Whole-Genome Shotgun Sequencing for Nasopharyngeal Microbiome in Pre-school Children With Recurrent Wheezing

Yuping Song[1†], Jinpao Hou[2†], Jamie Sui Lam Kwok[2†], Haoyi Weng[3], Man Fung Tang[1,4], Maggie Haitian Wang[3], Agnes Sze Yin Leung[1], Kin Pong Tao[1,5], Gary Wing Kin Wong[1], Renee Wan Yi Chan[1,4,5], Stephen Kwok Wing Tsui[2*] and Ting Fan Leung[1,4,5*]

[1] Department of Pediatrics, Prince of Wales Hospital, The Chinese University of Hong Kong, Sha Tin, Hong Kong SAR, China, [2] School of Biomedical Sciences, The Chinese University of Hong Kong, Sha Tin, Hong Kong SAR, China, [3] Jockey Club School of Public Health and Primary Care, Prince of Wales Hospital, The Chinese University of Hong Kong, Sha Tin, Hong Kong SAR, China, [4] Hong Kong Hub of Paediatric Excellence, The Chinese University of Hong Kong, Sha Tin, Hong Kong SAR, China, [5] The Chinese University of Hong Kong-University Medical Center Utrecht Joint Research Laboratory of Respiratory Virus and Immunobiology, The Chinese University of Hong Kong, Sha Tin, Hong Kong SAR, China

Microbiome mediates early life immune deviation in asthma development. Recurrent wheeze (RW) in pre-school years is a risk factor for asthma diagnosis in school-age children. Dysbiosis exists in asthmatic airways, while its origin in pre-school years and relationship to RW is not clearly defined. This study investigated metagenomics of nasopharyngeal microbiome in pre-school children with RW. We applied whole-genome shotgun sequencing and human rhinovirus (HRV) detection on nasopharyngeal samples collected from three groups of pre-school children: (i) RW group: 16 children at-risk for asthma who were hospitalized for RW, (ii) inpatient control (IC): 18 subjects admitted for upper respiratory infection, and (iii) community control (CC): 36 children without respiratory syndromes. Sequence reads were analyzed by MetaPhlAn2 and HUMAnN2 algorithm for taxonomic and functional identification. Linear discriminant analysis effect size (LEfSe) analysis was used to identify discriminative features. We identified that *Moraxella catarrhalis* and *Dolosigranulum pigrum* were predominant species in nasopharynx. RW had lower alpha diversity (Shannon diversity index) than CC (0.48 vs. 1.07; $P_{adj} = 0.039$), characterized by predominant Proteobacteria. LEfSe analysis revealed *D. pigrum* was the only discriminative species across groups (LDA = 5.57, $P = 0.002$), with its relative abundance in RW, IC, and CC being 9.6, 14.2, and 37.3%, respectively ($P < 0.05$). LEfSe identified five (ribo)nucleotides biosynthesis pathways to be group discriminating. Adjusting for HRV status, pre-school children with RW have lower nasopharyngeal biodiversity, which is associated with Proteobacteria predominance and lower abundance of *D. pigrum*. Along with discriminative pathways found in RW and CC, these microbial biomarkers help to understand RW pathogenesis.

Keywords: dysbiosis, metagenomics, nasopharynx, pre-schooler, wheeze

Abbreviations: CC, community control; IC, inpatient control; RW, recurrent wheeze; HRV, human rhinovirus; CCC, community control negative for HRV; CCV, community control with HRV; ICC, inpatient control negative for HRV; ICV, inpatient control with HRV; RWV, recurrent wheezing with HRV; FNPS, flocked nasopharyngeal swab; NPA, nasopharyngeal aspirate; NPM, nasopharyngeal microbiome; SDI, Shannon diversity index; URTI, upper respiratory tract infection; WGS, whole genome shotgun.

INTRODUCTION

Pre-school age is a critical period for the occurrence of recurrent wheeze (RW) which is a strong risk factor for asthma. Mucosal microbiota serves as a mediator between host immunity and the environment, and microbial dysbiosis was found in many immune-mediated diseases (Kozik and Huang, 2019). Nonetheless, the association between allergies and airway microbiome was underexplored due to limitations with low microbial biomass (Quince et al., 2017). Upper airway samples were widely used in pediatric studies owing to the difficulty in obtaining lower airway samples.

Upper airway dysbiosis is longitudinally related to recurrent wheezing and asthma. Seminal work from Bisgaard et al. (2007) first noticed that asymptomatic hypopharyngeal colonization of pathogenic bacteria including *Streptococcus, Haemophilus,* and *Moraxella* in neonates was related to asthma risk at 5 years old. Similarly, enrichment of *Moraxella, Haemophilus,* or *Streptococcus*, in the nasal microbiota from hospitalized bronchiolitis infants was reported being longitudinally related to increased risk for recurrent wheezing (Mansbach et al., 2020; Zhang et al., 2020). Nasal dominance of *Moraxella* and *Staphylococcus* in school-age asthmatic children increased exacerbation risk (McCauley et al., 2019). Nonetheless, there is limited evidence on airway microbiome during acute episodes of RW or asthma exacerbations, particularly in the pre-school age group. Fazlollahi et al. (2018) reported the asthma exacerbation in adult patients is associated with nasal enrichment of *Prevotella, Alkanindiges,* and *Gardnerella*. Robinson et al. (2019) conducted pioneering investigation on pre-school children referred for bronchoscope for severe RW, and reported that the dominance of *Moraxella* in bronchoalveolar lavage (BAL) is related to neutrophilic pulmonary inflammation. Nevertheless, BAL microbiome profile was not related to clinical wheezing phenotypes in this study, though a trend of episodic viral wheeze and *Moraxella*-profile was noticed.

Regarding the technological aspect, metagenomics studies are more powerful than 16S rRNA sequencing to delineate microbiome due to their higher resolution and functional profiling capability. A whole genome shotgun (WGS) study reported enriched *Streptococcus pneumoniae* and sphingolipid metabolism pathway in nasopharyngeal aspirate (NPA) to be associated with severe bronchiolitis in infants (Stewart et al., 2017).

We hypothesized that specific markers of nasopharyngeal microbiome (NPM) from pre-school wheezers at risk of developing asthma can discriminate them from non-wheeze inpatient and community controls (CCs). Besides, we hypothesized that human rhinovirus (HRV) infection might influence the association between NPM and pre-school wheezing. This study applied WGS to characterize the NPM in pre-school children hospitalized for RW who were positive for asthma predictive index (API), and explore the differences in NPM compositions and functional capabilities between RW cases and control groups and the possible effect of HRV infection on such associations between NPM and pre-school wheezing.

MATERIALS AND METHODS

Subjects

Three groups of Chinese pre-school children aged 2–5 years old were recruited: (a) children hospitalized for RW who were positive for stringent-asthma predictive index (S-API) (Castro-Rodríguez et al., 2000); (b) inpatient control (IC) children hospitalized for upper respiratory tract infection (URTI) with no respiratory distress, normal auscultation and chest radiograph, and no history of wheeze and asthma. These IC subjects were matched for gender and admission time (hospitalized within 1 week from the RW case recruitment) with RW patients; (c) CC children from nurseries and kindergartens who participated in our influenza surveillance study (Leung et al., 2017). These subjects were free from URTI for $\geq$ 4 weeks and had no history of asthma, allergic rhinitis, eczema, and food allergy. All subjects must not receive antibiotics within 4 weeks before study. Chinese University of Hong Kong – New Territories East Cluster Clinical Research Ethics Committee approved this study.

Following informed consent, NPA was collected from hospitalized subjects (RW and IC) upon admission; flocked nasopharyngeal swab (FNPS; Copan Diagnostics, Corona, CA, United States) was obtained from CC subjects at the kindergartens. All samples were placed immediately at 4°C in 2 ml viral transport medium containing 0.1 mg/ml gentamicin, 500 IU/ml penicillin, 500 IU/ml streptomycin, and 2.5 μg/ml fungizon. As HRV was the most important trigger for RW (Leung et al., 2010), 50 μl of each sample was aliquoted for HRV detection by molecular assays according to our published method (Mak et al., 2011). The remaining portion was stored at -80°C until microbiome analysis in one batch.
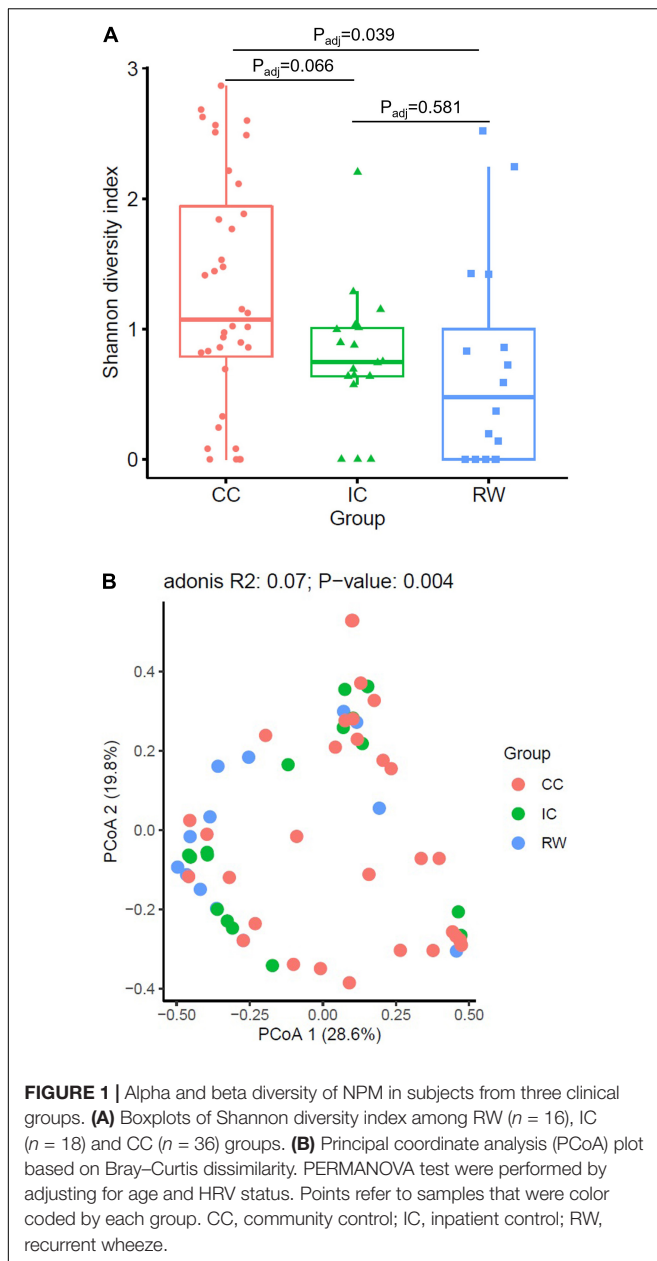
DNA Extraction and Library Preparation

Total genomic DNA was extracted from 300 μl of respiratory secretion using MO BIO PowerSoil DNA Isolation Kit (Mo Bio Laboratories; Carlsbad, CA, United States). The concentration and purity of extracted DNA was measured by NanoDrop® ND-2000 (Thermo Fisher Scientific, Waltham, MA, United States), with all samples having $\geq$ 20 ng DNA and acceptable OD260/280 ratio. DNA integrity was confirmed by 1% agarose gel electrophoresis (120 V, 45 min).

High-Throughput Sequencing and Sequence Data Pre-processing

Whole genome shotgun sequencing was performed by Groken Bioscience (Hong Kong) on HiSeq X Ten platform (Illumina, San Diego, CA, United States) to generate paired-end outputs (2 × 150 bp). The raw reads were pre-processed using KneadData version 0.7.2 for quality control and host sequence decontamination (Bowtie2 version 2.3.4.3 with human reference genome GRCh37/hg19). Please refer to online supplement for the details.

Taxonomic and Functional Analysis

Taxonomy assignment and functional prediction were performed using MetaPhlAn2 version 2.7.7 and HUMAnN2 version

FIGURE 1 | Alpha and beta diversity of NPM in subjects from three clinical groups. **(A)** Boxplots of Shannon diversity index among RW (*n* = 16), IC (*n* = 18) and CC (*n* = 36) groups. **(B)** Principal coordinate analysis (PCoA) plot based on Bray–Curtis dissimilarity. PERMANOVA test were performed by adjusting for age and HRV status. Points refer to samples that were color coded by each group. CC, community control; IC, inpatient control; RW, recurrent wheeze.

0.11.2 with default settings. Based on the reference of clade-specific marker genes embracing viruses, archaea, bacteria, and eukaryotes, MetaPhlAn2 classifies reads belonging to these kingdoms down to species level (Truong et al., 2015). HUMAnN2 is a three-tire pipeline designed for annotating the functional capabilities of microbes from metagenomic datasets based on MetaCyc database (Franzosa et al., 2018). Results were visualized by R version 3.5.0 (see online supplement).

Statistical Analysis

To determine the differences in clinical variables across groups, Wilcoxon rank-sum test were used for comparing categorical variables between two groups; Kruskal–Wallis test and Chi-square test were performed for comparison of continuous and categorical variables across more than two groups, respectively. Alpha diversity represented by Shannon diversity index (SDI) was compared across multiple groups using Kruskal–Wallis test. NPM composition (beta diversity) were compared using PERMANOVA adonis test in vegan R package with 999 permutations. Regarding taxonomic and functional profiles, Statistical Analysis Metagenomic Profiles (STAMP) version 2.0 (Parks et al., 2014) was used to compare between-group differences. Unless stated otherwise, pairwise comparison between groups were performed using Dunn's *post hoc* test coupled with Benjamini–Hochberg false discovery rate (FDR) procedure for multiple testing correction. Linear discriminant analysis effect size (LEfSe) algorithm was used to identify discriminating features that were significant both statistically and biologically (biomarker discovery). Logarithmic LDA score > 2 was set as the threshold to estimate effect size. Considering the complexity of metagenomic profiles, we replicated findings from LEfSe by analysis of composition of microbiomes (ANCOM) (Mandal et al., 2015) which has sensitivity on complexed microbiome profile and advantage of low FDR (see online supplement). Considering possible impact of HRV infection and on microbiota (Molyneaux et al., 2013; Korten et al., 2016), HRV status was included as confounders in our data analysis. This study considered potential microbial biomarkers microbes or their metabolic pathways that displayed consistent statistical significance from all three analyses above and remained group-discriminating after HRV adjustment. $P < 0.05$ was set as the significance threshold.

RESULTS

Subjects

Nasopharyngeal microbiome profiles were successfully recovered from a total of 70 samples, including 16 RW, 18 IC, and 36 CC. Eleven subjects (8 for RW and 3 IC) were excluded as non-human metagenomic reads and their samples were deemed "100% unclassified" after WGS analysis. **Supplementary Table 1** summarizes clinical characteristics of analyzable study participants. We found that the age of CC subjects (median age: 5.0 years old) is higher than IC (median 3.3 years, $P_{adj} < 0.01$) and RW (median 3.5 years, $P_{adj} < 0.01$) groups. All RW samples (RWV) were positive for HRV (HRV+), whereas HRV was detected in 33.3% (6/18) of IC and 22.2% (8/36) of CC. These two control groups were further divided into ICV (positive for HRV; $n = 6$), ICC (negative for HRV; $n = 12$), CCV (positive for HRV; $n = 8$), and CCC (negative for HRV; $n = 28$) subgroups.

The Community Controls Had Higher Nasopharyngeal Microbiome Biodiversity Than Patients Hospitalized for Recurrent Wheezing

Whole genome shotgun yielded an average throughput of 14 million 150-bp raw reads per sample, among which 98% passed quality filtering using standard parameters and > 90% of read bases met the Q30 criteria. Approximately 0.1–0.3 million pairs

FIGURE 2 | Taxonomic profiles of NPM in all 70 subjects. **(A)** Heatmap showing relative abundances of the top 11 most abundant species in individual samples. Samples were hierarchically clustered based on Bray–Curtis dissimilarity of species abundances, and color coded according to group or HRV status as indicated. **(B)** Taxonomic cladogram illustrating the phylogenetic diversity of microbial species (bacteria and viruses) found in all samples. Size of the node was proportional to the relative abundance for the corresponding taxon. Family and genus-level annotations were labeled on the tree, with some of which annotated using X–Z and a–i keys to avoid image overlapping. Species with minimal relative abundance threshold of 1% were annotated using A–W as keys. Order-level microbes are indicated by different colors shown in the legend.

of reads (1.4–4%) considered non-human were subjected to downstream analysis. Since SDI was positively correlated with age (Spearman $R = 0.29$, $P = 0.016$), age was a confounder of the relationship between RW and SDI. The median SDI for RW, IC, and CC groups were 0.48, 0.75, and 1.07 respectively (**Figure 1A**), with SDI being higher in CC than RW (Benjamini–Hochberg FDR, $P_{adj} = 0.039$) after adjustment for age. This difference between subgroups remained significant after adjustment for HRV status (Kruskal–Wallis test, $P = 0.048$, **Supplementary Figure 1A**). However, no significant difference in SDI was detected by pairwise comparisons between subgroups ($P > 0.05$ for all). This was consistent with our finding that SDI was not associated with HRV status ($P = 0.191$).

Discriminative Taxonomic Profiles in Clinical Groups

Based on Bray–Curtis dissimilarity, we observed that NPM composition was associated with HRV status (PERMANOVA, $R^2 = 0.052$, $P = 0.003$) but not with age ($R^2 = 0.015$, $P = 0.446$). PERMANOVA analysis indicated that NPM composition in CC group is different from that of IC group ($R^2 = 0.06$, $P_{adj} = 0.006$) and RW group ($R^2 = 0.067$, $P_{adj} = 0.006$) after controlling for HRV status (**Figure 1B**). Furthermore, subjects in CC who were negative for HRV infection (CCC) had distinct NPM structure compared to those in IC with HRV infection ($R^2 = 0.096$, $P_{adj} = 0.02$ for RWV; $R^2 = 0.095$, $P_{adj} = 0.035$ for ICV; **Supplementary Figure 1B**).

At the species level, WGS identified 28 bacteria, 12 viruses, and one fungus (*Aspergillus fumigatus*). **Supplementary Figure 2** shows relative abundance (RA) of all identified microbes in all 70 samples in each kingdom. The 90th percentile RA across all samples for each species was calculated and ranked to identify the topmost abundant species, which is the default parameter implemented in MetaPhlAn2. **Supplementary Table 2** summarizes the top 11 most abundant species, and **Figure 2A** illustrates heatmap of their RA, in which samples were hierarchical clustered based on Bray–Curtis dissimilarity. *Moraxella catarrhalis* and *Dolosigranulum pigrum* were the predominant species in our nasopharyngeal samples, with RA > 70% being found in ~10% of our samples. **Figure 2B** shows the cladogram on phylogenetic relationship among the presenting species. Of note, among 12 viruses detected by WGS, *porcine type C oncovirus* ranked the third based on 90th percentile of RA and was predominantly detected in CC group (**Supplementary Figure 3**; $P = 0.044$).

Thirty-three species [21 bacteria, 11 viruses, and one eukaryote (*Aspergillus fumigatus*)] had different RA among three groups (**Supplementary Table 3**). Among these species, five bacteria and one virus had an eta squared effect size (η^2) > 0.06, suggesting medium to high biological relevance to the clinical diagnosis (**Table 1**; Watson, 2018). After adjusting for HRV status, *D. pigrum* was the only species that remained group-discriminating in LEfSe analysis (**Supplementary Figure 4**). Because bacteria accounted for nearly 90% of non-human reads, subsequent taxonomical and functional analyses focused on bacterial sequences. LEfSe revealed seven microbial taxa as

significant in discriminating between groups (LDA score > 2, $P < 0.05$; **Supplementary Figure 5A**). **Supplementary Figure 5B** depicts their phylogenetic relationship. Proteobacteria was the only phylum representative for RW (LDA score 5.13, $P = 0.026$), although we did not find any discriminative genus or species within this phylum. In contrast, *D. pigrum*, GCF_000245815 (*the strain ATCC 51524 belonging to D. pigrum*), the family Carnobacteriaceae (*Dolosigranulum*), and the phylum Actinobacteria, were identified as the discriminative taxa in CC. Adjusted for HRV status, *D. pigrum* was the only discriminative species across three groups with lower abundance in RW and IC than in CC (**Supplementary Figure 5C**).

To confirm the above findings, ANCOM was performed on bacterial taxa that were present in ≥ 10% of all samples. *D. pigrum* and *Haemophilus influenzae* as well as Proteobacteria phylum were differentially abundant between cases and controls ($P < 0.05$; **Figures 3A–C**). Specifically, *D. pigrum* had a higher RA in CC (37.3 ± 35.3%) compared to IC (14.2 ± 24.5%, $P_{adj} = 0.037$) and RW (9.6 ± 21%, $P_{adj} = 0.027$, Benjamini–Hochberg FDR). *H. influenzae* was more abundant in IC (14.6 ± 32.9%) than the other two groups (6.3 ± 25.8% in RW and 3.0 ± 14.1% in CC). Consistent with LEfSe analysis, Proteobacteria was more abundant in RW than CC ($P = 0.049$). After adjusting for HRV status and age, *D. pigrum* (IC vs. CC, $P_{adj} = 0.042$) and Proteobacteria (IC vs. CC, $P_{adj} = 0.022$) remained differentially abundant across samples (**Supplementary Figures 6A,C**), while that of *H. influenza* became insignificant. Interestingly, *H. influenza* was primarily detected in HRV-positive IC samples (**Supplementary Figure 6B**), suggesting this species may contribute to URTI-related hospitalization.

Functional Potential of Nasopharyngeal Microbiome

HUMAnN2 recognized a total of 301 pathways. **Supplementary Table 4** summarizes the top 20 most abundant pathways (ranked by 90th percentile of RA across all samples), whereas **Figure 4A** illustrates the heatmap of their corresponding RA in individual samples. Purine deoxyribonucleotide *de novo* biosynthesis (3.07 ± 3.05% for both adenosine and guanosine, respectively) and glycolysis IV (5.3 ± 1%) were the topmost abundant pathways having the highest 90th percentile of RA. STAMP analyses identified 32 pathways to be significantly different (corrected $P < 0.05$) with a moderate to high eta-squared effect size (η^2 > 0.06) across groups (**Supplementary Table 5**). Seven of these 32 significant pathways were also among the top 20 most abundant: glycolysis IV (plant cytosol), superpathway of adenosine nucleotides *de novo* biosynthesis II, S-adenosyl-L-methionine cycle I, guanosine ribonucleotides *de novo* biosynthesis, superpathway of guanosine nucleotides *de novo* biosynthesis I, superpathway of adenosine nucleotides *de novo* biosynthesis I and L-valine biosynthesis. Hierarchical clustering of samples according to pathway abundances showed a clear pattern that PWY-1042: glycolysis IV (plant cytosol) (q-value = 0.048, η^2 = 0.11) and GLYCOGENSYNTH-PWY: glycogen biosynthesis I (from ADP-D-glucose) (q-value = 0.044,

TABLE 1 | Microbial species with medium to high effect size across different subject groups identified by STAMP.

Species	Relative abundance			q-value	η^2 effect size
	RW	**CC**	**IC**		
Dolosigranulum pigrum	6.35 ± 16.80	34.66 ± 34.49	12.24 ± 22.21	0.0008	0.18
Bifidobacterium longum	0.25 ± 0.88	1.47 ± 2.74	0.02 ± 0.11	0.013	0.11
Bifidobacterium breve	0.78 ± 2.58	3.89 ± 7.49	0.11 ± 0.49	0.014	0.10
Akkermansia muciniphila	0.89 ± 2.98	5.86 ± 14.34	0.11 ± 0.50	0.029	0.07
Corynebacterium pseudodiphtheriticum	0.80 ± 3.46	5.64 ± 13.14	0.87 ± 2.28	0.029	0.06
Porcine type C oncovirus	1.23 ± 4.11	5.87 ± 11.43	0.36 ± 1.62	0.015	0.09
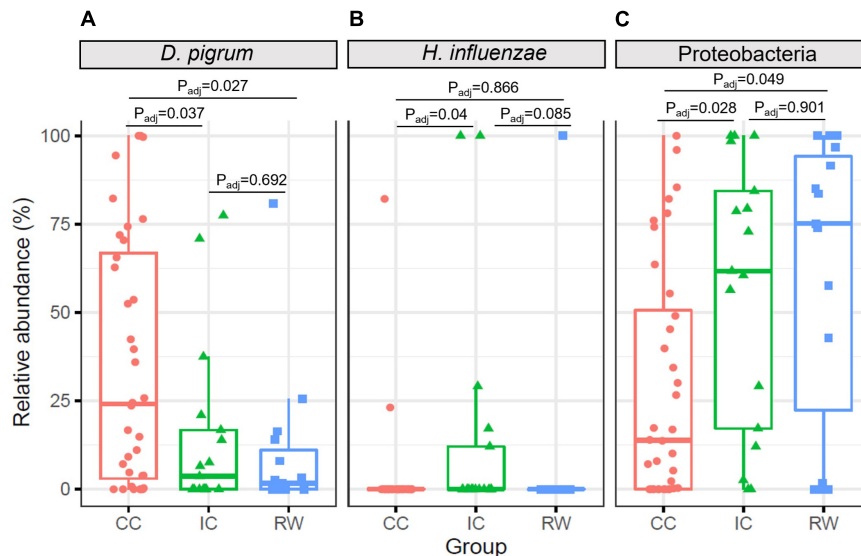
FIGURE 3 | Boxplots showing relative abundances of differentially abundant taxa at different taxonomic levels identified by ANCOM. **(A)** Left panel, *D. pigrum* species had higher abundance in CC group but lower abundance in IC and RW groups. **(B)** Middle panel, *H. influenzae* was more abundant in IC than CC group. **(C)** Right panel, *Proteobacteria* was more abundant in RW and IC than CC. CC, community control; IC, inpatient control; RW, recurrent wheeze.

$\eta^2 = 0.06$) were the only two pathways being enriched in RW. By contrast, purine metabolism such as guanosine ribonucleotides *de novo* biosynthesis (PWY-7221, $P_{adj} = 0.02$, $\eta^2 = 0.12$), and superpathway of guanosine nucleotides *de novo* biosynthesis I (PWY-7228, $P_{adj} = 0.025$, $\eta^2 = 0.09$) pathways were significantly more abundant in CC (PWY-7221, 3.26 ± 3.3%; PWY-7228, 2.88 ± 2.65%) as compared to IC (PWY-7221, 1.51 ± 0.95%; PWY-7228, 1.66 ± 1.08%) and RW (PWY-7221, 1.42 ± 1.7%; PWY-7228, 1.48 ± 1.87%) groups.

When analyzed by LEfSe (**Figure 4B**), RW was enriched in adenosine ribonucleotides *de novo* biosynthesis (LDA = 4.06, $P = 0.006$). Bacterial enzymes involved in inosine 5 phosphate biosynthesis were more abundant in IC than the other groups (LDA = 3.49, $P < 0.001$) whereas guanosine ribonucleotides *de novo* biosynthesis was enriched in CC (LDA = 3.91, $P = 0.004$). Fifteen of these 20 pathways overlapped with those identified using STAMP, and the remaining five pathways also belonged to three categories mentioned above (i.e., Biosynthesis, Degradation/utilization/assimilation, and Generation of Precursor Metabolite and Energy). Adjusting for HRV status, five discriminative MetaCyc pathways remained significant even by

using the more stringent "all-against-all" strategy in multi-class comparison implemented in LEfSe (**Figure 4C**). Adenosine ribonucleotides *de novo* biosynthesis (MetaCyc pathway PWY-7219) was the only representative pathway in RW (LDA = 4.25, $P = 0.006$), while four discriminative components were found in CC (**Table 2**).

Linking Metabolic Pathways to Microbes

To understand the species contributing to a particular functional pathway, contributional diversity bar plots illustrating the linkage between metabolic pathway and attributable species were generated for the five discriminative pathways (**Supplementary Figure 7**). The discriminative pathways representing CC group (i.e., PWY-7229, PWY-6126, and PWY-6125), being present only in a small number of samples, were contributed mainly by *Streptococcus spp.* and *H. influenzae* (**Supplementary Figures 7A–C**). In contrast, PWY-7219 being enriched in RW was present in 82.9% of all samples. This pathway was contributed mainly by *M. catarrhalis* in RW and IC groups and by *D. pigrum* and *Akkermansia muciniphila* in the CC group (**Supplementary Figure 7D**). This result was consistent with our previous finding
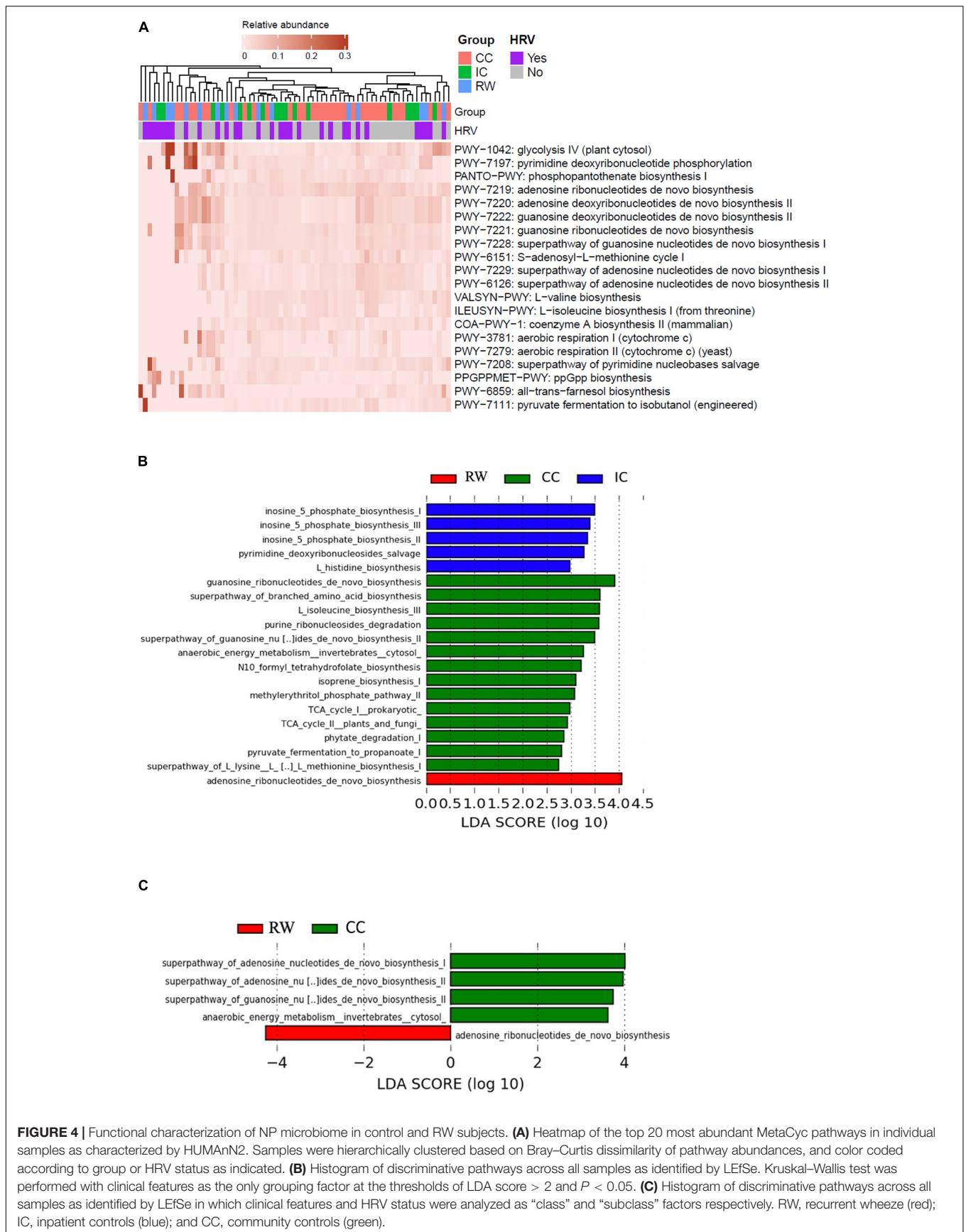
FIGURE 4 | Functional characterization of NP microbiome in control and RW subjects. (A) Heatmap of the top 20 most abundant MetaCyc pathways in individual samples as characterized by HUMAnN2. Samples were hierarchically clustered based on Bray–Curtis dissimilarity of pathway abundances, and color coded according to group or HRV status as indicated. (B) Histogram of discriminative pathways across all samples as identified by LEfSe. Kruskal–Wallis test was performed with clinical features as the only grouping factor at the thresholds of LDA score > 2 and P < 0.05. (C) Histogram of discriminative pathways across all samples as identified by LEfSe in which clinical features and HRV status were analyzed as "class" and "subclass" factors respectively. RW, recurrent wheeze (red); IC, inpatient controls (blue); and CC, community controls (green).

of over-representation of *M. catarrhalis* in RW and *D. pigrum* in CC. To further dissect the functional potential of *D. pigrum*, we identified nine MetaCyc pathways in which *D. pigrum* was involved (**Supplementary Table 6**). The contributional bar plots revealed that, compared to RW and IC groups, S-adenosyl-L-methionine cycle I (PWY-6151) and purine ribonucleosides degradation (PWY0-1296) were more abundant in CC and mainly contributed by *D. pigrum* (**Supplementary Figure 8**).

Of interest, several pathways identified from our dataset were related to genes from other kingdoms. Specifically, we found pathways specific to *protozoa* and *fungi* among the control subjects. These pathways included PWY-7383: anaerobic energy metabolism (invertebrates, cytosol) ($q = 0.013$, $\eta^2 = 0.09$) and PWY-5690: TCA cycle II (plants and fungi) ($q = 0.025$, $\eta^2 = 0.07$) in CC group, and PWY-7357: thiamin formation from pyrithiamine and oxythiamine (yeast) ($q = 0.038$, $\eta^2 = 0.07$) in IC group (**Supplementary Table 5**). Besides, the previously mentioned pathway PWY-1042 enriched in RW is a glycolysis pathway contributed by *Viridiplantae* kingdom.

DISCUSSION

This WGS-based study revealed that *M. catarrhalis* (Proteobacteria phylum) and *D. pigrum* (Firmicutes phylum) were predominant species in nasopharyngeal samples of pre-school children. This was concordant with other studies on NPM among infants and children (Bogaert et al., 2011; Teo et al., 2015). We discovered that the NPM of RW with asthma predisposition was characterized by low biodiversity characterizing with enriched phylum Proteobacteria being the only representative taxa. Conversely, *D. pigrum* was overrepresented in CC. Adjusted for HRV status, *D. pigrum* was the only discriminative species across RW, IC, and CC groups. LEfSe analyses identified five discriminative MetaCyc pathways, including adenosine ribonucleotides *de novo* biosynthesis in RW and four bacterial pathways involved in purine and adenosine biosynthesis, and energy metabolism for CC group.

In the present study, two approaches were applied for collecting nasopharyngeal samples from hospitalized and community subjects. For hospitalized patients, NPAs were obtained as standard practice for pathogen detection upon admission (Leung et al., 2010). On the other hand, FNPSs were collected from CC subjects as such approach (versus NPA) obviates the need for negative-pressure facilities during the procedure (Chan et al., 2008). In practice, Tunsjø et al. (2015) reported that these two sample types did not differ in virus detection among children. Previous studies have also indicated that FNPS samples represented a less invasive and more reliable option than NPA in terms of the sensitivity and specificity for respiratory virus detection especially for children below 5 years old (Abu-Diab et al., 2008; Lambert et al., 2008). These findings supported our use of FNPS in CC subjects and NPA for inpatients and would not introduce any substantial difference in nasopharyngeal microbial communities from the methodological point of view.

Our study reported that hospitalized children with RW (and positive S-API) or URTI had lower microbial biodiversity in upper airway than CC. Specifically, pre-school children with RW had higher abundance of Proteobacteria while the latter group possessed higher abundance of phylum Actinobacteria and genus *Dolosigranulum* in their nasopharynx. These observations were consistent with previous studies for wheeze and asthma that the asthmatic upper airway had lower biodiversity but more abundant Proteobacteria (Depner et al., 2017; Fazlollahi et al., 2018). Particularly, *Moraxella* is a genus under Proteobacteria that was repeatedly reported to be related to asthma and wheeze. Children hospitalized for severe bronchiolitis who carried more *Moraxella* in upper airway had increased risk for RW at later childhood (Mansbach et al., 2020; Zhang et al., 2020). Similar observations between nasal *Moraxella* and asthma exacerbation were found in schoolchildren (McCauley et al., 2019; Zhou et al., 2019). McCauley et al. (2019) reported nasal domination of *Moraxella spp.* was related with type-2 inflammation in nasal epithelium. All our RW cases were positive for S-API which reflected familial and personal history of allergic diseases. Further investigation is warranted to link the host inflammatory profile to NPM dysbiosis during acute wheezing episodes.

Moreover, Robinson et al. reported that pre-school episodic viral wheezers had *Moraxella*-dominating lung microbiome during the stable phase. Their results also suggested that *Moraxella*-enriched microbial profile was related to local but not systemic neutrophilic inflammation (Robinson et al., 2019). Our study extended the relationship between Proteobacteria and wheezing illnesses into pre-school age group. Additionally, our finding supported that NPM dysbiosis shared similar patterns with that found in BAL from severe pre-school wheezers. This observation was consistent with a previous study that upper airway microbiota in young children was generally representative for lung disease-related microbial patterns in the lower airway (Marsh et al., 2016).

On the contrary, Actinobacteria and *Dolosigranulum* might protect against asthma (Biesbroek et al., 2014; Ta et al., 2018). Both taxa were normal upper airway commensals that colonized infants' nasopharyngeal niche within the first month of life (Bosch et al., 2017; Man et al., 2017). The predominance of *Dolosigranulum* in the airway was associated with lower risk for asthma and viral respiratory diseases (Bosch et al., 2017; Teo et al., 2018). *Dolosigranulum* and *Corynebacterium* (Actinobacteria) also played an important role in respiratory health (Laufer et al., 2011; Kozik and Huang, 2019). Our study extended such relationship into the pre-school age group. Our results suggested that RW-related dysbiosis in the upper airway was present prior to the onset of asthma among pre-school children. This finding may provide an investigative approach for asthma prediction in young children.

The ability of WGS sequencing to taxonomically identify microbes down to the species level (e.g., *M. catarrhalis* in RW and *D. pigrum* in CC) is one of its strengths over 16S rRNA sequencing. Our metagenomic data identified that *D. pigrum* was the only species that remained group-discriminating after adjustment for HRV status. This bacterium was first described as a new lineage of lactic acid bacterium (Aguirre et al., 1993),

and its whole-genome sequence was announced in 2017 (Mukhopadhyay et al., 2017). Recent studies revealed an inverse correlation in RAs between *D. pigrum* and *Staphylococcus aureus*, a pathobiont in the respiratory tract (Liu et al., 2015; Escapa et al., 2018). Furthermore, *D. pigrum* was reported to inhibit the *in vitro* growth of *S. aureus* (Brugger et al., 2020) and decrease the production of epithelial pro-inflammatory cytokines induced by *S. aureus* which maintain epithelial barrier integrity against disruptive interleukin-4 (De Boeck et al., 2017).

In addition, another strength of WGS is its capability to identify all kingdoms of microbes that use DNA as their genetic materials while 16S rRNA profiling could only reconstruct phylogeny of bacteria and archaea. This WGS study uncovered six viral families including *Siphoviridae*, *Herpesviridae, Anelloviridae, Papillomaviridae, Polyomaviridae,* and *Retroviridae* from nasopharyngeal samples that accounted for nearly 10% of non-human reads. Among them, *Herpesviridae, Polyomaviridae,* and *Anelloviridae* were reported in children with community-acquired pneumonia (Li Y. et al., 2019). We found *porcine type C oncovirus* of the Retroviridae family to be more abundant in CC (**Supplementary Figure 3**), and existing data suggested low risk of *in vivo* infection of this virus in human (Łopata et al., 2018).

Another remarkable advantage of WGS related to its ability to characterize the functional potential of a metagenome. Accumulating evidence indicated that mucosal microbiome impacted allergy and asthma development through perturbation of the host immunity (Kemter and Nagler, 2019). We found a higher functional capacity for purine nucleotide *de novo* biosynthesis (PWY-7229, PWY-6126, and PWY-6125) in CC subjects (**Table 2**), suggesting nasopharyngeal microbes with greater proliferation potential could contribute to immune homeostasis in healthy individuals. Nevertheless, these discriminative metabolic pathways were present in only less than one quarter of all CC samples (**Supplementary Figures 7A–C**). Hence, the immunological implications of overrepresentation of these pathways required further investigation integrating metatranscriptomics and metabolomics. In this regard, Stewart et al. (2017) interrogated the microbiome and metabolome in NPA and explored their relationship to bronchiolitis in infants. In school-age children with asthma, serum metabolomic biomarkers were linked to oropharyngeal microbiome to illustrate the interaction between microbes and host immunity (Chiu et al., 2020). Similar exploration leveraging multi-omics approaches are warranted to implement in pre-school children with recurrent wheezing, which would further our understanding

on how functionally aberrant airway microbiome modulates host immune response during this critical phase and leads to later asthma.

As discussed above, *D. pigrum* has been shown to inhibit *in vitro* growth of *S. aureus* (Brugger et al., 2020). Among the nine *D. pigrum*-involved pathways, PWY-7219 represented the only discriminative function enriched in RW by LEfSe analysis (and HRV status as a subclass grouping factor). Interestingly, this pathway was mainly contributed by *M. catarrhalis* in RW but primarily by *D. pigrum* in CC (**Supplementary Figure 7D**). S-adenosyl-L-methionine cycle I (PWY-6151) responsible for amino acid metabolism was another *D. pigrum*-involved pathway. By providing methyl in living cells (Caspi et al., 2017), this pathway was important in the biosynthesis of multiple antibiotics including neomycin B (Kudo et al., 2014) and anthracycline (Grocholski et al., 2015), which might partially explain inhibition of *S. aureus* growth by *D. pigrum* in co-culture. Moreover, some asthmatic children present with chronic cough (Morice et al., 2014). Of interest, lung-derived ATP interacted with P2X receptor (evolutionarily conserved ATP-gated ion channels) to stimulate hypersensitive cough through peripheral nerves (Turner and Birring, 2019). Endogenous ATP could be produced locally in response to tissue stresses such as inflammation or epithelial damage (Ford and Undem, 2013). Our metagenomic data revealed that the NPM from CC group was enriched for purine and adenosine biosynthesis. Whether this finding may be related to ATP-P2X pathway require further investigation.

Persistent bacterial bronchitis (PBB) is another leading cause of chronic cough in young children. Differentiating these two conditions is important as the treatment plan and prognosis are distinct. Kazachkov et al. (2018) investigated microbiome of BAL samples from pre-school children with chronic cough. They found that asthmatic BAL was enriched for *Rothia, Gemellaceae*, and *Granulicatella*, while that of PBB was characterized by increased abundance for *Prevotella*. Another study reported bronchial dysbiosis in pre-school PBB to be associated with enrichment of *Haemophilus* and *Neisseria* when compared to healthy controls (Cuthbertson et al., 2017). Similarly, Bao et al. (2018) extended this finding to infants by showing that those with PBB had enriched *Haemophilus* and *Bacteroides* but decreased *Lactococcus* and *Lactobacillus*.

Some studies suggested that upper airway samples, such as throat swabs, demonstrated representative microbial communities for lower airways as assessed by BAL or bronchial brushes (Ahmed et al., 2018).

TABLE 2 | Discriminative pathways for RW and CC groups as identified by LEfSe analysis.

Pathway	Group discriminated	*q*-value	LDA score
PWY-7219: adenosine ribonucleotides *de novo* biosynthesis	RW	0.006	4.25
PWY-7229: superpathway of adenosine nucleotides *de novo* biosynthesis I	CC	0.012	4.03
PWY-6126: superpathway of adenosine nucleotides *de novo* biosynthesis II	CC	0.020	3.98
PWY-6125: superpathway of guanosine nucleotides *de novo* biosynthesis II	CC	0.014	3.75
PWY-7383: anaerobic energy metabolism (invertebrates, cytosol)	CC	0.001	3.63

LEfSe analysis was performed by Kruskal–Wallis test and pairwise Wilcoxon test.

This finding supported the use of upper airway samples for studying microbiome in different respiratory diseases in children as it would be too invasive to obtain BAL samples by bronchoscopy in most cases. Studies using upper airway surrogates are warranted to unveil potential microbial signals for early recognition of asthma-predisposed signals among children with chronic cough.

Of interest, this study identified pathways (PWY-7383, PWY-5690, and PWY-7357) specific to *fungi* among the control subjects. This supported the hypothesis that fungi in NPM may play a role in the pathogenesis of recurrent wheezing (Li E. et al., 2019). Fujimura et al. (2016) reported higher RAs of specific fungi (*Candida* and *Rhodotorula*) in the gut pose a high risk of childhood asthma. Fungal dysbiosis characteristic of *Pneumocystis* overrepresentation in the lung was associated with severe asthma in children (Barcik et al., 2020). More studies focusing on airway fungiome are needed to deepen our understanding of the mechanisms through which fungi influence respiratory health.

There are several limitations in this study. Firstly, we could not analyze the effect of HRV on RW as all our RW cases were HRV positive. Our local findings found that the large majority of hospitalized children with RW and asthma exacerbation had HRV infection (Leung et al., 2010; Mak et al., 2011). However, we were still able to adjust for HRV by having HRV-negative subjects in the IC and CC groups. Secondly, this was a cross-sectional study on pre-school recurrent wheezers which attempted to apply NPM to predict later development of asthma. The results will need to be validated in prospective cohorts that followed RW group into later childhood. In addition, WGS is able to reveal the metabolic potentials embedded in the microbial genes. The integration of multi-omics information allows us to explore the mechanisms through which microbial dysbiosis affects respiratory health, but *in vitro* and animal experiments are required to confirm the functionality of these microbes. Besides, NPM analyses in this study was based on NPA from hospitalized children in the RW and IC groups and FNPS from children in the CC group. There has not been any study that compared microbiome compositions in these upper airway samples, so it is unknown if such might introduce bias in the bacterial community. A previous local study revealed that these two sampling methods had comparable power to detect respiratory viruses by molecular assays (Chan et al., 2008). Further comparative studies are required to analyze upper airway microbiota in NPA and FNPS samples.

CONCLUSION

We identified *M. catarrhalis* and *D. pigrum* to be the predominant species in nasopharynx of pre-school children. Hospitalized children with RW and positive for S-API had lower nasopharyngeal biodiversity than CC, while Proteobacteria was the only representative taxa in RW. On the other hand, *D. pigrum* was overrepresented in CC. Adjusted for HRV status,

D. pigrum and five MetaCyc pathways were discriminative across groups. These microbial biomarkers provide insights into RW pathogenesis in pre-school children.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/ **Supplementary Material**.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Joint Chinese University of Hong Kong – New Territories East Cluster Clinical Research Ethics Committee. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

TL obtained research funding, designed this study, collected clinical data, and drafted this manuscript. YS collected clinical data, performed laboratory experiments, analyzed laboratory and WGS data, and drafted this manuscript. JH and JK conducted bioinformatics analysis, interpreted WGS data, and contributed to the manuscript. ST supervised bioinformatic data analysis and results interpretation. HW and MW contributed as statistical consultation. MT contributed to laboratory experiments. KT and RC conducted experiments for HRV detection. AL and GW analyzed clinical data. All authors reviewed and approved the manuscript.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2021.792556/full#supplementary-material

REFERENCES

Abu-Diab, A., Azzeh, M., Ghneim, R., Ghneim, R., Zoughbi, M., Turkuman, S., et al. (2008). Comparison between pernasal flocked swabs and nasopharyngeal aspirates for detection of common respiratory viruses in samples from children. *J. Clin. Microbiol.* 46, 2414–2417. doi: 10.1128/jcm.00369-08

Aguirre, M., Morrison, D., Cookson, B. D., Gay, F. W., and Collins, M. D. (1993). Phenotypic and phylogenetic characterization of some Gemella-like organisms from human infections: description of Dolosigranulum pigrum gen. nov., sp. nov. *J. Appl. Bacteriol.* 75, 608–612. doi: 10.1111/j.1365-2672.1993.tb01602.x

Ahmed, B., Cox, M. J., Cuthbertson, L., James, P. L., Cookson, W. O. C., Davies, J. C., et al. (2018). Comparison of the upper and lower airway microbiota in children with chronic lung diseases. *PLoS One* 13:e0201156. doi: 10.1371/journal.pone.0201156

Bao, Y., Li, Y., Qiu, C., Wang, W., Yang, Z., Huang, L., et al. (2018). Bronchoalveolar lavage fluid microbiota dysbiosis in infants with protracted bacterial bronchitis. *J. Thorac. Dis.* 10, 168–174. doi: 10.21037/jtd.2017.12.59

Barcik, W., Boutin, R. C. T., Sokolowska, M., and Finlay, B. B. (2020). The role of lung and gut microbiota in the pathology of asthma. *Immunity* 52, 241–255. doi: 10.1016/j.immuni.2020.01.007

Biesbroek, G., Bosch, A. A. T. M., Wang, X., Keijser, B. J. F., Veenhoven, R. H., Sanders, E. A. M., et al. (2014). The impact of breastfeeding on nasopharyngeal microbial communities in infants. *Am. J. Respir. Crit. Care Med.* 190, 298–308. doi: 10.1164/rccm.201401-0073OC

Bisgaard, H., Hermansen, M. N., Buchvald, F., Loland, L., Halkjaer, L. B., Bønnelykke, K., et al. (2007). Childhood asthma after bacterial colonization of the airway in neonates. *N. Engl. J. Med.* 357, 1487–1495. doi: 10.1056/NEJMoa052632

Bogaert, D., Keijser, B., Huse, S., Rossen, J., Veenhoven, R., Van Gils, E., et al. (2011). Variability and diversity of nasopharyngeal microbiota in children: a metagenomic analysis. *PLoS One* 6:e17035. doi: 10.1371/journal.pone.0017035

Bosch, A. A. T. M., Piters, W. A. A. D. S., Van Houten, M. A., Chu, M. L. J. N., Biesbroek, G., Kool, J., et al. (2017). Maturation of the infant respiratory microbiota, environmental drivers, and health consequences. a prospective cohort study. *Am. J. Respir. Crit. Care Med.* 196, 1582–1590. doi: 10.1164/rccm.201703-0554OC

Brugger, S. D., Eslami, S. M., Pettigrew, M. M., Escapa, I. F., Henke, M. T., Kong, Y., et al. (2020). *Dolosigranulum pigrum* cooperation and competition in human nasal microbiota. *mSphere* 5:e852-20. doi: 10.1128/mSphere.00852-20

Caspi, R., Billington, R., Fulcher, C. A., Keseler, I. M., Kothari, A., Krummenacker, M., et al. (2017). The MetaCyc database of metabolic pathways and enzymes. *Nucleic Acids Res.* 46, D633–D639.

Castro-Rodríguez, J. A., Holberg, C. J., Wright, A. L., and Martinez, F. D. (2000). A clinical index to define risk of asthma in young children with recurrent wheezing. *Am. J. Respir. Crit. Care Med.* 162, 1403–1406. doi: 10.1164/ajrccm.162.4.9912111

Chan, K. H., Peiris, J. S., Lim, W., Nicholls, J. M., and Chiu, S. S. (2008). Comparison of nasopharyngeal flocked swabs and aspirates for rapid diagnosis of respiratory viruses in children. *J. Clin. Virol.* 42, 65–69. doi: 10.1016/j.jcv.2007.12.003

Chiu, C. Y., Chou, H. C., Chang, L. C., Fan, W. L., Dinh, M. C. V., Kuo, Y. L., et al. (2020). Integration of metagenomics-metabolomics reveals specific signatures and functions of airway microbiota in mite-sensitized childhood asthma. *Allergy* 75, 2846–2857. doi: 10.1111/all.14438

Cuthbertson, L., Craven, V., Bingle, L., Cookson, W., Everard, M. L., and Moffatt, M. F. (2017). The impact of persistent bacterial bronchitis on the pulmonary microbiome of children. *PLoS One* 12:e0190075. doi: 10.1371/journal.pone.0190075

De Boeck, I., Wittouck, S., Wuyts, S., Oerlemans, E. F. M., Van Den Broek, M. F. L., Vandenheuvel, D., et al. (2017). Comparing the healthy nose and nasopharynx microbiota reveals continuity as well as niche-specificity. *Front. Microbiol.* 8:2372. doi: 10.3389/fmicb.2017.02372

Depner, M., Ege, M. J., Cox, M. J., Dwyer, S., Walker, A. W., Birzele, L. T., et al. (2017). Bacterial microbiota of the upper respiratory tract and childhood asthma. *J. Allergy Clin. Immunol.* 139, 826.e813–834.e813. doi: 10.1016/j.jaci.2016.05.050

Escapa, I. F., Chen, T., Huang, Y., Gajare, P., and Lemon, K. P. (2018). New insights into human nostril microbiome from the expanded Human Oral 1 Microbiome Database (eHOMD): a resource for species-level identification of 2 microbiome data from the aerodigestive tract 3 4. *mSystems* 3:e187-18. doi: 10.1101/347013

Fazlollahi, M., Lee, T. D., Andrade, J., Oguntuyo, K., Chun, Y., Grishina, G., et al. (2018). The nasal microbiome in asthma. *J. Allergy Clin. Immunol.* 142, 834.e832–843.e832. doi: 10.1016/j.jaci.2018.02.020

Ford, A. P., and Undem, B. J. (2013). The therapeutic promise of ATP antagonism at P2X3 receptors in respiratory and urological disorders. *Front. Cell Neurosci.* 7:267. doi: 10.3389/fncel.2013.00267

Franzosa, E. A., Mciver, L. J., Rahnavard, G., Thompson, L. R., Schirmer, M., Weingart, G., et al. (2018). Species-level functional profiling of metagenomes and metatranscriptomes. *Nat. Methods* 15, 962–968. doi: 10.1038/s41592-018-0176-y

Fujimura, K. E., Sitarik, A. R., Havstad, S., Lin, D. L., Levan, S., Fadrosh, D., et al. (2016). Neonatal gut microbiota associates with childhood multisensitized atopy and T cell differentiation. *Nat. Med.* 22, 1187–1191. doi: 10.1038/nm.4176

Grocholski, T., Dinis, P., Niiranen, L., Niemi, J., and Metsä-Ketelä, M. (2015). Divergent evolution of an atypical S-adenosyl-l-methionine-dependent monooxygenase involved in anthracycline biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* 112, 9866–9871. doi: 10.1073/pnas.1501765112

Kazachkov, M., Kapoor, B. C., Malecha, P. W., Wu, B. G., Li, Y., Levine, J., et al. (2018). Aerodigestive dysbiosis in children with chronic cough. *Pediatr. Pulmonol.* 53, 1288–1298. doi: 10.1002/ppul.24115

Kemter, A. M., and Nagler, C. R. (2019). Influences on allergic mechanisms through gut, lung, and skin microbiome exposures. *J. Clin. Invest.* 129, 1483–1492. doi: 10.1172/JCI124610

Korten, I., Mika, M., Klenja, S., Kieninger, E., Mack, I., Barbani, M. T., et al. (2016). Interactions of respiratory viruses and the nasal microbiota during the first year of life in healthy infants. *mSphere* 1:e312-16.

Kozik, A. J., and Huang, Y. J. (2019). The microbiome in asthma: role in pathogenesis, phenotype, and response to treatment. *Ann. Allergy Asthma Immunol.* 122, 270–275. doi: 10.1016/j.anai.2018.12.005

Kudo, F., Hoshi, S., Kawashima, T., Kamachi, T., and Eguchi, T. (2014). Characterization of a radical S-adenosyl-L-methionine epimerase, NeoN, in the last step of neomycin B biosynthesis. *J. Am. Chem. Soc.* 136, 13909–13915. doi: 10.1021/ja507759f

Lambert, S. B., Whiley, D. M., O'neill, N. T., Andrews, E. C., Canavan, F. M., Bletchly, C., et al. (2008). Comparing nose-throat swabs and nasopharyngeal aspirates collected from children with symptoms for respiratory virus identification using real-time polymerase chain reaction. *Pediatrics* 122:e615-20. doi: 10.1542/peds.2008-0691

Laufer, A. S., Metlay, J. P., Gent, J. F., Fennie, K. P., Kong, Y., and Pettigrew, M. M. (2011). Microbial communities of the upper respiratory tract and otitis media in children. *mBio* 2:e245-10. doi: 10.1128/mBio.00245-10

Leung, T. F., Chan, R. W. Y., Kwok, A., Ho, W. C. S., Tao, M. K. P., Hon, K. L., et al. (2017). School-based surveillance for influenza vaccine effectiveness during 2014-2015 seasons in Hong Kong. *Influenza Other Respir. Viruses* 11, 319–327. doi: 10.1111/irv.12455

Leung, T. F., To, M. Y., Yeung, A. C. M., Wong, Y. S., Wong, G. W. K., and Chan, P. K. S. (2010). Multiplex Molecular Detection of Respiratory Pathogens in Children With Asthma Exacerbation. *Chest* 137, 348–354. doi: 10.1378/chest.09-1250

Li, E., Knight, J. M., Wu, Y., Luong, A., Rodriguez, A., Kheradmand, F., et al. (2019). Airway mycosis in allergic airway disease. *Adv. Immunol.* 142, 85–140. doi: 10.1016/bs.ai.2019.05.002

Li, Y., Fu, X., Ma, J., Zhang, J., Hu, Y., Dong, W., et al. (2019). Altered respiratory virome and serum cytokine profile associated with recurrent respiratory tract infections in children. *Nat. Commun.* 10:2288. doi: 10.1038/s41467-019-10294-x

Liu, C. M., Price, L. B., Hungate, B. A., Abraham, A. G., Larsen, L. A., Christensen, K., et al. (2015). Staphylococcus aureus and the ecology of the nasal microbiome. *Sci. Adv.* 1:e1400216. doi: 10.1126/sciadv.1400216

Łopata, K., Wojdas, E., Nowak, R., Łopata, P., and Mazurek, U. (2018). Porcine endogenous retrovirus (PERV) - molecular structure and replication strategy in the context of retroviral infection risk of human cells. *Front. Microbiol.* 9:730. doi: 10.3389/fmicb.2018.00730

Mak, R. K. Y., Tse, L. Y., Lam, W. Y., Wong, G. W. K., Chan, P. K. S., and Leung, T. F. (2011). Clinical spectrum of human rhinovirus infections in hospitalized

hong kong children. *Pediatr. Infect. Dis. J.* 30, 749–753. doi: 10.1097/INF. 0b013e31821b8c71

Man, W. H., De Steenhuijsen Piters, W. A. A., and Bogaert, D. (2017). The microbiota of the respiratory tract: gatekeeper to respiratory health. *Nat. Rev. Microbiol.* 15, 259–270. doi: 10.1038/nrmicro.2017.14

Mandal, S., Van Treuren, W., White, R. A., Eggesbø, M., Knight, R., and Peddada, S. D. (2015). Analysis of composition of microbiomes: a novel method for studying microbial composition. *Microb. Ecol. Health Dis.* 26:27663. doi: 10. 3402/mehd.v26.27663

Mansbach, J. M., Luna, P. N., Shaw, C. A., Hasegawa, K., Petrosino, J. F., Piedra, P. A., et al. (2020). Increased *Moraxella* and Streptococcus species abundance after severe bronchiolitis is associated with recurrent wheezing. *J. Allergy Clin. Immunol.* 145, 518.e8–527.e8. doi: 10.1016/j.jaci.2019.10.034

Marsh, R. L., Kaestli, M., Chang, A. B., Binks, M. J., Pope, C. E., Hoffman, L. R., et al. (2016). The microbiota in bronchoalveolar lavage from young children with chronic lung disease includes taxa present in both the oropharynx and nasopharynx. *Microbiome* 4:37. doi: 10.1186/s40168-016-0182-1

McCauley, K., Durack, J., Valladares, R., Fadrosh, D. W., Lin, D. L., Calatroni, A., et al. (2019). Distinct nasal airway bacterial microbiotas differentially relate to exacerbation in pediatric patients with asthma. *J. Allergy Clin. Immunol.* 144, 1187–1197. doi: 10.1016/j.jaci.2019.05.035

Molyneaux, P. L., Mallia, P., Cox, M. J., Footitt, J., Willis-Owen, S. A. G., Homola, D., et al. (2013). Outgrowth of the bacterial airway microbiome after rhinovirus exacerbation of chronic obstructive pulmonary disease. *Am. J. Respir. Crit. Care Med.* 188, 1224–1231. doi: 10.1164/rccm.201302-0341OC

Morice, A. H., Millqvist, E., Belvisi, M. G., Bieksiene, K., Birring, S. S., Chung, K. F., et al. (2014). Expert opinion on the cough hypersensitivity syndrome in respiratory medicine. *Eur. Respir. J.* 44, 1132–1148. doi: 10.1183/09031936. 00218613

Mukhopadhyay, R., Joaquin, J., Hogue, R., Fitzgerald, S., Jospin, G., Mars, K., et al. (2017). Complete Genome Sequence of Dolosigranulum pigrum from a Patient with Interstitial Lung Disease Using Single-Molecule Real-Time Sequencing Technology. *Genome Announc.* 5:e317-17. doi: 10.1128/genomeA. 00317-17

Parks, D. H., Tyson, G. W., Hugenholtz, P., and Beiko, R. G. (2014). STAMP: statistical analysis of taxonomic and functional profiles. *Bioinformatics* 30, 3123–3124. doi: 10.1093/bioinformatics/btu494

Quince, C., Walker, A. W., Simpson, J. T., Loman, N. J., and Segata, N. (2017). Shotgun metagenomics, from sampling to analysis. *Nat. Biotechnol.* 35, 833–844.

Robinson, P. F. M., Pattaroni, C., Cook, J., Gregory, L., Alonso, A. M., Fleming, L. J., et al. (2019). Lower airway microbiota associates with inflammatory phenotype in severe preschool wheeze. *J. Allergy Clin. Immunol.* 143, 1607.e3–1610.e3. doi: 10.1016/j.jaci.2018.12.985

Stewart, C. J., Mansbach, J. M., Wong, M. C., Ajami, N. J., Petrosino, J. F., Camargo, C. A. Jr., et al. (2017). Associations of nasopharyngeal metabolome and microbiome with severity among infants with bronchiolitis. A multiomic analysis. *Am. J. Respir. Crit. Care Med.* 196, 882–891. doi: 10.1164/rccm.201701-0071OC

Ta, L. D. H., Yap, G. C., Tay, C. J. X., Lim, A. S. M., Huang, C. H., Chu, C. W., et al. (2018). Establishment of the nasal microbiota in the first 18 months of life: correlation with early-onset rhinitis and wheezing. *J. Allergy Clin. Immunol.* 142, 86–95. doi: 10.1016/j.jaci.2018.01.032

Teo, S. M., Mok, D., Pham, K., Kusel, M., Serralha, M., Troy, N., et al. (2015). The infant nasopharyngeal microbiome impacts severity of lower respiratory infection and risk of asthma development. *Cell Host Microbe* 17, 704–715. doi: 10.1016/j.chom.2015.03.008

Teo, S. M., Tang, H. H. F., Mok, D., Judd, L. M., Watts, S. C., Pham, K., et al. (2018). Airway microbiota dynamics uncover a critical window for interplay of pathogenic bacteria and allergy in childhood respiratory disease. *Cell Host Microbe* 24, 341.e5–352.e5. doi: 10.1016/j.chom.2018.08.005

Truong, D. T., Franzosa, E. A., Tickle, T. L., Scholz, M., Weingart, G., Pasolli, E., et al. (2015). MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nat. Methods* 12, 902–903. doi: 10.1038/nmeth.3589

Tunsjø, H. S., Berg, A. S., Inchley, C. S., Røberg, I. K., and Leegaard, T. M. (2015). Comparison of nasopharyngeal aspirate with flocked swab for PCR-detection of respiratory viruses in children. *APMIS* 123, 473–477. doi: 10.1111/apm. 12375

Turner, R. D., and Birring, S. S. (2019). Chronic cough: ATP, afferent pathways and hypersensitivity. *Eur. Respir. J.* 54:1900889. doi: 10.1183/13993003.00889-2019

Watson, P. (2018). *Rules of Thumb on Magnitudes of Effect Sizes.* Available online at: http://imaging.mrc-cbu.ac.uk/statswiki/FAQ/effectSize (accessed December 14, 2019).

Zhang, X., Zhang, X., Zhang, N., Wang, X., Sun, L., Chen, N., et al. (2020). Airway microbiome, host immune response and recurrent wheezing in infants with severe respiratory syncytial virus bronchiolitis. *Pediatr. Allergy Immunol.* 31, 281–289. doi: 10.1111/pai.13183

Zhou, Y., Jackson, D., Bacharier, L. B., Mauger, D., Boushey, H., Castro, M., et al. (2019). The upper-airway microbiota and loss of asthma control among asthmatic children. *Nat. Commun.* 10:5714. doi: 10.1038/s41467-019-13698-x