



OPEN ACCESS

EDITED BY

Richard Allen White III,
University of North Carolina at Charlotte,
United States

REVIEWED BY

Miguel Ignacio Uyaguari-Diaz,
University of Manitoba,
Canada

Alberto Sola-Leyva,
University of Granada,
Spain

Alessandra Borgognone,
IrsiCaixa, Spain

*CORRESPONDENCE

Wen Sun
sunwen@bucm.edu.cn

SPECIALTY SECTION

This article was submitted to
Virology,
a section of the journal
Frontiers in Microbiology

RECEIVED 24 May 2022

ACCEPTED 27 July 2022

PUBLISHED 16 August 2022

CITATION

Huang L, Wu X, Guo S, Lv Y, Zhou P,
Huang G, Duan Z and Sun W (2022)
Metagenomic-based characterization of
the gut virome in patients with polycystic
ovary syndrome.
Front. Microbiol. 13:951782.
doi: 10.3389/fmicb.2022.951782

COPYRIGHT

© 2022 Huang, Wu, Guo, Lv, Zhou, Huang,
Duan and Sun. This is an open-access
article distributed under the terms of the
[Creative Commons Attribution License \(CC
BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in
other forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Metagenomic-based characterization of the gut virome in patients with polycystic ovary syndrome

Liansha Huang¹, Xiaoling Wu¹, Shumin Guo¹, Ying Lv¹,
Peng Zhou², Guangrong Huang³, Zuzhen Duan³ and
Wen Sun^{4,5,6*}

¹Department of Reproductive Health, Shenzhen Bao'an Chinese Medicine Hospital, Guangzhou University of Chinese Medicine, Shenzhen, China, ²Department of Acupuncture, Shenzhen Bao'an Chinese Medicine Hospital, Guangzhou University of Chinese Medicine, Shenzhen, China, ³Department of Gynecology, Shenzhen Bao'an Chinese Medicine Hospital, Guangzhou University of Chinese Medicine, Shenzhen, China, ⁴Key Laboratory of Health Cultivation of the Ministry of Education, Beijing University of Chinese Medicine, Beijing, China, ⁵Beijing key Laboratory of Health Cultivation, Beijing University of Chinese Medicine, Beijing, China, ⁶School of Traditional Chinese Medicine, Beijing University of Chinese Medicine, Beijing, China

Background: Polycystic ovary syndrome (PCOS) is a complex disease that afflicts women of reproductive age, and its pathological mechanism has not been well explained. The gut microbiota is believed to be closely related to the development of PCOS. Although an important component of the gut microbiome, the role of the gut virome in the development of PCOS is still unclear.

Methods: In this study, we profiled and compared the gut viral community of 50 patients with PCOS and 43 healthy women based on the analysis of their fecal whole-metagenome dataset.

Results: The gut virome of PCOS patients exhibited a significant decrease in within-sample viral diversity and a remarkable alteration of the overall virome composition compared with that of healthy controls. At the family level, *Siphoviridae* was significantly depleted in the gut virome of patients, while *Quimbyviridae* was enriched. We identified 1,089 viral operational taxonomic units (vOTUs) that differed in relative abundance between the two groups, of which 455 vOTUs were enriched in PCOS patients (including numerous *Bacteroidaceae* phages) and 634 were enriched in controls (including numerous viruses predicted to infect *Oscillospiraceae*, *Prevotellaceae*, and *Ruminococcaceae*). Functional comparison of the PCOS-enriched and control-enriched vOTUs uncovered the viral functional signatures associated with PCOS. Furthermore, we demonstrated gut viral signatures for disease discrimination and achieved an area under the receiver operator characteristic curve (AUC) of 0.938, demonstrating the potential of the gut virome in the prediction of PCOS.

Conclusion: Our findings reveal specific alterations in viral diversity and taxonomic and functional compositions of the gut virome of PCOS patients.

Further studies on the etiology of PCOS and the gut viral community will offer new prospects for treating and preventing PCOS and its related diseases.

KEYWORDS

polycystic ovary syndrome, gut virome, viral dysbiosis, gut bacteriome, whole-metagenome sequencing

Introduction

Polycystic ovary syndrome (PCOS) is a syndromic female endocrine disease characterized by a disorder of androgen secretion (Patel, 2018), which affects 6–20% of women of reproductive age worldwide (Meier, 2018). The syndrome leads to an increased risk of infertility, high blood pressure, obesity, metabolic disorders, and insulin resistance (Gilbert et al., 2018). Presently, many studies have identified a variety of causes of PCOS, including genetic and environmental factors (Goodarzi et al., 2011; Cui et al., 2013; McAllister et al., 2015; Merkin et al., 2016; Kshetrimayum et al., 2019). However, due to the complexity and heterogeneity of this disease, its exact pathogenesis is still not well known. The gut microbiota has been shown to be involved in host energy metabolism, insulin secretion, and inflammatory responses (Visconti et al., 2019; Lee et al., 2020; Ruff et al., 2020), suggesting that it may play an important role in the development and progression of a variety of diseases (Carding et al., 2015; Fan and Pedersen, 2021). Recent studies have highlighted the potential role of the gut microbiota in influencing the onset and development of PCOS (Qi et al., 2019). Mechanistically, the gut microbiota of PCOS patients was characterized by a marked elevation of *Bacteroides vulgatus*, this bacterium could influence the metabolism of bile acids (e.g., glycodeoxycholic acid) and thus reduce the secretion of immune cytokines such as interleukin-22 (IL-22) that are beneficial to PCOS phenotype. The finding of this gut microbiota-bile acid-IL-22 axis also suggests that modulation of gut microbiota could be of great value for the treatment of PCOS (Qi et al., 2019).

As an important component of the gut microbiota, the viral community (referred to as the “virome”) is closely related to the health state of humans (Cadwell, 2015). In the gut, viral biological entities are thought to be in the same range as bacterial populations (Kim et al., 2011; Cadwell, 2015). They consist of a large number of prokaryotic viruses (mainly bacteriophages) and eukaryotic viruses (Shkorporov et al., 2019). The healthy human gut virome is highly specific to the individual and temporally stable (Reyes et al., 2010; Minot et al., 2013; Shkorporov et al., 2019). Gut viruses coevolved with gut bacteria to play an important role in intestinal homeostasis (Virgin, 2014; De Sordi et al., 2017). Recent studies have confirmed that enterovirus populations are associated with autoimmune diseases and inflammatory bowel diseases (Liang et al., 2021). Gut viruses may be the direct cause of diseases or indirectly affect the development of diseases by regulating the structure of the bacteria (Carding

et al., 2017; Beller and Matthijnsens, 2019; Cao et al., 2022). In some diseases where the gut microbiota is seriously maladjusted, variation and maladjustment of the gut virome can also be observed (Santiago-Rodriguez and Hollister, 2019). The structure of the gut bacterial microbiota in PCOS patients is significantly different from that in healthy people (Liu et al., 2017; Torres et al., 2018; Qi et al., 2019), which may also lead to changes in the gut virome. Investigating these changes will help us understand how the gut virome relates to the gut microbiota. With the combination of metagenomic sequencing and virus sequence recognition algorithms, re-mining past metagenomic data may help deepen our understanding of the link between the gut virome and PCOS.

In this study, focusing on the alteration of the gut viral community in PCOS, we reanalyzed the deep fecal whole-metagenome dataset from a recent PCOS study comprising 50 patients and 43 healthy subjects (Qi et al., 2019). We profiled the whole gut virome from fecal metagenomes and identified the variations in gut viral composition and functions between patients and healthy individuals. We also performed correlation analyses between PCOS-associated gut viruses and bacteria. Our findings may contribute to a better understanding of PCOS and the development of novel therapeutic strategies.

Materials and methods

Data acquisition and processing

The metagenomic dataset of 93 samples from 50 PCOS patients and 43 healthy controls was downloaded from the Sequence Read Archive (NCBI-SRA) database under project accession no. PRJNA530971. Patients and controls were matched in their age, body mass index, and waist-hip ratio, as described in the original study (Qi et al., 2019). Quality control of raw metagenomic reads was performed using fastp (Chen et al., 2018). Briefly, low-quality (>45 bases with quality score <20, or >5 “N” bases), low-complexity, and adapter-containing reads were removed, and the remaining reads were trimmed at the tails for low-quality (<Q20) or “N” bases. Human genomic reads were removed by mapping to the reference human genome (GRCh38) using Bowtie2 (Langmead and Salzberg, 2012).

Gut virome profiling and analyses

To quantify the composition of the gut viral community in fecal metagenomic samples, we used a non-redundant gut virus catalog derived from over 10,000 publicly available fecal metagenomes from the Chinese population as a reference. Briefly, raw metagenomic reads were downloaded from public databases and assembled into contigs using MEGAHIT (Li et al., 2015) with the options “--kmer 21,33,55” (for samples with read length < 100 bp or less) or “--kmer 21,33,55,77” (for samples with read length > 100 bp). Potential viral sequences were identified from the contigs based on any of the following criteria: (1) contig whose viral genes were more than the number of microbial genes in CheckV (contigs with length < 10 kbp and estimated as low-quality or undetermined contigs were removed; Nayfach et al., 2021b); (2) contig with value of $p < 0.01$ and score > 0.90 in DeepVirFinder v1.0 (Ren et al., 2017); and (3) contig identified by VIBRANT v1.2.1 (Kieft et al., 2020) with default options. Viral sequences which shared 95% nucleotide identity across 75% of the sequence were then clustered into a viral operational taxonomic unit (vOTU) using the custom scripts. The resulting gut virus catalog contained the genome sequences of a total of 67,096 vOTUs with completeness > 50%, as estimated by CheckV (Nayfach et al., 2021b). Taxonomy annotation and prokaryotic host prediction of the vOTUs were performed according to the methods described in the previous studies (Yan et al., 2021; Li et al., 2022).

High-quality metagenomic reads of the samples of PCOS patients and healthy controls were mapped into the reference gut virus catalog using Bowtie2 with a nucleotide similarity threshold of 95% (Langmead and Salzberg, 2012). The abundance profile of vOTUs in each sample was generated by aggregating the number of reads mapped to each vOTU. The relative abundance of vOTUs was divided by the number of total mapped reads in each sample. The relative abundance profile at the viral family level was generated by aggregating the relative abundance of vOTUs assigned to the same family.

We calculated two diversity indexes to assess the richness and evenness of the vOTU composition for each sample. The number of observed vOTUs was defined as the count of unique vOTUs in each sample. Shannon's diversity index was calculated using the *vegan* package (function *diversity*) in the R platform.

Viral proteins of vOTUs were predicted using Prodigal v2.6.3 (Hyatt et al., 2010). Functional annotation of viral proteins was performed based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (Kanehisa et al., 2017) using DIAMOND with the options “--query-cover 50 --subject-cover 50 -e 1e-5 --min-score 50 --max-target-seqs 50.” Each protein was assigned a KEGG orthology (KO) on the basis of the best-hit gene in the database. The viral auxiliary metabolic genes (AMGs) were identified according to the method described by a previous study (Kieft et al., 2020).

Gut bacteriome profiling

The comprehensive Unified Human Gastrointestinal Genome (UHGG; Almeida et al., 2021) database of the human gut

microbiome was used as a reference for bacteriome profiling of fecal metagenomes. The gut species of the UHGG database were taxonomically annotated using the Genome Taxonomy Database (GTDB; Chaumeil et al., 2019). High-quality metagenomic reads of the samples of PCOS patients and healthy controls were mapped into the UHGG database using Bowtie2 with a nucleotide similarity threshold of 95% (Langmead and Salzberg, 2012). To improve the comparison of samples with vastly different read counts, we randomly subsampled 20,000,000 mapped reads per sample to recalculate the read count of each gut species. The relative abundance of each gut species in every sample was its read count divided by 20,000,000. For genus-level profiles, the relative abundances of species with the same genus were added together to calculate the abundance for the corresponding genus.

Statistical analyses

Statistical analyses were performed based on the R 4.0.1 platform. Principal coordinate analysis (PCoA) was performed with the R *vegan* package, based on the Bray–Curtis dissimilarity, and visualized via the R *ade4* package. Permutational multivariate analysis of variance (PERMANOVA, also known as *adonis* analysis) was performed with the R *vegan* package, and the *adonis* p value was generated based on 1,000 permutations. Random forest models were analyzed with the R *randomForest* package (1,000 trees). The performance of the predictive model was evaluated using receiver operator characteristic (ROC) analysis, which was performed with the R *pROC* package.

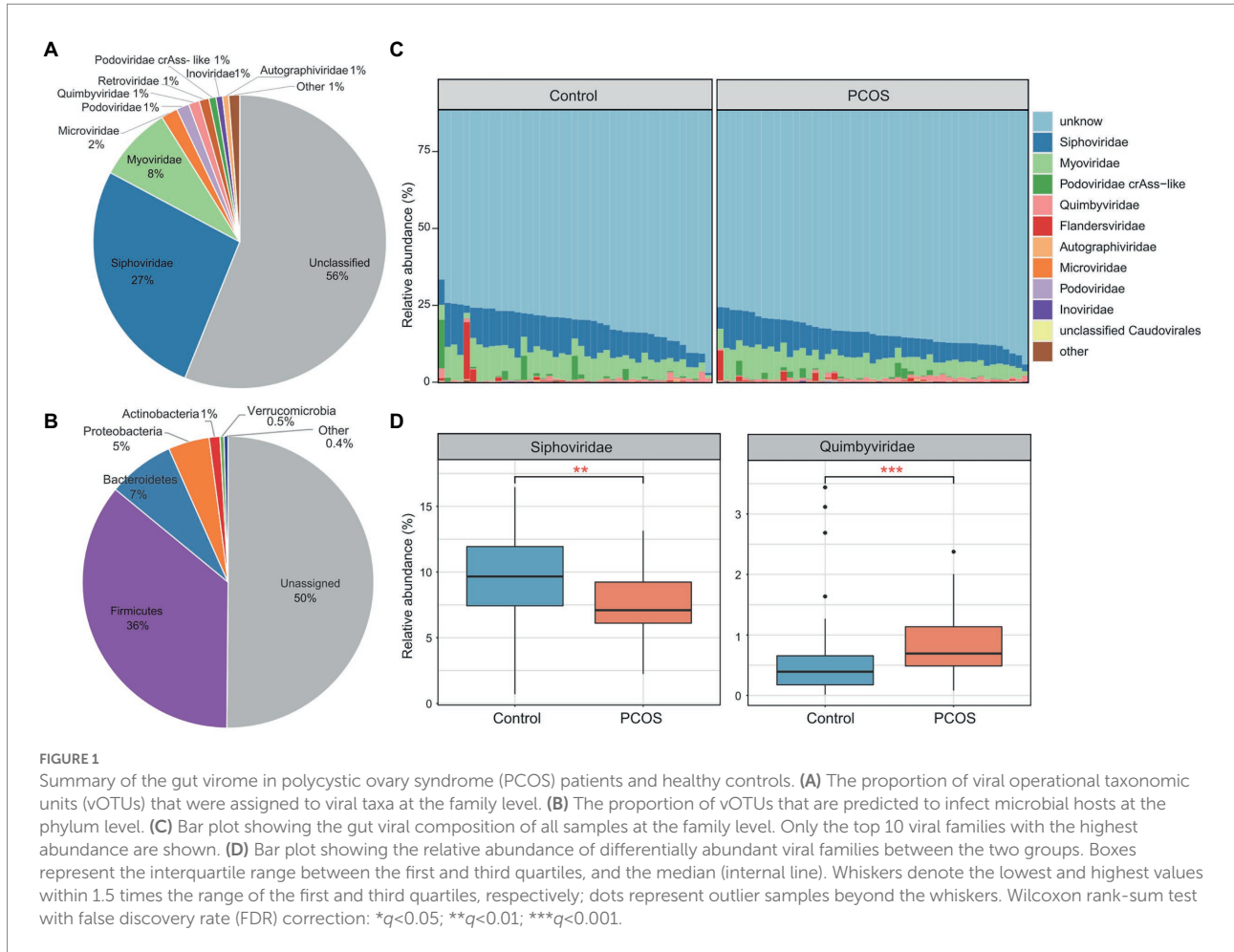
The Wilcoxon rank-sum test was used to measure significant differences between two different groups. The q value was used to evaluate the false discovery rate (FDR) for correction of multiple comparisons and was calculated based on the R *fdrtool* package (Strimmer, 2008), and q -values < 0.05 were considered statistically significant.

Spearman's correlation analysis was used to analyze the correlation between viruses and bacteria. Correlations with an absolute correlation coefficient $p > 0.60$ and Spearman's correlation test $q < 0.05$ were shown in the correlation network. The network was visualized using Cytoscape (Su et al., 2014).

Results

Overview of the gut virome

To investigate the gut viral characteristics in PCOS patients, we analyzed the metagenomic sequencing dataset from fecal samples of 50 patients and 43 healthy individuals (Qi et al., 2019). For all samples, the gut viral compositions were profiled by mapping the metagenomic reads to a non-redundant gut virus catalog constructed from publicly available fecal metagenomes of Chinese populations (see “Materials and methods”). A total of 21,204 vOTUs were observed and quantified in the dataset, which captured, on average, 10.1 ± 3.7 and $9.8\% \pm 4.0\%$ of metagenomic



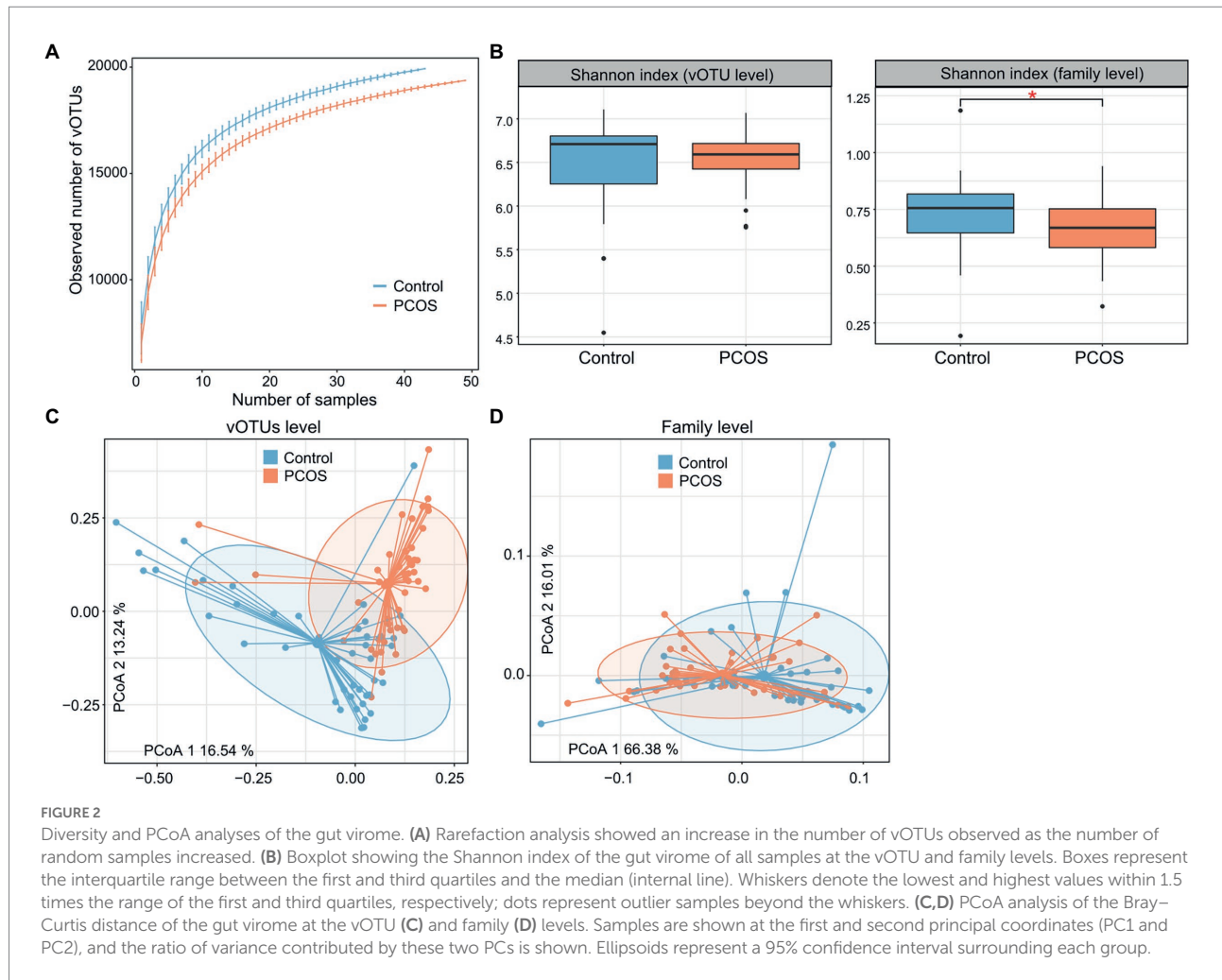
reads from the samples of the patient and control groups, respectively. A total of 43.9% (9,308/21,204) of these vOTUs could be taxonomically assigned into 16 viral families. *Siphoviridae* and *Myoviridae* were dominant families, comprising 27 and 8% of all vOTUs, respectively (Figure 1A). The remaining vOTUs were members of *Microviridae*, *Podoviridae*, *Quimbyviridae*, *Retroviridae*, *Podoviridae_crAss-like*, *Inoviridae*, *Autographiviridae*, and a small number of *unclassified_Caudovirales*, *Gratiaviridae*, *Flandersviridae*, *Drexelviridae*, *Metaviridae*, *Mitoviridae*, and *Herelleviridae* viruses. Additionally, we predicted the prokaryotic hosts of the vOTUs based on the comprehensive UHGG database (Almeida et al., 2021) and CRISPR interval sequences (see Methods). A total of 49.9% (10,572/21,204) of the vOTUs could be assigned to one or more microbial hosts (Figure 1B). Most of the predicted hosts were members of *Firmicutes* (dominated by *Ruminococcaceae* and *Lachnospiraceae* at the family level), *Bacteroidaceae* (*Bacteroidaceae* and *Prevotellaceae*), *Proteobacteria* (mainly *Enterobacteriaceae*), and *Actinobacteria* (mainly *Coriobacteriaceae*).

In the viral compositional profile, an average of only 20% of the total viral sequences could be captured by vOTUs belonging to known viral families (Figure 1C), suggesting that most viruses in the human gut are uncultured and need to be described. As

expected, *Siphoviridae* (average relative abundance $9.5 \pm 3.4\%$) and *Myoviridae* ($7.6 \pm 2.7\%$) were the most dominant families in all metagenomic samples, followed by *Podoviridae_crAss-like* ($1.0 \pm 2.4\%$), *Quimbyviridae* ($0.8 \pm 0.7\%$), and *Flandersviridae* ($0.7 \pm 2.5\%$). Compared with those of the healthy controls, the viral communities of the PCOS patients showed a significant increase in *Quimbyviridae* abundance and a significant decrease in *Siphoviridae* abundance (Wilcoxon rank-sum test, $q < 0.05$; Figure 1D).

Characteristics of the gut virome in PCOS patients

Rarefaction curve analysis showed that (1) the curve was approximately saturated under 20 samples in each group, and (2) the vOTU richness was significantly lower in the PCOS patients than in healthy controls at the same sample sizes ($p = 0.008$; Figure 2A). We then assessed the within-sample diversity of the gut viromes using Shannon's diversity index. PCOS patients showed a slightly but not significantly lower Shannon index than healthy subjects at the vOTU level, whereas, at the viral family level, this index was significantly lower in PCOS patients than in controls (Figure 2B; Supplementary Table S1).



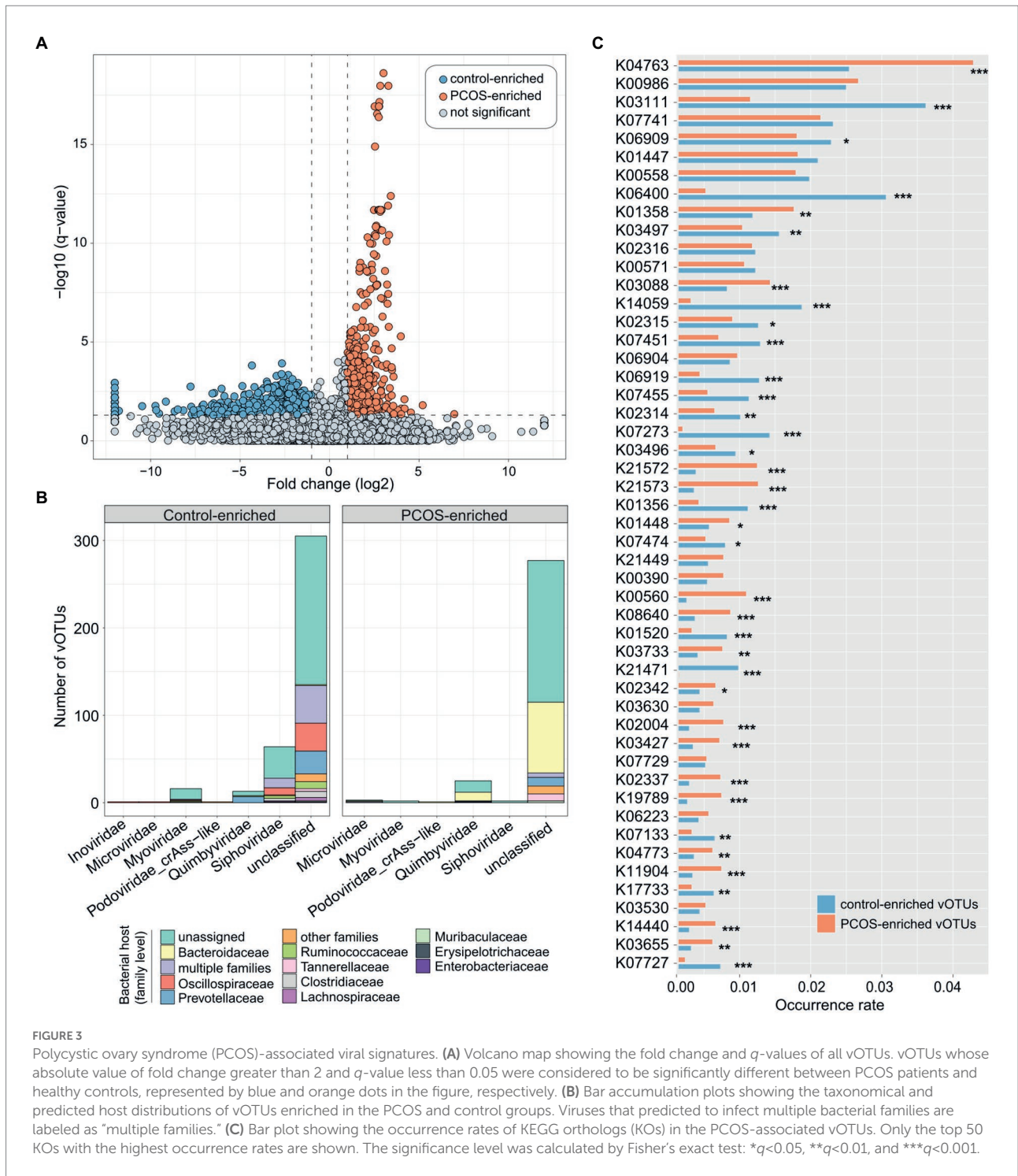
Next, we carried out PCoA based on the Bray–Curtis distance to further investigate the differences in gut viral communities between PCOS patients and healthy controls and revealed a clear separation between them at both the vOTU and family levels (Figures 2C,D). PERMANOVA also supported that the gut virome was significantly different between PCOS patients and healthy controls, with effect sizes of 7.7% ($adonis\ p < 0.001$) and 5.9% ($adonis\ p < 0.001$) at the vOTU and family levels, respectively.

Identification of PCOS-associated gut viral signatures

To explore the gut viral signatures of PCOS, we compared the viral composition between PCOS patients and healthy controls at the vOTU level. A total of 1,089 vOTUs were identified with significant differences in relative abundance between the two groups (Wilcoxon rank-sum test with FDR-correction $q < 0.05$, $|\text{fold-change}| > 2$; Supplementary Table S2). Among these, 455 vOTUs were enriched in the virome of PCOS patients, and 634 vOTUs were enriched in controls (Figure 3A). Most (81.5%) of these PCOS-associated vOTUs could not be classified into known viral families, while the remaining

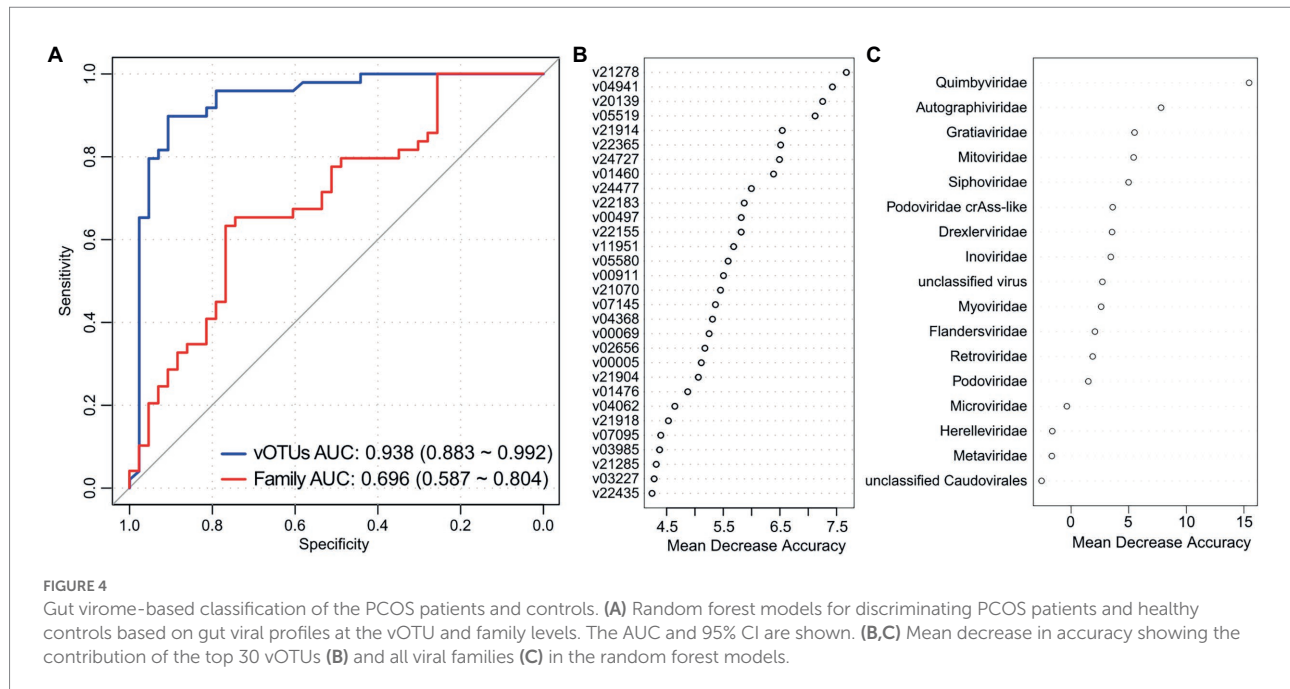
taxonomically annotated vOTUs were members of *Siphoviridae* (containing 113 vOTUs), *Quimbyviridae* (50 vOTUs), *Myoviridae* (23 vOTUs), *Microviridae* (8 vOTUs), *Podoviridae_crAss-like* (5 vOTUs), *Gratiaviridae* (1 vOTU), and *Inoviridae* (1 vOTU; Figure 3B). Thirty-four of 50 PCOS-associated *Quimbyviridae* vOTUs were enriched in the gut virome of PCOS patients compared with those in the gut virome of controls, while only 16 *Quimbyviridae* vOTUs were depleted in patients. Conversely, 108 of 113 PCOS-associated *Siphoviridae* vOTUs and 21 of 23 PCOS-associated *Microviridae* vOTUs were enriched in the gut virome of healthy controls. Regarding the host, most of the differentially abundant gut vOTUs were bacteriophages that were predicted to infect bacterial hosts such as *Bacteroidaceae*, *Oscillospiraceae*, *Prevotellaceae*, and *Ruminococcaceae* (Figure 3B; Supplementary Table S2). The PCOS-enriched vOTUs were concentrated in *Bacteroidaceae* phages (35.2% of 455 PCOS-enriched vOTUs), while the control-enriched vOTUs were more likely to be distributed in *Oscillospiraceae* (10.7% of 634 control-enriched vOTUs) and *Prevotellaceae* (9.5%) phages.

To elucidate the functional potential of the PCOS-associated viruses, we performed functional annotation of the vOTUs based on the KEGG database (Kanehisa et al., 2017). A total of 87,357 protein-coding genes were predicted from 1,089 PCOS-associated



vOTUs, and 16.8% (14,691/87,357) of these genes matched the corresponding KEGG orthologs. We then focused on the top 50 KOs (which represented 41.6% of all annotated genes) with the highest occurrence rates in all PCOS-associated vOTUs. Thirty-seven of 50 KOs had significantly different occurrence rates between PCOA and healthy control groups, of which 18 were more frequent in PCOS-enriched vOTUs and 19 in control-enriched vOTUs (Figure 3C). The PCOS-enriched vOTUs

encoded functions, such as integrase/recombinase (K04763), ATP-dependent protease (K01358), RNA polymerase sigma-70 factor (K03088), starch-binding outer membrane proteins (K21572/K21573), and thymidylate synthase (K00560), while the control-enriched vOTUs encoded more functions involving single-stranded DNA-binding protein (K03111), site-specific DNA recombinase (K06400), chromosome-partitioning protein (K03497), integrase (K14059), 5-methylcytosine-specific



restriction enzyme (K07451), and DNA primase/helicase (K06919; [Supplementary Table S3](#)).

Classification of PCOS state using gut viral signatures

We used the random forest regression model with 5-fold cross-validation to assess the performance of gut viral signatures in recognizing PCOS status. A model trained based on the relative abundances of the PCOS-associated vOTUs obtained high identification ability in classifying PCOS patients and healthy controls, with an AUC of 0.938 (95% CI, 0.993–0.992, [Figure 4A](#)). Several PCOS-enriched vOTUs, such as v21278, v21914, v05519, v20139, and v04941, featured the highest score for the discrimination of patients and controls ([Figure 4B](#)), and the former 4 vOTUs were predicted to be Bacteroidaceae phages. On the other hand, a model trained by all viral families resulted in a significant decline in discrimination power, with an AUC of 0.696 (95% CI, 0.587–0.804). The PCOS-enriched family *Quimbyviridae* had the highest contribution in this model ([Figure 4C](#)). Taken together, these findings suggested that gut viral signatures have the potential for the differentiation of PCOS patients from controls.

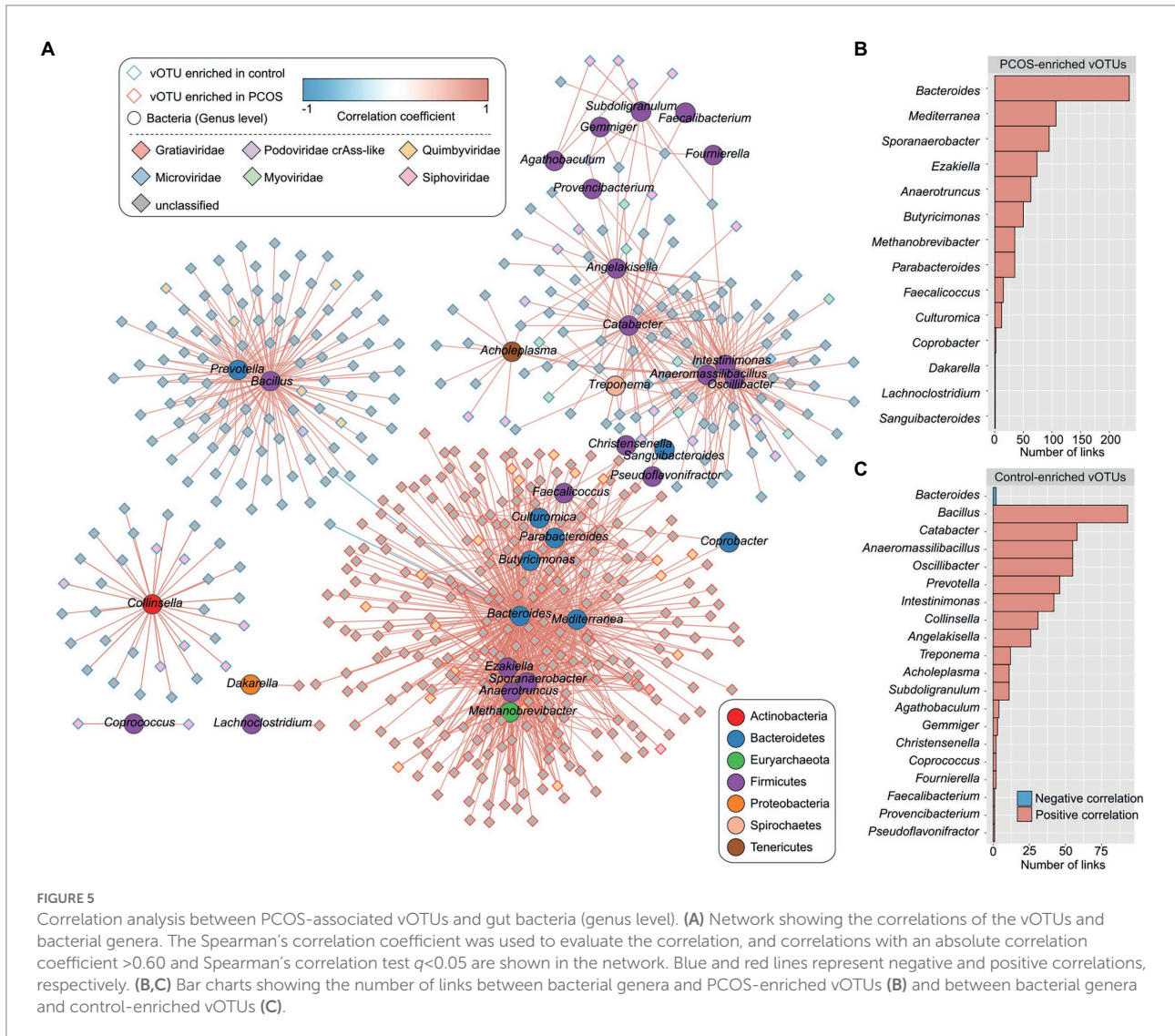
Correlations between gut viral signatures and bacteria

To explore the relationship between viral signatures and bacteria in the gut ecosystem, we mapped the metagenomic sequencing dataset of PCOS patients and healthy controls to the gut prokaryotic UHGG database (see “Materials and methods”) and obtained a genus-level bacterial profile of all samples

(representing 85.8% of metagenomic reads for all samples). Spearman’s correlation coefficient analysis was used to evaluate the interactions between 1,089 PCOS-associated vOTUs and 287 bacterial genera. This analysis revealed a large virus–bacterium interaction network ([Figure 5A](#)) consisting of a total of 1,184 co-occurrence correlations between 514 vOTUs and 33 bacterial genera ([Supplementary Table S4](#)). Several bacterial taxa, such as *Bacteroides*, *Mediterranea*, and *Sporanaerobacter*, were positively correlated with the highest number of PCOS-enriched vOTUs in the network, while other bacteria, such as *Bacillus*, *Catabacter*, *Anaeromassilibacillus*, *Oscillibacter*, *Prevotella*, and *Intestinimonas*, were positively correlated with the highest number of healthy control-enriched vOTUs ([Figures 5B,C](#)); these findings suggest the potential central roles of these bacterial genera in terms of interacting with PCOS-associated viruses.

Discussion

Polycystic ovary syndrome (PCOS) has been described as a serious health problem for women of childbearing age. There is abundant evidence that the gut microbiota contributes significantly to the development of PCOS ([Liu et al., 2017](#); [Torres et al., 2018](#); [Qi et al., 2019](#)). Deep metagenomic sequencing provides an opportunity to explore the links between the alteration of the gut viral community and disease ([Gregory et al., 2020](#); [Camarillo-Guerrero et al., 2021](#); [Nayfach et al., 2021b](#)). In this study, using a metagenomic-based methodology, we characterized the diversity, composition, and functional repertoire of the gut viromes of 50 PCOS patients and 43 healthy women. Multivariate analyses revealed that the overall structure of the gut virome was changed in PCOS patients, with considerably high effect size, consistent with previous observations for the PCOS bacteriome ([Guo et al.,](#)



2016; Liu et al., 2017; Qi et al., 2019; Belkova et al., 2020). Our study strengthened the previous bacteriome study on PCOS (Qi et al., 2019) by adding new information about the virome. Furthermore, we observed significant differences in viral community dysbiosis, taxonomic shifts, and functional changes between PCOS and control viromes.

The within-sample diversity index of the gut virome of PCOS patients was significantly lower than that of healthy women, which may be related to the decrease in the diversity and total amount of bacteria in these patients (Qi et al., 2019). Reduction of gut viral diversity and alteration of virome structure had also been observed in patients with other disorders, such as inflammatory bowel disease (Clooney et al., 2019; Zuo et al., 2019), colorectal cancer (Nakatsu et al., 2018), or liver diseases (Lang et al., 2020). Thus, our findings highlighted a remarkable dysbiosis of the gut virome in PCOS patients.

Siphoviridae, the most dominant viral family of the normal human gut viral community (Zuo et al., 2020; Tomofuji et al., 2021; Lee et al., 2022), was found to be less distributed in the gut virome of PCOS patients than in that of controls. *Siphoviridae* is

a type of double-stranded DNA (dsDNA) virus that is known to be rich in temperate phages and infects a wide range of gut bacteria, such as *Bacillus*, *Enterococcus*, *Clostridium*, *Streptococcus*, *Lactobacillus*, and *Escherichia* (Sekulovic et al., 2014; Adriaenssens et al., 2015; Chehoud et al., 2016). At the vOTU level, 108 *Siphoviridae* vOTUs were identified as enriched in the gut virome of healthy controls and were predicted to infect *Oscillospiraceae*, *Prevotellaceae*, and *Ruminococcaceae*, whereas only five *Siphoviridae* vOTUs were enriched in PCOS patients. On the other hand, the most representative PCOS-enriched viral taxon was *Quimbyviridae*, which accounted for over 2-fold of the relative abundance in the gut virome of PCOS patients and contained 34 vOTUs that were significantly enriched in these patients. *Quimbyviridae* is a recently described viral family with a highly abundant and wide prevalence in the human gut and is suspected to be an obligate lytic phage (Benler et al., 2021). The known bacterial hosts of PCOS-enriched *Quimbyviridae* members were mostly Bacteroidaceae. Collectively, these findings suggested that a low level of *Siphoviridae* viruses and overrepresentation of

Bacteroidaceae-infected *Quimbyviridae* viruses in the gut environment may be considered risk factors for PCOS.

Functional comparison between PCOS-enriched and control-enriched vOTUs identified a large number of viral functions that are potentially associated with the disease. In accordance with previous observations (Nayfach et al., 2020), most of these functions were typical viral functions, such as integrase, recombinase, and RNA polymerase. Additionally, some functions involving viral auxiliary metabolism (e.g., thymidylate synthase K00560) were also identified with a significantly different frequency between the PCOS- and control-enriched vOTUs. Based on this information, however, the potential correlation between viral function and PCOS still needs to be explored.

Our study further provided gut viral biomarkers for PCOS discrimination and achieved an AUC of 0.938 for identifying disease status based on the differentially abundant vOTUs. This discriminatory power was higher than that of the prediction models based on genotypic or phenotypic markers (Cui et al., 2013; McAllister et al., 2015; Kodipalli and Devi, 2021). Thus, the gut virome showed good potential for the prediction and early diagnosis of PCOS; however, systematic investigations of key viral or functional markers identified here might be helpful in the future.

There were limitations in this study. One of these was the lack of information on the individuals' phenotypic factors, such as sex, age, body mass index, diet, and lifestyle. The relationship between host factors and the gut virome is still unclear (Shkoporov et al., 2019); thus, a larger well-described cohort of PCOS patients is needed to further understand the relationship between the gut virome and PCOS. On the other hand, as PCOS is a highly complex and heterogeneous disease, it is currently infeasible to draw any conclusions about the causal relationships between the gut virome and PCOS, and direct experimental studies (e.g., animal model studies) are needed to show the causality of proposed viral or functional signatures.

Conclusion

To our knowledge, this is the first study investigating the gut virome in patients with PCOS. Our findings revealed additional important information on the correlations between the gut viral community and PCOS and extended previous knowledge. Our results suggested that the gut virome is dysregulated in patients with PCOS and may play a potential role in the pathogenesis of the disease. The viral taxonomic signatures, virus–bacteria associations, and functional signatures identified in this study provided potential markers for PCOS prediction and intervention.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

LH and WS proposed the idea, designed the study, and wrote and improved the manuscript. LH, XW, SG, YL, and WS performed the study. LH, PZ, GH, ZD, and WS processed the experimental data and performed the analysis and interpretation of the results. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Science and Technology Innovative Project of Bao'an, Shenzhen (2019JD318) and COVID-19 Research & Application Project, Shenzhen Bao'an Chinese Medicine Developing Fund (2020KJ CX-KTYJ-76).

Acknowledgments

We thank Qiulong Yan (Dalian Medical University) for his technical assistance and professional guidance.

Conflict of interest

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

Publisher's note

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

Supplementary material

The supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2022.951782/full#supplementary-material>

References

- Adriaenssens, E. M., Edwards, R., Nash, J. H. E., Mahadevan, P., Seto, D., Ackermann, H. W., et al. (2015). Integration of genomic and proteomic analyses in the classification of the Siphoviridae family. *Virology* 477, 144–154. doi: 10.1016/j.virol.2014.10.016
- Almeida, A., Nayfach, S., Boland, M., Strozzi, F., Beracochea, M., Shi, Z. J., et al. (2021). A unified catalog of 204,938 reference genomes from the human gut microbiome. *Nat. Biotechnol.* 39, 105–114. doi: 10.1038/s41587-020-0603-3
- Belkova, N., Klimenko, E., Vilson, N., Sambialova, A., Markova, D., Igumnov, I., et al. (2020). Metagenome datasets from women with polycystic ovary syndrome from Irkutsk, eastern Siberia, Russia. *Data Br.* 32:106137. doi: 10.1016/j.dib.2020.106137
- Beller, L., and Matthijssens, J. (2019). What is (not) known about the dynamics of the human gut virome in health and disease. *Curr. Opin. Virol.* 37, 52–57. doi: 10.1016/j.coviro.2019.05.013
- Benler, S., Yutin, N., Antipov, D., Rayko, M., Shmakov, S., Gussow, A. B., et al. (2021). Thousands of previously unknown phages discovered in whole-community human gut metagenomes. *Microbiome* 9, 78. doi: 10.1186/s40168-021-01017-w
- Cadwell, K. (2015). The virome in host health and disease. *Immunity* 42, 805–813. doi: 10.1016/j.immuni.2015.05.003
- Camarillo-Guerrero, L. F., Almeida, A., Rangel-Pineros, G., Finn, R. D., and Lawley, T. D. (2021). Massive expansion of human gut bacteriophage diversity. *Cell* 184, 1098–1109.e9. doi: 10.1016/j.cell.2021.01.029
- Cao, Z., Sugimura, N., Burgermeister, E., Ebert, M. P., Zuo, T., and Lan, P. (2022). The gut virome: A new microbiome component in health and disease. *EBioMedicine* 81:104113. doi: 10.1016/j.ebiom.2022.104113
- Carding, S., Verbeke, K., Vipond, D. T., Corfe, B. M., and Owen, L. J. (2015). Dysbiosis of the gut microbiota in disease. *Microb. Ecol. Health Dis.* 26:26191. doi: 10.3402/mehd.v26.26191
- Carding, S. R., Davis, N., and Hoyles, L. (2017). Review article: the human intestinal virome in health and disease. *Aliment. Pharmacol. Ther.* 46, 800–815. doi: 10.1111/apt.14280
- Chaumeil, P. A., Mussig, A. J., Hugenholtz, P., and Parks, D. H. (2019). GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. *Bioinformatics*. doi: 10.1093/bioinformatics/btz848
- Chehoud, C., Dryga, A., Hwang, Y., Nagy-Szakal, D., Hollister, E. B., Luna, R. A., et al. (2016). Transfer of viral communities between human individuals during fecal microbiota transplantation. *MBio* 7:e00322. doi: 10.1128/mBio.00322-16
- Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890. doi: 10.1093/bioinformatics/bty560
- Clooney, A. G., Sutton, T. D. S., Shkorporov, A. N., Holohan, R. K., Daly, K. M., O'Regan, O., et al. (2019). Whole-Virome analysis sheds light on viral dark matter in inflammatory bowel disease. *Cell Host Microbe* 26, 764–778.e5. doi: 10.1016/j.chom.2019.10.009
- Cui, L., Zhao, H., Zhang, B., Qu, Z., Liu, J., Liang, X., et al. (2013). Genotype-phenotype correlations of PCOS susceptibility SNPs identified by GWAS in a large cohort of Han Chinese women. *Hum. Reprod.* 28, 538–544. doi: 10.1093/humrep/des424
- De Sordi, L., Khanna, V., and Debarbieux, L. (2017). The gut microbiota facilitates drifts in the genetic diversity and infectivity of bacterial viruses. *Cell Host Microbe* 22, 801–808.e3. doi: 10.1016/j.chom.2017.10.010
- Fan, Y., and Pedersen, O. (2021). Gut microbiota in human metabolic health and disease. *Nat. Rev. Microbiol.* 19, 55–71. doi: 10.1038/s41579-020-0433-9
- Gilbert, E. W., Tay, C. T., Hiam, D. S., Teede, H. J., and Moran, L. J. (2018). Comorbidities and complications of polycystic ovary syndrome: an overview of systematic reviews. *Clin. Endocrinol.* 89, 683–699. doi: 10.1111/cen.13828
- Goodarzi, M. O., Dumesic, D. A., Chazenbalk, G., and Azziz, R. (2011). Polycystic ovary syndrome: etiology, pathogenesis and diagnosis. *Nat. Rev. Endocrinol.* 7, 219–231. doi: 10.1038/nrendo.2010.217
- Gregory, A. C., Zablocki, O., Zayed, A. A., Howell, A., Bolduc, B., and Sullivan, M. B. (2020). The gut Virome database reveals age-dependent patterns of Virome diversity in the human gut. *Cell Host Microbe* 28, 724–740.e8. doi: 10.1016/j.chom.2020.08.003
- Guo, Y., Qi, Y., Yang, X., Zhao, L., Wen, S., Liu, Y., et al. (2016). Association between polycystic ovary syndrome and gut microbiota. *PLoS One* 11:e0153196. doi: 10.1371/journal.pone.0153196
- Hyatt, D., Chen, G.-L., LoCasio, P. F., Land, M. L., Larimer, F. W., and Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinform.* 11, 1–11. doi: 10.1186/1471-2105-11-119
- Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y., and Morishima, K. (2017). KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* 45, D353–D361. doi: 10.1093/nar/gkw1092
- Kieft, K., Zhou, Z., and Anantharaman, K. (2020). VIBRANT: automated recovery, annotation and curation of microbial viruses, and evaluation of viral community function from genomic sequences. *Microbiome* 8:90. doi: 10.1186/s40168-020-00867-0
- Kim, M.-S., Park, E.-J., Roh, S. W., and Bae, J.-W. (2011). Diversity and abundance of single-stranded DNA viruses in human feces. *Appl. Environ. Microbiol.* 77, 8062–8070. doi: 10.1128/AEM.06331-11
- Kodipalli, A., and Devi, S. (2021). Prediction of PCOS and mental health using fuzzy inference and SVM. *Front. Public Health* 9:789569. doi: 10.3389/fpubh.2021.789569
- Kshetrimayum, C., Sharma, A., Mishra, V. V., and Kumar, S. (2019). Polycystic ovarian syndrome: environmental/occupational, lifestyle factors: An overview. *J. Turk. Ger. Gynecol. Assoc.* 20, 255–263. doi: 10.4274/jtgga.galenos.2019.2018.0142
- Lang, S., Demir, M., Martin, A., Jiang, L., Zhang, X., Duan, Y., et al. (2020). Intestinal virome signature associated with severity of nonalcoholic fatty liver disease. *Gastroenterology* 159, 1839–1852. doi: 10.1053/j.gastro.2020.07.005
- Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923
- Lee, C. J., Sears, C. L., and Maruthur, N. (2020). Gut microbiome and its role in obesity and insulin resistance. *Ann. N. Y. Acad. Sci.* 1461, 37–52. doi: 10.1111/nyas.14107
- Lee, C. Z., Zoqratt, M. Z. H. M., Phipps, M. E., Barr, J. J., Lal, S. K., Ayub, Q., et al. (2022). The gut virome in two indigenous populations from Malaysia. *Sci. Rep.* 12, 1–10. doi: 10.1038/s41598-022-05656-3
- Li, D., Liu, C. M., Luo, R., Sadakane, K., and Lam, T. W. (2015). MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 31, 1674–1676. doi: 10.1093/bioinformatics/btv033
- Li, S., Guo, R., Zhang, Y., Li, P., Chen, F., Wang, X., et al. (2022). A catalogue of 48, 425 nonredundant viruses from oral metagenomes expands the horizon of the human oral virome. *iScience* 25:104418. doi: 10.1016/j.isci.2022.104418
- Liang, G., Cobián-Güemes, A. G., Albenberg, L., and Bushman, F. (2021). The gut virome in inflammatory bowel diseases. *Curr. Opin. Virol.* 51, 190–198. doi: 10.1016/j.coviro.2021.10.005
- Liu, R., Zhang, C., Shi, Y., Zhang, F., Li, L., Wang, X., et al. (2017). Dysbiosis of gut microbiota associated with clinical parameters in polycystic ovary syndrome. *Front. Microbiol.* 8, 324. doi: 10.3389/fmicb.2017.00324
- McAllister, J. M., Legro, R. S., Modi, B. P., and Strauss, J. F. (2015). Functional genomics of PCOS: from GWAS to molecular mechanisms. *Trends Endocrinol. Metab.* 26, 118–124. doi: 10.1016/j.tem.2014.12.004
- Meier, R. K. (2018). Polycystic ovary syndrome. *Nurs. Clin. North Am.* 53, 407–420. doi: 10.1016/j.cnur.2018.04.008
- Merkin, S. S., Phy, J. L., Sites, C. K., and Yang, D. (2016). Environmental determinants of polycystic ovary syndrome. *Fertil. Steril.* 106, 16–24. doi: 10.1016/j.fertnstert.2016.05.011
- Minot, S., Bryson, A., Chehoud, C., Wu, G. D., Lewis, J. D., and Bushman, F. D. (2013). Rapid evolution of the human gut virome. *Proc. Natl. Acad. Sci. U. S. A.* 110, 12450–12455. doi: 10.1073/pnas.1300833110
- Nakatsu, G., Zhou, H., Wu, W. K. K., Wong, S. H., Coker, O. O., Dai, Z., et al. (2018). Alterations in enteric Virome are associated with colorectal cancer and survival outcomes. *Gastroenterology* 155, 529–541.e5. doi: 10.1053/j.gastro.2018.04.018
- Nayfach, S., Camargo, A. P., Schulz, F., Eloe-Fadros, E., Roux, S., and Kyrpides, N. C. (2020). Check V assesses the quality and completeness of metagenome-assembled viral genomes. *Nat. Biotechnol.* 39, 578–585. doi: 10.1038/s41587-020-00774-7
- Nayfach, S., Paez-Espino, D., Call, L., Low, S. J., Sberro, H., Ivanova, N. N., et al. (2021b). Metagenomic compendium of 189,680 DNA viruses from the human gut microbiome. *Nat. Microbiol.* 6, 960–970. doi: 10.1038/s41564-021-00928-6
- Patel, S. (2018). Polycystic ovary syndrome (PCOS), an inflammatory, systemic, lifestyle endocrinopathy. *J. Steroid Biochem. Mol. Biol.* 182, 27–36. doi: 10.1016/j.jsbmb.2018.04.008
- Qi, X., Yun, C., Sun, L., Xia, J., Wu, Q., Wang, Y., et al. (2019). Gut microbiota-bile acid-interleukin-22 axis orchestrates polycystic ovary syndrome. *Nat. Med.* 25, 1225–1233. doi: 10.1038/s41591-019-0509-0
- Ren, J., Ahlgren, N. A., Lu, Y. Y., Fuhrman, J. A., and Sun, F. (2017). VirFinder: a novel k-mer based tool for identifying viral sequences from assembled metagenomic data. *Microbiome* 5:69. doi: 10.1186/s40168-017-0283-5
- Reyes, A., Haynes, M., Hanson, N., Angly, F. E., Heath, A. C., Rohwer, F., et al. (2010). Viruses in the faecal microbiota of monozygotic twins and their mothers. *Nature* 466, 334–338. doi: 10.1038/nature09199

- Ruff, W. E., Greiling, T. M., and Kriegel, M. A. (2020). Host-microbiota interactions in immune-mediated diseases. *Nat. Rev. Microbiol.* 18, 521–538. doi: 10.1038/s41579-020-0367-2
- Santiago-Rodriguez, T. M., and Hollister, E. B. (2019). Human Virome and disease: high-throughput sequencing for virus discovery, identification of phage-bacteria Dysbiosis and development of therapeutic approaches with emphasis on the human gut. *Viruses* 11:656. doi: 10.3390/v11070656
- Sekulovic, O., Garneau, J. R., Neron, A., and Fortier, L. C. (2014). Characterization of temperate phages infecting *Clostridium difficile* isolates of human and animal origins. *Appl. Environ. Microbiol.* 80, 2555–2563. doi: 10.1128/AEM.00237-14
- Shkoporov, A. N., Clooney, A. G., Sutton, T. D. S., Ryan, F. J., Daly, K. M., Nolan, J. A., et al. (2019). The human gut Virome is highly diverse, stable, and individual specific. *Cell Host Microbe* 26, 527–541.e5. doi: 10.1016/j.chom.2019.09.009
- Strimmer, K. (2008). Fdrtool: a versatile R package for estimating local and tail area-based false discovery rates. *Bioinformatics* 24, 1461–1462. doi: 10.1093/bioinformatics/btn209
- Su, G., Morris, J. H., Demchak, B., and Bader, G. D. (2014). Biological network exploration with Cytoscape 3. *Curr. Protoc. Bioinformatics* 47, 8.13.1–8.13.24. doi: 10.1002/0471250953.bi0813s47
- Tomofuji, Y., Kishikawa, T., Maeda, Y., Ogawa, K., Nii, T., Okuno, T., et al. (2021). Whole gut virome analysis of 476 Japanese revealed a link between phage and autoimmune disease. *Ann. Rheum. Dis.* 81, 278–288. doi: 10.1136/annrheumdis-2021-221267
- Torres, P. J., Siakowska, M., Banaszewska, B., Pawelczyk, L., Duleba, A. J., Kelley, S. T., et al. (2018). Gut microbial diversity in women With polycystic ovary syndrome correlates With Hyperandrogenism. *J. Clin. Endocrinol. Metab.* 103, 1502–1511. doi: 10.1210/jc.2017-02153
- Virgin, H. W. (2014). The virome in mammalian physiology and disease. *Cell* 157, 142–150. doi: 10.1016/j.cell.2014.02.032
- Visconti, A., Le Roy, C. I., Rosa, F., Rossi, N., Martin, T. C., Mohney, R. P., et al. (2019). Interplay between the human gut microbiome and host metabolism. *Nat. Commun.* 10:4505. doi: 10.1038/s41467-019-12476-z
- Yan, Q., Wang, Y., Chen, X., Jin, H., Wang, G., Guan, K., et al. (2021). Characterization of the gut DNA and RNA Viromes in a cohort of Chinese residents and visiting Pakistanis. *Virus Evol.* 7:veab022. doi: 10.1093/ve/veab022
- Zuo, T., Lu, X. J., Zhang, Y., Cheung, C. P., Lam, S., Zhang, F., et al. (2019). Gut mucosal virome alterations in ulcerative colitis. *Gut* 68, 1169–1179. doi: 10.1136/gutjnl-2018-318131
- Zuo, T., Sun, Y., Wan, Y., Yeoh, Y. K., Zhang, F., Cheung, C. P., et al. (2020). Human-gut-DNA Virome variations across geography, ethnicity, and urbanization. *Cell Host Microbe* 28, 741–751.e4. doi: 10.1016/j.chom.2020.08.005