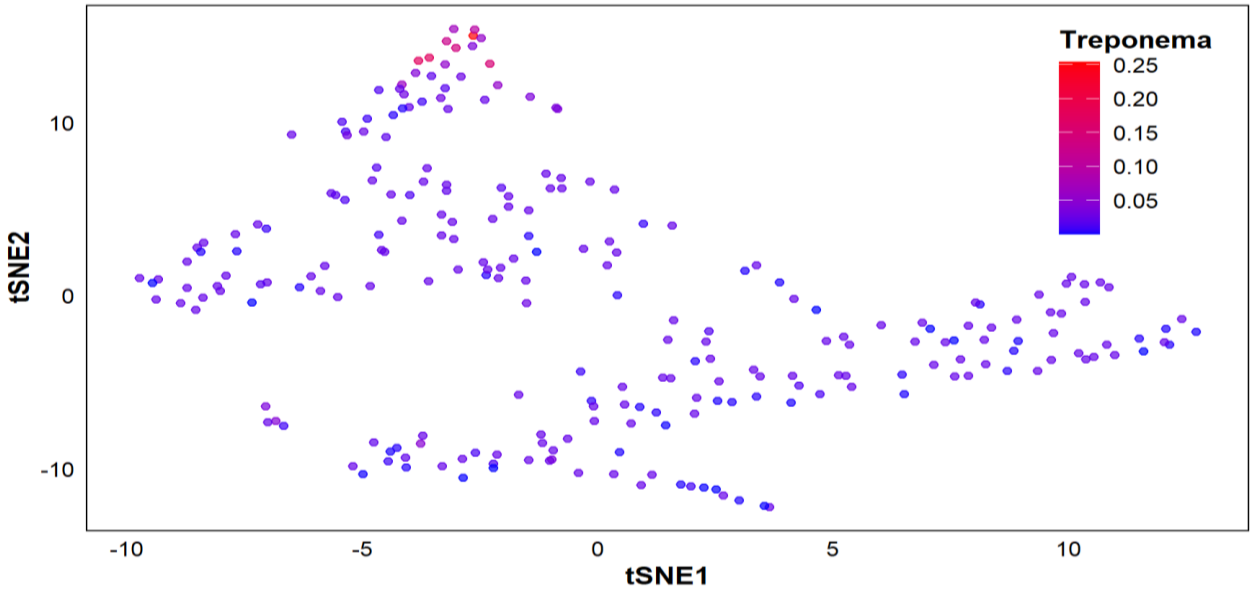


Fig 4 (e): The t-SNE plot is color-coded by the relative abundance of genus *Treponema*.

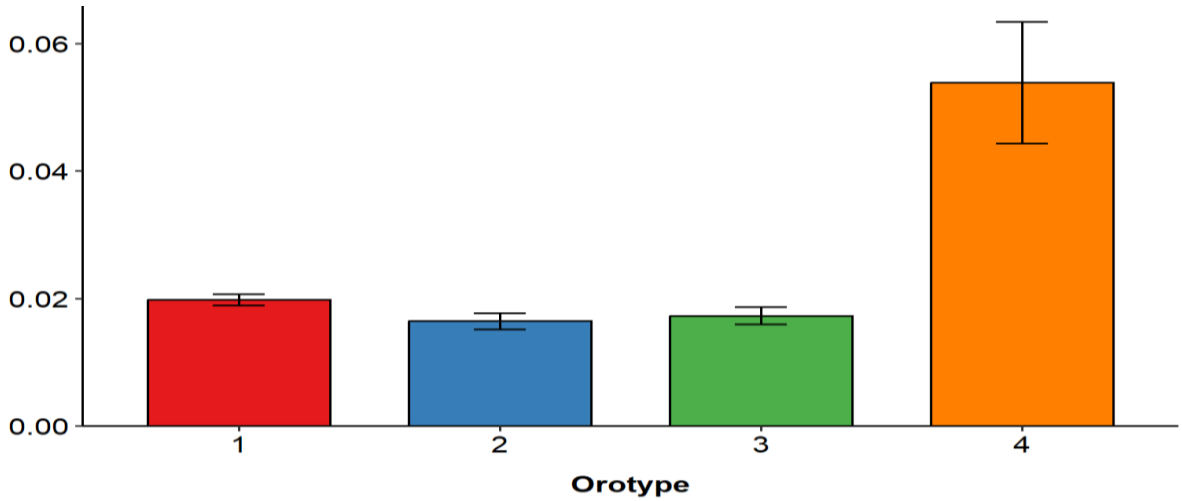


Fig 4 (f): Mean and standard error of the relative abundance of genus *Treponema* in orotypes 1 to 4. *P* = 0.033by Jonckheere-Terpstra trend test.

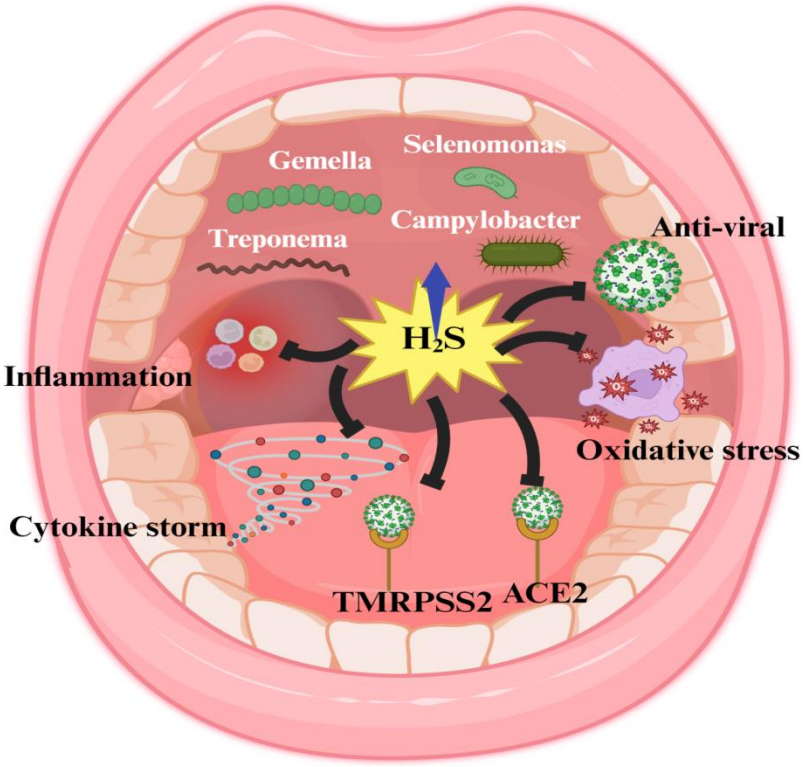
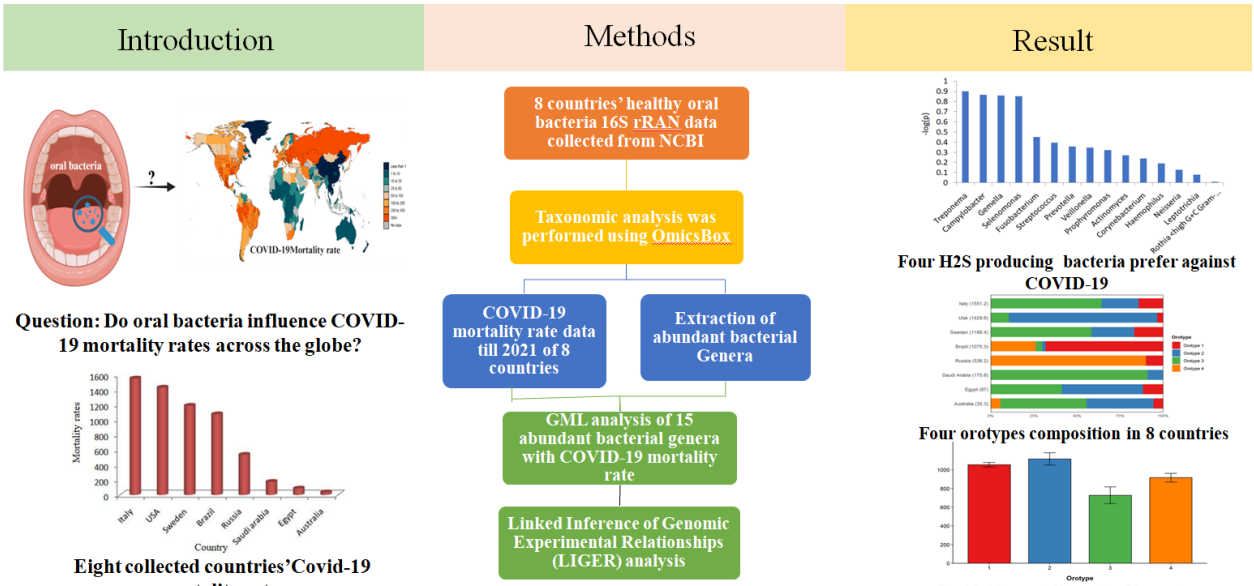


Fig 5: Proposed mechanisms of oral bacteria-produced hydrogen sulfide (H_2S) in protecting against SARS-CoV-2 infection. H_2S, produced by oral bacteria including Treponema, Gemella, Campylobacter, and Selenomonas, exhibits multifaceted protective effects against SARS-CoV-2. These include: (a) Direct antiviral activity, (b) Modulation of viral entry by altering ACE2 and TMPRSS2 receptors, (c) Anti-inflammatory effects and prevention of cytokine storm, and (d) Enhancement of antioxidant defenses through the Nrf2 pathway. These mechanisms, previously observed in other RNA viral infections, likely contribute to the potential protective role of H_2S-producing oral bacteria against severe COVID-19 outcomes.

343 **Graphical Abstract**



344

345

Supplementary Information

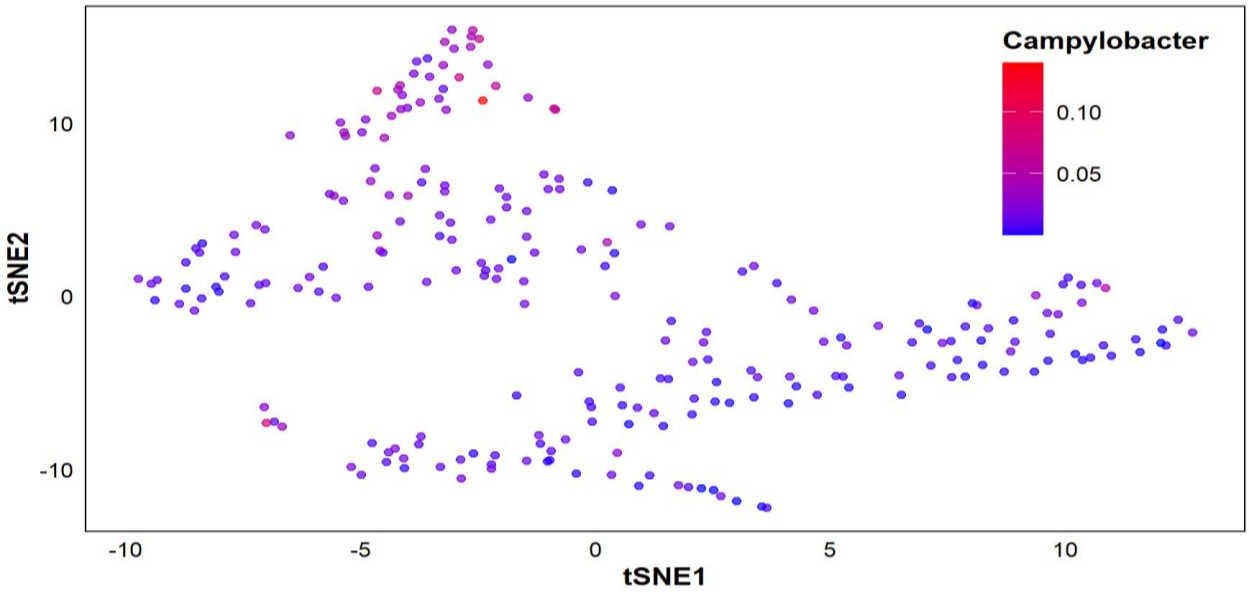


Fig S1a: The t-SNE plot is color-coded by the relative abundance of genus *Campylobacter*.

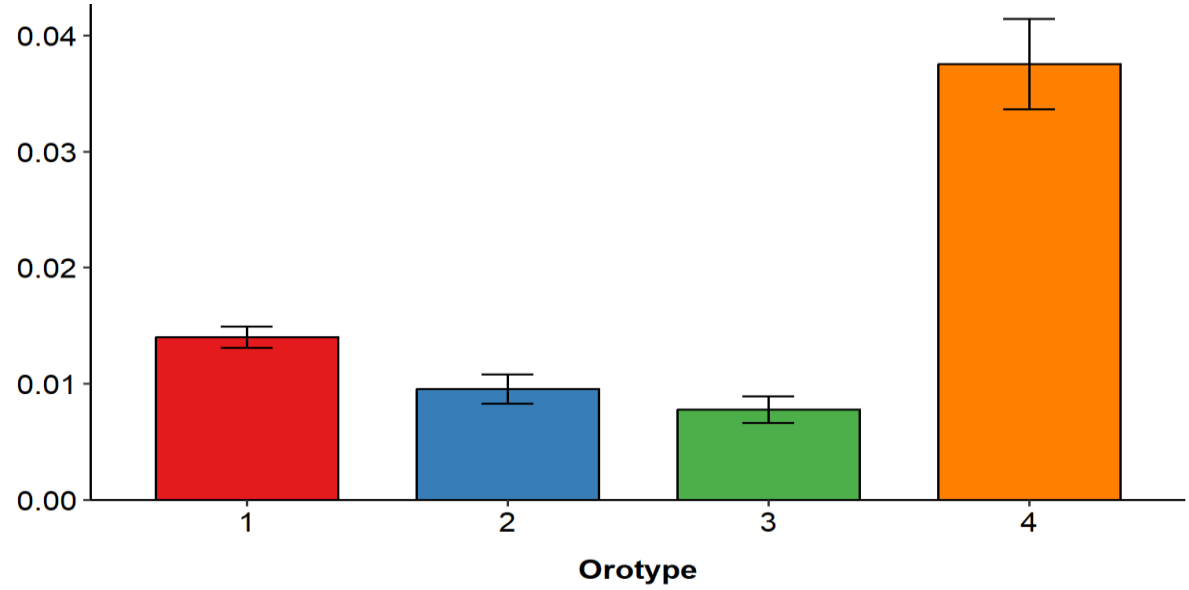


Fig S1b: Mean and standard error of the relative abundance of genus *Campylobacter* in orotypes 1 to 4. *P* = 0.236 by Jonckheere-Terpstra trend test.

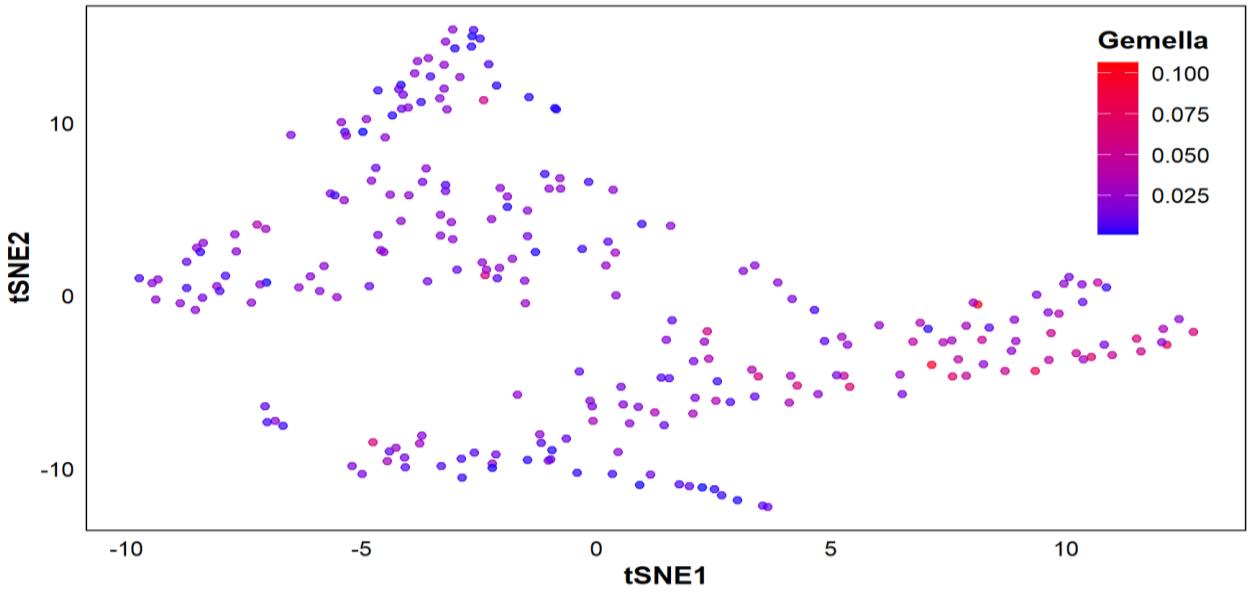


Fig S1c: The t-SNE plot is color-coded by the relative abundance of genus *Gemella*.

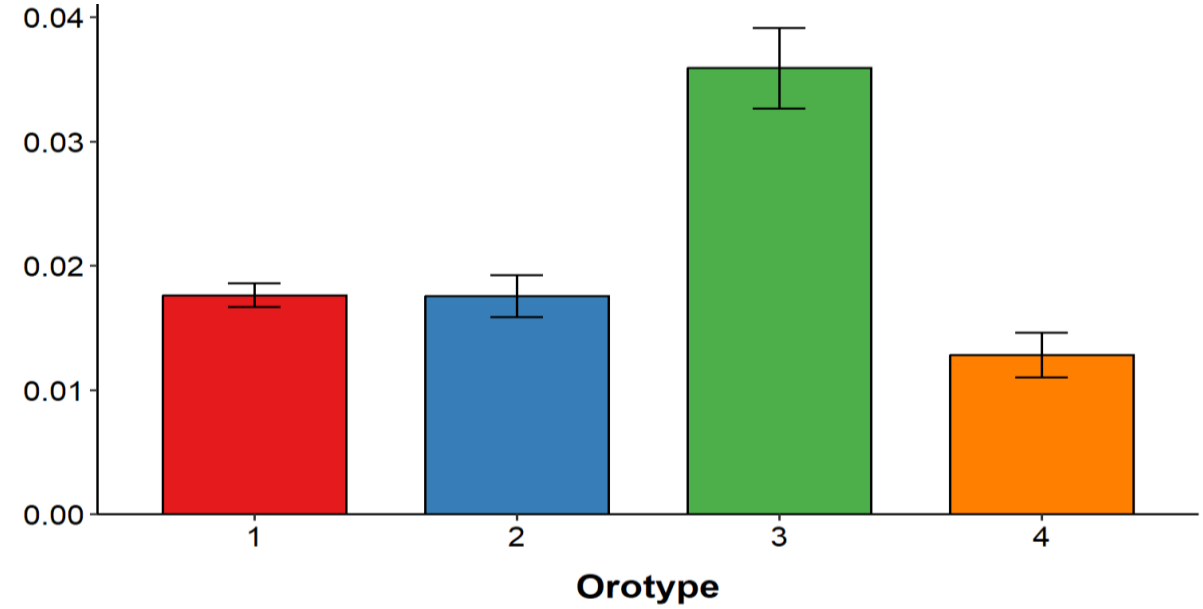
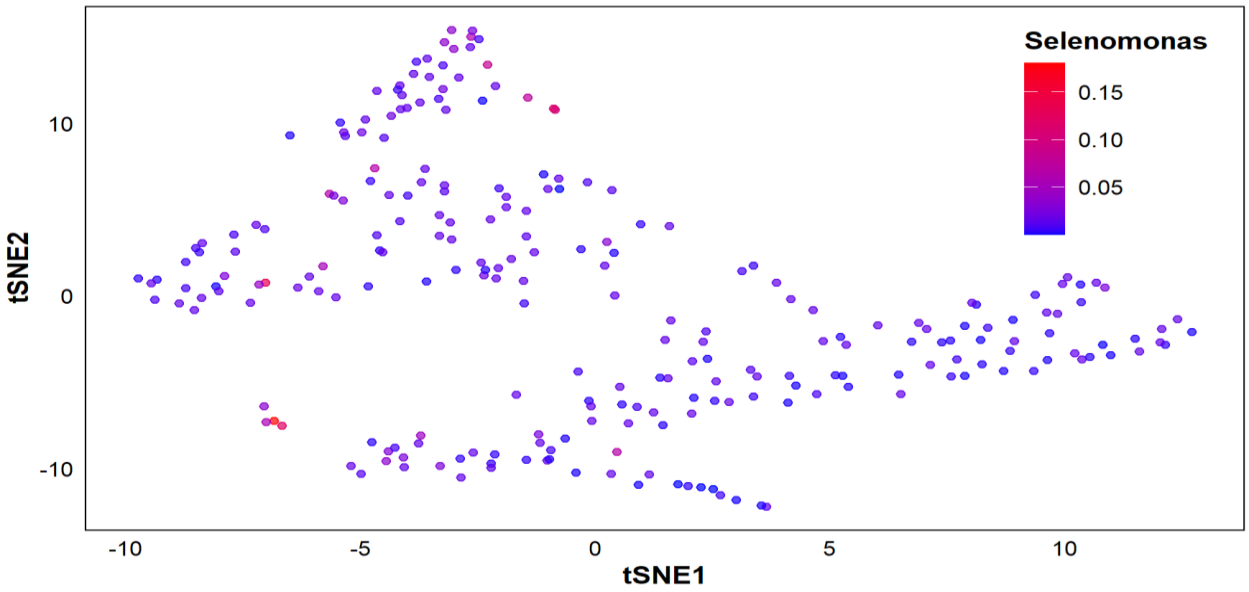


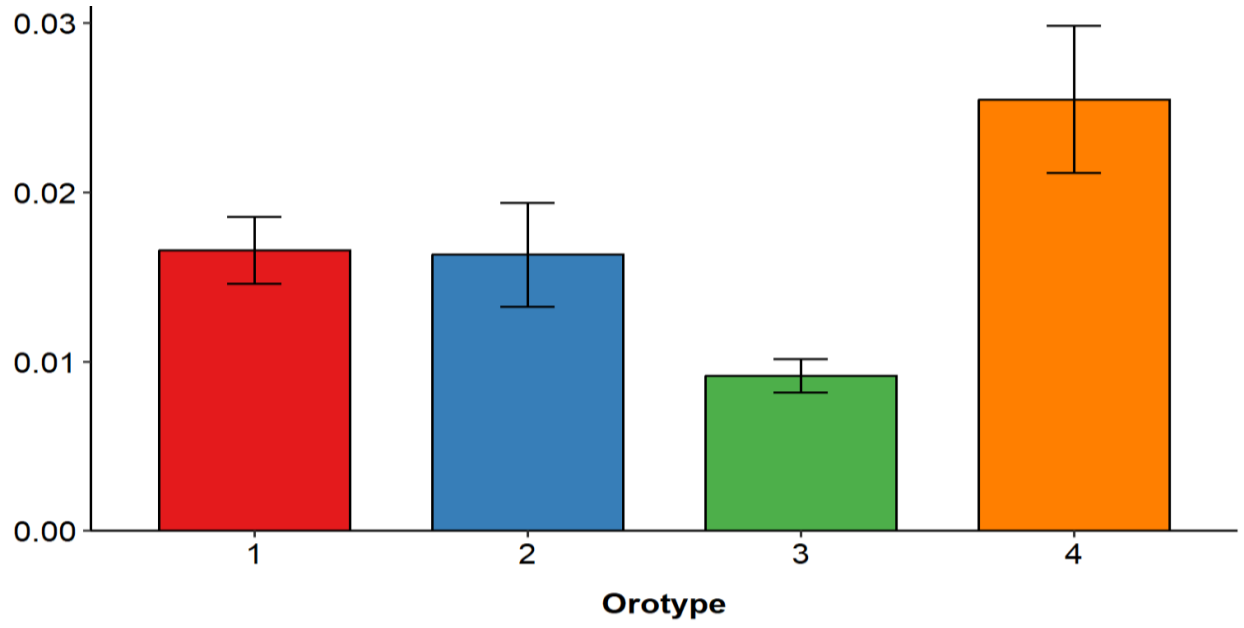
Fig S1d: Mean and standard error of the relative abundance of genus *Gemella* in orotypes 1 to 4. *P* = 0.652 by Jonckheere-Terpstra trend test.



365

366 **Fig S1e: The t-SNE plot is color-coded by the relative abundance of**
367 **genus *Selenomonas*.**

368



369

370 **Fig S1f: Mean and standard error of the relative abundance of**
371 **genus *Selenomona* in orotypes 1 to 4. *P* = 0.859 by Jonckheere-Terpstra trend**
372 **test.**

373

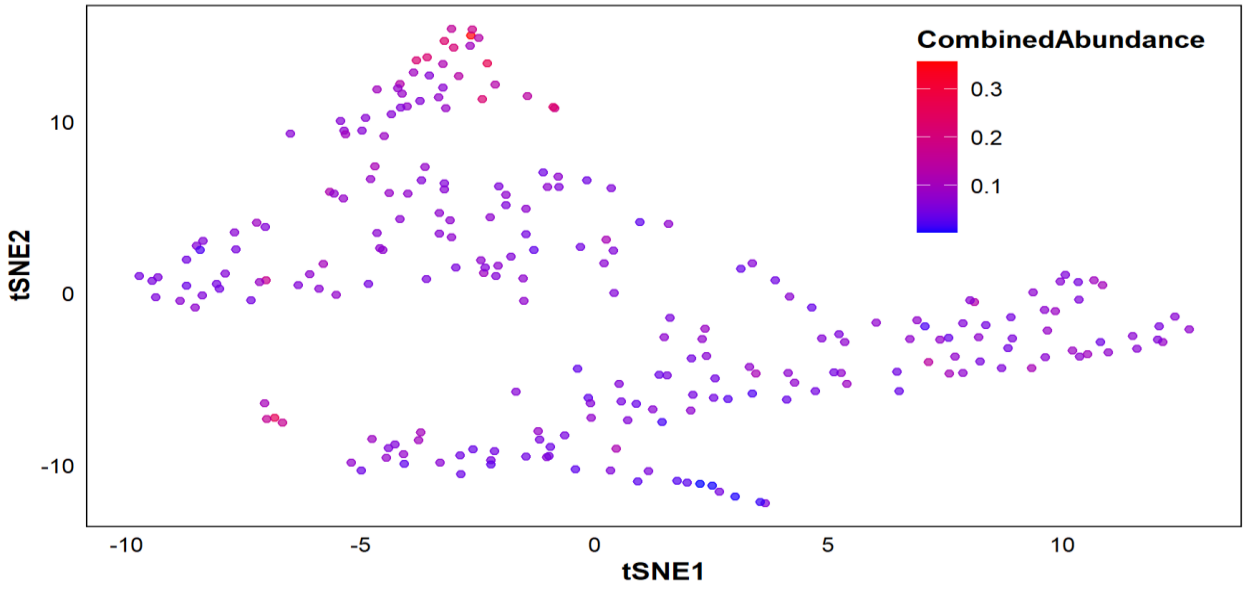


Fig S1g: The t-SNE plot is color-coded by the relative combine abundance of genus *Treponema, Gemella, Campylobacter, Selenomonas*.

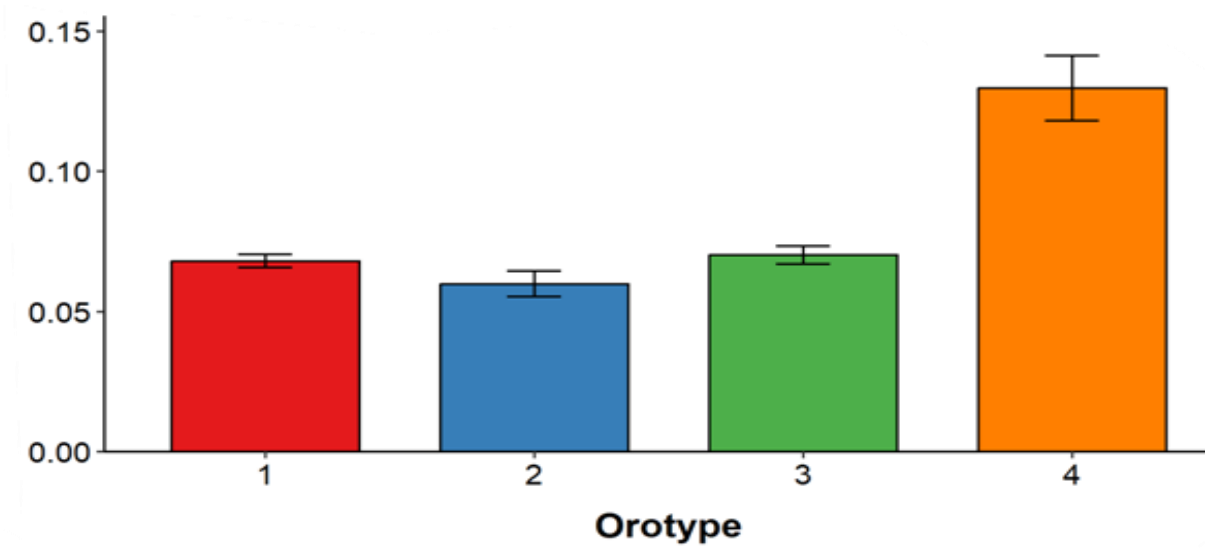


Fig S1h: Mean and standard error of the relative combine abundance of genus *Treponema, Gemella, Campylobacter, Selenomonas* in orotypes 1 to 4. p = 7.198e-05 by Jonckheere-Terpstra trend test.

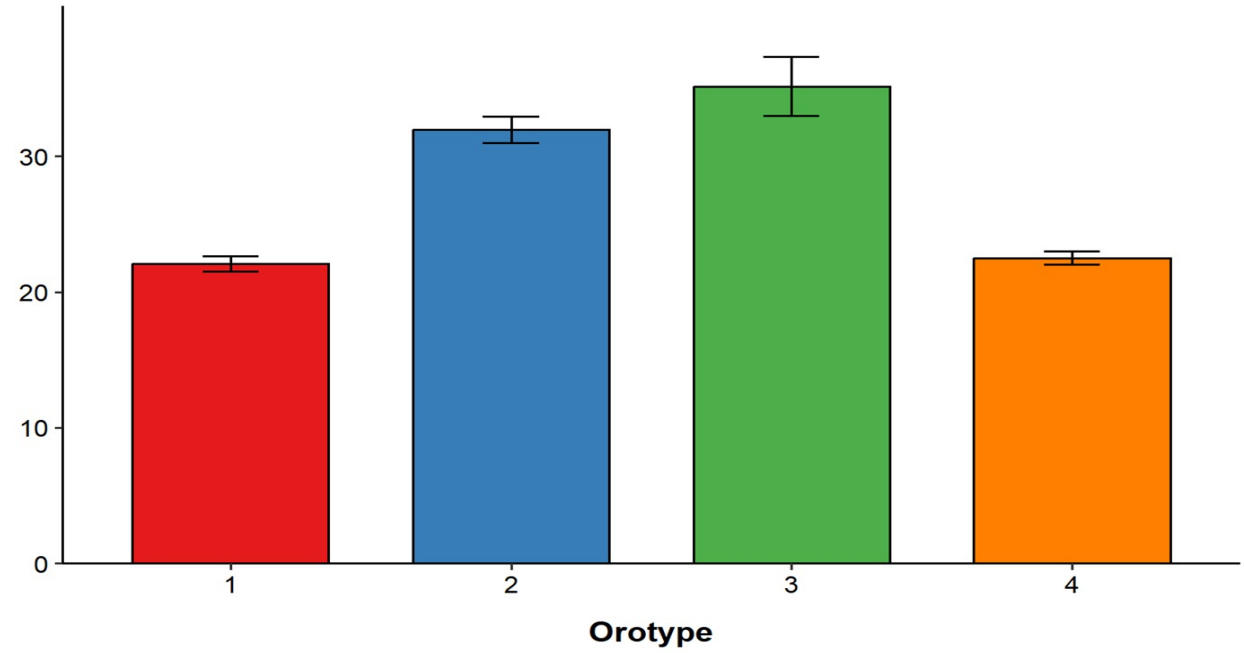


Fig S2: Mean and standard error of the Influenza mortality rates in orotypes 1 to 4. *p = 4e - 04* by Jonckheere-Terpstra trend test.

Supplementary Table 1. Generalized linear model (GLM) to predict the COVID-19 mortality rates with 15 oral bacteria

Genera	P-value	Relative abundance (%)[a]
Streptococcus	0.401	21.34
Fusobacterium	0.354	11.18
Veillonella	0.45	8.76
Prevotella	0.439	11.84
Hemophilus	0.644	3.86
Rothia<high G+CGram-positivebacteria>	0.977	5.09
Actinomyces	0.536	3.67
Leptotrichia	0.829	2.97
Prophyromonas	0.474	3.86
Neisseria	0.742	3.26
Corynebacterium	0.577	2.18
Gemella	0.138	2.08
Campylobacter	0.136	1.49
Selenomonas	0.14	1.63
Treponema	0.125	2.35

[a]Average relative abundance in 244 healthy subjects in eight countries

Supplementary Table 2: Most abundance genera in each orotype

orotype	Top 5 Genera
1	*Streptococcus, Prevotella, Veillonella, Fusobacterium, Haemophilus*
2	*Streptococcus, Fusobacterium, Prevotella, Rothia (high G+C Gram-positive bacteria), Veillonella*
3	*Streptococcus, Fusobacterium, Prevotella, Veillonella, Actinomyces*
4	*Streptococcus, Veillonella, Prevotella, Fusobacterium, Megasphaera*

400 **Supplementary Table 3. The mean relative abundances of 15 most prevalent**
401 **genera for each orotype**

402

Genus	Orotype 1	Orotype 2	Orotype 3	Orotype 4
Streptococcus	18.00%	35.33%	42.79%	5.60%
Fusobacterium	13.91%	5.12%	4.22%	46.79%
Veillonella	6.99%	19.50%	7.53%	5.76%
Prevotella	9.86%	26.60%	10.87%	6.69%
Haemophilus	4.51%	3.25%	10.03%	3.29%
Rothia <high G+C Gram-positive bacteria>	6.61%	4.24%	10.60%	2.77%
Actinomyces	10.66%	2.07%	1.58%	3.56%
Leptotrichia	5.64%	3.25%	2.56%	2.69%
Porphyromonas	5.04%	3.13%	5.37%	8.68%
Neisseria	5.00%	2.66%	7.17%	2.60%
Corynebacterium	5.43%	1.53%	1.38%	1.97%
Gemella	2.69%	2.14%	4.58%	1.59%
Campylobacter	2.14%	1.12%	0.98%	4.64%
Selenomonas	2.42%	1.90%	1.17%	3.03%
Treponema	3.12%	1.86%	2.24%	6.78%

403

404

List of legend

Supplementary information of legend

Fig S1: (a) The t-SNE plot is color-coded by the relative abundance of genus *Campylobacter*. (b) Mean and standard error of the relative abundance of genus *Campylobacter* in orotypes 1 to 4. $P = 0.236$ by Jonckheere-Terpstra trend test. (c)The t-SNE plot is color-coded by the relative abundance of genus *Gemella*. (d)Mean and standard error of the relative abundance of genus *Gemella* in orotypes 1 to 4. $P = 0.652$ by Jonckheere-Terpstra trend test. (e)The t-SNE plot is color-coded by the relative abundance of genus *Selenomonas*. (f)Mean and standard error of the relative abundance of genus *Selenomona* in orotypes 1 to 4. $P = 0.859$ by Jonckheere-Terpstra trend test. (g)The t-SNE plot is color-coded by the relative combine abundance of genus *Treponema, Gemella, Campylobacter, Selenomonas*. (h)Mean and standard error of the relative combine abundance of genus *Treponema, Gemella, Campylobacter, Selenomonas* in orotypes 1 to 4. p = 7.198e-05 by Jonckheere-Terpstra trend test.

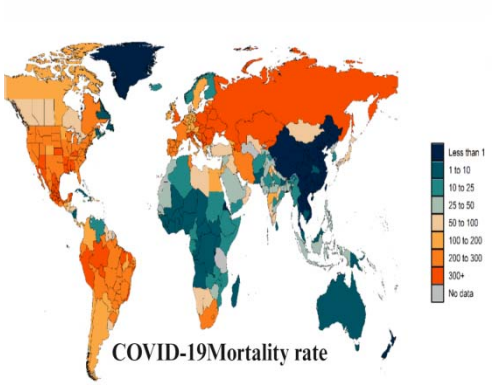
Fig S2: Mean and standard error of the Influenza mortality rates in orotypes 1 to 4. $p = 4e - 04$ by Jonckheere-Terpstra trend test

Supplementary Table 1: Generalized linear model (GLM) to predict the COVID-19 mortality rates with 15 oral bacteria

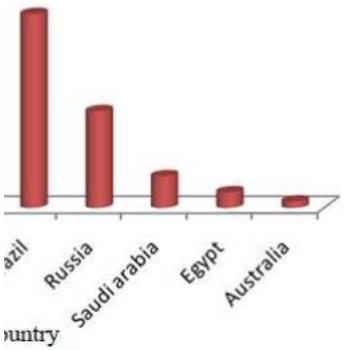
Supplementary Table 2: Most abundance genera in each orotype

Supplementary Table 3: The mean relative abundances of 15 most prevalent genera for each orotype

ıction



teria influence COVID-
s across the globe?



countries'Covid-19
ality rates

Methods

8 countries' healthy oral bacteria 16S rRAN data collected from NCBI

Taxonomic analysis was performed using OmicsBox

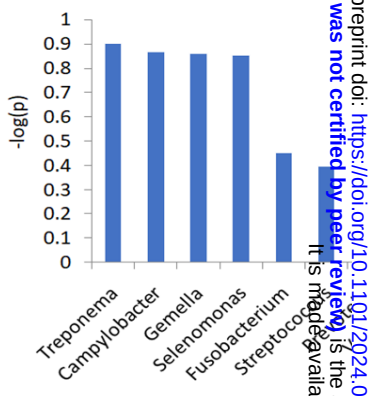
COVID-19 mortality rate data till 2021 of 8 countries

Extraction of abundant bacterial Genera

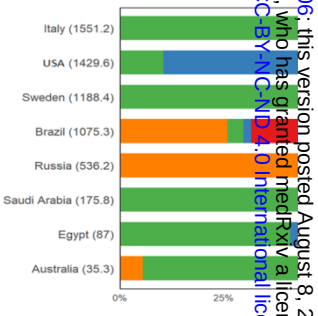
GML analysis of 15 abundant bacterial genera with COVID-19 mortality rate

Linked Inference of Genomic Experimental Relationships (LIGER) analysis

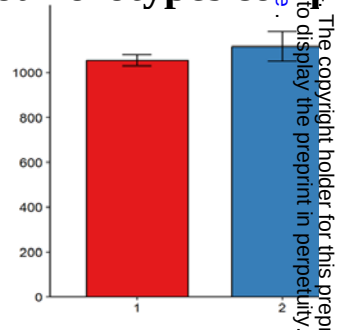
Resu



Four H2S producing
COV



Four orotypes com



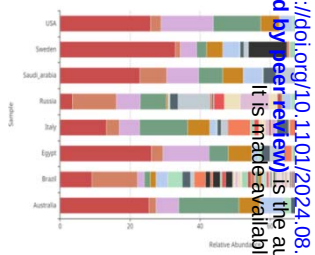
Covid-19 mortali

logy for this
e (a) Healthy
was obtained
Center for
ation (NCBI).
processed and
Box software.
lant bacterial
. (d) COVID-
ta from Our
corporated (e)
(Generalized
ed to analyze
en the 15 most
d COVID-19
ized technique
ally used for
cell RNA
s applied to
obiota of 244
o four distinct
es".

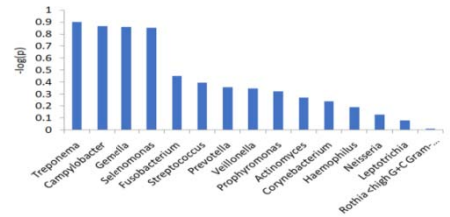
8 countries healthy oral microbiome data collected from NCBI

Taxonomic analysis was performed using OmicsBox

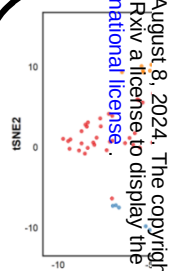
Collection of COVID-19 mortality rate data till 2021 of 8 countries. (https://ourworldindata.org/)



Extraction abundan bacterial Gen



GLM analysis of 15 abundant bacterial genera with COVID-19 mortality rate



Lin Gen Rela

Table 1: Eight 16S rRNA-seq datasets from Eight countries and the mortality rates of COVID-19.

Dataset	City, Country	Accession number	Mortality rate per million[a]	The number of samples	Sequencing
1	Australia	PRJDB5153	35.26	18	16S rRNA (V3-V4)
2	Oklahoma, USA	PRJNA292800	1429.6	57	16S rRNA
3	Trento, Italy	PRJNA227796	1551.22	14	16S rRNA
4	Sweden	PRJNA598825	1188.41	12	16S rRNA (V3-V4)
5	Jazan, Saudi arabia	PRJNA267483	175.84	12	16S rRNA
6	Brazil, Sao Paolo, Piracicaba	PRJNA606501	1075.33	104	16S rRNA (V1-3 and V4-5)
7	Egypt	PRJNA384402	86.95	17	16S rRNA
8	Russia, Moscow	PRJNA256234	536.22	10	16S rRNA

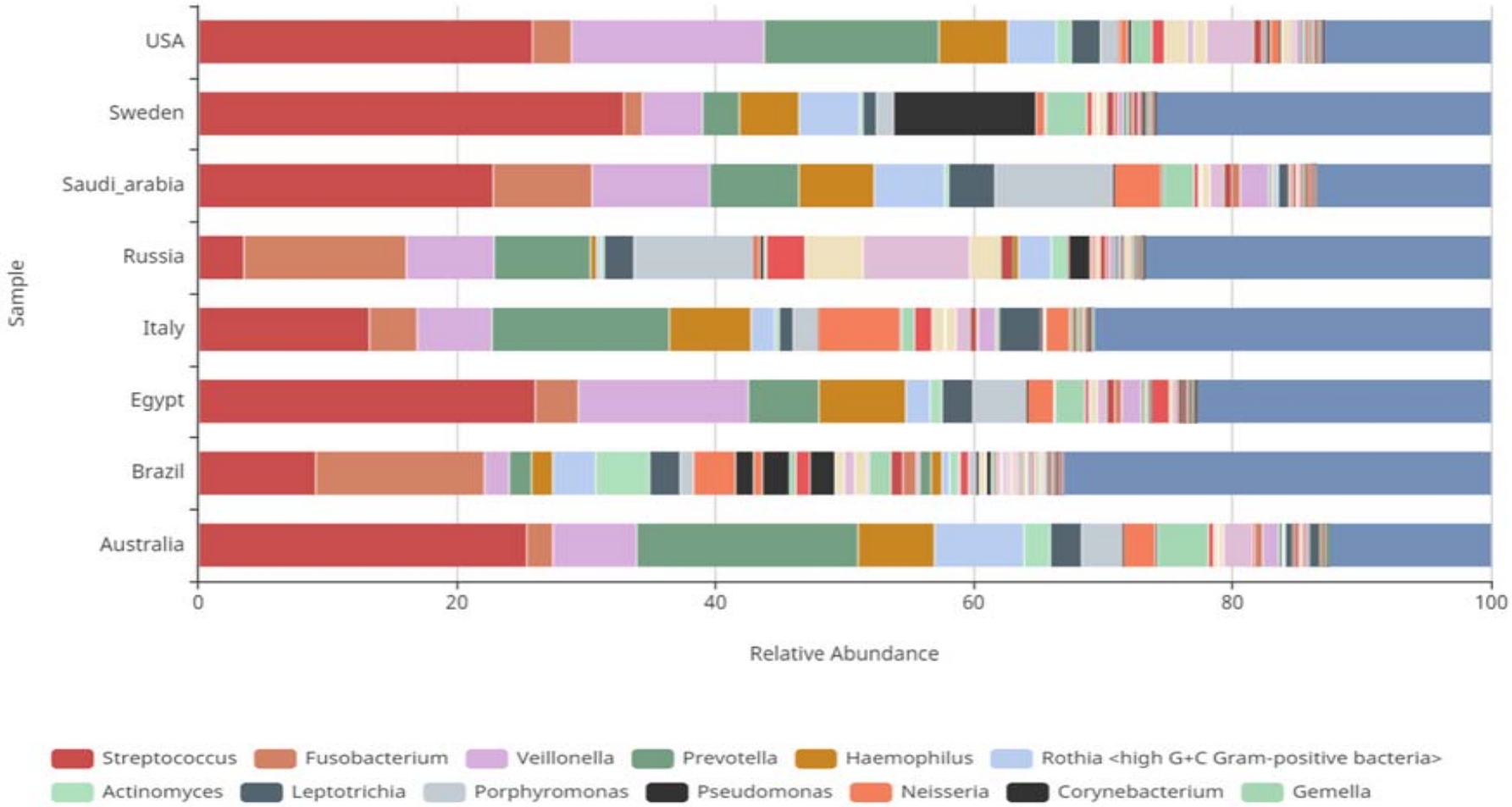


Fig 2: Genera level abundance of health oral metagenomic

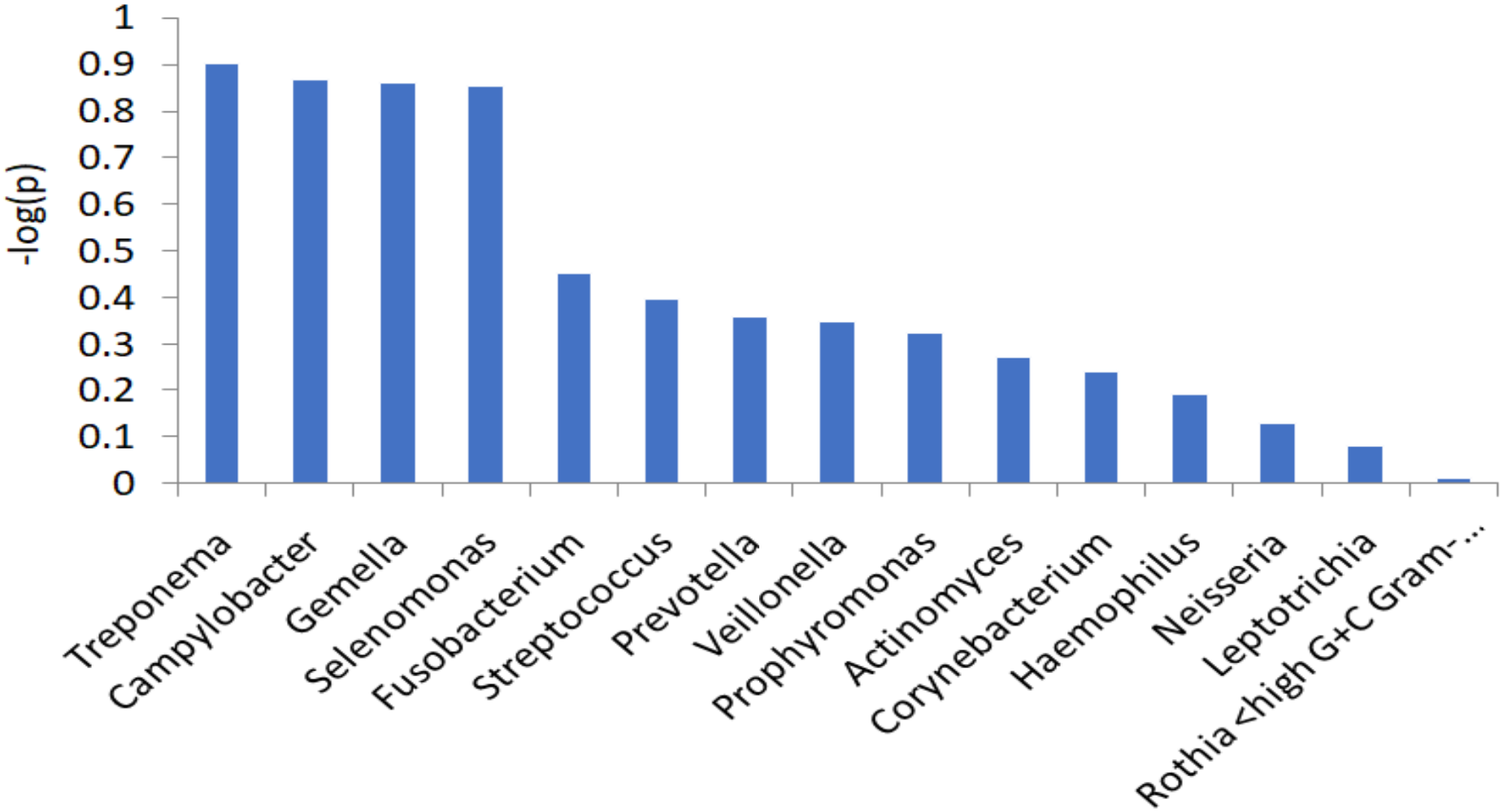


Fig 3: Plot of p-values of 15 genera in a generalized linear model (GLM) to predict the COVID-19 mortality rates.

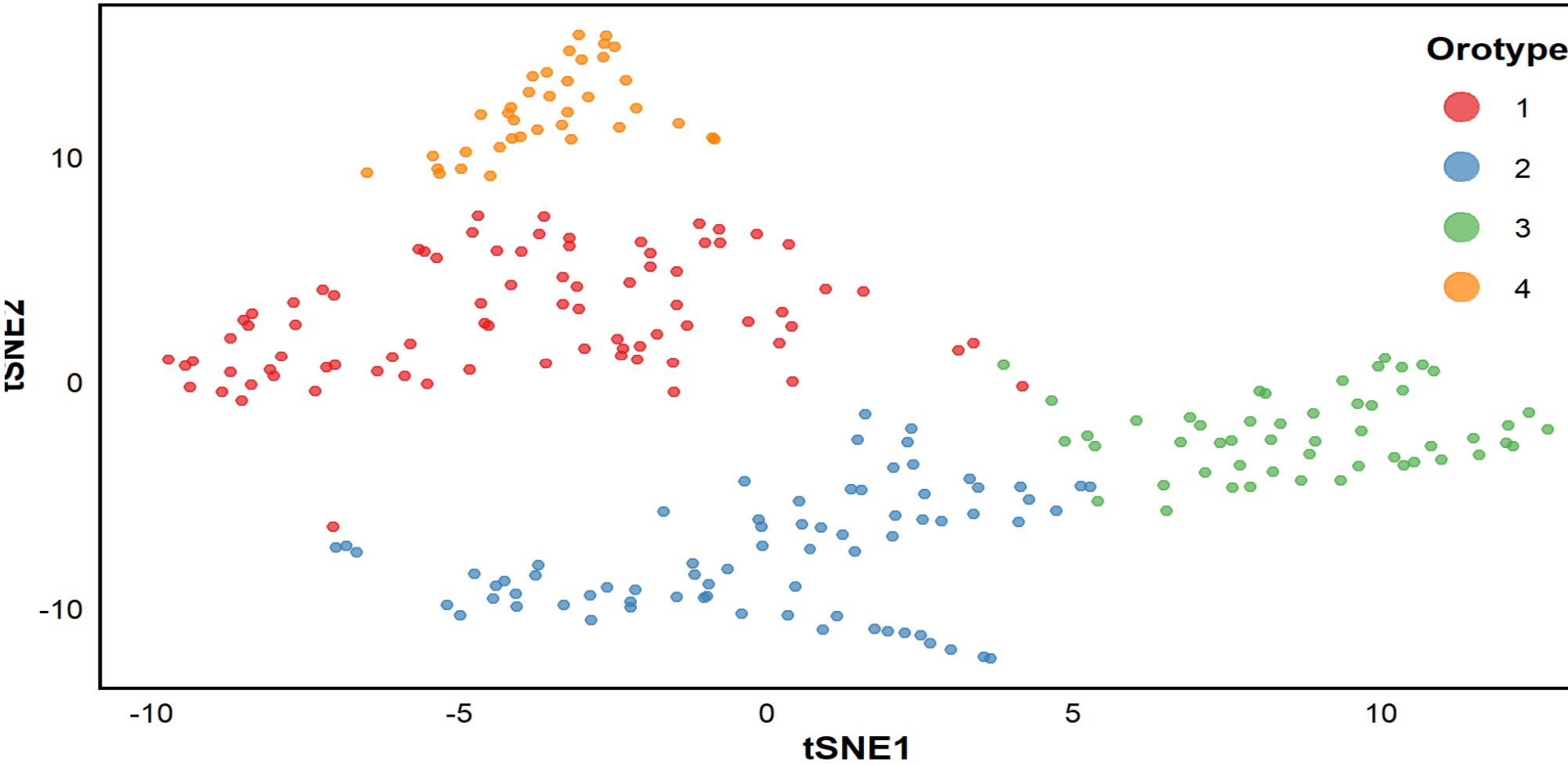


Fig 4 (a):Unsupervised clustering of oral microbiota in 244 healthy subjects in ten countries by LIGER generated 4 orotypes. Each subject is plotted with t-SNE and is color-coded by its orotype.

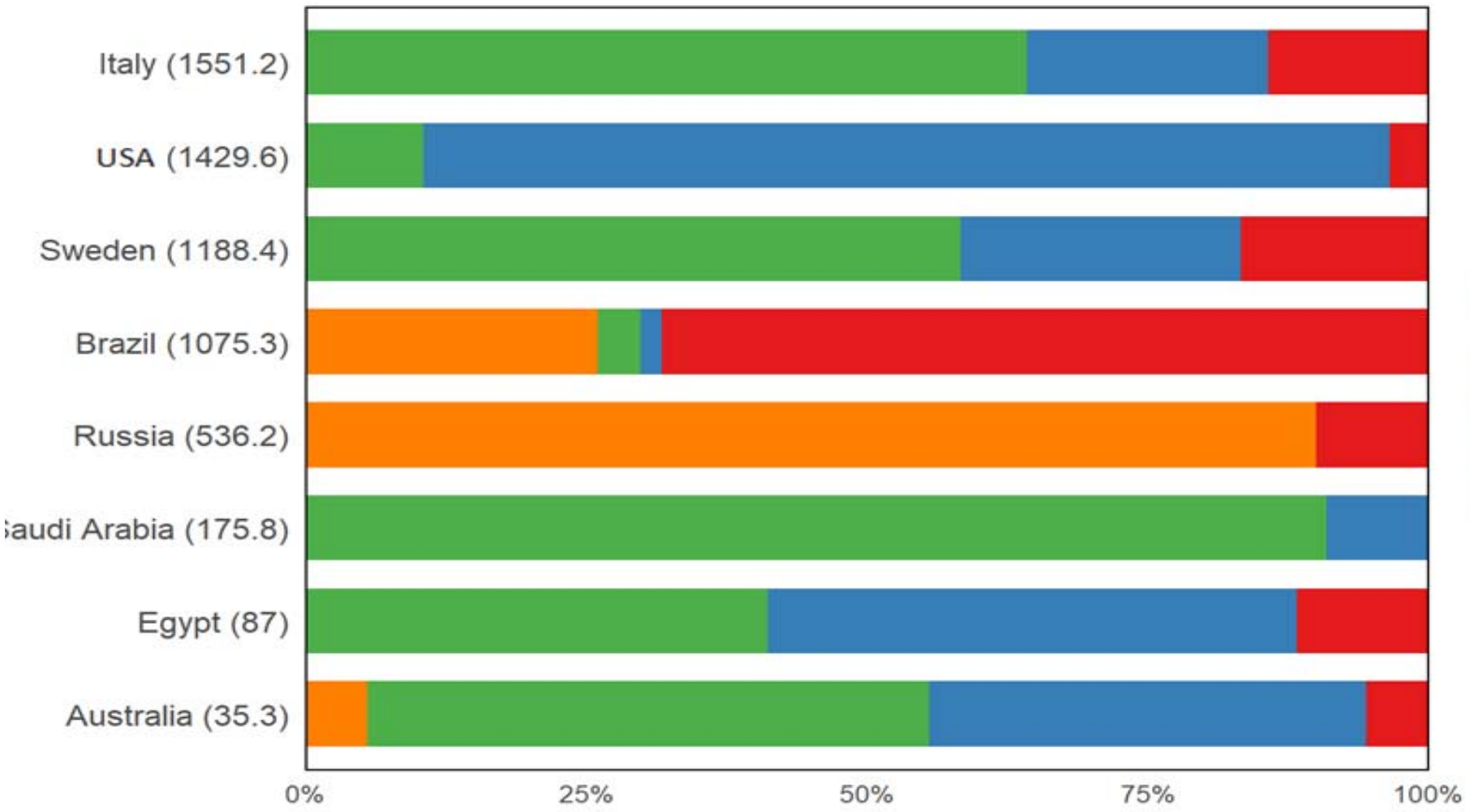


Fig 4 (b): Fractions of orotypes 1 to 4 in eight countries. eight countries are sorted in descending order of the COVID-19 mortality rates per million, which are indicated in parentheses.

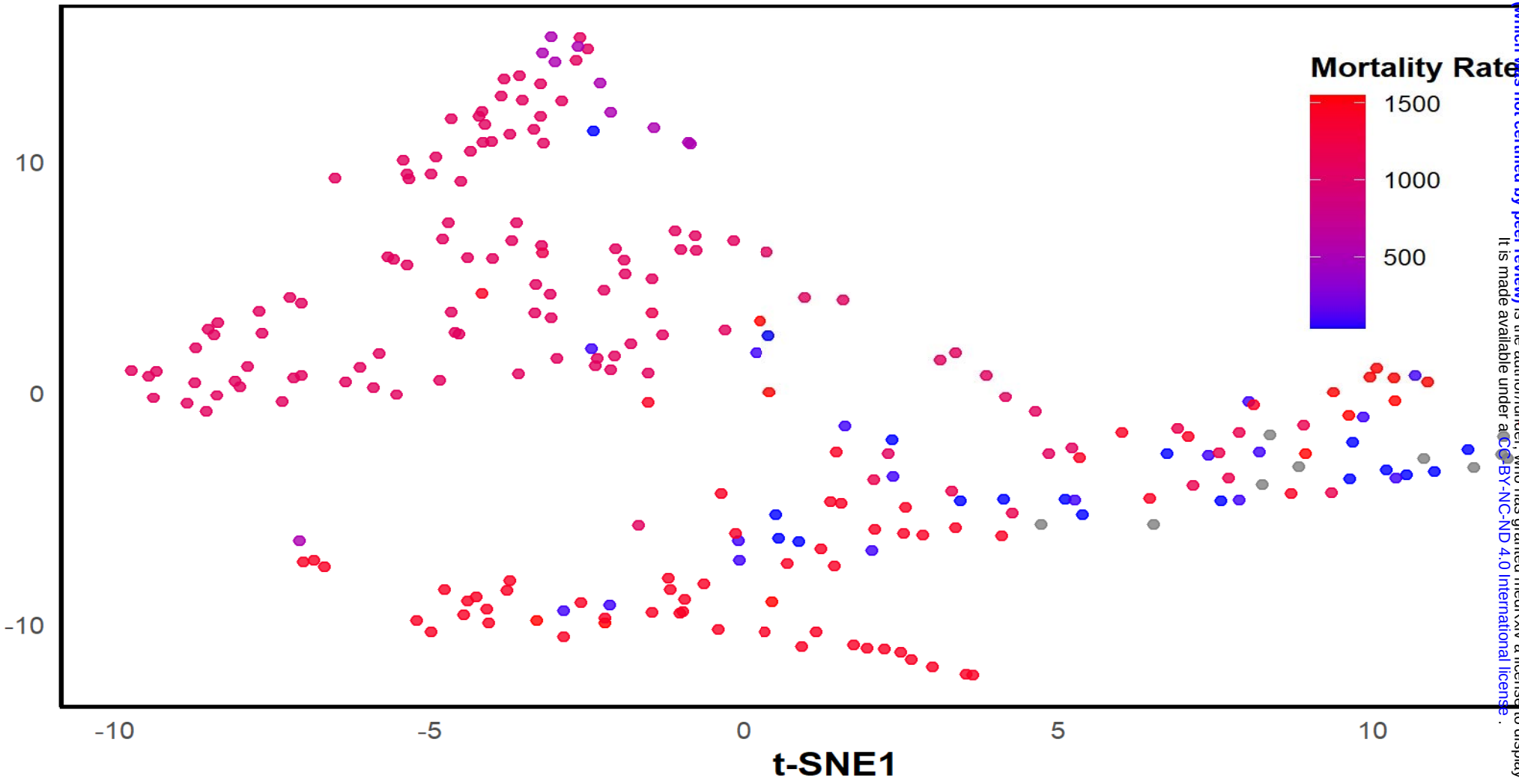


Fig 4 (c): The t-SNE plot is color-coded by the COVID-19 mortality rates in Eight countries.

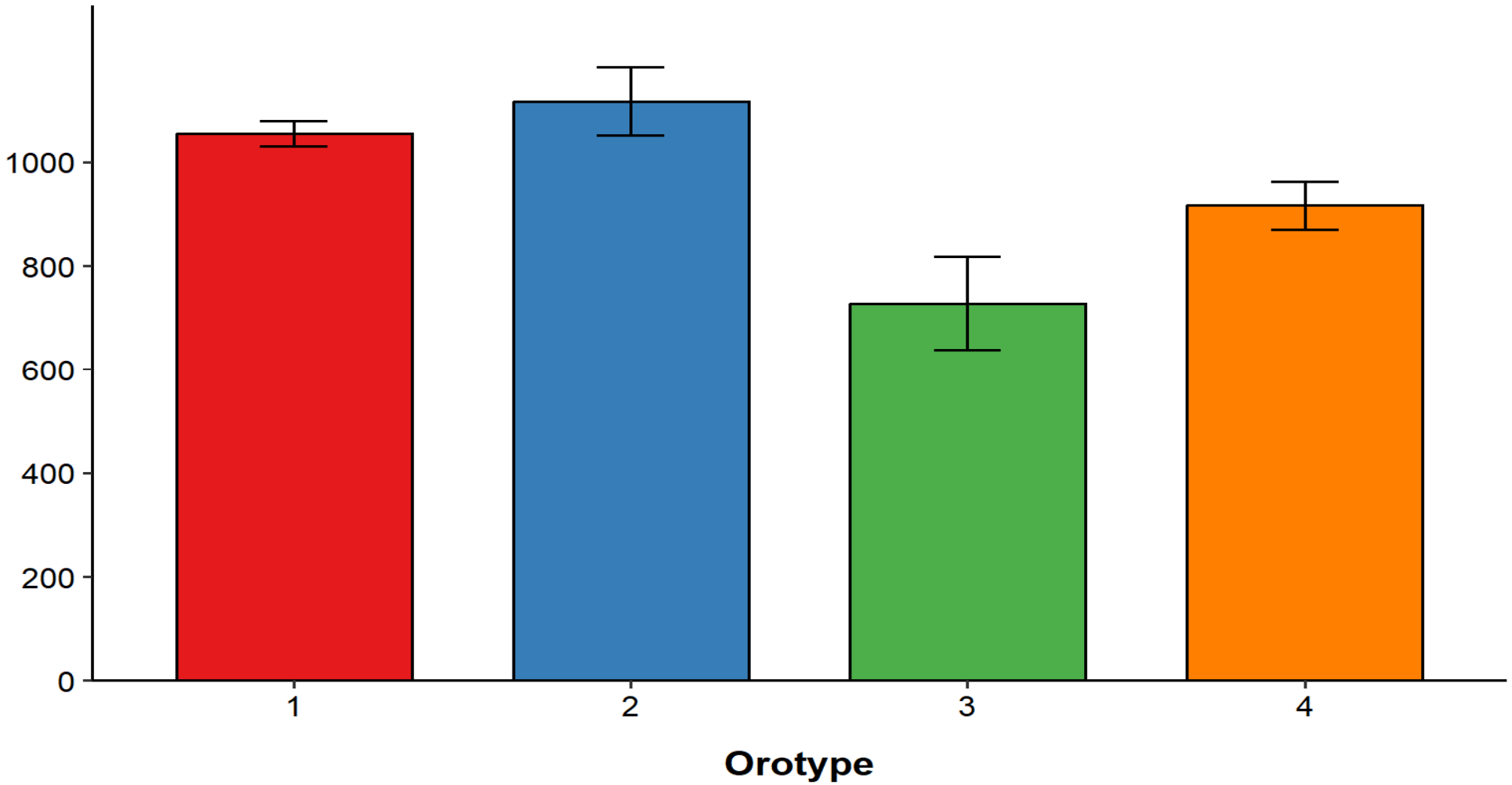


Fig 4 (d): Mean and standard error of the COVID-19 mortality rates in orotypes 1 to 4. *p = 0.203* by Jonckheere-Terpstra trend test.

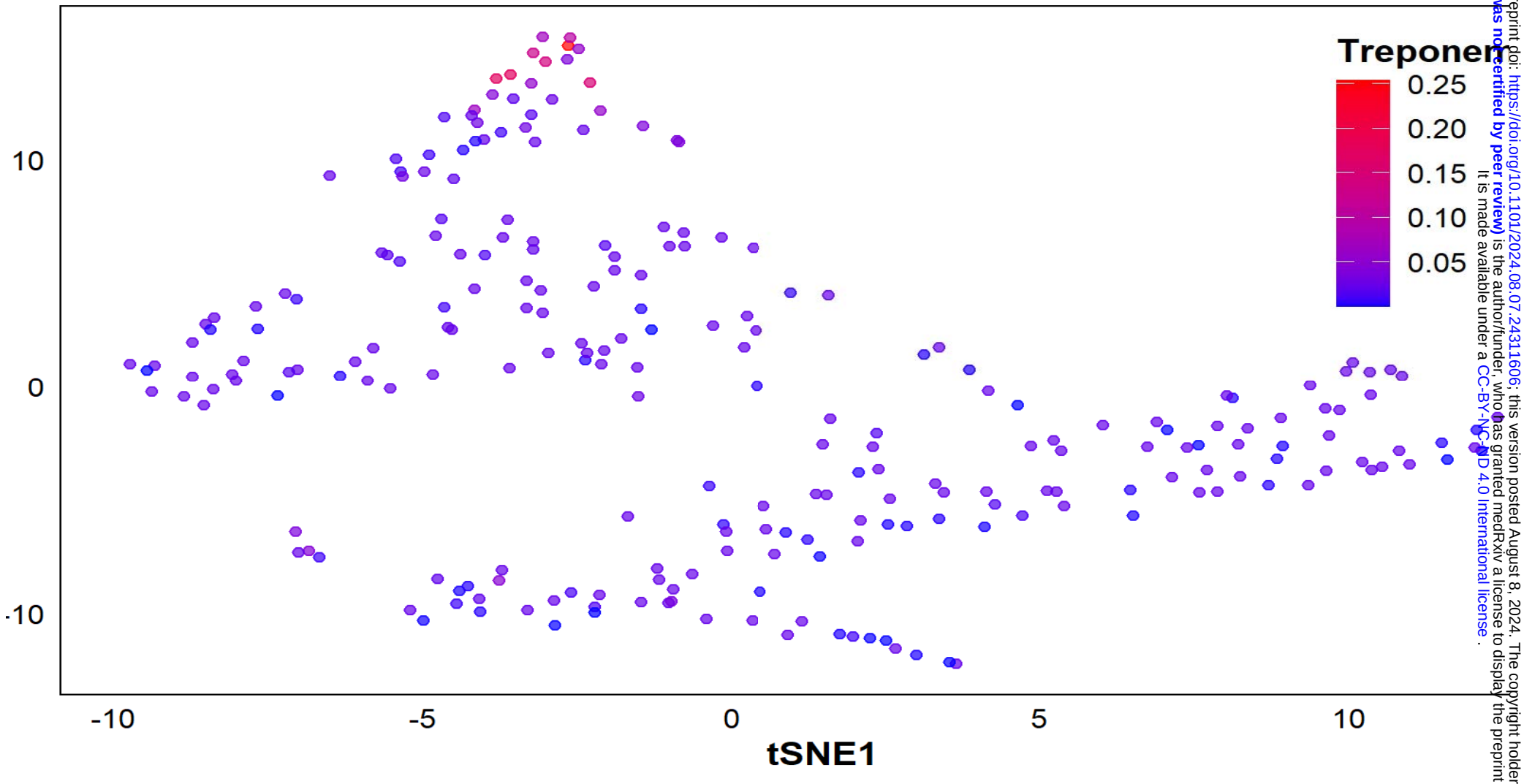


Fig 4 (e): The t-SNE plot is color-coded by the relative abundance of genus *Treponema*.

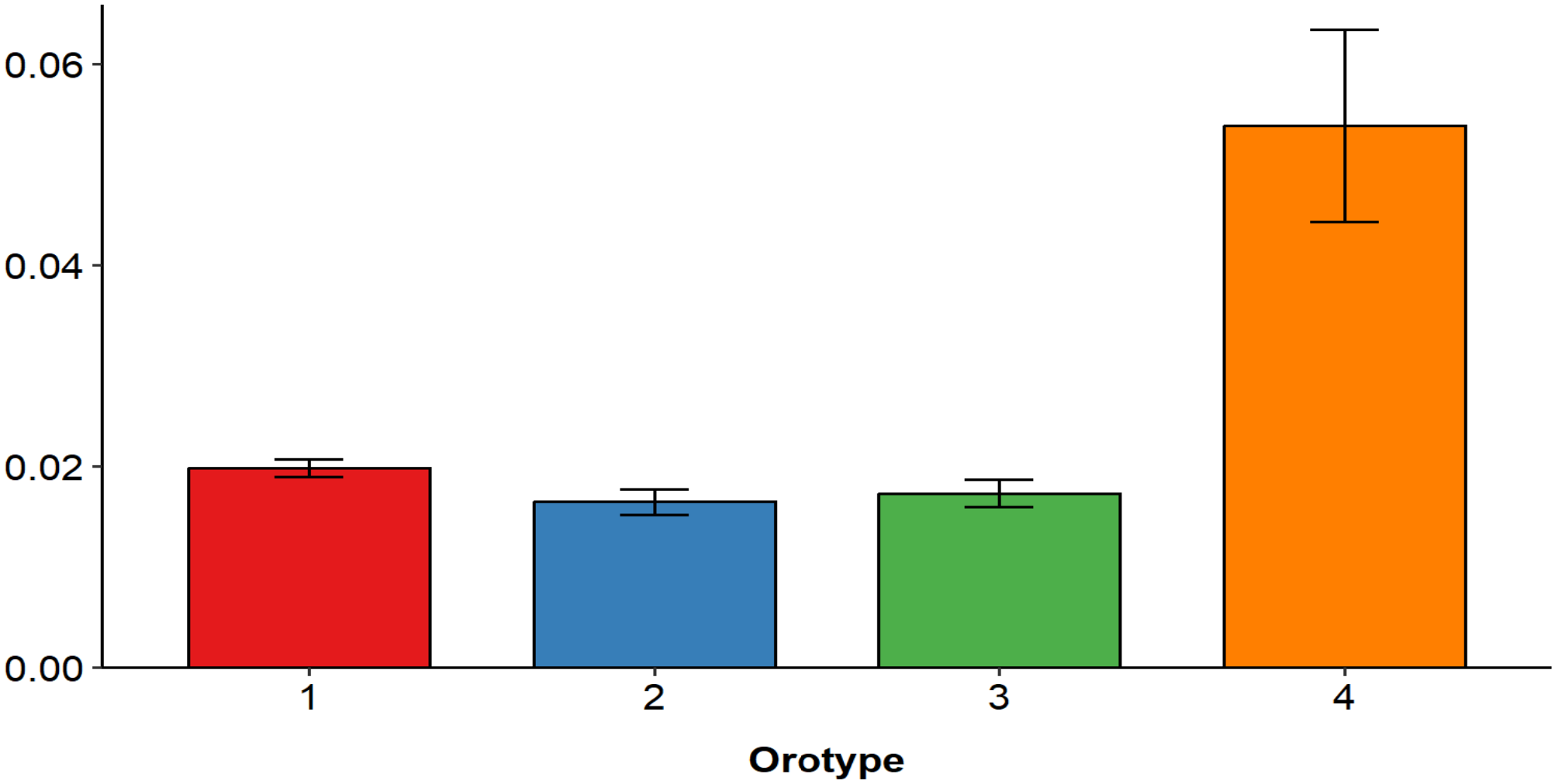


Fig 4 (f): Mean and standard error of the relative abundance of genus *Treponema* in orotypes 1 to 4. *P* = 0.033by Jonckheere-Terpstra trend test.

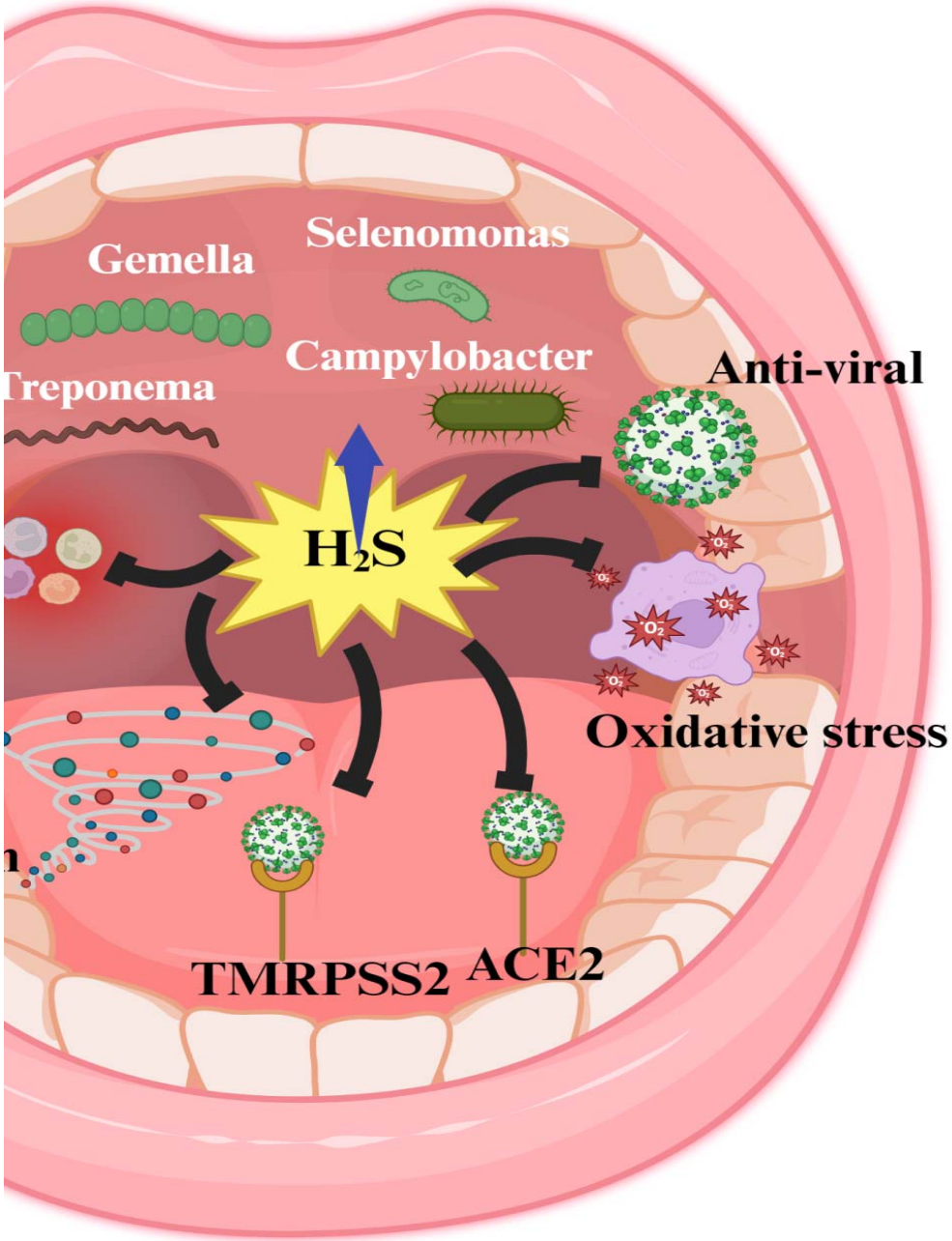


Fig 5: Proposed mechanisms of oral hydrogen sulfide (H_2S) in protecting again infection.H_2S, produced by oral bacteria incl Gemella, Campylobacter, and Selenom multifaceted protective effects against SAR include: (a) Direct antiviral activity, (b) Mo entry by altering ACE2 and TMPRSS2 re inflammatory effects and prevention of cyto Enhancement of antioxidant defenses th pathway. These mechanisms, previously RNA viral infections, likely contribute protective role of H_2S-producing oral bacte COVID-19 outcomes.

List of legend

odology for this bioinformatics pipeline (a) Healthy oral microbiome data was obtained from the National nformation (NCBI). (b) This data was processed and analyzed using OmicsBox software. (c) The most abu were identified. (d) COVID-19 mortality rate data from Our World in Data was incorporated (e) A statistic ıear Model) was used to analyze the relationship between the 15 most abundant genera and COVID-19 mor ıique called LIGER, typically used for analyzing single-cell RNA sequencing data, was applied to classify 4 healthy individuals into four distinct groups, termed "orotypes".

6S rRNA-seq datasets from Eight countries and the mortality rates of COVID-19.

vel abundance of health oral metagenomic

values of 15 genera in a generalized linear model (GLM) to predict the COVID-19 mortality rates.

rvised clustering of oral microbiota in 244 healthy subjects in ten countries by LIGER generated 4 orotype l with t-SNE and is color-coded by its orotype. (b) Fractions of orotypes 1 to 4 in eight countries. eight cou ding order of the COVID-19 mortality rates per million, which are indicated in parentheses. (c) The t-SNE p OVID-19 mortality rates in Eight countries. (d) Mean and standard error of the COVID-19 mortality rates b y Jonckheere-Terpstra trend test. (e) The t-SNE plot is color-coded by the relative abundance of genus *Tre* ndard error of the relative abundance of genus *Treponema* in orotypes 1 to 4. $P = 0.033$by Jonckheere-Terp

ed mechanisms of oral bacteria-produced hydrogen sulfide (H_2S) in protecting against SARS-CoV-2 infec
oral bacteria including Treponema, Gemella, Campylobacter, and Selenomonas, exhibits multifaceted p
t SARS-CoV-2. These include: (a) Direct antiviral activity, (b) Modulation of viral entry by altering A
ceptors, (c) Anti-inflammatory effects and prevention of cytokine storm, and (d) Enhancement of an
ıgh the Nrf2 pathway. These mechanisms, previously observed in other RNA viral infections, likely co
rotective role of H_2S-producing oral bacteria against severe COVID-19 outcomes.