Front. Environ. Sci. Eng. 2024, 18(10): 129
https://doi.org/10.1007/s11783-024-1889-z

RESEARCH ARTICLE

Socioeconomic drivers of the human microbiome footprint in global sewage

Minglei Ren[1,2,#], Shaojuan Du[3,#], Jianjun Wang (✉)[1,2]

1 Key Laboratory of Lake and Watershed Science for Water Security, Nanjing Institute of Geography and Limnology,
Chinese Academy of Sciences, Nanjing 210008, China
2 State Key Laboratory of Lake Science and Environment, Nanjing Institute of Geography and Limnology,
Chinese Academy of Sciences, Nanjing 210008, China
3 Nanjing Institute of Environmental Sciences, Ministry of Ecology and Environment, Nanjing 210042, China

HIGHLIGHTS

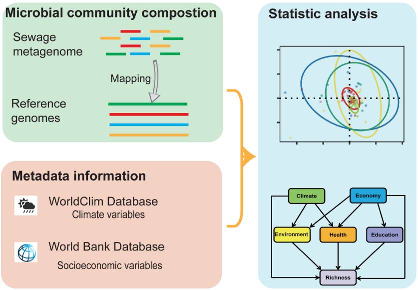
- We built a read-mapping framework to profile human microbes from sewages (HSM).
- There were 95.03% human microbial species successfully recaptured from sewages.
- The HSM composition showed a distance-decay pattern at a global scale.
- The HSM communities from developed regions were separated from developing regions.
- Economy was the key socioeconomic factors driving the HSM diversity.

ARTICLE INFO

Article history:
Received 18 February 2024
Revised 25 March 2024
Accepted 26 March 2024
Available online 1 August 2024

Keywords:
Human sewage microbiome
Biogeography
Socioeconomic factors
Climate factors

GRAPHIC ABSTRACT



ABSTRACT

The human microbiome leaves a legacy in sewage ecosystems, also referred to as the human sewage microbiomes (HSM), and could cause potential risk to human health and ecosystem service. However, these host-associated communities remain understudied, especially at a global scale, regarding microbial diversity, community composition and the underlying drivers. Here, we built a metagenomic read mapping-based framework to estimate HSM abundance in 243 sewage samples from 60 countries across seven continents. Our approach revealed that 95.03% of human microbiome species were identified from global sewage, demonstrating the potential of sewage as a lens to explore these human-associated microbes while bypassing the limitations of human privacy concerns. We identified significant biogeographic patterns for the HSM community, with species richness increasing toward high latitudes and composition showing a distance-decay relationship at a global scale. Interestingly, the HSM communities were mainly clustered by continent, with those from Europe and North America being separated from Asia and Africa. Furthermore, global HSM diversity was shown to be shaped by both climate and socioeconomic variables. Specifically, the average annual temperature was identified as the most important factor for species richness (33.18%), whereas economic variables such as country export in goods and services contributed the most to the variation in community composition (27.53%). Economic and other socioeconomic variables, such as education, were demonstrated to have direct effects on the HSM, as indicated by structural equation modeling. Our study provides the global biogeography of human sewage microbiomes and highlights the economy as an important socioeconomic factor driving host-associated community composition.

1 Introduction

Anthropogenic activities in cities usually leave chemical or biological footprints in sewage (Matus et al., 2019), such as the release of toxic chemicals and the migration of microorganisms from human-related sources into sewage ecosystems (Bibby et al., 2010; Chen et al., 2020;

✉ Corresponding author
E-mail: jjwang@niglas.ac.cn

# These authors contributed equally to this work.

Special Issue—Microbiological Contaminants in Water Environment: Occurrence and Control (Responsible Editors: Xin Yu, Hong Chen & Yunho Lee)

Gudda et al., 2020). For example, human pathogenic viruses could be detected from wastewater, including SARS-CoV-2, influenza virus, and monkeypox viruses (Li et al., 2022; Tisza et al., 2023). The wastewater-based surveillance has also shown that antibiotic-resistant bacterial pathogens have been detected in global wastewater systems (Nowrotek et al., 2019; Zhao et al., 2022; Tiwari et al., 2024), which are proposed as hotspots for environmental antibiotic-resistance (Nadimpalli et al., 2020). In terms of community composition, global sewage microbes resemble the human-related microbiome more than other habitats (García-Aljaro et al., 2019; Hendriksen et al., 2019; LaMartina et al., 2021). At the species level, a human bacteria population of *Blautia* obtained from sewage show remarkable similarity as one in fecal samples (Eren et al., 2014). Considering this close relatedness between sewage and human-related microbiomes, the sewage microorganisms derived from human domestic sources are defined as the human sewage microbiome, referred to the HSM hereafter, as proposed previously (Cai et al., 2014).

The possibility of exploring the HSM has been demonstrated by a series of studies across from cities to regions (Cai et al., 2014; Newton et al., 2015). For example, human gut microbial communities are found to be the dominant force of shaping the influent sewage bacterial profiles in a Hong Kong Wastewater Plant using the 16S rRNA gene sequencing (Cai et al., 2014). The traditional culture-based technologies demonstrate that the human fecal indicator bacteria, such as *Escherichia coli* and *Streptococci* are often monitored in various aquatic environments associated with sewage pollution (Sidhu et al., 2012; Staley et al., 2012). Moreover, the comparison of oligotying sequence reveal that nearly 97% of human fecal taxa are recaptured from sewage communities in 71 American cities (Newton et al., 2015). The investigation of sewage and human gut microbes reveals the consistent presence of 14 dominant pathogens shared by both sewage influents and human gut using metagenomic approach (Li et al., 2015), suggesting that sewage microbial composition could well reflect the average bacterial pathogen in human guts. Taken together, these studies demonstrate that the analysis of the human-related microbiome from sewages could efficiently explore and compare the microbial communities from human populations at different demographic scales, e.g., city, country, or even continent. However, there is a lack of efficient approaches for monitoring the HSM at a global scale and comprehensive evaluations of climate and human-related factors influencing the communities.

Our understanding of the influence of socioeconomic factors on the HSM remains elusive. The human microbiome over the life course could be affected by not only local and climate drivers (Pasolli et al., 2019; Gunawan et al., 2023), but also the socioeconomic factors, such as family income, education and occupational status (Bradley and Corwyn, 2002; Dowd and Renson, 2018; Ahn and Hayes, 2021). For example, a lower area-level socioeconomic status is associated with reduced alpha diversity, more abundant Bacteroides, and less abundant Prevotella in the colonic microbiota according to a study of 44 healthy volunteers in Chicago (Miller et al., 2016). Socioeconomic status is also one of the main drivers of differences in the gut microbiota of Indonesian schoolchildren (Amaruddin et al., 2020) and in the saliva microbiome (Belstrøm et al., 2014). These studies focused on human gut microbes rather than the HSM, however, an overall evaluation of the influence of climatic and socioeconomic factors on the HSM is urgently needed.

Here, we performed a comprehensive study of the HSM at a global scale to evaluate their distribution and to determine the effects of anthropogenic activities on both the diversity and community structure of the HSM. We first built a read-mapping framework to profile the 4930 curated human-associated species-level microbial genomes across 243 sewage metagenomes from 74 cities in 60 countries across seven continents (Fig. 1). Our study aimed to answer three questions: 1) What is the global biogeographical pattern of HSM? 2) How are the microbial characteristics of the HSM such as species abundance, diversity and community composition affected by climate and socio-economic factors? 3) What is the relative importance of these types of factors on the diversity and community composition of the HSM? Our findings highlight the important roles of anthropogenic factors, such as economic and health variables, in shaping the human sewage microbiome.

2 Materials and methods

2.1 Sequence downloads and analyses

To identify and profile the human-associated microbiome across sewage environments (HSM), a custom pipeline was developed and described in detail as below. First, we downloaded the sequence data sets, including the 4930 species-level human microbial genomes of diverse populations spanning body site, geography and lifestyle (Pasolli et al., 2019) and 234 metagenomic reads of sewage samples around the world (Hendriksen et al., 2019). For sewage communities, 243 metagenomic samples of untreated sewage from 79 sites in 60 countries across seven continents were used to represent the diversity of microbes in sewage environments. Notably, three samples with missing latitude and longitude data were discarded from our study. For human microbes, 154723 genomes were first reconstructed from 9428 metagenomic samples, then clustered at 95% sequence similarity, resulting in 4930 species-level reference genomes (Pasolli et al., 2019). These high-quality

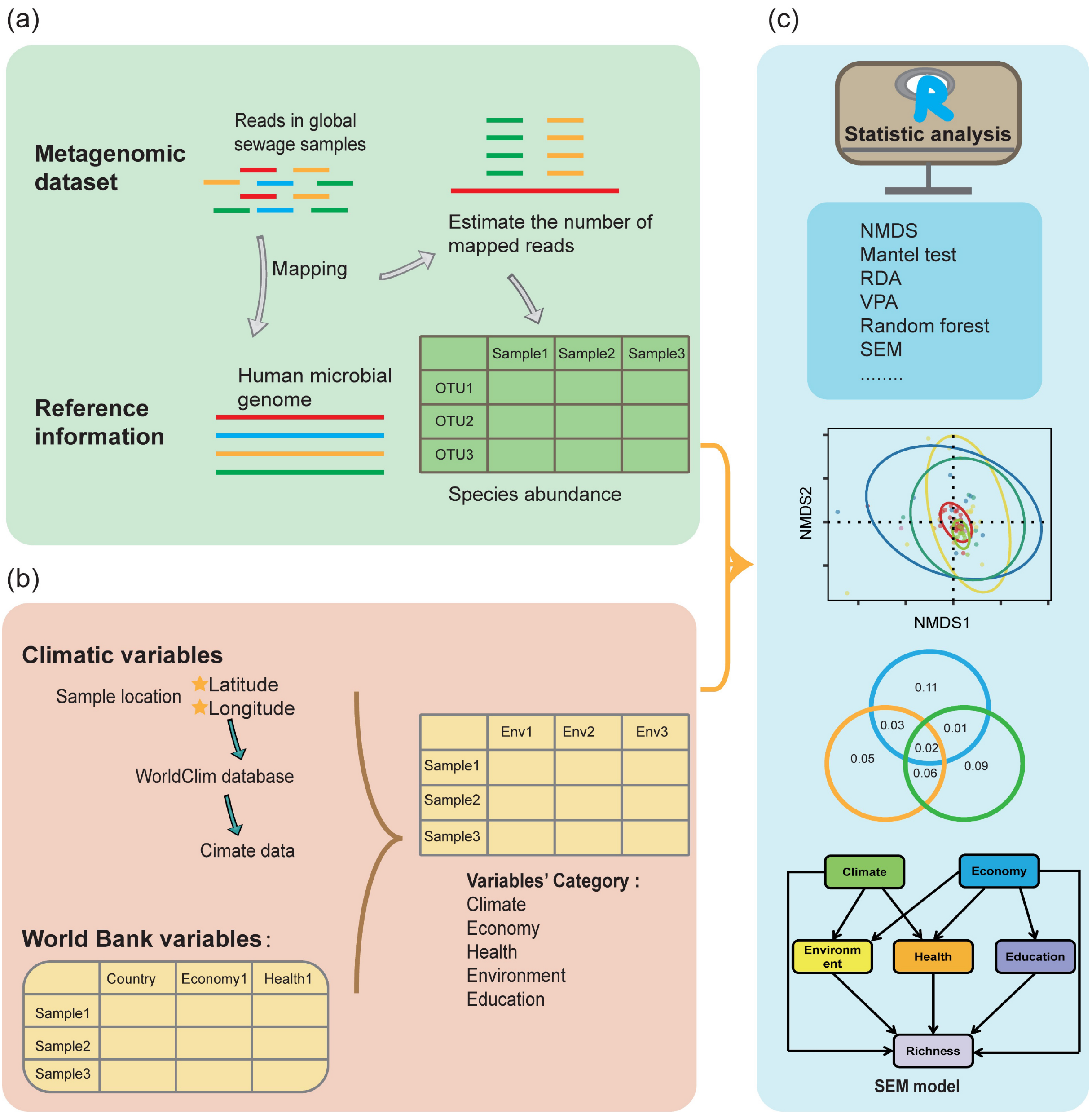


Fig. 1 The overview of the framework for exploring HSM. The framework includes three steps highlighted in panels with distinct colors: (a) The taxonomic profiling of the human-related microbiome in sewage (HSM), (b) The collection of the relevant variables from the curated databases, (c) Statistical analyses.

genomes spanned six different body sites, including the stool, oral cavity, skin, airways, nasal cavity, and vagina. Therefore, these genomes could well represent the great diversity of the human microbiome and were used as reference genomes to detect human-associated microorganisms. Second, we estimated the abundance of the human reference genomes through a metagenome read recruitment approach. The quality of raw reads was checked using FastQC v0.11.8 (Andrews 2010), indicating no sequencing adapters in the samples. The raw sequencing reads were trimmed using Trimmomatic v0.39 (Bolger et al., 2014), discarding the reads with an average Phred quality lower than 25 using a 4-base-wide sliding window and shorter than 50 bp. The longest contig from each reference genome was extracted as the representative sequence, then merged into the reduced human microbial genome data set. The clean reads from each sample were mapped against with the genome data set using Bowtie v2.3.5.1 (Langmead and Salzberg 2012), and the total bases of the reads uniquely mapped onto the representative sequence with sequence identity greater than 99% were counted using the 'depth' module of SAMtools v1.3.1 (Li et al., 2009). The short sequence showing $\geqslant$ 99% similarity against a species' genome could be assigned to the sequence belonging to the species of the interest. This relative strict threshold could be used to target the sequence which are more similar to the human-related microbial genomes, rather than their distantly relatives in sewage. The abundance of each representative across each sewage sample was initially denoted by the total number of bases mapped to the sites of the sequence. Finally, we normalized the abundance of each reference genome in a sample using the length of the representative sequence and the total sequencing depth of the sample. The abundances of the reference genomes across sewage samples and their taxonomic information were shown in the Supplementary Tables S1 and S2, respectively.

2.2 Screening socioeconomic and climate factors

We collected two types of factors, namely the socioeconomic and climate factors, to examine their influence on the diversity and composition of the HSM (Amaruddin et al., 2020, Catania et al., 2021). For the socioeconomic factors, 178 country-specific variables were collected from the World Bank database and published elsewhere (Nowak et al., 2015). The variables that were highly correlated with other ones (Pearson's correlation coefficient $r > 0.80$) were removed to avoid the collinearity among variables. The remaining variables were further selected through the "bioenv" function in the vegan package v2.6.4 (Dixon 2003), which considers the correlation coefficient between the community structure distance matrix and the environmental factor distance matrix. These selected socioeconomic variables were then classified into four categories, namely economy, health, environment, and education according to the classification system in the World Bank database. For the climate factors, the local average annual precipitation and average annual temperature were extracted from the WorldClim database according to geological locations of sewage samples (Hijmans et al., 2005). Finally, 40 variables from five categories, namely, climate, economy, health, environment, and education, were selected and used to estimate their contributions to the variations in human sewage microbial communities. Detailed information on the climate and socioeconomic variables across sewage samples was shown in Table S3.

2.3 Statistical analyses

The taxonomic alpha diversity indices, including species richness, Pielou's evenness, the Chao1 index and the Shannon index, were calculated using the 'diversity' function from the vegan package (Dixon, 2003). Nonmetric multidimensional scaling (NMDS) analysis was performed to evaluate microbial community compositions across seven continents using the Bray-Curtis distance matrix. To explore the distance-decay relationship (DDR), the dissimilarity between samples or groups of samples was calculated by using the Bray-Curtis distance metric. Geographic distance was calculated using the "distm" function from the "geosphere" package v1.5.18. The slope of the DDR was calculated via linear least squares regression on the relationship between geographical distance and community composition dissimilarity. The statistical significance was determined using Mantel tests with 999 permutations and the Spearman's correlation method and the observed slopes were compared between continents.

A series of statistical tests were performed to determine the main drivers of the diversity and community composition of the global HSM. The relationships between species richness of HSM and each of the climate and socioeconomic variables were tested by linear regressions. A Mantel test was used to examine the correlation between each variable and community composition represented by the first axis of NMDS with 999 permutations using the vegan package (Dixon, 2003). We constructed a correlation matrix and network details of HSM species abundance and the relevant climate and socioeconomic variables by calculating all possible pairwise Spearman's rank correlations between species and variables using the "Hmisc" package v5.0.1. Only the factors that showed robust (Spearman's rho > 0.5 or rho < −0.5) and statistically significant ($P < 0.01$) correlations with microbial taxa were considered in the network analysis. The visualization of the network diagram and the calculation of topology characteristics (e.g., degree, betweenness centrality, and closeness centrality) of the network were implemented using the Gephi platform (Bastian et al., 2009).

Redundancy analysis (RDA) was used to estimate the effects of each explanatory variable on species distribution of HSM. Species abundance data were Hellinger-transformed. Forward selection with 999 permutations was selected to retain only the significant environmental variables in the final model ($P < 0.05$) using the "ordistep" function in the vegan package. In addition, the random forest model was further used to calculate the relative importance of the indicators. We constructed each model containing 22000 decision trees to predict climate and socioeconomic indicator variables for HSM species richness and community composition. The variables with relative contributions less than 5% were removed to obtain the best model. The constructed random forest regression model ranked the variables according to their degree of contribution. Random forest modeling analysis was conducted using the 'randomForestSRC' package v3.2.2.

Structural equation modeling (SEM) was applied to explore the direct and indirect relationships between the factors of interest and the diversity and community composition of the HSM. The 40 indicators selected above were classified into the five composite variables through multiple regression, namely climate, economy, health, environment and education. These composite variables were used to account for the joint influence of five categories of factors on community characteristics. We first considered a hypothetical model containing all reasonable paths, then transformed the model into a regression equation, and eliminated the nonsignificant paths until the final model contained the significant paths ($P < 0.05$). The final optimal model with the lowest Akaike information criterion (AIC) value was selected after comparing the AIC values of all alternative models (Hu et al., 2020). The SEM analysis was conducted using the "lavaan" R package v0.6.16. All the statistical analyses described above were performed in R v4.0.2.

3 Results and discussion

3.1 Global patterns of the human sewage microbiomes (HSM)

The read-mapping-based framework revealed that there were 4685 human-associated species identified from sewage, accounting for 95.03% of the reference species (Fig. 1). The results showed the great potential of the framework to recapture and profile the human-associated microbes from sewage (HSM). Among all the human microbiome species identified above, 77.59% were derived from stool (Fig. S1a), indicating that HSMs are dominated by human stool microbes. Our metagenomic-based results are consistent with previous findings that the human gut microbiome is the dominant force shaping influent sewage bacterial profiles based on high-throughput 16S rRNA gene sequence data (Cai et al., 2014; Newton et al., 2015). Moreover, the dominant stool bacteria in the sewage included Firmicutes (50.4%), Proteobacteria (23.7%), Actinobacteria (16.9%), and Bacteroidetes (7.05%, Fig. S1b). As expected, these dominant phyla are among the most common and abundant taxa in human intestines (Qin et al., 2010), highlighting the close relatedness between the HSM and human gut microbes.

Anthrophonic activities produce more than 300 km^3 of wastewater per year, equaling one-seventh of the global river volume (Mateo-Sagasta et al., 2015; Wu et al., 2019) and nearly half of the global wastewater production is released to the environment untreated (Jones et al., 2021). A large volume of sewage receives microbial assemblages from multiple sources, especially from human domestic waste such as feces, urine, washing and bathing (Cai et al., 2014). Therefore, the high proportion of human-related microbiome successfully identified from global sewage was not surprising, highlighting the ubiquity of human-associated microbiomes in the sewage ecosystem. Recent studies further show that wastewater metagenomics sequencing surveillance facilitate the identification and source-tracking of viruses, such as SARS-CoV-2 during the COVID-19 pandemic (Crits-Christoph et al., 2020; Ji et al., 2021; Tisza et al., 2023), potentially pathogenic microorganisms (Schneeberger et al., 2019), and even traces of human mitochondrial DNA (Pipek et al., 2019). The predominant occurrence of human-associated microbes across global sewage in our study and elsewhere indicates the potential health consequences that the HSM imposes upon the exposed human populations.

Microbial diversity of the HSM showed global geological variation across latitudes and continents. The relationship between the HSM species richness and latitude was examined, revealing that species richness of the HSM increased toward high latitudes (Fig. 2(a)). In addition, the spatial distribution of HSM determined by the distance-decay relationship analysis (DDR) showed that the HSM community dissimilarity increased significantly with increasing of geographic distance at the global scale (Mantel test, $r = 0.18$, $P = 0.004$). However, the DDRs of the HSM depended on spatial scale, which was supported by the finding that the DDR pattern was nonsignificant at the continental scale (Fig. 2(b)). We also found that species richness of HSM in Europe and North America was greater than those in the developing regions, such as Asia and the Middle East, although the variation was nonsignificant (Fig. S2). Nonmetric multidimensional scaling analysis revealed that community compositions of the samples from Europe ($n = 27$) and North America ($n = 13$) were more similar than those from underdeveloped regions of Asia and Africa (Fig. 2(c)). These clustering patterns for Europe and North America suggest that microbial composition of the HSM in developed regions are distinct from those in other regions, which might be related to the local socio-economic factors. For example, the composition of human gut microbiome has been widely demonstrated to be closely associated with diet (Parizadeh and Arrieta, 2023), geological location, human lifestyle (Yatsunenko et al., 2012), and ethnicity (Deschasaux et al., 2018). The alterations in the human microbiome in industrialized regions might arise from modern lifestyle practices limiting bacterial dispersal (Martínez et al., 2015). Therefore, the patterns observed in human-related microbes in sewage are explained by the alteration in human gut bacteria associated with socioeconomic factors.

3.2 Socioeconomic factors determine the diversity and composition of the HSM

To determine the underlying factors associated with the HSM at the global scale, we collected two types of factors, namely the socioeconomic and climate variables (see the Method section, Table S3). Socioeconomic variables were carefully selected from the World Bank Database and divided into four categories, economic (n = 14), health (n = 10), environment (n = 9), and education (n = 5). Climate variables, including the average annual

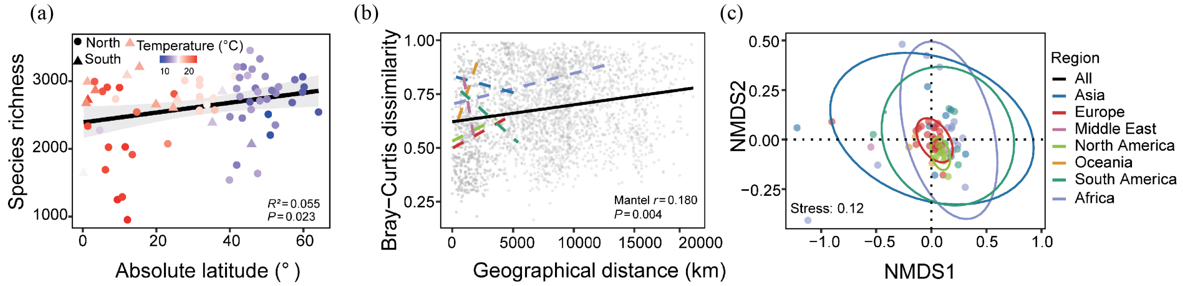


Fig. 2 Global distribution and diversity of the HMS. (a) Latitudinal distribution of the HMS diversity. The points represent the HMS species richness. (b) Distance-decay relationships based on the Bray-Curtis dissimilarity of the HMS. (c) NMDS ordinations showing the distribution patterns of HMS communities across global continents.

temperature and precipitation were retrieved from the WorldClim database based on the geographic location of each site (Fig. 1).

The relationships between species abundance and both two types of variables were first investigated using network analysis (Fig. 3). The nodes of the network represent the relative abundance of each species in the HMS, whereas the edges represent the relationships between species abundance and both climate and socioeconomic variables. Only the edges with significant relationships (Spearman's rho > 0.5, $P < 0.05$) were considered, resulting in a network comprising 584 nodes and 3295 edges (Fig. 3(a)). The network results showed that the adequacy of social insurance programs (economy), the number of wage and salaried workers (economy), electric power consumption of environment category (environment), and open defecation practices (health) were significantly correlated with the majority of HSM species with the greatest network degrees. Among these significant correlations, the climate, economy and health showed distinct relationships, including the positive (n = 2, 94, 83) and negative (n = 152, 65, 66) links to the HSM species, respectively (Fig. 3(b)). Species richness of the HSM was also significantly associated with most of these variables, as indicated by linear regression analysis, including average annual temperature, import and export trade, salary and social security, sanitation, sickness, and health (t-test, $P < 0.05$, Fig. S3). Moreover, the healthy factors, such as antiretroviral therapy coverage for prevention of mother to child transmission (PMTCT), condom use for young women, and the tuberculosis case detection rate were significantly correlated with HSM community dissimilarity (Mantel test, $P < 0.05$, Fig. S4). In addition, a total of 16 variables selected by the redundancy analysis (RDA) explained 28.8% of the community variation, with 16.6% and 12.2% of the total variance explained by RDA1 and RDA2, respectively (Fig. 4). Interestingly, the health variables such as the use of antiretroviral drugs had a significant effect on HSM communities (Fig. 4(c)). This finding is consistent with previous reports that distinct antiretroviral combinations could partially alter the human gut microbial composition (Nowak et al., 2015).

We further quantified the relative contributions of both groups of drivers to the diversity and community composition of the HSM by a random forest model. Consistent with the linear model results, a wide range of socioeconomic factors were attributed to the variation in HSM species richness (Fig. 4(a)), including average annual temperature (climate, 33.18%), the export values for communications (economy, 16.75%), food imports (economy, 14.72%), aquaculture production (environment, 7.36%), open defecation practices (health, 6.07%), and adequacy of social insurance programs (economy, 5.94%). For community composition, the export values for communications (economy), proportion of cities with electricity, import value index and open defecation practices were attributed to 27.53%, 15.62%, 11.37%, and 10.47%, respectively (Fig. 4(b)).

The direct and indirect effects of both groups of factors on species richness and community composition of the HSM were determined through structural equation modeling (SEM, Fig. 5). Consistent with the linear model and random forest results above, economy had the strongest direct effect on the species richness (standardized path coefficient, $\beta = 0.402$, $P < 0.001$), followed by education ($\beta = 0.272$, $P < 0.01$). Climate variables had no significant effect on the first axes of NMDS representing the community structure community composition, but the environment, economy, education, and health had direct and positive impacts (Fig. 5(b)). Consequently, anthropogenic activity, especially economic factors, serves as a significant promoter to structure the abundance, diversity, and community composition of the HSM worldwide. Here, we linked the observed differences in the HSM around the world to global socioeconomic factors and demonstrated their nonnegligible effect on the HSM community structure variation. To our knowledge, this was for the first time to explore how both climate and socioeconomic factors affect human microbial communities in the sewage ecosystem at a global scale. Earlier studies focus on the effects of socioeconomic factors on the human gut microbiome, or the HSM constrained at the regional scale (Newton et al., 2015; Bowyer et al.,

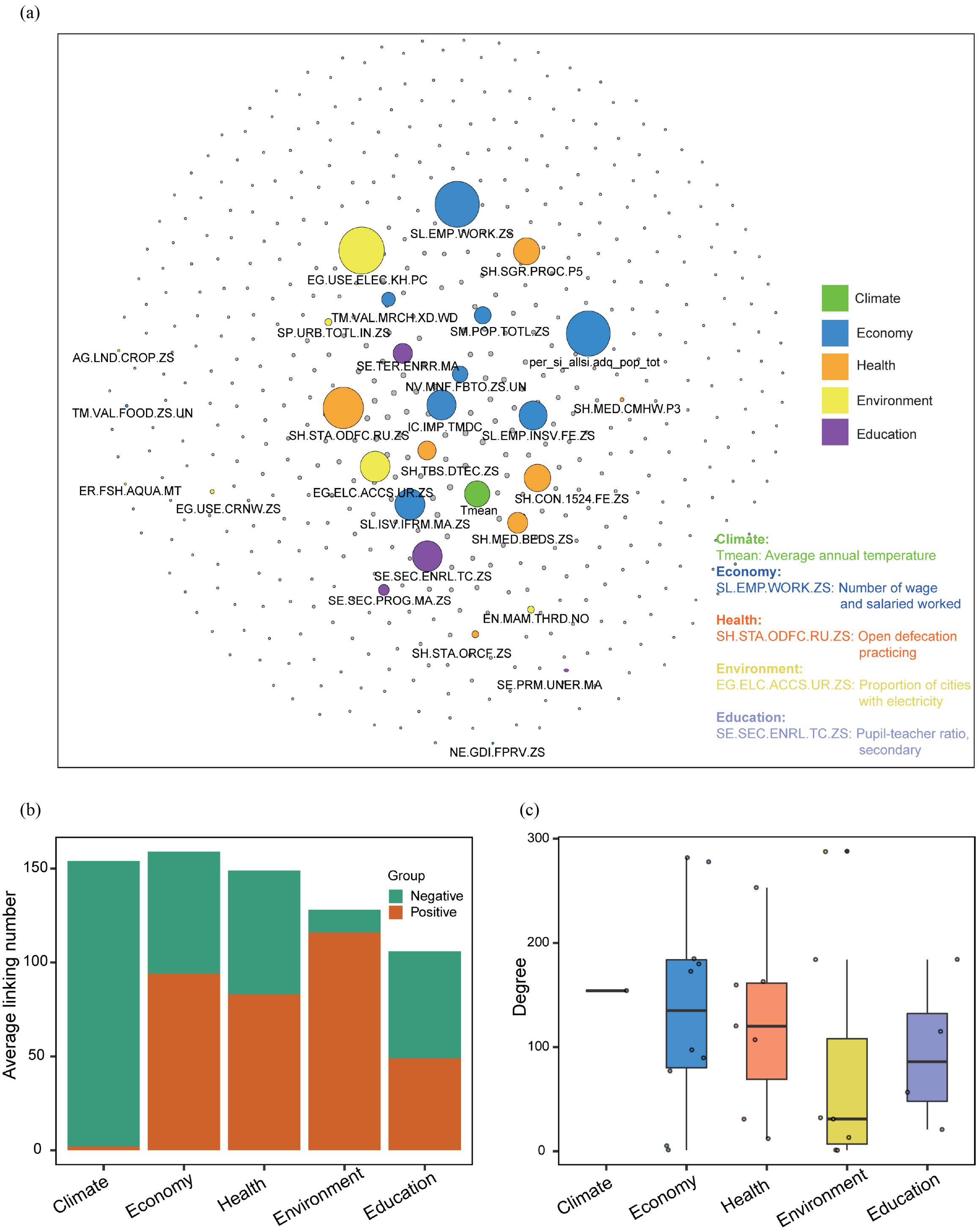


Fig. 3 The correlation network of the HMS and variables. The network represents the relationships between species abundance and both two types of variables (a). The distributions of average linking numbers and degrees across different types of variables were shown in panels (b, c).

2019). For example, there are strong correlations between human gut microbial diversity of twin cohort of UK and socioeconomic parameters, such as income, area-level socioeconomic status, and education level (Bowyer et al., 2019). The human demographic patterns, such as city population percent obesity are correlated the variation in human fecal communities in sewage influent 71 American cities (Newton et al., 2015). Specifically, the countries with low socioeconomic condition likely face the challenges related with poor health-related behaviors,

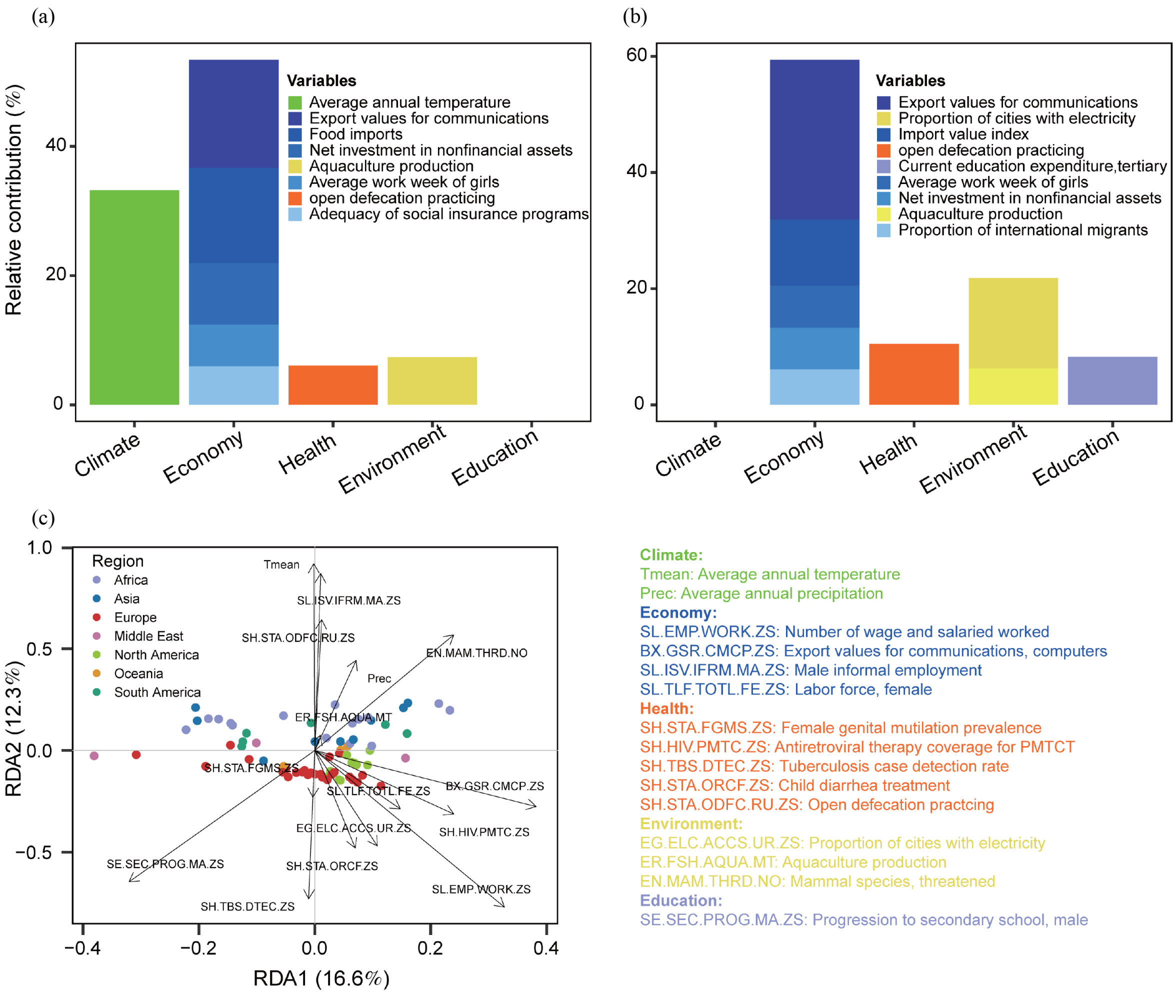


Fig. 4 The relative importance of the variables to HMS diversity. The relative contributions of the variables to HMS species richness and community composition based on the random forest model results (a, b). The effects of the significant variables on the HMS were shown based on redundancy analysis (c).

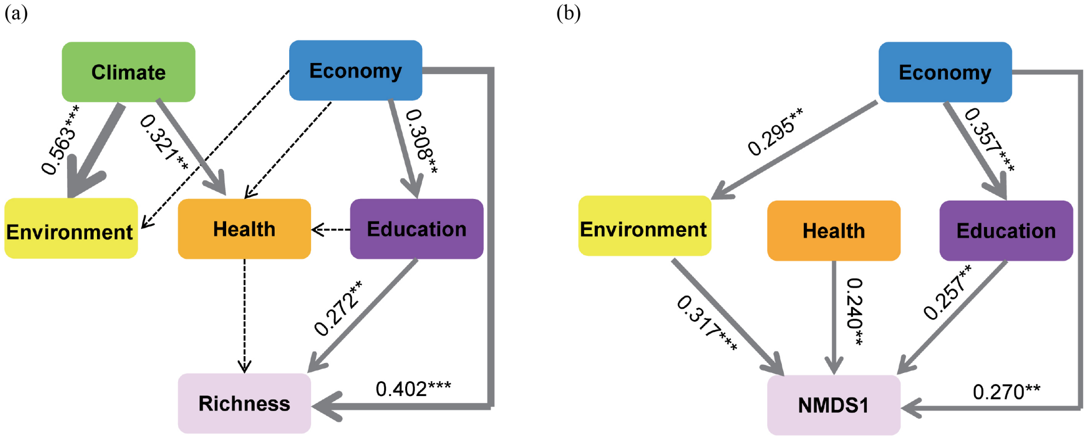


Fig. 5 The drivers of the HMS diversity revealed through SEMs. The SEMs showed the relationships among the variables and HMS species richness and community composition (a, b). The width of the arrows represents the strength of influence between variables, with numbers denoting the standardized path coefficients.

higher rates of morbidity and mortality (Ahn and Hayes, 2021). Together, these characteristics associated with socioeconomic factors and lifestyles are an important driver of human sewage microbial composition based on a large number of samples around the world combined with the consideration of a wide range of regional socioeconomic factors.

4 Conclusions

In conclusion, our study provides an efficient approach for comprehensively exploring the biogeographical patterns of the human microbial community dispersed in sewage (HSM) from 60 countries across seven continents and determining the potential socioeconomic-associated drivers within the context of theoretical ecological frameworks. Our characterization of the persistence, biogeography and underlying driving factors of the human-associated microbiomes from sewage at the continental scale is particularly relevant to global public health. Nearly half of the global wastewater production is released to the environment untreated (Jones et al., 2021), from which potential antibiotic-resistant pathogens and viruses are increasingly being detected by the wastewater-based surveillance. The human-related microbes, especially stool microbes were detected at high abundance in global sewage serving as a complex species pool, using the framework implemented here. These findings highlight the use of sewage as a lens to explore the fecal microbiota from millions of people and its potential to elucidate microbiome patterns associated with socioeconomic factors such as economic or sanitary conditions or human demographics.

Acknowledgements We sincerely thank Prof. Zongjun Du from Shandong University for helpful comments about the early version of the manuscript. This study was financially supported by the National Natural Science Foundation of China (Nos. 42225708, 42372353, 92251304, 92351303, and 42002304).

Conflict of Interests 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.

Electronic Supplementary Material Supplementary material is available in the online version of this article at https://doi.org/10.1007/s11783-024-1889-z and is accessible for authorized users.

References

Ahn J, Hayes R B (2021). Environmental influences on the human microbiome and implications for noncommunicable disease. Annual Review of Public Health, 42(1): 277–292

Amaruddin A I, Hamid F, Koopman J P R, Muhammad M, Brienen E A T, van Lieshout L, Geelen A R, Wahyuni S, Kuijper E J, Sartono E, et al. (2020). The bacterial gut microbiota of schoolchildren from high and low socioeconomic status: a study in an urban area of Makassar, Indonesia. Microorganisms, 8(6): 961

Andrews S (2010). FastQC: a Quality Control Tool for High Throughput Sequence Data. Cambridge: Babraham Institute

Bastian, M, Heymann S, Jacomy M (2009). Gephi: an open source software for exploring and manipulating networks. Proceedings of the International AAAI Conference on Web and Social Media, 3(1): 361–362

Belstrøm D, Holmstrup P, Nielsen C H, Kirkby N, Twetman S, Heitmann B L, Klepac-Ceraj V, Paster B J, Fiehn N E (2014). Bacterial profiles of saliva in relation to diet, lifestyle factors, and socioeconomic status. Journal of Oral Microbiology, 6(1): 09

Bibby K, Viau E, Peccia J (2010). Pyrosequencing of the 16S rRNA gene to reveal bacterial pathogen diversity in biosolids. Water Research, 44(14): 4252–4260

Bolger A M, Lohse M, Usadel B (2014). Trimmomatic: a flexible trimmer for illumina sequence data. Bioinformatics, 30(15): 2114–2120

Bowyer R C, Jackon M A, Le Roy C I, Lochlainn M N, Spector T D, Dowd J B, Steves C J (2019) Socioeconomic status and the gut microbiome: a TwinsUK cohort study. Microorganisms, 7(1): 17

Bradley R H, Corwyn R F (2002). Socioeconomic status and child development. Annual Review of Psychology, 53(1): 371–399

Cai L, Ju F, Zhang T (2014). Tracking human sewage microbiome in a municipal wastewater treatment plant. Applied Microbiology and Biotechnology, 98(7): 3317–3326

Catania F, Baedke J, Fábregas-Tejeda A, Nieves Delgado A, Vitali V, Long L A N (2021). Global climate change, diet, and the complex relationship between human host and microbiome: towards an integrated picture. BioEssays, 43(6): 2100049

Chen X, Chen X, Zhao Y, Zhou H, Xiong X, Wu C (2020). Effects of microplastic biofilms on nutrient cycling in simulated freshwater systems. Science of the Total Environment, 719: 137276

Crits-Christoph A, Kantor R S, Olm M R, Whitney O N, Al-Shayeb B, Lou Y C, Flamholz A, Kennedy L C, Greenwald H, Hinkle A J (2020). Genome sequencing of sewage detects regionally prevalent SARS-CoV-2 variants. mBio, 12(1): e02703-20

Deschasaux M, Bouter K E, Prodan A, Levin E, Groen A K, Herrema H, Tremaroli V, Bakker G J, Attaye I, Pinto-Sietsma S J, et al. (2018). Depicting the composition of gut microbiota in a population with varied ethnic origins but shared geography. Nature Medicine, 24(10): 1526–1531

Dixon P (2003). VEGAN, a package of R functions for community ecology. Journal of Vegetation Science, 14: 927-930

Dowd J B, Renson A (2018). “Under the skin” and into the gut: social epidemiology of the microbiome. Current Epidemiology Reports, 5(4): 432–441

Eren A M, Sogin M L, Morrison H G, Vineis J H, Fisher J C, Newton R J, McLellan S L (2015). A single genus in the gut microbiome reflects host preference and specificity. ISME Journal, 9(1): 90–100

García-Aljaro C, Blanch A R, Campos C, Jofre J, Lucena F (2019). Pathogens, faecal indicators and human-specific microbial source-tracking markers in sewage. Journal of Applied Microbiology, 126(3): 701–717

Gudda F O, Waigi M G, Odinga E S, Yang B, Carter L, Gao Y (2020). Antibiotic-contaminated wastewater irrigated vegetables pose resistance selection risks to the gut microbiome. Environmental Pollution, 264: 114752

Gunawan W B, Abadi M N P, Fadhillah F S, Nurkolis F, Pramono A (2023). The interlink between climate changes, gut microbiota, and aging processes. Human Nutrition & Metabolism, 32: 200193

Hendriksen R S, Munk P, Njage P, van Bunnik B, McNally L, Lukjancenko O, Röder T, Nieuwenhuijse D, Pedersen S K, Kjeldgaard J, et al. (2019). Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage. Nature Communications, 10(1): 1124

Hijmans, R J, Cameron S E, Parra J L, Jones P G, Jarvis A (2005). Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology, 25: 1965–1978

Hu, A, Wang J, Sun H, Niu B, Si G, Wang J, Yeh C, Zhu X, Lu X, Zhou J, et al. (2020). Mountain biodiversity and ecosystem functions: interplay between geology and contemporary environments. ISME Journal, 14: 1–14

Ji M, Liu Z, Sun K, Li Z, Fan X, Li Q (2021). Bacteriophages in water pollution control: advantages and limitations. Frontiers of Environmental Science & Engineering, 15(5): 84

Jones E R, van Vliet M T H, Qadir M, Bierkens M F P (2021). Country-level and gridded estimates of wastewater production, collection, treatment and reuse. Earth System Science Data, 13(2): 237–254

LaMartina E L, Mohaimani A A, Newton R J (2021). Urban wastewater bacterial communities assemble into seasonal steady states. Microbiome, 9(1): 116

Langmead B, Salzberg S L (2012). Fast gapped-read alignment with Bowtie 2. Nature Methods, 9(4): 357–359

Li B, Ju F, Cai L, Zhang T (2015). Profile and fate of bacterial pathogens in sewage treatment plants revealed by high-throughput metagenomic approach. Environmental Science & Technology, 49(17): 10492–10502

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R (2009). The sequence alignment/map format and SAMtools. Bioinformatics, 25(16): 2078–2079

Li M, Song G, Liu R, Huang X, Liu H (2022). Inactivation and risk control of pathogenic microorganisms in municipal sludge treatment: a review. Frontiers of Environmental Science & Engineering, 16(6): 70

Martínez I, Stegen J C, Maldonado-Gómez M X, Eren A M, Siba P M, Greenhill A R, Walter J (2015). The gut microbiota of rural Papua New Guineans: composition, diversity patterns, and ecological processes. Cell Reports, 11(4): 527–538

Mateo-Sagasta J, Raschid-sally L, Thebo A (2015). Global wastewater and sludge production, treatment and use. In: Drechsel P, Qadir M, Wichelns D, eds. Wastewater. Dordrecht: Springer

Matus M, Duvallet C, Soule M K, Kearney S M, Endo N, Ghaeli N, Brito I, Ratti C, Kujawinski E B, Alm E J (2019). 24-hour multi-omics analysis of residential sewage reflects human activity and informs public health. bioRxiv, doi:10.1101/728022

Miller G E, Engen P A, Gillevet P M, Shaikh M, Sikaroodi M, Forsyth C B, Mutlu E, Keshavarzian A (2016). Lower neighborhood socioeconomic status associated with reduced diversity of the colonic microbiota in healthy adults. PLoS One, 11(2): e0148952

Nadimpalli M L, Marks S J, Montealegre M C, Gilman R H, Pajuelo M J, Saito M, Tsukayama P, Njenga S M, Kiiru J, Swarthout J, et al. (2020). Urban informal settlements as hotspots of antimicrobial resistance and the need to curb environmental transmission. Nature Microbiology, 5(6): 787–795

Newton R J, Mclellan S, Dila D K, Vineis J H, Morrison H G, Eren A M, Sogin M L (2015). Sewage reflects the microbiomes of human populations. mBio, 6(2): e02574

Nowak P, Troseid M, Avershina E, Barqasho B, Neogi U, Holm K, Hov J R, Noyan K, Vesterbacka J, Svärd J, et al. (2015). Gut microbiota diversity predicts immune status in HIV-1 infection. AIDS, 29(18): 2409–2418

Nowrotek M, Jałowiecki Ł, Harnisz M, Płaza G A (2019). Culturomics and metagenomics: in understanding of environmental resistome. Frontiers of Environmental Science & Engineering, 13: 40

Parizadeh M, Arrieta M C (2023). The global human gut microbiome: genes, lifestyles, and diet. Trends in Molecular Medicine, 29(10): 789–801

Pasolli E, Asnicar F, Manara S, Zolfo M, Karcher N, Armanini F, Beghini F, Manghi P, Tett A, Ghensi P, et al. (2019). Extensive unexplored human microbiome diversity revealed by over 150,000 genomes from metagenomes spanning age, geography, and lifestyle. Cell, 176(3): 649–662.e20

Pipek O A, Medgyes-Horváth A, Dobos L, Stéger J, Szalai-Gindl J, Visontai D, Kaas R S, Koopmans M, Hendriksen R S, Aarestrup F M, et al. (2019). Worldwide human mitochondrial haplogroup distribution from urban sewage. Scientific Reports, 9(1): 11624

Qin J, Li R, Raes J, Arumugam M, Burgdorf K S, Manichanh C, Nielsen T, Pons N, Levenez F, Yamada T J, et al. (2010). A human gut microbial gene catalogue established by metagenomic sequencing. Nature, 464(7285): 59–65

Schneeberger P H H, Fuhrimann S, Becker S L, Pothier J F, Duffy B, Beuret C, Frey J E, Utzinger J (2019). Qualitative microbiome profiling along a wastewater system in Kampala, Uganda. Scientific Reports, 9(1): 17334

Sidhu J P S, Hodgers L, Ahmed W, Chong M N, Toze S (2012). Prevalence of human pathogens and indicators in stormwater runoff in Brisbane, Australia. Water Research, 46(20): 6652–6660

Staley C, Reckhow K H, Lukasik J, Harwood V J (2012). Assessment of sources of human pathogens and fecal contamination in a Florida freshwater lake. Water Research, 46(17): 5799–5812

Tisza M, Javornik C S, Avadhanula V, Zhang P, Ayvaz T, Feliz K, Hoffman K L, Clark J R, Terwilliger A, Ross M C, et al. (2023). Wastewater sequencing reveals community and variant dynamics of the collective human virome. Nature Communications, 14(1): 6878

Tiwari A, Krolicka A, Tran T T, Räisänen K, Ásmundsdóttir Á M, Wikmark O G, Lood R, Pitkänen T (2024). Antibiotic resistance monitoring in wastewater in the Nordic countries: a systematic review. Environmental Research, 246: 118052

Wu L, Ning D, Zhang B, Li Y, Zhang P, Shan X, Zhang Q, Brown M R, Li Z, Van Nostrand J D, et al. (2019). Global diversity and biogeography of bacterial communities in wastewater treatment plants. Nature Microbiology, 4(7): 1183–1195

Yatsunenko T, Rey F E, Manary M J, Trehan I, Dominguez-Bello M G, Contreras M, Magris M, Hidalgo G, Baldassano R N, Anokhin A P (2012). Human gut microbiome viewed across age and geography. Nature, 486(7402): 222–227

Zhao J, Li B, Lv P, Hou J, Qiu Y, Huang X (2022). Distribution of antibiotic resistance genes and their association with bacteria and viruses in decentralized sewage treatment facilities. Frontiers of Environmental Science & Engineering, 16(3): 35