


RESEARCH

Open Access



Integrated large-scale metagenome assembly and multi-kingdom network analyses identify sex differences in the human nasal microbiome

Yanmei Ju^{1,2,3†} , Zhe Zhang^{1,2†}, Mingliang Liu^{1,2,3†}, Shutian Lin^{1,2,3†}, Qiang Sun^{1,2,6}, Zewei Song¹, Weiting Liang^{1,2,3}, Xin Tong^{1,2}, Zhuye Jie^{1,2}, Haorong Lu⁷, Kaiye Cai¹, Peishan Chen¹, Xin Jin¹, Wenwei Zhang¹, Xun Xu¹, Huanming Yang^{1,8}, Jian Wang¹, Yong Hou¹, Liang Xiao^{1,4,5}, Huijue Jia^{9,10*}, Tao Zhang^{2,11*} and Ruijin Guo^{1,2,11*}

[†]Yanmei Ju, Zhe Zhang, Mingliang Liu, and Shutian Lin contributed equally to this work.

*Correspondence: jiahuijue@ipm-gba.org.cn; tao.zhang@genomics.cn; guoruijin@genomics.cn

¹ BGI Research, Shenzhen 518083, China
⁹ School of Life Sciences, Fudan University, Shanghai 200433, China

¹¹ BGI Research, Wuhan 430074, China

Full list of author information is available at the end of the article

Abstract

Background: Respiratory diseases impose an immense health burden worldwide. Epidemiological studies have revealed extensive disparities in the incidence and severity of respiratory tract infections between men and women. It has been hypothesized that there might also be a nasal microbiome axis contributing to the observed sex disparities.

Results: Here, we study the nasal microbiome of healthy young adults in the largest cohort to date with 1593 individuals, using shotgun metagenomic sequencing. We compile the most comprehensive reference catalog for the nasal bacterial community containing 4197 metagenome-assembled genomes and integrate the mycobiome, to provide a valuable resource and a more holistic perspective for the understudied human nasal microbiome. We systematically evaluate sex differences and reveal extensive sex-specific features in both taxonomic and functional levels in the nasal microbiome. Through network analyses, we capture markedly higher ecological stability and antagonistic potentials in the female nasal microbiome compared to the male's. The analysis of the keystone bacteria reveals that the sex-dependent evolutionary characteristics might have contributed to these differences.

Conclusions: In summary, we construct the most comprehensive catalog of metagenome-assembled-genomes for the nasal bacterial community to provide a valuable resource for the understudied human nasal microbiome. On top of that, comparative analysis in relative abundance and microbial co-occurrence networks identify extensive sex differences in the respiratory tract community, which may help to further our understanding of the observed sex disparities in the respiratory diseases.

Keywords: Nasal microbiome, Metagenome-assembled genomes, Sex differences, Respiratory health, Network analysis, Keystone, Ecological stability, Genetic evolutionary forces, Biosynthetic gene cluster



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Background

Respiratory diseases impose an immense health burden worldwide, affecting billions of people's lives and accounting for over 10% of all disability-adjusted life-years (DALY) as of 2019 according to the Global Burden of Diseases (GBD) study [1–4], let alone the catastrophic impact of the COVID-19 pandemic. Sex is a significant factor in many diseases. Epidemiological studies have revealed extensive disparities in the incidence and severity of respiratory tract infections (RTIs) between males and females. Males are generally more commonly and severely affected by most RTIs than females across all age groups [4–6]. A greater mortality rate for males was also observed in COVID-19 [7–9]. Sex-specific differences in immunity mediated by sex chromosome complement, genes, and sex hormones can play important roles in the observed disparity [6–8]. Nevertheless, the mechanism remains unclear. The respiratory tract microbiome has been implicated in different respiratory diseases [10–17]. It is recently hypothesized that there might also be a nasal microbiome axis contributing to the observed sex disparities [18].

The nasal bacterial community is characterized by a high prevalence of *Corynebacterium* spp., *Cutibacterium* spp., and *Staphylococcus* spp., with most of the components belonging to phyla Actinobacteria, Firmicutes, and Proteobacteria [19, 20]. In addition to bacterial colonizers, the nasal cavity also harbors a mycobiota [21–23], as well as the presence of viruses. Nevertheless, the nasal microbiome studies are hitherto limited to small sample sizes or conservative gene-based sequencing (16S rRNA, 18S rRNA, ITS) [20, 22–28]. Sex differences in the nasal microbiome have never been systematically evaluated. This is not surprising considering that even in the most researched gut microbiome, sex differences only came to light very recently through large-scale population-level studies [29–31]. It is increasingly recognized that the nasal microbiome might function as a gatekeeper in respiratory health [32, 33]. The nasal cavity is featured by limited nutrients and adhesion surfaces [34] and also represents a major reservoir for opportunistic pathogens, such as *Staphylococcus aureus*, *Streptococcus pneumoniae*, and *Haemophilus influenzae* [35]. The microbes in this niche are hence in constant competition, and sometimes form cooperative relations, to gain self-fitness [36–40]. Extensive antimicrobial substance productions have been identified in nasal microbes, which can be potential mediators of the interactions [38–44]. The competitive (antagonistic) and cooperative (synergistic) interactions influence both the initial colonization of pathogens and the thereafter dynamics. Network-based approaches have been shown helpful in deciphering complex interactions and are increasingly applied in the microbial field. Understanding the nature of microbial co-occurrence and correlation patterns within and across domains may provide insights into the ecological systems as well as related human diseases. Through network-based analyses, researchers studying bronchiectasis exacerbations found that patients with different exacerbation risks featured distinct microbial interaction networks [45]. Instead of the implicated pathobiont *Pseudomonas* alone, it is the interaction network that is associated with the exacerbation risk. While cross-domain interactions are rarely explored, Tipton et al. recently showed that compared to single-domain networks, bacteria-fungi combined networks had higher overall connectivity and increased attack robustness [46]. More importantly, network analyses can help elucidate and prioritize the keystones of a community, which may not be the species dominant in abundance, and sometimes can even be unknown as “microbial dark matter” [47, 48].

In this work, we study the nasal microbiome of healthy young adults in a so far largest cohort with 1593 individuals based on deep shotgun metagenomic sequencing. De novo assembly is performed to catalog the nasal bacterial colonizers/residents, which has also identified and therefore accounted for uncharacterized components of the community. We then characterized the composition of the nasal bacterial and fungal community in this cohort. Unsupervised clustering of the weighted similarity matrix integrating the bacterial and fungal community reveals clearly separable patterns between the two sexes, implying a distinct structure of the nasal microbiome between males and females, which is further confirmed by PERMANOVA and random forest analysis. Following this link, we systematically evaluate sex differences for the first time and reveal extensive sex-specific features in both compositional and ecological aspects in the nasal microbiome. Through network analyses, we capture markedly higher stability and antagonistic potentials in the nasal microbiome of females than that of males, in the shaping of which, sex-dependent evolutionary characteristics might have played a role as revealed by the keystone bacteria of the communities.

Results

Characterizing the nasal bacteriome and mycobiome

To characterize the nasal microbiome of healthy young adults, we performed deep shotgun metagenomic sequencing on anterior nares samples of 1593 individuals from the 4D-SZ cohort [49–53]. The workflow for this study is shown in Additional file 2: Fig. S1. The study cohort included 439 males, 807 females, and 347 individuals with sex information missing (Additional file 1: Table S1a). The average age was 29.9 (± 5.13) years old (males: 30.4 (± 5.25); female: 29.6 (± 5.03)), and body mass index (BMI) was in the range of 21.9 (± 3.36) (males: 23.3 (± 3.26); females: 21.1 (± 3.15)) (Additional file 1: Table S1b, Table S5a). In total, 128.21 terabases raw data were generated with an average of 80.48 gigabases for each sample (Additional file 1: Table S2). A single-sample assembly, and single-sample binning strategy (see the “Methods” section) was employed to reconstruct genomes from the ultra-deeply sequenced metagenomic data. A total of 4197 MAGs were assembled at a threshold for quality control of $>50\%$ completeness and $<10\%$ contamination. To compile a non-redundant MAGs catalog, we performed de-replication with 99% of the average nucleotide identity (ANI). In the end, a catalog of 974 non-redundant MAGs for human nasal-associated bacteria was retained, which included 718 high-quality (completeness $>90\%$ and contamination $<5\%$) and 256 medium-quality (completeness $>50\%$ and contamination $<10\%$) ones (Fig. 1a; Additional file 2: Fig. S2). 16S rRNA genes had been detected in about 45% of the 974 MAGs (Fig. 1a). To explore the taxonomic coverage of this catalog, we classified the MAGs according to 95% average nucleotide identity. Overall, we obtained 232 species from 13 known phyla, with 150 annotated to known genomes in the GTDB database, and the other 82 as newly identified (unknown) (Fig. 1b; Additional file 1: Table S3a). The unknown species spanned over the 12 phyla, with the largest number from Bacteroidota. For four phyla, including Fusobacteriota, Eremiobacterota, Deinococcota, and Bdellovibrionota, only unknown species were discovered. This suggests that the habitat of the nasal cavity featured drastically distinct characteristics from other habitats where the species of these phyla are often identified, e.g. *Fusobacterium nucleatum* of phylum Fusobacteriota is often detected in

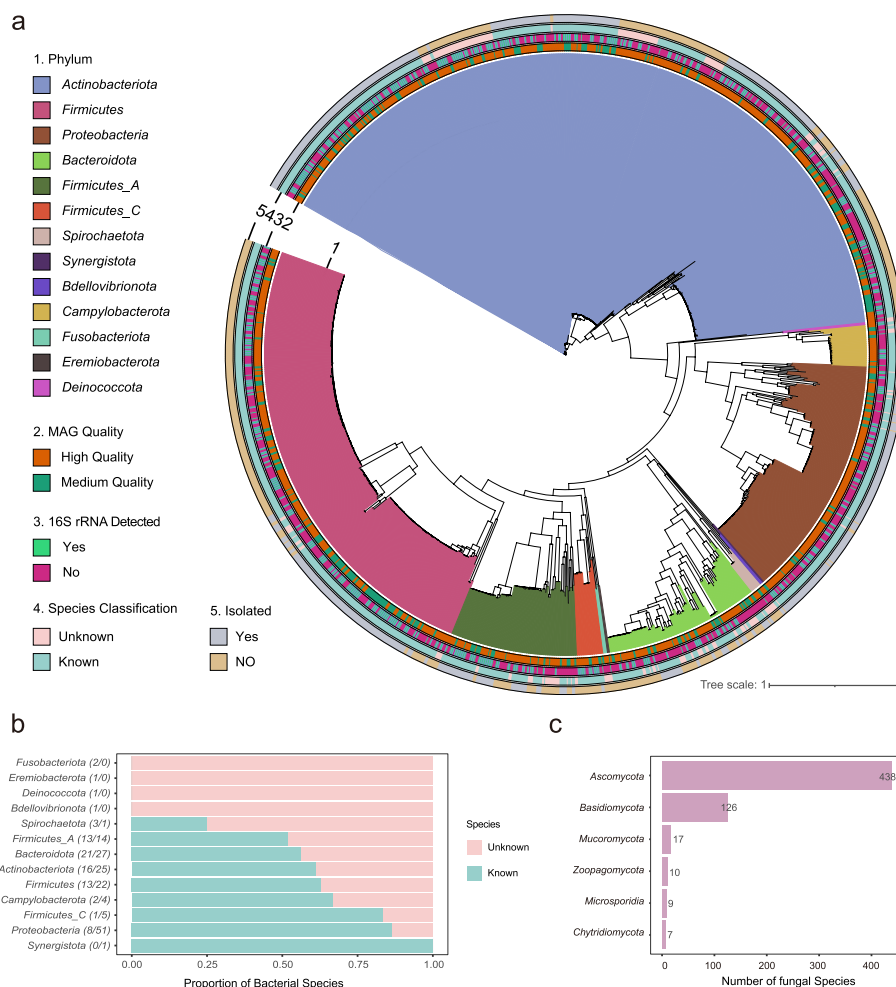


Fig. 1 Overall representation of the microbes in anterior nares of healthy young adults. **a** Phylogeny of 974 non-redundant bacterial MAGs (metagenome-assembled genomes) detected in anterior nares. It constituted five layers representing respectively: 1 for phylum, 2 for MAGs quality, 3 for if 16s rRNA detected, 4 for if classified in the species level, and 5 for if isolated as depicted in the GTDB database. **b** Proportion of unknown and known bacterial species in each phylum with the absolute number indicated in the brackets respectively. **c** Number of fungal species in each phylum

oral and fecal samples. Additionally, we identified six novel genera from phylum Proteobacteria, Bacteroidota, and Firmicutes_A, and one novel family from phylum Eremiobacterota, which cannot be assigned to any known taxa in a finer level at the respective phylogenetic distance cut-offs (Additional file 1: Table S3a). Notably, our data also improved the genome completeness of a singleton taxon, namely *QFNR01 sp003248485* (90.42% completeness, compared with 75.15% completeness in GTDB; Additional file 1: Table S3a). Overall, the majority of the MAGs in the catalog belonged to Actinobacteriota, Proteobacteria, and Firmicutes, which is typical for the human nasal microbiome [19, 20, 35, 54].

Next, we explored the functional potential encoded in the nasal-associated non-redundant MAGs catalog. Core functional pathways reconstructed with KEGG modules showed that species from different phyla all characterized functions for degrading

the three major nutrient sources including (poly-)saccharides, proteins, and lipids (Additional file 2: Fig. S3). Proteobacteria species displayed more diverse pathways for drug efflux transporter/pump and drug resistance. Two-component systems (TCSs), as a major mechanism for bacteria to sense and respond to environmental stresses [55], showed high heterogeneity among different phyla. The MtrAB TCS commonly found in Actinobacteria [56], was identified in most MAGs of this phyla with near-complete coverage. The VicRK TCS, a conserved two-component transcriptional regulatory system in several streptococcal species of the human microbiota [57], was also detected in many *Staphylococcus* spp. from the same Firmicutes phylum.

Natural products of human microbiota are increasingly recognized as important mediators for a variety of microbe-host and microbe-microbe interactions, which in turn can be explored for potential pharmaceutical applications [42, 58, 59]. As an example, a nasal isolate of *Staphylococcus lugdunensis* has recently been shown to produce a novel antibiotic, lugdunin, a non-ribosomally synthesized bioactive natural product, which is bactericidal against major human pathogens and prohibits the colonization of *S. aureus* in the nasal cavity [40]. We therefore screened for the presence of secondary metabolites biosynthetic gene clusters (BGCs) encoded within the 974 non-redundant MAGs using antiSMASH [60] (Additional file 1: Table S3b). In total, we detected 2921 BGCs, which were primarily inferred as synthesized terpenes, nonribosomal peptides (NRPs), types I polyketide synthases (PKSs), siderophores and other unspecified ribosomally synthesized and post-translationally modified peptide products (RiPPs). Notably, 514 of them were screened from MAGs newly identified from the nasal microbiome in this cohort (Additional file 2: Fig. S4a). In addition, 1975 (67.2%) of the detected gene clusters were novel clusters, most of which were from Actinobacteriota, Firmicutes, and Proteobacteria (Additional file 2: Fig. S4b). These data, in particular the high number and proportion of novel clusters, suggest that the nasal microbiota may serve as a rich reservoir for new antibiotics or other pharmaceuticals.

To profile the nasal microbiome composition, we mapped the metagenome data to the constructed non-redundant nasal bacterial MAGs catalog and a manually curated database of fungi genomes, and then filtered out species with low prevalence and low relative abundance to generate the bacterial and fungal profiles (see the “[Methods](#)” section). In total, 122 bacteria and 131 fungal species with high confidence were retained. The nasal bacteriome composition was highly variable among the young adults in the cohort that the relative abundance of genus *Corynebacterium*, which characterized the most abundant genus across the cohort (mean relative abundance 55.2%) and within 94.9% of the individuals, varied from 0.39 to 85.3% (Fig. 2a). The top 10 species accounted for an accumulative mean relative abundance of 74.7% (Additional file 1: Table S4a), suggesting that the nasal bacterial community was dominated by a few taxa (Fig. 2a). Consistent with former studies, the most abundant species were mainly from the genera *Corynebacterium*, *Staphylococcus*, *Moraxella*, *Cutibacterium*, *Dolosigranulum*, and *Lawsonella* [33, 35, 54]. In patients with nasal polyps, however, *Corynebacterium*, *Staphylococcus*, *Moraxella*, and *Dolosigranulum* were found to be significantly reduced compared to controls [12]. The fungal species in this cohort were mostly from phyla Ascomycota and Basidiomycota (Fig. 1c). Compared to the bacterial community, the mycobiome was more evenly distributed, taking 35 species to account for 74.8% of the overall fungal

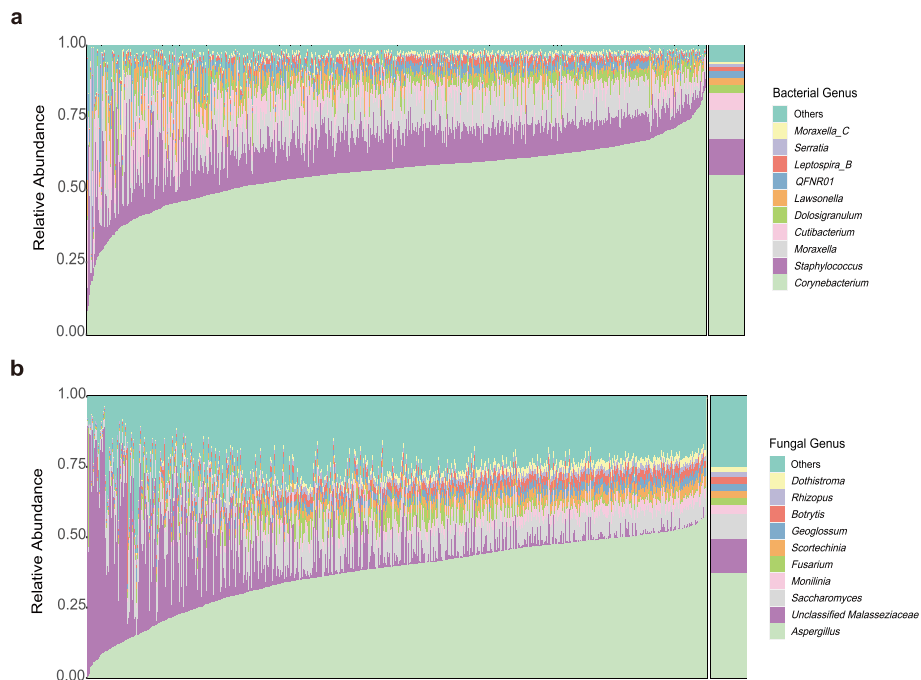


Fig. 2 Relative abundance of top 10 abundant fungal and bacterial taxa in genus level. Bar chart showing the individual relative abundance and the mean relative abundance of top 10 abundant bacterial (a) and fungal (b) taxa in genus level of the nasal microbiome

mycobiome composition (Additional file 1: Table S4c). *Aspergillus* and an unclassified *Malasseziaceae* genus made the two pillars of the fungal community in the nasal cavity of this cohort (Fig. 2b; Additional file 1: Table S4d). In a recent study in chronic obstructive pulmonary disease (COPD) patients, *Malassezia* was found to be the most abundant genus in the nasal mycobiome, but *Aspergillus* ranked much lower [23].

Unsupervised clustering helps uncover sex differences in the nasal microbiome composition

To gain a holistic perspective of the microbial structure, we integrated the Bray–Curtis dissimilarity matrices of the bacterial and fungal community with a weighted similarity network fusion (WSNF) approach [45] (see the “Methods” section). Unsupervised clustering of the resultant similarity matrix classified the cohort into three clusters, with clusters 2 and 3 comprising samples almost exclusively from each single sex, and cluster 1 featuring two separable patterns in the similarity matrix corresponding to the two sexes (Additional file 2: Fig. S5a; see the “Methods” section). Furthermore, multivariate permutational multivariate analysis of variance (PERMANOVA) between cluster and age, sex, or BMI showed that the variance explained by the cluster had a substantial decrease when adjusted by sex and only a minimal decrease when adjusted by age or BMI (Additional file 2: Fig. S5b). We also observed a distinct difference between males and females in WSNF similarity (Fig. 3a). This gives a strong indication that the young adult males and females in this cohort probably characterized distinctive structures in their nasal microbiome.

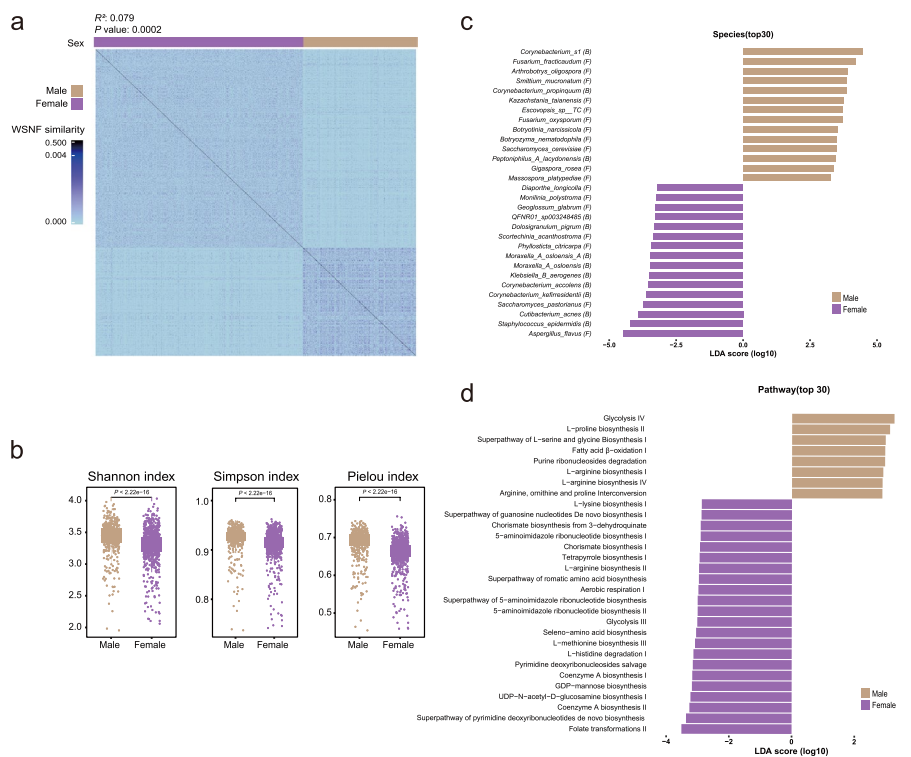


Fig. 3 Unsupervised clustering and sex differences in the nasal microbiome composition. **a** Heatmap illustrating WSNF similarity scores stratified by unsupervised clustering, with sex information indicated by the bar on the top. WSNF similarity represents the fused similarity score. R^2 and P value were calculated by PERMANOVA with WSNF similarity matrices **b** Box plot showing male and female nasal microbial alpha diversity indices including Shannon Simpson and Pielou calculated with merged profile. P values were obtained from two-sided Wilcoxon rank-sum tests. **c** The comparison of LDA effect size (LEfSe) between males (green) and females (brown) illustrating discriminative species. Only the top 30 discriminative species by LDA score are shown, which are significantly different (BH-adjusted P value < 0.05 and LDA score > 2.5). **d** The comparison of LDA effect size between males and females illustrating discriminative pathway. Only the top 30 discriminative pathways by LDA score are shown. (BH-adjusted P value < 0.05 and LDA score > 2.5)

Following these links that unsupervised clustering and PERMANOVA uncovered, we next systematically evaluated the sex differences in the nasal microbiome composition. Considering the discrepancies in age and BMI between males and females might potentially confound the observed variance (Additional file 1: Table S5a), multivariate PERMANOVA analysis was applied and confirmed that sex was a significant covariant for the bacteria-fungi integrated microbiome while age and BMI were not. For single-domain PERMANOVA analyses, sex accounted for higher variance in the mycobiome than in the bacteriome (Additional file 1: Table S5b). To assess the discriminatory potential of sex in the nasal microbiome, random forest analysis was carried out and showed that, the nasal microbiome achieved an AUC of 0.9779 in differentiating males and females (0.9686 for bacteria and 0.9846 for fungi) (Additional file 2: Fig. S6b). In the bacteria-fungi merged profile, we observed a significantly higher alpha diversity including the Shannon index, Simpson and Pielou indices in males than in females (Fig. 3b). And such a significant difference was also detected in the bacteria and fungi single domain communities except for Shannon index (Additional file 2: Fig. S6a). For individual microbial taxon, we performed linear discriminative analysis effect size (LEfSe)

and identified considerable significant associations between the relative abundances and sex. Specifically, at the species level 34 bacteria and 57 fungi, and at the genus level 23 bacteria and 45 fungi, were significantly different (BH-adjusted P value < 0.05 and LDA score > 2.5) in abundance between males and females (Fig. 3c; Additional file 1: Table S6). Interestingly, the taxa number enriched in females almost doubled that in males. Notably, *Corynebacterium accolens* and *Dolosigranulum pigrum*, as the most abundant species among others in this cohort, were significantly more abundant in females. These two species have been shown to inhibit the growth of *S. aureus*, a commonly known opportunistic pathogen in the nasal cavity [23]. Additionally, *Lactobacillus* spp., typically found in the female vagina, has been recently reported to have a niche in the human nose and may exert a beneficial effect [34, 61]. In our cohort, however, we did not detect a female-biased colonization of *Lactobacillus* spp. in the nasal cavity. Additionally, the most abundant fungal species in this cohort, i.e., *Aspergillus flavus*, was also significantly enriched in females. Consistent with the observations on the taxonomic level, extensive differences were identified in both overall functional capacity and specific pathways of the nasal microbiome (Additional file 1: Table S5c, Table S6e). For instance, folate transformations, biosynthesis for pyrimidine, coenzyme A, UDP-N-acetyl-D-glucosamine, GDP-mannose, and pyrimidine deoxyribonucleosides salvage were significantly more abundant in males, whereas the pathways for biosynthesis of proline, L-serine, glycine, and L-arginine were enriched in females (Fig. 3d).

Additionally, since we have applied two different assemblers in generating the bacterial MAGs catalog based on which the bacterial profile was derived, we further carried out analyses to assess the influence of the assembler especially on sex differences. The results confirmed that the assembler does have a significant effect on the bacterial composition (Additional file 1: Table S7a). Nevertheless, after adjusting for assembler, sex remained as a significant covariate for the nasal bacterial community (Additional file 1: Table S7b). More importantly, stratified analyses showed that sex differences were consistently observed within the two assembler-divided subgroups, on both the community level and specific individual taxa level (Additional file 1: Table S7d). This underscores the significance and robustness of the observed sex differences.

Network analyses capture markedly higher ecological stability and antagonistic potentials in the nasal microbiome of females than that of males

Having uncovered extensive sex differences in the microbial composition, next, we aimed to determine if the nasal microbiome featured distinct characteristics in ecological relationships between males and females. To characterize the microbial interactions within each sex, we employed an integrated approach combining COAT (composition-adjusted thresholding) [62], HUGE (High-dimensional Undirected Graph Estimation) [63], MI (mutual information), and Bray–Curtis dissimilarity to construct the co-occurrence networks (Fig. 4a; see the “Methods” section). The inferred interactions between microbes, as nodes in the graphs, were represented by signed edges (P value < 0.01) in the network, with positive for cooperative/synergistic relation and negative for competitive/antagonistic relation. The total number of interactions (edges) was very close between the two sexes with a marginally higher number of negative interactions in females. Splitting the entire network into three sub-networks, i.e., within the bacteria domain, within the fungi

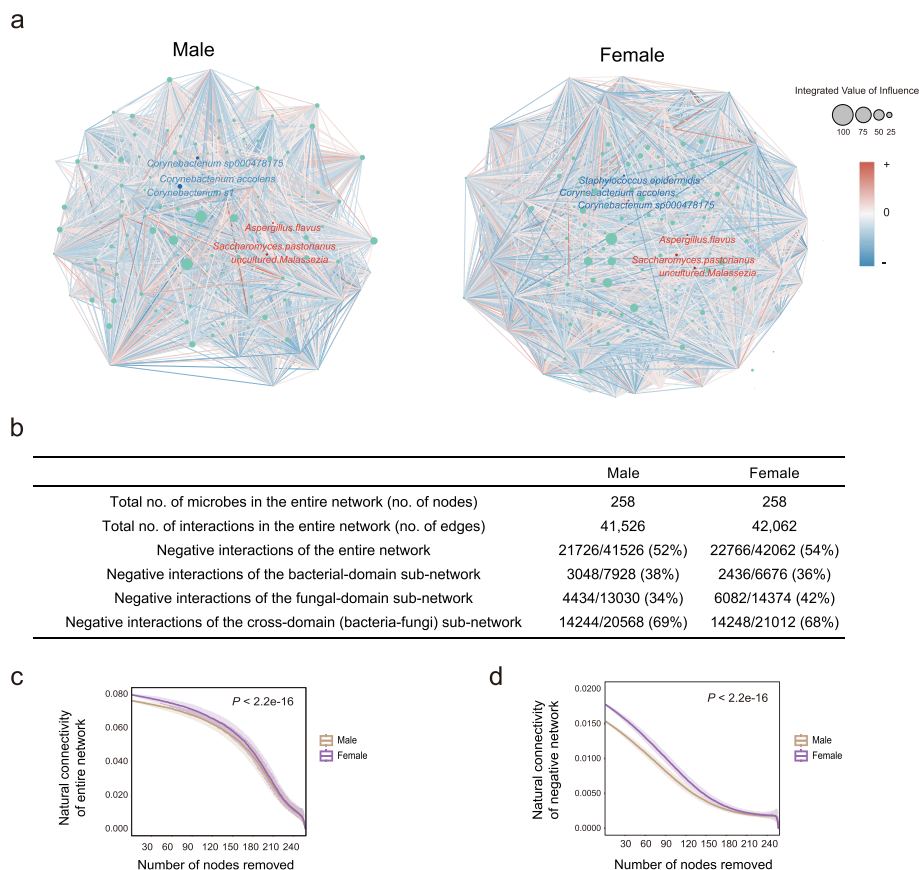


Fig. 4 Network characterization of the nasal microbiome for males and females. **a** Nasal microbial interaction network of males and females. Node size represents each taxon’s integrated value of influence (IVI). Red and blue lines indicate positive and negative interactions respectively. The top 3 bacteria and fungi by relative abundance are annotated with blue and red fonts respectively. **b** Summary table of network characteristics of males and females. **c–d** Attack robustness of the entire network of total interaction (**c**) and negative interaction (**d**) for males (green) and females (brown) as measured by natural connectivity. Line and box reflect the median and IQRs. A statistical measure of P value is described in the methods section

domain, and cross bacteria-fungi domain, revealed that the cross-domain sub-network accounted for over half of the negative interactions (Fig. 4b; Additional file 1: Table S8).

The functioning of complex networks largely relies on their robustness [64], a better understanding of which can provide valuable insights into RTI susceptibility and pathologies. We thus adopted a sensitive and reliable measure, namely natural connectivity [65–67], to quantify the stability of the inferred networks. To simulate the influence of microbes’ loss on the network, we performed random attacks and assessed the stability of the remaining network [68] (see the “Methods” section). Intriguingly, the network robustness was much higher for females than for males (Fig. 4c). While the human nasal microbiome is increasingly regarded as a gatekeeper of respiratory health, opportunistic pathogens do often present even in healthy individuals. Thus, the negative/antagonistic interactions are of particular interest. When only considering the negative interactions, we observed that much higher robustness for females still held (Fig. 4d). Significant separations of the natural connectivity plot were observed between the networks of males and females, for both the entire network and the negative network (P value $< 2.2e - 16$;

Fig. 4c and d). Higher overall natural connectivity for females largely remained until over half of the species were removed, further confirming that females characterized a more stable network with more intensive interactions and higher antagonistic potentials which may provide stronger resistance against opportunistic pathogens. Nevertheless, among the three sub-networks, only the fungal domain had the same direction as the entire network in terms of higher robustness in females than in males, and the bacterial domain and bacteria-fungi sub-networks were more robust in males (P value $< 2.2e - 16$; Additional file 2: Fig. S7). Interestingly, a recent study also suggested that fungi played a stabilizing role in the lung and skin microbial ecosystems [46].

Sex-dependent genetic evolutionary forces in the shaping of keystones in the nasal microbial community

Network analysis can be a powerful tool for inferring keystone taxa of the microbial communities [47, 48, 64, 69]. To this end, we adopted a novel influential node detection method, integrated value of influence (IVI), which captures all topological dimensions of the networks, to assess the importance of individual taxon in the ecosystem [70]. Notably, the IVIs of most taxa derived from the entire networks of males and females were considerably different (Additional file 1: Table S9), indicating different levels of importance of the respective taxa potentially eliciting in the microbial community of each sex. Moreover, IVI only weakly correlated with relative abundance (Additional file 2: Fig. S8), suggesting that the most abundant taxa may not necessarily exert the strongest influences from an ecological perspective [47, 48]. Keystone microbes represent the ones contributing the most to the robustness of the community. With a permutational approach (see the “Methods” section) we derived the keystone sets for males and females, which included 13 and 10 taxa respectively (Fig. 5a). Intriguingly, the keystone sets for males and females both contained taxa from bacterial and fungal domains, but with completely different specific components and remarkably different IVIs between the two sexes for each keystone.

Evolution is important for ecological dynamics in bacterial communities. To illuminate the genetic evolutionary characteristics of the keystones for each sex, we utilized the bacterial MAGs and evaluated the selection of environmental pressures for the keystone bacteria by estimating the pN/pS ratio within each genome for each sample [71, 72]. The results showed that the pN/pS ratios varied among different species, but were mostly below one for both males and females (Fig. 5b). This suggested that the evolution of the keystone bacteria was largely predominated by long-term purifying selection. On the other hand, the pN/pS ratios differed significantly between males and females in some of the keystone bacteria, including male-specific keystone *Staphylococcus warneri* and female-specific keystone *Anaerococcus provencensis*, *Stenotrophomonas geniculata*, and *Finegoldia s1* (Fig. 5b). This can potentially be in relation to sex-specific evolutionary constraints confronted by the microbes in the nasal cavity of males and females, such as different levels of immunoinflammatory characteristics.

On the gene level, however, we observed considerable deviations in pN/pS ratios of the same keystone taxa between males and females, indicating sex-dependent selective pressures and genetic adaptations (Fig. 5c, Additional file 2: Fig. S9; Additional file 1: Table S10). For instance, the nasal cavity is noted for limited resources available, such as

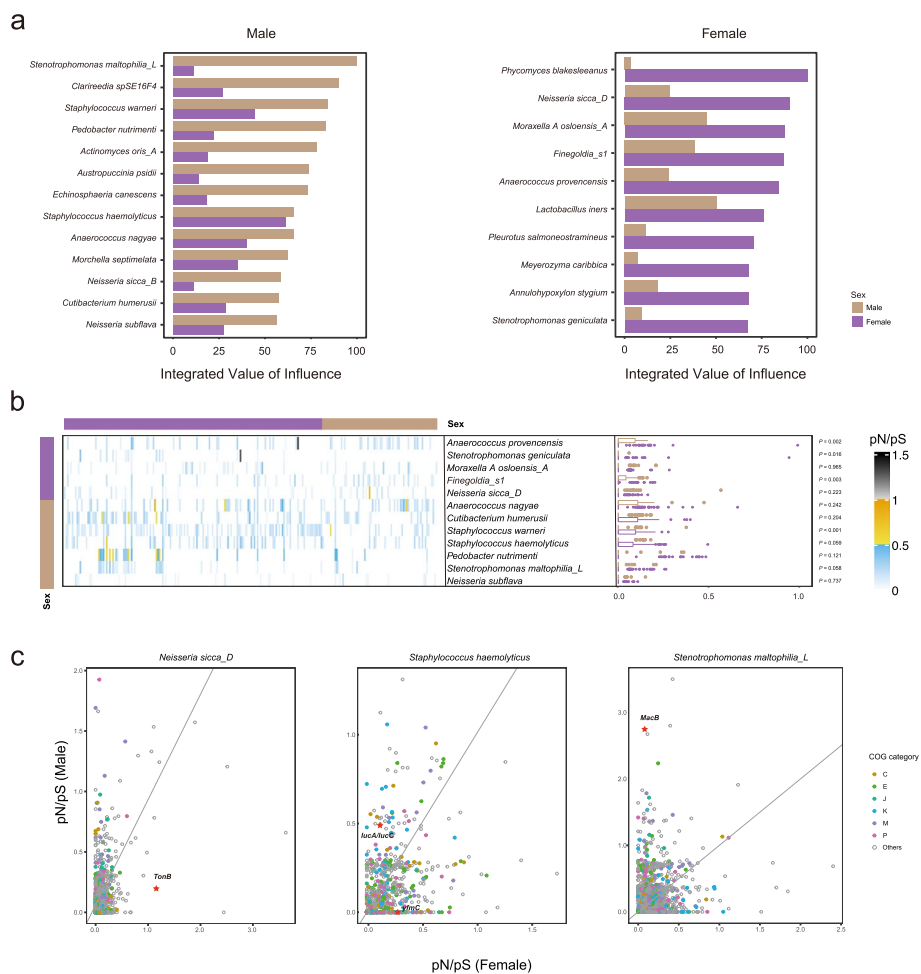


Fig. 5 Characteristics of the keystone taxa identified in male and female nasal microbial interaction networks. **a** The integrated value of influence (IVI) of the keystone taxa of males (left) and females (right). Green and brown bars represent the IVI of the respective taxa in the male and female networks respectively. **b** The pN/pS ratio of keystone bacteria for individuals shown by heatmap and boxplot. Green and brown represent male and female respectively (vertical bar: keystone belongs to male or female network; horizontal bar: male and female individuals; boxplot: pN/pS ratios for male and female individuals). **c** The pN/pS ratio for 3 bacterial keystones in gene levels of male (y-axis) and female (x-axis) participants with COG category. The stars represent the genes that have been described in detail in the main text. One-letter abbreviations for the functional categories: C, energy production and conversion; E, amino acid metabolism and transport; J, translation, including ribosome structure and biogenesis; K, transcription; M, cell wall structure and biogenesis and outer membrane; P, inorganic ion transport and metabolism

iron limitation [16, 34, 73]. Notably, the 974 nasal bacterial MAGs encoded remarkably more siderophores (Additional file 2: Fig. S4a), one of the main mechanisms for bacterial iron sequestering, compared to that detected in the large collection of human gut bacterial MAGs derived from over 10,000 samples [74]. Although *Neisseria sicca*, a common nasopharyngeal commensal, does not encode siderophores [75], we found the female keystone *N. sicca_D* underwent positive adaptation in genes encoding TonB-dependant siderophore receptors (mean pN/pS ratio of 1.168) in females, with which the bacteria can exploit siderophores produced by other members of the community for iron sequestering. In contrast, in males, it was subjected to purifying selection in these genes with a mean pN/pS ratio of 0.195. In a male keystone bacterium, *Staphylococcus haemolyticus*,

we also observed that genes encoding IucA/IucC family siderophore biosynthesis protein showed relaxed purifying selection in males (mean pN/pS ratio of 0.496) but tight purified selection in females (mean pN/pS ratio of 0.116). Gene *yfmC* in *S. haemolyticus*, which encodes Fe⁽³⁺⁾-citrate-binding protein involved in iron transport, was purged in males (mean pN/pS ratio of 0) whereas showed tight purified selection (mean pN/pS ratio of 0.259) in females. Antibiotics represent another major category of stresses for bacteria, for which resistance evolves over time. As a global emerging multidrug-resistant organism, *Stenotrophomonas maltophilia* has been most commonly associated with respiratory infections in humans [76] and isolated predominantly in elderly males of hospitalized lower RTI patients [77]. Like the other MacA-MacB-TolC tripartite efflux pumps, *S. maltophilia* MacB has been previously revealed to drive resistance to a variety of antibiotics, such as macrolides, aminoglycosides, and polymyxins, in concert with MacA adaptor protein and TolC outer membrane exit duct [78–80]. Interestingly, we found in *S. maltophilia_L* the gene coding for MacB exhibited strong positive selection in males, but tight negative selection in females (mean pN/pS ratio: 2.73 vs. 0.08). Together, the keystone bacteria exhibited highly sex-specific genetic evolutionary characteristics in niche-specific or sex-biased stress-related functional units, which were in close relation to their role in the respective network of each sex. This suggests that the genetic evolutionary forces might have played a role in the shaping of the keystones of the nasal microbial community of each sex. The effect might even be mutual, such that interactions of the keystones spurred evolution which in turn reinforced their role as keystone, or the other way around.

Discussion

With advances in sequencing technologies, microbial research is no more restricted to cultivation. Great efforts have since been made to characterize the human microbiome. However, most of the studies rely on 16S rDNA amplicon-based or gene-centric microbial community characterization, which is heavily skewed by microbes that are easily cultivatable or the most researched habitats' residents, such as the human gut microbiome [81–86]. Recently, genome-resolved metagenomics through de novo assembly has transformed our understanding of the microbiome composition, which can meanwhile provide valuable knowledge of individual species for deciphering their biological roles. The human microbiome has a strong niche specialization both within and among individuals [19]. Large reference genome catalogs have been constructed for the human gut and oral microbiome and massively expanded the known species repertoire of the respective habitats [52, 74, 87–89]. Here in this work, we leveraged ultra-deeply sequenced metagenome data from a large cohort of healthy young adults and constructed a non-redundant nasal-associated bacterial MAGs catalog among which about 1/3 species were newly identified. This underscores the uniqueness of the nasal microbiome from other often studied human microbiomes and the power of metagenomic sequencing data. It represents the first endeavor in cataloging the human nasal microbial reference genome and makes a great contribution to the global effort for characterizing the human microbiome. The catalog provides a valuable resource for profiling the nasal microbiome and developing new antibiotics or other pharmaceuticals in future studies. Meanwhile, it makes it possible for uncovering potentially important unknown taxa in

this ecosystem. Further, we characterized nasal microbial composition in the healthy young adults in this so far largest cohort and found that the most abundant bacteria in this cohort largely agreed with former reports in the bacteriome [33, 35, 54]. As for the even less studied mycobiome, it was more evenly distributed compared with the bacterial community, and *Aspergillus* and an unclassified *Malasseziaceae* genus made the most abundant fungi in this cohort. The human microbiome has been implicated in a wide array of diseases, including the nasal microbiome, for which the nasal bacteriome and mycobiome featured some differences in conditions such as polyps and COPD, respectively [12, 23]. Additionally, airway microbiome exhibits seasonal variation which was found to be associated with childhood asthma exacerbations [90]. In this cohort, however, most of the samples were collected in summer and we didn't observe such a phenomenon (Additional file 1: Table S5d).

Respiratory health is of vital importance for human beings. The COVID-19 pandemic has made it unprecedentedly clear. Sex biases have been widely noted in different types of respiratory diseases, including COVID-19. Recently it has been argued that the nasal microbiome might also play a role in the observed disparities between males and females, but unfortunately lacked support and evidence [18]. Previously, Liu et al. identified seven community state types (CSTs) of the nasal bacterial community in a cohort of 86 twin pairs above 50 years old, but found no significant difference in the CST distribution between the two sexes despite of higher microbial loads in the nasal cavity of males [20]. In this work, unsupervised clustering of the nasal microbiota revealed clearly separable patterns between healthy young males and females. This led us to further evaluate the sex differences systematically in this community and uncovered extensive sex-specific features. Interestingly, the gut microbiome also demonstrated sex-specific aging trajectories, where the sex differences were especially evident between premenopausal female adults (approximately below ~50 years old) and age-matched male adults, and gradually diminished after 50 years of age [30]. The disparity regarding sex differences between this study and Liu et al.'s study may be explained by the population's demographic characteristics, especially the age. On the species level, we found females exhibited higher abundances of numerous taxa, including *Staphylococcus aureus*, a well-known opportunistic pathogen found in the respiratory tract, and *Corynebacterium accolens* and *Dolosigranulum pigrum* which can inhibit *Staphylococcus aureus* [25, 32, 33, 38, 39, 91]. The nasal microbiome we identified in this cohort shares similarities with the skin microbiome in their major components [92, 93]. In a study investigating the skin microbiome shifts in healthy children transitioning through puberty, the researchers found that the microbial changes in both bacterial and fungal communities appear to be sex-specific [94]. On the functional level, biosynthesis of proline, glycine, and arginine was significantly lower in males than in females. Smoking has previously been linked to the decrease of these in the lower respiratory tract microbiome [95]. Additionally, glycine is known to decrease the activation of inflammatory cells to avoid the development of chronic inflammation and was shown to improve the status of cystic fibrosis patients in a pilot randomized trial [96, 97]. The pathways enriched in males are primarily linked to the synthesis of purine and pyrimidine, suggesting higher DNA replication activities [98], which may be in connection with the reported higher microbial load in males [20].

The interaction networks exhibited distinctive characteristics between males and females from an ecological perspective. Females featured higher robustness and stronger antagonistic interaction potentials than males. Interestingly, Coyte et al. also concluded that competitive, rather than cooperative, interactions promote the stability of the microbial communities based on ecological theory deduction [99]. The connection of such characteristics with lower susceptibility and severity of RTIs in females compared to males warrants further investigation. There are growing interests in exploring bacterial-fungal interactions, often with several isolates in consideration [100, 101]. While bacterial interaction networks are widely studied and cross-domain interactions are rarely explored, our work integrated the bacteriome and mycobiome and gained a more holistic perspective of the community. Our results suggested that the mycobiome might play an important stabilizing role, in an echo of a former study [46]. Further, through network analysis, we identified sex-specific keystone microbes, which also included formerly unknown taxa, demonstrating the power and necessity of cataloging the community through de novo assembly. Different from comparative analysis in relative abundance, keystone taxa were determined based on their ecological importance through network analysis. The sex-dependent evolutionary characteristics of the keystone bacteria strongly correlated with their role played in the microbial community of each sex, i.e., as a keystone for one sex but not for the other, suggesting a role of the evolutionary forces in the shaping of the keystones, which may have further contributed to the formation of the communities. For instance, the nasopharyngeal commensal *N. sicca_D*, acts as the most influential keystone in females while undergoing positive genetic adaptation in response to niche-specific stress conditions with respect to limited iron, which might have contributed to the formation of the more stable nasal microbial communities against infections. Additionally, Pierce et al. found in their study that fungal species of different environments, including cheese rind, soil, and skin, consistently modulated the availability of iron to bacterial species, alleviating the requirement of *E. coli* for its own siderophore [100]. In our data, the mycobiome appeared to have a higher power than the bacteriome in differentiating the males and females in both composition and networks. It would be interesting in the future to investigate if there is any connection between this observation and fungal modulation of nutrients availability to bacterial species in the high-stress, low-resource environment of the nasal cavity. In another case, *S. maltophilia_L*, a male-prone respiratory infection associated with multidrug-resistant organisms [76, 77], acts as the most influential keystone in males and exhibits strong positive selection for antibiotic resistance-relevant efflux pumps, which may further predispose males to higher susceptibility to infections.

Our study features several limitations. The findings are limited to mathematical modeling and inference, and experimental validation is desired in the future. While MEGAHIT and metaSPAdes both are well-established assemblers and their performances regarding assembly were benchmarked [102], we found that they can additionally affect the compositional profile when derived from mapping to the MAGs they generated. Nevertheless, indirect adjustment and direct stratified analyses confirmed significant sex differences in the nasal microbiota largely consistent with the original results. Besides, interactions between viruses and bacteria widely exist, such as the synergism between influenza virus and *S. pneumoniae* [103–105]. Though females are less contracted with

most types of RTIs, they are indeed more vulnerable to certain respiratory viral pathogens, such as influenza [106]. While we are in short of reliably profiled virome data, antagonistic potentials against influenza as well as other specific pathogens require further investigation.

Conclusion

In conclusion, we leveraged in this work the most advanced techniques in the microbiome research field and applied deep shotgun whole metagenome sequencing, de novo assembly, and network analyses to explore the understudied human nasal microbiome in the largest cohort as of today. Based on that, we constructed a non-redundant nasal bacterial MAGs catalog and revealed extensive sex differences in the nasal microbiome of healthy young adults. The results provide valuable insights into the observed discrepancies between males and females in respiratory tract diseases and will help further our understanding of the microbial roles in pathology and etiology.

Methods

Collection of the nasal microbiome samples

Extensive metadata and different biological samples were collected during physical examination in the 4D-SZ cohort as previously reported [52]. By means of a health questionnaire, no complex disease in personal medical history was found in this young adults cohort. In this study, we collected anterior nares swabs from 1593 individuals of this cohort in 2018 in the city of Shenzhen, with an average age of 29.9 (± 5.13) years old, and sex information obtained for 439 males and 807 females. Demographic characteristics of the participants were provided in Additional file 1: Table S1, and the missing values were shown as NA. The participants without sex information were not included in the analysis of sex differences. The study was approved by the Institutional Review Boards (IRB) at BGI-Shenzhen, and all participants provided written informed consent at enrollment.

The anterior nares samples were self-collected by the volunteers following three steps. First, the sterile swab was moistened with sterile water before use. Then the pre-moistened swab rotated three times around the inside of each nostril with approximately constant pressure. Last, dropping the swab into the 2 ml BGI stabilizing reagent [107] for the preservation of metagenome at room temperature and then stored at -80°C for long-term storage.

DNA extraction, sequencing, and quality control

DNA extraction of the stored samples was performed using the MagPure Stool DNA KF Kit B (MD5115, Magen) [108]. Metagenomic sequencing was performed on the DNBSEQ platform (BGI, Shenzhen, China) [22, 109] with 150 bp of paired-end reads, which generated 854.7 billion pairs of raw reads (on average 536.5 million paired reads per sample, 159.6 million pairs of standard deviation). The metapi pipeline (<https://github.com/ohmeta/metapi>) was used to process the sequencing data. Quality control was first performed with strict standards for filtering and trimming the reads (average Phred quality score ≥ 20 and length ≥ 30) using fastp v0.20.1 [110, 111]. Human reads were then removed using Bowtie2 2.4.2 [112] (human genome GRCh38). In total, 4.2

terabases of high-quality paired-end reads were retained with an average 96.35% host ratio (Additional file 1: Table S2).

Recovery of the bacterial community

A single sample assembly and single sample binning strategy were employed to reconstruct bacterial genomes from the preprocessed data using the metapi pipeline. Specifically, the high-quality reads of each sample were individually assembled by applying MEGAHIT v1.2.9 [113] or SPAdes v3.15.2 [114] (`-meta`). BWA-MEM v0.7.17 [115] with default parameters was then used to map reads back to the contigs, and the contig depth was calculated by `jgi_summarize_bam_contig_depths` [116]. Metagenomic binning was performed with DAS Tool 1.1.2 [117], combining CONCOCT v1.1.0 [118], MaxBin v2.2.7 [119], and MetaBAT2 v 2.15 [116] for each sample individually. CheckM v1.1.3 [120] was used to assess the quality of the MAGs. Bins with $\geq 80\%$ completeness and $\leq 10\%$ contamination were retained for further analysis [121]. All of the MAGs were then together dereplicated by dRep v3.0.1 (`-pa 0.9 -sa 0.99 -nc 0.30 -cm larger -p 25`) [122], in which the primary cluster using MASH with 90% ANI and the secondary cluster using ANImf with 99% ANI, resulting in 974 non-redundant MAGs. The 16S rRNA sequences in the MAGs were searched by Barrnap v0.9 (`-reject 0.01 -evaluate 1e-3`, <https://github.com/tseemann/barrnap>), and tRNA sequences in the MAGs were searched by tRNAscan-SE 2.0.7 [123] with default parameters. Taxonomic classification of the 974 non-redundant MAGs was assigned using GTDB-Tk v1.5.1 [124] to classify workflow with external Genome Taxonomy Database release 95. The phylogenetic tree of the 974 MAGs was built using GTDB-Tk v1.5.1. Genome-wide functional annotation was performed using EggNOG mapper v2.1.3 [125] based on EggNOG v5.0 database [126]. The bacterial biome profile was then generated using CoverM 0.6.1 with genome mode (`-min-covered-fraction 0`) (<https://github.com/wwood/CoverM>) based on the non-redundant nasal bacterial MAGs catalog. Then we filtered the bacteria species with relative abundance greater than $1e-4$ and a prevalence greater than 10% among the 1593 individuals. Finally, 122 bacteria were retained.

Characterization of fungal community composition

High-quality cleaned reads were mapped to a manually curated database using Kraken2 with default parameters to generate the fungal biome profile. This database contained 39,559 species in total, including human genome GRCh38, GTDB r95, fungi, and protists from NCBI. We filtered the fungi species with relative abundance greater than $1e-3$ and a prevalence greater than 10% among the 1593 individuals. Finally, 131 fungi were retained.

Unsupervised clustering

Similarity network fusion (SNF) [127] can construct the fused sample similarity matrix from multiple types of data to represent the characteristics of the samples. The weighted similarity network fusion (WSNF) analysis [45] can integrate multi-biome data and cluster samples into distinct groups using each biome's taxonomic richness as the SNF's weight. A Bray–Curtis similarity matrix of the samples was first created for each biome data (vegan 2.5–7 package). The WSNF pipeline was then used to integrate the similarity

matrices of different into a single similarity network. The respective weights of each biome were assigned based on the richness of the biome. The optimal number of clusters was determined using the eigengap method and the value of K nearest neighbors was set based on the optimal silhouette width. Three clusters were derived with WSNF from the filtered dataset. Other parameters are set as default.

Co-occurrence analysis of microbial interaction

To mitigate the influence of spurious and artifactual correlation, a modified co-occurrence analysis based on ensemble methods was implemented [128] which developed an ensemble approach that can assess nonparametrically for statistical significance while mitigating the compositionality bias by bootstrap and renormalization. Based on the original co-occurrence analysis, Aogáin et al. constructed a microbial network using this ensemble method with some modifications [45]. In this study, we made a further modification of this co-occurrence analysis by replacing some methods in the ensemble considering adjusting compositionality bias and relaxing the normality assumption. First, we implemented COAT (composition-adjusted thresholding) [62] instead of Spearman and Person correlation with default parameters except soft which was set to 0.2. Then we replaced HUGE (High-dimensional Undirected Graph Estimation) [63] with GBLM (generalized boosted linear models) [128] with default parameters except for n lambda which was set to 100. Last, the sign of the correlation depends on COAT and HUGE. The ensemble contained MI (mutual information), Bray–Curtis dissimilarity, COAT, and HUGE. The final interaction score aggregated the normalized absolute edge scores, and the sign was assigned based on COAT and HUGE. The final P value was merged using the weighted Simes test. This analysis was performed on the merged profile which was obtained by integrating and renormalizing the bacteria and fungi profile after filtering. The final edge of network was filtered by merged P value (P value < 0.01).

With the modified ensemble method, we conducted a co-occurrence analysis on the filtered nasal microbiome dataset (as described in the “[Unsupervised clustering](#)” section) for males and females. Filtering out of the low abundance and low prevalence taxa of the microbiome data helped to avoid artificial interactions resulting from random noises though at the expense of sensitivity loss for weak signals. Following the co-occurrence analysis, the nasal microbial interaction networks were established with a threshold of P value lower than $1e - 3$ for males and females.

Stability of microbial co-occurrence network

Natural connectivity is a robustness measure of complex networks [129]. Higher natural connectivity indicates higher network stability. In this study, we performed a random attack by removing randomly selected nodes for 1000 times and assessed normalized natural connectivity for each remaining network (R package pulsar). The number of nodes removed was sequentially increased from 1 to all the nodes. The P value of robustness between male and female networks was calculated following two steps. First, for each attacked network, compare the 1000 natural connectivity between the two sexes with the Wilcoxon rank-sum test. Second, a merged P value was measured using the weighted Simes test, with the number of remaining nodes as the weight.

Selection of keystone taxa in the co-occurrence networks

We selected keystones based on the IVI (Integrated Value of Influence) [70] by influential v2.2.6 R package, which is a novel influential node detection method. It integrates the most important and commonly used network centrality measures in an unbiased way which can capture all topological dimensions of the network and improves the performance of current tools and accurately detects influential nodes. To determine the keystones of each network, we utilized a permutational approach by comparing each robustness attack along the IVI decreasing axis with random attacks (as control). The keystone set was then decided based on a P value < 0.001 calculated from the 1000 permutations. Through analysis, we got 13 and 10 key players of male and female networks, respectively.

pN/pS ratios

SNVs of nonsynonymous and synonymous variants at the gene and genome levels were identified for the keystone taxa using inStrain v1.5.4 [130]. The pN/pS ratio was calculated using the formula $((\text{nonsynonymous SNVs}/\text{nonsynonymous sites})/(\text{synonymous SNVs}/\text{synonymous sites}))$.

BGCs prediction

BGCs (biosynthetic gene clusters) type and location of non-redundant MAGs were predicted using AntiSMASH 6.0.0 [60] (`-cb-knownclusters`). Novel BGCs were defined which did not match the Minimum Information about a Biosynthetic Gene cluster (MIBiG) database.

Statistical analysis and data visualization

Linear discriminant analysis effect size

For discriminant analysis of the microbiome between males and females, the LEfSe was implemented using the webtool available at <http://huttenhower.sph.harvard.edu/galaxy/>. LEfSe uses the Kruskal–Wallis test to identify taxa as well as KEGG pathways whose relative abundances are significantly different between males and females. And then LDA is applied to taxa that meet the significance threshold (0.05) to estimate their effect size. We filtered bacteria and fungi with a prevalence greater than 10%; P value results obtained from the Kruskal–Wallis test were adjusted with the Benjamini–Hochberg method; taxa were considered significantly differentially abundant between males and females at BH-adjusted P value < 0.05 and LDA score > 2.5

PERMANOVA analysis

Univariate ADONIS (permutational multivariate analysis of variance using distance matrices) testing between the observed clusters was performed using R package ‘vegan’ v2.5–7 with 4999 permutations based on the WSNF similarity matrix. The effect of covariant metadata on the microbiome composition was calculated with multivariate ADONIS, which was performed on the microbial dissimilarity matrices, including Bray–Curtis, Jaccard, Euclidean, and Jensen-Shannon divergence distance matrix calculated using relative abundances of microbial species and KEGG pathway, and assess the marginal effects of the terms for each phenotype with 4999 permutations. We also applied

the multivariate ADONIS test to assess which factor had the greatest impact in decreasing the variance explained by the clusters by the changed in the R^2 in three taxa profiles based on the Bray–Curtis distance matrix. P value results within groups were using the Benjamini–Hochberg correction to control multiple testing. Results were considered significant if the BH-adjusted P value < 0.05 .

Anosim (Analysis of similarities)

To analyze whether three seasonal groups are significantly different for nasal microbiota, for which the sample size is rather unbalanced, anosim analysis was performed using the “anosim” function within R package “vegan” v2.5–7 including bacteria, fungi, and merged profile based on the Bray–Curtis distance matrix.

Random forest analysis

To explore the discriminatory potential of sex factors in the nasal microbiome, random forest analysis was performed on bacteria, fungi, and merged species profiles, using the R package “RandomForest” v4.7–1.1. Function “createFolds” in R package “caret” v6.0–94 was used to perform 10 repeats of tenfold cross-validation for each data set. All microbial features in each profile were included in the input. ROC analysis was performed using the “pROC” package v1.18.0.

Diversity analysis

The nasal microbiome α -diversity (within-sample diversity) was calculated using the Shannon index, Simpson index, and Pielou index at the species level (R package “vegan”). The differences between males and females were assessed with the Wilcoxon rank-sum test.

Correlation of IVI and relative abundance

The correlation between IVI and relative abundance of the keystone taxa was measured by Spearman’s correlation.

Visualization

The co-occurrence network was visualized using Cytoscape 3.9.0. The heatmap of the similarity score was drawn by ComplexHeatmap (2.10.0). The boxplot was drawn by ggpubr (2.10.0).

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13059-024-03389-2>.

Additional file 1: Contains Supplementary Table S1 - S10.

Additional file 2: Contains Supplementary Figures S1 - S9.

Additional file 3: Contains the review history.

Acknowledgements

We sincerely thank the support provided by the China National Gene Bank. We thank all the volunteers for their time and contribution.

Peer review information

Andrew Cosgrove was the primary editor of this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Review history

The review history is available as Additional file 3.

Authors' contributions

R.G., T.Z., and H.J. conceived and directed the project. J.W. initiated the overall health project. H.Y., X.X., X.J., K.C., P.C., Y.H., and L.X. contributed to the organization of the cohort, the sample collection, and questionnaire collection. W.L., W.Z., X.T., and Z.J. helped check the phenotypes. H.L. led the DNA extraction and sequencing. R.G. led the bioinformatic analyses. Y.J., M.L., S.L., and Z.S. performed the bioinformatic analyses and prepared figures and texts for the manuscript. Z.Z. and R.G. interpreted the results. Z.Z., R.G., and Y.J. wrote the manuscript. Q.S. contributed to the revision of the manuscript. All authors read and approved the final manuscript.

Funding

Not applicable.

Availability of data and materials

Metagenomic data have been deposited into the Genome Sequence Archive (GSA) with accession number CRA006819 [131] and CNGB Sequence Archive (CNSA) [132] of China National GeneBank DataBase (CNGBdb) [133] with accession number CNP0002487. The source code is available at https://github.com/Leonn369/nasal_PJ/ [134] and Zenodo [135], under a GPL 3.0 license. The other data supporting the findings of this study are available within the paper and additional files.

Declarations

Ethics approval and consent to participate

The study was approved by the Institutional Review Board (IRB) of BGI-Shenzhen. (Nos. BGI-IRB19121). All subjects have given written informed consent for participation and publication. All experimental methods comply with the Helsinki Declaration.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹BGI Research, Shenzhen 518083, China. ²Shenzhen Key Laboratory of Human Commensal Microorganisms and Health Research, BGI Research, Shenzhen 518083, China. ³College of Life Sciences, University of Chinese Academy of Sciences, Beijing 100049, China. ⁴Shenzhen Engineering Laboratory of Detection and Intervention of Human Intestinal Microbiome, BGI Research, Shenzhen 518083, China. ⁵Qingdao-Europe Advanced Institute for Life Sciences, BGI Research, Qingdao 266555, China. ⁶Department of Statistical Sciences, University of Toronto, 700 University Ave, Toronto, ON M5G 1Z5, Canada. ⁷China National Genebank, BGI Research, Shenzhen 518210, China. ⁸James D. Watson Institute of Genome Sciences, Hangzhou 310013, China. ⁹School of Life Sciences, Fudan University, Shanghai 200433, China. ¹⁰Greater Bay Area Institute of Precision Medicine, Guangzhou 511458, China. ¹¹BGI Research, Wuhan 430074, China.

Received: 27 August 2022 Accepted: 6 September 2024

Published online: 08 October 2024

References

1. Ferkol T, Schraufnagel D. The global burden of respiratory disease. *Ann Am Thorac Soc*. 2014;11:404–6.
2. Vos T, Lim SS, Abbafati C, Abbas KM, Abbasi M, Abbasifard M, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet*. 2020;396:1204–22.
3. Viegli G, Maio S, Fasola S, Baldacci S. Global burden of chronic respiratory diseases. *J Aerosol Med Pulm Drug Deliv*. 2020;33:171–7.
4. Jin X, Ren J, Li R, Gao Y, Zhang H, Li J, et al. Global burden of upper respiratory infections in 204 countries and territories, from 1990 to 2019. *eClinicalMedicine*. 2021;37:100986.
5. Falagas ME, Mourtzoukou EG, Vardakas KZ. Sex differences in the incidence and severity of respiratory tract infections. *Respir Med*. 2007;101:1845–63.
6. Jacobsen H, Klein SL. Sex differences in immunity to viral infections. *Front Immunol*. 2021;12: 720952.
7. Vahidy FS, Pan AP, Ahnstedt H, Munshi Y, Choi HA, Tiruneh Y, et al. Sex differences in susceptibility, severity, and outcomes of coronavirus disease 2019: cross-sectional analysis from a diverse US metropolitan area. *PLoS One*. 2021;16: e0245556.
8. Scully EP, Haverfield J, Ursin RL, Tannenbaum C, Klein SL. Considering how biological sex impacts immune responses and COVID-19 outcomes. *Nat Rev Immunol*. 2020;20:442–7.
9. Klein SL, Dhakal S, Ursin RL, Deshpande S, Sandberg K, Mauvais-Jarvis F. Biological sex impacts COVID-19 outcomes. *PLoS Pathog*. 2020;16:e1008570.

10. Fazlollahi M, Lee TD, Andrade J, Oguntuyo K, Chun Y, Grishina G, et al. The nasal microbiome in asthma. *J Allergy Clin Immunol*. 2018;142:834–843.e2.
11. Mahdavinia M, Keshavarzian A, Tobin MC, Landay AL, Schleimer RP. A comprehensive review of the nasal microbiome in chronic rhinosinusitis (CRS). *Clin Exp Allergy*. 2016;46:21–41.
12. Gan W, Zhang H, Yang F, Liu S, Liu F, Meng J. The influence of nasal microbiome diversity and inflammatory patterns on the prognosis of nasal polyps. *Sci Rep*. 2021;11:6364.
13. Wu BG, Sulaiman I, Wang J, Shen N, Clemente JC, Li Y, et al. Severe obstructive sleep apnea is associated with alterations in the nasal microbiome and an increase in inflammation. *Am J Respir Crit Care Med*. 2019;199:99–109.
14. Rhee RL, Lu J, Bittinger K, Lee J, Mattei LM, Sreih AG, et al. Dynamic changes in the nasal microbiome associated with disease activity in patients with granulomatosis with polyangiitis. *Arthritis Rheumatol*. 2021;73:1703–12.
15. Ramakrishnan VR, Frank DN. Microbiome in patients with upper airway disease: moving from taxonomic findings to mechanisms and causality. *J Allergy Clin Immunol*. 2018;142:73–5.
16. Kumpitsch C, Koskinen K, Schöpf V, Moissl-Eichinger C. The microbiome of the upper respiratory tract in health and disease. *BMC Biol*. 2019;17:87.
17. Dickson RP, Erb-Downward JR, Martinez FJ, Huffnagle GB. The microbiome and the respiratory tract. *Annu Rev Physiol*. 2016;78:481–504.
18. Shah V. Letter to the editor: microbiota in the respiratory system—a possible explanation to age and sex variability in susceptibility to SARS-CoV-2. *Microbiol Insights*. 2021;14:117863612098860.
19. The Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. *Nature*. 2012;486:207–14.
20. Liu CM, Price LB, Hungate BA, Abraham AG, Larsen LA, Christensen K, et al. *Staphylococcus aureus* and the ecology of the nasal microbiome. *Sci Adv*. 2015;1: e1400216.
21. Jung WH, Croll D, Cho JH, Kim YR, Lee YW. Analysis of the nasal vestibule mycobiome in patients with allergic rhinitis. *Mycoses*. 2015;58:167–72.
22. Wagner Mackenzie B, Chang K, Zoing M, Jain R, Hoggard M, Biswas K, et al. Longitudinal study of the bacterial and fungal microbiota in the human sinuses reveals seasonal and annual changes in diversity. *Sci Rep*. 2019;9:17416.
23. Alvarez Baumgartner M, Li C, Kuntz TM, Nurhussien L, Synn AJ, Sun WY, et al. Differences of the nasal microbiome and mycobiome by clinical characteristics of COPD patients. *Chronic Obstr Pulm Dis J COPD Found*. 2022;9:309–24.
24. De Steenhuijsen Piters WAA, Binkowska J, Bogaert D. Early life microbiota and respiratory tract infections. *Cell Host Microbe*. 2020;28:223–32.
25. Yan M, Pamp SJ, Fukuyama J, Hwang PH, Cho D-Y, Holmes S, et al. Nasal microenvironments and interspecific interactions influence nasal microbiota complexity and *S. aureus* carriage. *Cell Host Microbe*. 2013;14:631–40.
26. Earl JP, Adappa ND, Krol J, Bhat AS, Balashov S, Ehrlich RL, et al. Species-level bacterial community profiling of the healthy sinonasal microbiome using Pacific Biosciences sequencing of full-length 16S rRNA genes. *Microbiome*. 2018;6:190.
27. Dai W, Wang H, Zhou Q, Li D, Feng X, Yang Z, et al. An integrated respiratory microbial gene catalogue to better understand the microbial aetiology of *Mycoplasma pneumoniae* pneumonia. *GigaScience*. 2019;8: gjz093.
28. Kaul D, Rathnasinghe R, Ferrer M, Tan GS, Barrera A, Pickett BE, et al. Microbiome disturbance and resilience dynamics of the upper respiratory tract during influenza A virus infection. *Nat Commun*. 2020;11:2537.
29. De La Cuesta-Zuluaga J, Kelley ST, Chen Y, Escobar JS, Mueller NT, Ley RE, et al. Age- and sex-dependent patterns of gut microbial diversity in human adults. *mSystems*. 2019;4:e00261-19.
30. Zhang X, Zhong H, Li Y, Shi Z, Ren H, Zhang Z, et al. Sex- and age-related trajectories of the adult human gut microbiota shared across populations of different ethnicities. *Nat Aging*. 2021;1:87–100.
31. Sinha T, Vich Vila A, Garmaeva S, Jankipersadsing SA, Imhann F, Collij V, et al. Analysis of 1135 gut metagenomes identifies sex-specific resistome profiles. *Gut Microbes*. 2019;10:358–66.
32. Dai W, Chen J, Xiong J. Concept of microbial gatekeepers: positive guys? *Appl Microbiol Biotechnol*. 2019;103:633–41.
33. Man WH, de Steenhuijsen Piters WAA, Bogaert D. The microbiota of the respiratory tract: gatekeeper to respiratory health. *Nat Rev Microbiol*. 2017;15:259–70.
34. Krismer B, Liebeke M, Janek D, Nega M, Rautenberg M, Hornig G, et al. Nutrient limitation governs *Staphylococcus aureus* metabolism and niche adaptation in the human nose. *PLoS Pathog*. 2014;10: e1003862.
35. Clark SE. Commensal bacteria in the upper respiratory tract regulate susceptibility to infection. *Curr Opin Immunol*. 2020;66:42–9.
36. Hardy BL, Merrell DS. Friend or foe: interbacterial competition in the nasal cavity. *J Bacteriol*. 2021;203:e00480.
37. Stubbendieck RM, Dissanayake E, Burnham PM, Zelasko SE, Temkin MI, Wisdorf SS, et al. *Rothia* from the Human nose inhibit *Moraxella catarrhalis* colonization with a secreted peptidoglycan endopeptidase. *mBio*. 2023;14:e00464-23.
38. Brugger SD, Eslami SM, Pettigrew MM, Escapa IF, Henke MT, Kong Y, et al. *Dolosigranulum pigrum* cooperation and competition in human nasal microbiota. *mSphere*. 2020;5:e00852-20.
39. De Boeck I, Wittouck S, Martens K, Spacova I, Cauwenberghs E, Allonsius CN, et al. The nasal mutualist *Dolosigranulum pigrum* AMBR11 supports homeostasis via multiple mechanisms. *iScience*. 2021;24:102978.
40. Zipperer A, Konnerth MC, Laux C, Berscheid A, Janek D, Weidenmaier C, et al. Human commensals producing a novel antibiotic impair pathogen colonization. *Nature*. 2016;535:511–6.
41. Donia MS, Cimermancic P, Schulze CJ, Wieland Brown LC, Martin J, Mitreva M, et al. A systematic analysis of biosynthetic gene clusters in the human microbiome reveals a common family of antibiotics. *Cell*. 2014;158:1402–14.
42. Donia MS, Fischbach MA. Small molecules from the human microbiota. *Science*. 2015;349: 1254766.
43. Iwase T, Uehara Y, Shinji H, Tajima A, Seo H, Takada K, et al. *Staphylococcus epidermidis* Esp inhibits *Staphylococcus aureus* biofilm formation and nasal colonization. *Nature*. 2010;465:346–9.
44. Janek D, Zipperer A, Kulik A, Krismer B, Peschel A. High frequency and diversity of antimicrobial activities produced by nasal *Staphylococcus* strains against bacterial competitors. *PLoS Pathog*. 2016;12: e1005812.

45. Mac Aogáin M, Narayana JK, Tiew PY, Ali NABM, Yong VFL, Jaggi TK, et al. Integrative microbiomics in bronchiectasis exacerbations. *Nat Med*. 2021;27:688–99.
46. Tipton L, Müller CL, Kurtz ZD, Huang L, Kleeerup E, Morris A, et al. Fungi stabilize connectivity in the lung and skin microbial ecosystems. *Microbiome*. 2018;6:12.
47. Banerjee S, Schlaeppli K, van der Heijden MGA. Keystone taxa as drivers of microbiome structure and functioning. *Nat Rev Microbiol*. 2018;16:567–76.
48. Zamkova T, Foster JS, De Crécy-Lagard V, Conesa A. A network approach to elucidate and prioritize microbial dark matter in microbial communities. *ISME J*. 2021;15:228–44.
49. Jie Z, Liang S, Ding Q, Li F, Tang S, Wang D, et al. A transomic cohort as a reference point for promoting a healthy human gut microbiome. *Med Microecol*. 2021;8: 100039.
50. Jie Z, Chen C, Hao L, Li F, Song L, Zhang X, et al. Life history recorded in the vagino-cervical microbiome along with multi-omes. *Genomics Proteomics Bioinformatics*. 2022;20:304–21.
51. Liu X, Tong X, Zou Y, Lin X, Zhao H, Tian L, et al. Mendelian randomization analyses support causal relationships between blood metabolites and the gut microbiome. *Nat Genet*. 2022;54:52–61.
52. Zhu J, Tian L, Chen P, Han M, Song L, Tong X, et al. Over 50,000 metagenomically assembled draft genomes for the human oral microbiome reveal new taxa. *Genomics Proteomics Bioinformatics*. 2022;20:246–59.
53. Chen C, Hao L, Zhang Z, Tian L, Zhang X, Zhu J, et al. Cervicovaginal microbiome dynamics after taking oral probiotics. *J Genet Genomics*. 2021;48:716–26.
54. Dimitri-Pinheiro S, Soares R, Barata P. The microbiome of the nose—friend or foe? *Allergy Rhinol*. 2020;11:215265672091160.
55. Jacob-Dubuisson F, Mechaly A, Betton J-M, Antoine R. Structural insights into the signalling mechanisms of two-component systems. *Nat Rev Microbiol*. 2018;16:585–93.
56. Qin X, Zhang K, Fan Y, Fang H, Nie Y, Wu X-L. The bacterial MtrAB two-component system regulates the cell wall homeostasis responding to environmental alkaline stress. *Microbiol Spectr*. 2022;10:e02311–22.
57. Moraes JJ, Stipp RN, Harth-Chu EN, Camargo TM, Höfling JF, Mattos-Graner RO. Two-component system VicRK regulates functions associated with establishment of *Streptococcus sanguinis* in biofilms. *Infect Immun*. 2014;82:4941–51.
58. Sugimoto Y, Camacho FR, Wang S, Chankhamjon P, Odabas A, Biswas A, et al. A metagenomic strategy for harnessing the chemical repertoire of the human microbiome. *Science*. 2019;366: eaax9176.
59. Milshcheyn A, Colosimo DA, Brady SF. Accessing bioactive natural products from the human microbiome. *Cell Host Microbe*. 2018;23:725–36.
60. Blin K, Shaw S, Kloosterman AM, Charlop-Powers Z, van Wezel GP, Medema MH, et al. antiSMASH 6.0: improving cluster detection and comparison capabilities. *Nucleic Acids Res*. 2021;49:W29–35.
61. De Boeck I, Van Den Broeck MFL, Allonsius CN, Spacova I, Wittouck S, Martens K, et al. Lactobacilli have a niche in the human nose. *Cell Rep*. 2020;31:107674.
62. Cao Y, Lin W, Li H. Large covariance estimation for compositional data via composition-adjusted thresholding. *J Am Stat Assoc*. 2019;114:759–72.
63. Zhao T, Liu H, Roeder K, Lafferty J, Wasserman L. The huge package for high-dimensional undirected graph estimation in R. *J Mach Learn Res*. 2012;13:1059–62. <https://www.scopus.com/record/display.uri?eid=2-s2.0-84860650411&origin=inward&txGid=ca8d50c77691a1178b299ca1e24dea73>.
64. Matchado MS, Lauber M, Reitmeier S, Kacprowski T, Baumbach J, Haller D, et al. Network analysis methods for studying microbial communities: a mini review. *Comput Struct Biotechnol J*. 2021;19:2687–98.
65. Jun W, Barahona M, Yue-Jin T, Hong-Zhong D. Natural connectivity of complex networks. *Chin Phys Lett*. 2010;27: 078902.
66. Zhang X-K, Wu J, Tan Y-J, Deng H-Z, Li Y. Structural robustness of weighted complex networks based on natural connectivity. *Chin Phys Lett*. 2013;30: 108901.
67. Peng G, Wu J. Optimal network topology for structural robustness based on natural connectivity. *Phys Stat Mech Its Appl*. 2016;443:212–20.
68. Otsuka M, Tsugawa S. Robustness of network attack strategies against node sampling and link errors. *PLoS One*. 2019;14: e0221885.
69. Layeghifard M, Hwang DM, Guttman DS. Disentangling interactions in the microbiome: a network Perspective. *Trends Microbiol*. 2017;25:217–28.
70. Salavaty A, Ramialison M, Currie PD. Integrated value of influence: an integrative method for the identification of the most influential nodes within networks. *Patterns*. 2020;1: 100052.
71. Schloissnig S, Arumugam M, Sunagawa S, Mitreva M, Tap J, Zhu A, et al. Genomic variation landscape of the human gut microbiome. *Nature*. 2013;493:45–50.
72. Garud NR, Pollard KS. Population genetics in the human microbiome. *Trends Genet*. 2020;36:53–67.
73. Stubbendieck RM, May DS, Chevrette MG, Temkin MI, Wendt-Pienkowski E, Cagnazzo J, et al. Competition among nasal bacteria suggests a role for siderophore-mediated interactions in shaping the human nasal microbiota. *Appl Environ Microbiol*. 2019;85:e02406–18.
74. Almeida A, Mitchell AL, Boland M, Forster SC, Gloor GB, Tarkowska A, et al. A new genomic blueprint of the human gut microbiota. *Nature*. 2019;568:499–504.
75. Marri PR, Paniscus M, Weyand NJ, Rendón MA, Calton CM, Hernández DR, et al. Genome sequencing reveals widespread virulence gene exchange among human *Neisseria* species. *PLoS One*. 2010;5: e11835.
76. Brooke JS. *Stenotrophomonas maltophilia*: an emerging global opportunistic pathogen. *Clin Microbiol Rev*. 2012;25:2–41.
77. Chawla K. *Stenotrophomonas maltophilia* in lower respiratory tract infections. *J Clin Diagn Res*. 2014. <https://doi.org/10.7860/JCDR/2014/10780.5320>.
78. Lin Y-T, Huang Y-W, Liou R-S, Chang Y-C, Yang T-C. MacABCsm, an ABC-type tripartite efflux pump of *Stenotrophomonas maltophilia* involved in drug resistance, oxidative and envelope stress tolerances and biofilm formation. *J Antimicrob Chemother*. 2014;69:3221–6.

79. Crow A, Greene NP, Kaplan E, Koronakis V. Structure and mechanotransmission mechanism of the MacB ABC transporter superfamily. *Proc Natl Acad Sci*. 2017;114:12572–7.
80. Greene NP, Kaplan E, Crow A, Koronakis V. Antibiotic resistance mediated by the MacB ABC transporter family: a structural and functional perspective. *Front Microbiol*. 2018;9: 950.
81. Woyke T. Beyond the census of human gut dwellers. *Nat Rev Microbiol*. 2019;17:401–401.
82. MetaHIT Consortium, Li J, Jia H, Cai X, Zhong H, Feng Q, et al. An integrated catalog of reference genes in the human gut microbiome. *Nat Biotechnol*. 2014;32:834–41.
83. Segata N, Waldron L, Ballarini A, Narasimhan V, Jousson O, Huttenhower C. Metagenomic microbial community profiling using unique clade-specific marker genes. *Nat Methods*. 2012;9:811–4.
84. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: high-resolution sample inference from illumina amplicon data. *Nat Methods*. 2016;13:581–3.
85. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol*. 2019;37:852–7.
86. Xie H, Guo R, Zhong H, Feng Q, Lan Z, Qin B, et al. Shotgun metagenomics of 250 adult twins reveals genetic and environmental impacts on the gut microbiome. *Cell Syst*. 2016;3:572–584.e3.
87. Nayfach S, Shi ZJ, Seshadri R, Pollard KS, Kyrpides NC. New insights from uncultivated genomes of the global human gut microbiome. *Nature*. 2019;568:505–10.
88. Pasolli E, Asnicar F, Manara S, Zolfo M, Karcher N, Armanini F, et al. Extensive unexplored human microbiome diversity revealed by over 150,000 genomes from metagenomes spanning age, geography, and lifestyle. *Cell*. 2019;176:649–662.e20.
89. Almeida A, Nayfach S, Boland M, Strozzi F, Beracochea M, Shi ZJ, et al. A unified catalog of 204,938 reference genomes from the human gut microbiome. *Nat Biotechnol*. 2021;39:105–14.
90. McCauley KE, Flynn K, Calatroni A, DiMassa V, LaMere B, Fadrosh DW, et al. Seasonal airway microbiome and transcriptome interactions promote childhood asthma exacerbations. *J Allergy Clin Immunol*. 2022;150:204–13.
91. Krüger B, Weidenmaier C, Zipperer A, Peschel A. The commensal lifestyle of *Staphylococcus aureus* and its interactions with the nasal microbiota. *Nat Rev Microbiol*. 2017;15:675–87.
92. Leung MHY, Chan KCK, Lee PKH. Skin fungal community and its correlation with bacterial community of urban Chinese individuals. *Microbiome*. 2016;4:46.
93. Flowers L, Grice EA. The skin microbiota: balancing risk and reward. *Cell Host Microbe*. 2020;28:190–200.
94. Park J, Schwarzt NH, Jo J-H, Zhang Z, Pillai V, Phang S, et al. Shifts in the skin bacterial and fungal communities of healthy children transitioning through puberty. *J Invest Dermatol*. 2022;142:212–9.
95. Li K, Chen Z, Huang Y, Zhang R, Luan X, Lei T, et al. Dysbiosis of lower respiratory tract microbiome are associated with inflammation and microbial function variety. *Respir Res*. 2019;20:272.
96. Aguayo-Cerón KA, Sánchez-Muñoz F, Gutierrez-Rojas RA, Acevedo-Villavicencio LN, Flores-Zarate AV, Huang F, et al. Glycine: the smallest anti-inflammatory micronutrient. *Int J Mol Sci*. 2023;24: 11236.
97. Vargas MH, Del-Razo-Rodríguez R, López-García A, Lezana-Fernández JL, Chávez J, Furuya MEY, et al. Effect of oral glycine on the clinical, spirometric and inflammatory status in subjects with cystic fibrosis: a pilot randomized trial. *BMC Pulm Med*. 2017;17:206.
98. Fuhrman JA, Azam F. Thymidine incorporation as a measure of heterotrophic bacterioplankton production in marine surface waters: evaluation and field results. *Mar Biol*. 1982;66:109–20.
99. Coyte KZ, Schluter J, Foster KR. The ecology of the microbiome: networks, competition, and stability. *Science*. 2015;350:663–6.
100. Pierce EC, Morin M, Little JC, Liu RB, Tannous J, Keller NP, et al. Bacterial–fungal interactions revealed by genome-wide analysis of bacterial mutant fitness. *Nat Microbiol*. 2020;6:87–102.
101. Li H, Goh BN, Teh WK, Jiang Z, Goh JPZ, Goh A, et al. Skin commensal *Malassezia globosa* secreted protease attenuates *Staphylococcus aureus* biofilm formation. *J Invest Dermatol*. 2018;138:1137–45.
102. Van Der Walt AJ, Van Goethem MW, Ramond J-B, Makhalanyane TP, Reva O, Cowan DA. Assembling metagenomes, one community at a time. *BMC Genomics*. 2017;18:521.
103. McCullers JA. Insights into the interaction between influenza virus and pneumococcus. *Clin Microbiol Rev*. 2006;19:571–82.
104. Bosch AATM, Biesbroek G, Trzcinski K, Sanders EAM, Bogaert D. Viral and bacterial interactions in the upper respiratory tract. *PLoS Pathog*. 2013;9: e1003057.
105. Korten I, Ramsey K, Mika M, Usemann J, Frey U, Hilty M, et al. Nasal microbiota and respiratory tract infections: the role of viral detection. *Am J Respir Crit Care Med*. 2019;199:919–22.
106. Klein SL, Hodgson A, Robinson DP. Mechanisms of sex disparities in influenza pathogenesis. *J Leukoc Biol*. 2012;92:67–73.
107. Han M, Hao L, Lin Y, Li F, Wang J, Yang H, et al. A novel affordable reagent for room temperature storage and transport of fecal samples for metagenomic analyses. *Microbiome*. 2018;6:43.
108. Yang F, Sun J, Luo H, Ren H, Zhou H, Lin Y, et al. Assessment of fecal DNA extraction protocols for metagenomic studies. *GigaScience*. 2020;9: g1aa071.
109. Li Q, Zhao X, Zhang W, Wang L, Wang J, Xu D, et al. Reliable multiplex sequencing with rare index mis-assignment on DNB-based NGS platform. *BMC Genomics*. 2019;20:215.
110. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*. 2018;34:884–90.
111. Chen S. Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. *iMeta*. 2023;2:e107.
112. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9:357–9.
113. Li D, Liu C-M, Luo R, Sadakane K, Lam T-W. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct *de Bruijn* graph. *Bioinformatics*. 2015;31:1674–6.
114. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. metaSPAdes: a new versatile metagenomic assembler. *Genome Res*. 2017;27:824–34.

115. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25:1754–60.
116. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, et al. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ*. 2019;7: e7359.
117. Sieber CMK, Probst AJ, Sharrar A, Thomas BC, Hess M, Tringe SG, et al. Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. *Nat Microbiol*. 2018;3:836–43.
118. Alneberg J, Bjarnason BS, De Bruijn I, Schirmer M, Quick J, Ijaz UZ, et al. Binning metagenomic contigs by coverage and composition. *Nat Methods*. 2014;11:1144–6.
119. Wu Y-W, Simmons BA, Singer SW. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics*. 2016;32:605–7.
120. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res*. 2015;25:1043–55.
121. Stewart RD, Auffret MD, Warr A, Wiser AH, Press MO, Langford KW, et al. Assembly of 913 microbial genomes from metagenomic sequencing of the cow rumen. *Nat Commun*. 2018;9:870.
122. Olm MR, Brown CT, Brooks B, Banfield JF. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J*. 2017;11:2864–8.
123. Chan PP, Lowe TM. tRNAscan-SE: searching for tRNA genes in genomic sequences. In: Kollmar M, editor. *Gene prediction*. New York: Springer New York; 2019. p. 1–14.
124. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. *Bioinformatics*. 2020;36:1925–7.
125. Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J. eggNOG-mapper v2: functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol Biol Evol*. 2021;38:5825–9.
126. Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, et al. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res*. 2019;47:D309–14.
127. Wang B, Mezlini AM, Demir F, Fiume M, Tu Z, Brudno M, et al. Similarity network fusion for aggregating data types on a genomic scale. *Nat Methods*. 2014;11:333–7.
128. Faust K, Sathirapongsasuti JF, Izard J, Segata N, Gevers D, Raes J, et al. Microbial co-occurrence relationships in the human microbiome. *PLoS Comput Biol*. 2012;8: e1002606.
129. Morone F, Makse HA. Influence maximization in complex networks through optimal percolation. *Nature*. 2015;524:65–8.
130. Olm MR, Crits-Christoph A, Bouma-Gregson K, Firek BA, Morowitz MJ, Banfield JF. inStrain profiles population microdiversity from metagenomic data and sensitively detects shared microbial strains. *Nat Biotechnol*. 2021;39:727–36.
131. Ju Y, Zhang Z, Liu M, Lin S, Sun Q, Song Z, et al. Integrated large-scale metagenome assembly and multi-kingdom network analyses identify sex differences in the nasal microbiome. *Datasets*. *Genome Sequence Archive*. 2024. <https://ngdc.cncb.ac.cn/gsa/search?searchTerm=CRA006819>.
132. Guo X, Chen F, Gao F, Li L, Liu K, You L, et al. CNSA: a data repository for archiving omics data. *Database*. 2020;2020:baaa055.
133. Chen FZ, You LJ, Yang F, Wang LN, Guo XQ, Gao F, et al. CNGBdb: China National GeneBank DataBase. *Yi Chuan*. 2020;42:799–809.
134. Ju Y, Zhang Z, Liu M, Lin S, Sun Q, Song Z, et al. Integrated large-scale metagenome assembly and multi-kingdom network analyses identify sex differences in the nasal microbiome. *Github*. 2024. https://github.com/Leonn369/nasal_PJ/.
135. Ju Y, Zhang Z, Liu M, Lin S, Sun Q, Song Z, et al. Integrated large-scale metagenome assembly and multi-kingdom network analyses identify sex differences in the nasal microbiome. *Zenodo*. 2024. <https://doi.org/10.5281/zenodo.13370585>.

Publisher's Note

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