

# Alterations of gut viral signals in atrial fibrillation: complex linkage with gut bacteriome

Kun Zuo<sup>1,\*</sup>, Jing Li<sup>1,\*</sup>, Chen Fang<sup>1</sup>, Jiuchang Zhong<sup>1</sup>, Li Xu<sup>1</sup>, Xinchun Yang<sup>1</sup>

<sup>1</sup>Heart Center and Beijing Key Laboratory of Hypertension, Beijing Chaoyang Hospital, Capital Medical University, Beijing 100020, China

\*Equal contribution

**Correspondence to:** Li Xu, Xinchun Yang; **email:** [imxuli@hotmail.com](mailto:imxuli@hotmail.com), <https://orcid.org/0000-0001-6406-9806>; [yxc6229@ccmu.edu.cn](mailto:yxc6229@ccmu.edu.cn), <https://orcid.org/0000-0001-9813-8347>

**Keywords:** atrial fibrillation, gut viral signals, metagenome, virus-bacteria linkages, predictive model

**Received:** April 28, 2022

**Accepted:** July 12, 2022

**Published:** August 18, 2022

**Copyright:** © 2022 Zuo et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/3.0/) (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

## ABSTRACT

The gut microbiota has a known complex association with atrial fibrillation (AF) progression, but the association of gut viruses with AF is undefined. Metagenomic data in a cohort of 50 AF patients and 50 matched controls were examined to profile the gut viral signals and determine their associations with intestinal bacteria and the AF phenotype. The gut viral alterations were examined, and the marked elevation of viral diversity, including increased Simpson, Shannon, and Pielou index, was revealed in AF patients. The specific alteration of the intestinal viral population, such as overgrowth of *Streptococcus virus DT1* and *Pseudomonas phage*, as well as imbalanced gut viral function, dominated by integral component of the membrane, and metal ion binding were detected in AF patients. Moreover, regarding co-occurrence networks connecting viruses and bacterial organisms, increasingly disordered virus-bacteria linkages were seen in AF cases with severe AF progression. Notably, the associations of *Synechococcus phage S-SM1* and *Cronobacter phage CR5* with bacterial species were very tight in control individuals but markedly dampened in AF cases. Furthermore, the viral score built by the selected discriminative taxa between AF cases with or without recurrence after ablation was still significantly associated with recurrence (HR = 2.959,  $P = 0.0085$ ), with a survival AUC of 0.878. We demonstrated for the first time that gut viral signatures are associated with AF, and suppressed viral-bacterial associations in AF suggest the gut virus might participate in AF progression, which has a potential value in predicting ablation outcomes.

## INTRODUCTION

Atrial fibrillation (AF), one of the most common arrhythmias health and socioeconomic burdens, has complex pathophysiology and significantly contributes to enhanced morbidity and mortality [1]. Recent evidence suggests a regulatory role for the gut bacteria in human health, including the influence on host metabolism and immune homeostasis in several cardiovascular diseases [2] such as hypertension [3], heart failure [4], atherosclerosis [5], and AF [6–8]. Notably, many AF patients (approximately 40%)

suffered from gastrointestinal comorbidities, predominantly dyspepsia [9]. Therefore, the role of the gut environment or microbiome deserves more attention in the context of AF pathogenesis.

However, researchers are still striving to determine what shapes the microbial community. Some changes are driven by the powerful external influence of the environment and lifestyle [5]. However, the internal micro-environment of gut bacteria is the gut microbiome, a complex ecosystem that contains communities of bacteria, archaea, unicellular eukaryotes, multi-

cellular eukaryotes, and viruses. Sharing the same host niches, these communities compete against, synergize with, and/or antagonize each other, potentially impacting the host [10].

With recent progress in high-throughput sequencing technology and data mining strategies, unculturable viral communities should not be underestimated as “dark matter”, whose number likely far surpasses those of bacterial populations [11–13]. The intestinal virome in healthy individuals includes eukaryote-infecting viruses and bacteriophages. Intestinal bacteriophages are substantially diverse among individuals and temporarily stable [11]. These viruses latently infect bacteria and co-evolve with intestinal bacteria, shaping bacterial ecology and acting as a driving factor in bacterial diversification and community composition through predation and horizontal gene transfer [10].

Currently, few reports have assessed the gut virome in AF patients. To examine the regulatory roles of intestinal viruses, it is urgent to determine their composition and functions. Hence, in this study, we performed analyses based on our previously published metagenomic dataset about AF, addressing the potential changes in virus composition and functions and the subsequent effects on clinical AF progression. More importantly, we proposed to reveal the importance of assessing the virus as an essential part of the microbiome ecological network based on the trans-kingdom interaction between virus and bacteria. In addition, this study evaluated the viral signal for its potential diagnostic potential in recurrent AF prediction after catheter ablation.

## RESULTS

### Data processing and taxonomic profiles of the gut viral signals in AF

Based on the metagenomic assembly, 857,736 contigs were obtained, with an average N50 length of 4,123 bp, ranging between 1,000 and 348,062 bp. Briefly, the comparison of bacterial and viral reads in non-AF controls and AF showed an elevation of relative proportions of the virome in the gut environment, with an average of 3,451,194.14 and 3,113,567.34 viral reads in AF cases and non-AF controls, respectively, and 49,676,139.2 and 44,817,408 bacterial-reads, respectively (Supplementary Figure 1A, 1B). Altogether, viral taxa, including 13 orders, 128 families, 756 genera, and 4,168 species, were annotated in the Kraken database, and relative abundance levels were calculated (Supplementary Figure 1C, 1D).

Virome differences could be driven by vertebrate viruses or bacteriophages. Next, the virome was examined at the order level for similarity to eukaryotic viruses or bacteriophages, based on the viral host information downloaded from the international committee on taxonomy of viruses (ICTV). Globally, 53.85% of all viral orders were known vertebrate-infecting viruses, and 15.38% were matched to bacteriophages’ reference genomes. Notably, AF was associated with a tendency towards the enrichment of phages (Supplementary Figure 1E). The above data indicated dysbiosis in bacteriophages is important in AF, with some vertebrate-infecting viruses showing an increased abundance in non-AF individuals, pointing to a potential linkage between bacterial dysbiosis and bacteriophage expansion in AF.

Further, we assessed whether these community bacteriophages were mostly lytic or temperate by detecting three markers of temperate phages in the (1) bacteria reference genome from the NCBI database, (2) prophage genes from the ACLAME database, and (3) Uniprot [14]. The abundances of lysogenic phages were calculated based on the species level, and we found that the proportion of lysogenic phages showed an increasing tendency throughout non-AF controls, PAF, and psAF (Supplementary Figure 1F). These findings might suggest that the intestinal virus mainly comprises temperate bacteriophages.

### Viral diversity in AF

Rarefaction analysis was carried out by enumerating viral contigs (Figure 1A) and viruses (Figure 1B) in sample pairs, and the results showed that the accumulation curves for the totality of specimens nearly plateaued, suggesting most viral contigs and viruses could be detected. Specifically, the scatter plot of nonmetric dimensional scaling (NMDS) based on the abundances of viruses at the species level separated control and AF patients ( $P = 0.0403$  for NMDS1;  $P = 1.25e-14$  for NMDS2, Figure 1C), indicating different gut viral structures between the two groups. Moreover,  $\alpha$  diversity (reflected by the Shannon index, Simpson index, and Pielou evenness at the species level) was markedly elevated in AF cases compared with control individuals ( $P = 0.0953$  for the number of viruses, Figure 1D;  $P = 1.74e-08$  for the Simpson index, Figure 1E;  $P = 7.05e-07$  for the Shannon index, Figure 1F;  $P = 2.05e-07$  for Pielou evenness, Figure 1G;  $P = 0.0953$  for the Chao1 index, Figure 1H).

To address whether the observed alteration of viral signals was influenced by traditional AF risk factors, multivariate linear regression analysis for age, BMI,

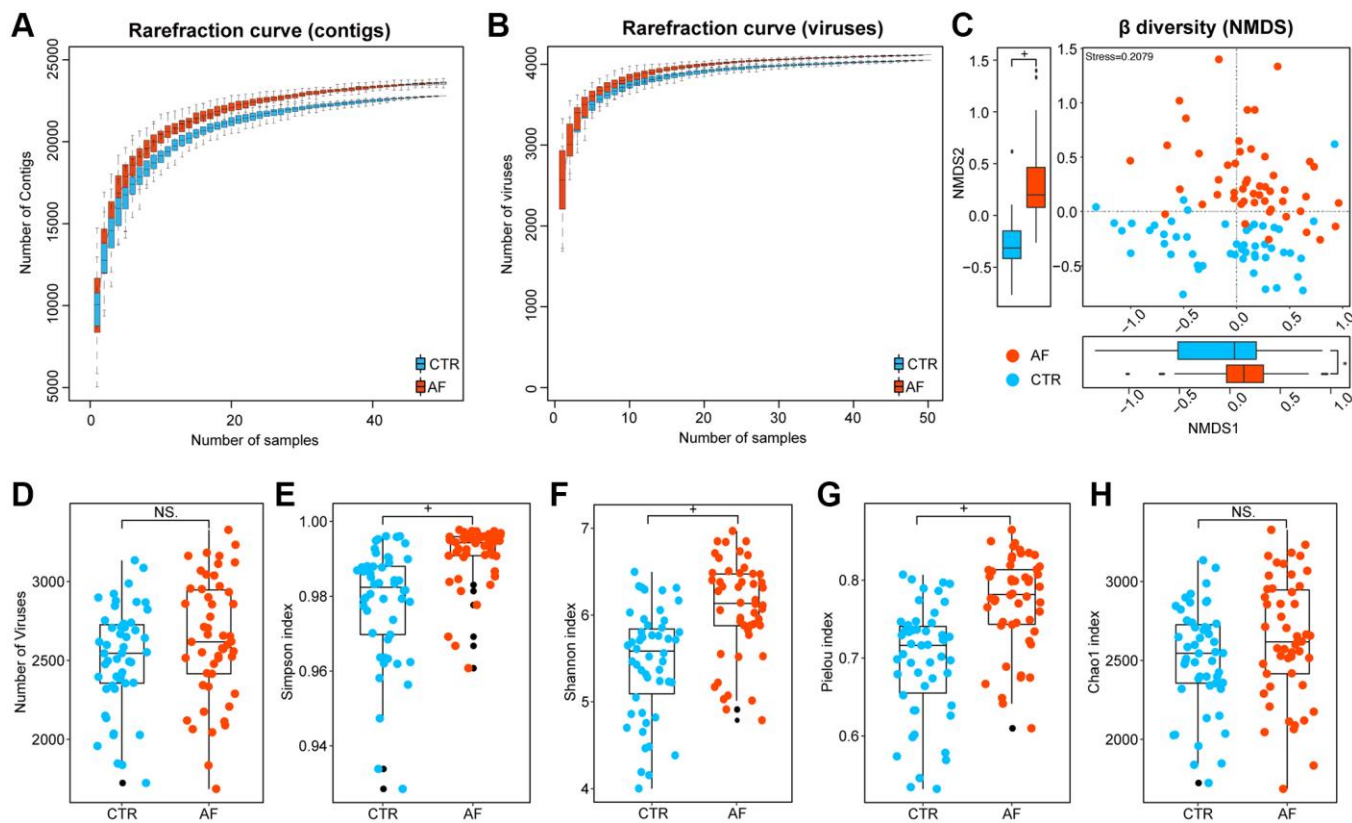
HTN, T2DM, TC, and medication was performed. Also, the independent strength of association between AF and the viral signatures of alpha diversity was examined. The results showed that AF was associated with increased viral Shannon index, Pielou evenness, and Simpson index independent of age, gender, BMI, TC, HTN, T2DM, or medication (Supplementary Table 1, Supplementary Figure 2). Therefore, it was concluded that the contribution of confounders to disordered viral signals was less than that of AF.

### Altered gut viral composition in AF

DESeq analysis was performed to identify taxa differentially with the standard of adjusted  $p < 0.05$  and  $|\text{Log}_2(\text{fold change})| > 2$  between controls and AF at the family, genus, and species levels. A total of 10 families, 134 genera, and 1,173 species were identified as differential taxa between non-AF controls and AF cases, and more taxa, including 7 families, 80 genera, and 694 species, were found to be significantly enriched in AF cases than in control patients (Figure 2A, Supplementary Table 2). There was a higher abundance of families such as *Secoviridae* and *Fimoviridae* in the gut of AF patients (Figure 2B).

### Functional changes in the intestinal viral signals in AF

To examine gut viral functions, a HUMANN2-based assessment of reads of viral contigs in the GO, Pfam protein family, and the Kyoto Encyclopedia of Genes and Genomes (KEGG) databases was conducted. We compared the predicted gut viral functions between non-AF controls and AF cases. Interestingly, AF patients had altered viral functions, with 505 GO (Figure 3A), 518 Pfam protein functions (Figure 3B), and 17 KEGG pathways (Figure 3C; Supplementary Table 3), where  $P < 0.05$  was considered statistically significant based on limma analysis. Of these, GO terms such as integral component of the membrane, cytoplasm, ATP-binding, DNA binding, and metal ion binding (Figure 3D), Pfam proteins such as ABC transporter, Histidine kinase-, and DNA gyrase B- (Figure 3E), as well as KEGG pathways such as Glycerophospholipid metabolism, Purine metabolism, Pyrimidine metabolism, RNA polymerase (Figure 3F), were the prominent proteins/functions, indicating the essential roles of viruses in virus-host interactions and the potential interconnection of gut viruses and bacterial organisms.



**Figure 1. Viral diversity in AF.** Rarefaction curves depicting viral contigs (A) and virus (B) numbers in control and AF specimens. Beta diversity by nonmetric dimensional scaling (NMDS) at the species level (C). Alpha diversity indexes, including the number of viruses (D), the Simpson index (E), the Shannon index (F), the Pielou index (G), and the Chao 1 index (H) at the species level in controls and AF cases. + denotes  $p$ -value  $< 0.01$ , \* $p < 0.05$ , and NS indicates no significant difference based on Wilcoxon rank-sum test.

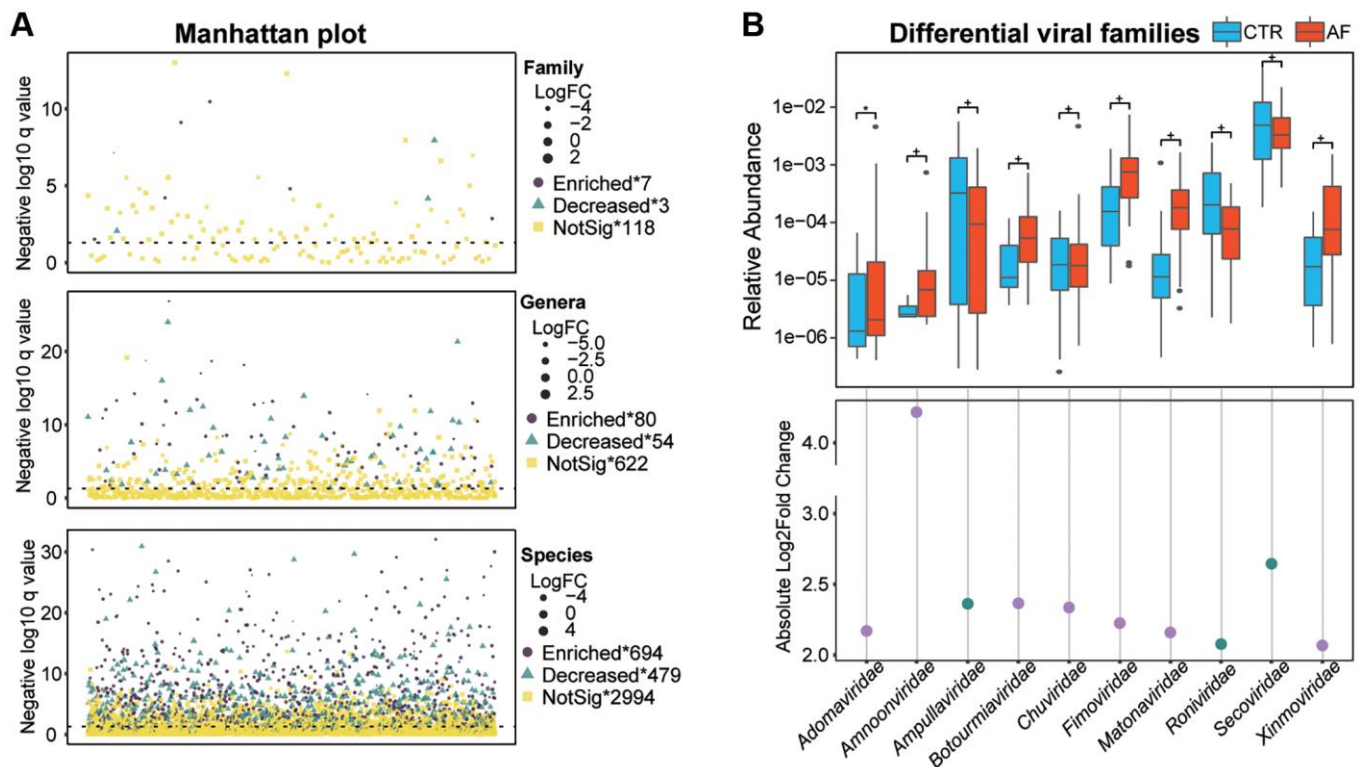
## Virus-bacteria co-occurrence networks

Virus-bacteria co-occurrence network analysis was carried out to determine the associations between viruses and bacterial organisms. We found that *Prevotella copri* CAG:164 and *Prevotella copri* were more abundant in non-AF control patients, and multiple viruses, including *Hepacivirus G*, were negatively associated with this group (Figure 4A). We also observed that *Bacteroidetes*, including *Bacteroides vulgatus*, *Bacteroides egerthii*, and *Bacteroides fragilis*, were positively correlated with *Synechococcus phage S-SM1* (enriched in control patients), suggesting *Synechococcus phage S-SM1* could infect these species (Figure 4B).

Structurally, the network of non-AF controls had 116 nodes, 183 edges, and 28 linkages, while that of AF cases encompassed 88 nodes, 117 edges, and 31 linkages (Supplementary Table 4). The densities of the networks in non-AF controls and AF cases were 0.0274 and 0.0306, respectively. The degrees of network distribution differed between the non-AF controls and AF cases (Figure 4C). Despite the higher number of linkages in the network of AF cases, its natural

connectivity was reduced than that of the network generated in non-AF controls (Figure 4D). Next, entropies were assessed for these networks, and the control network (1.7138) showed an elevated value compared with that of AF cases (1.4171), indicating reduced randomness of potential virus-bacteria interactions, while the linkage of viruses to bacteria was increased. Specifically, virus-bacteria linkages increased from 28 in control patients to 31 in AF cases (Figure 4E), suggesting the linkage is pervasive across non-AF individuals to AF patients.

To characterize the relationship between gut bacteriome and virome, the association of the  $\alpha$  diversity parameters of the bacteriome and the virome. In control patients, significant correlations were found between intra-kingdom and trans-kingdom  $\alpha$  diversity parameters. However, associations of intra- and trans-kingdom  $\alpha$  diversity parameters were attenuated in the AF group, being more severe from PAF to Pers AF <12 m and to Pers >12 m (Supplementary Figure 3A), indicating elevated dysbiosis of the gut microbiome in AF. Furthermore, the associations of bacterial species with viruses were examined in non-AF controls and AF patients. The altered virus-bacteria associations in AF

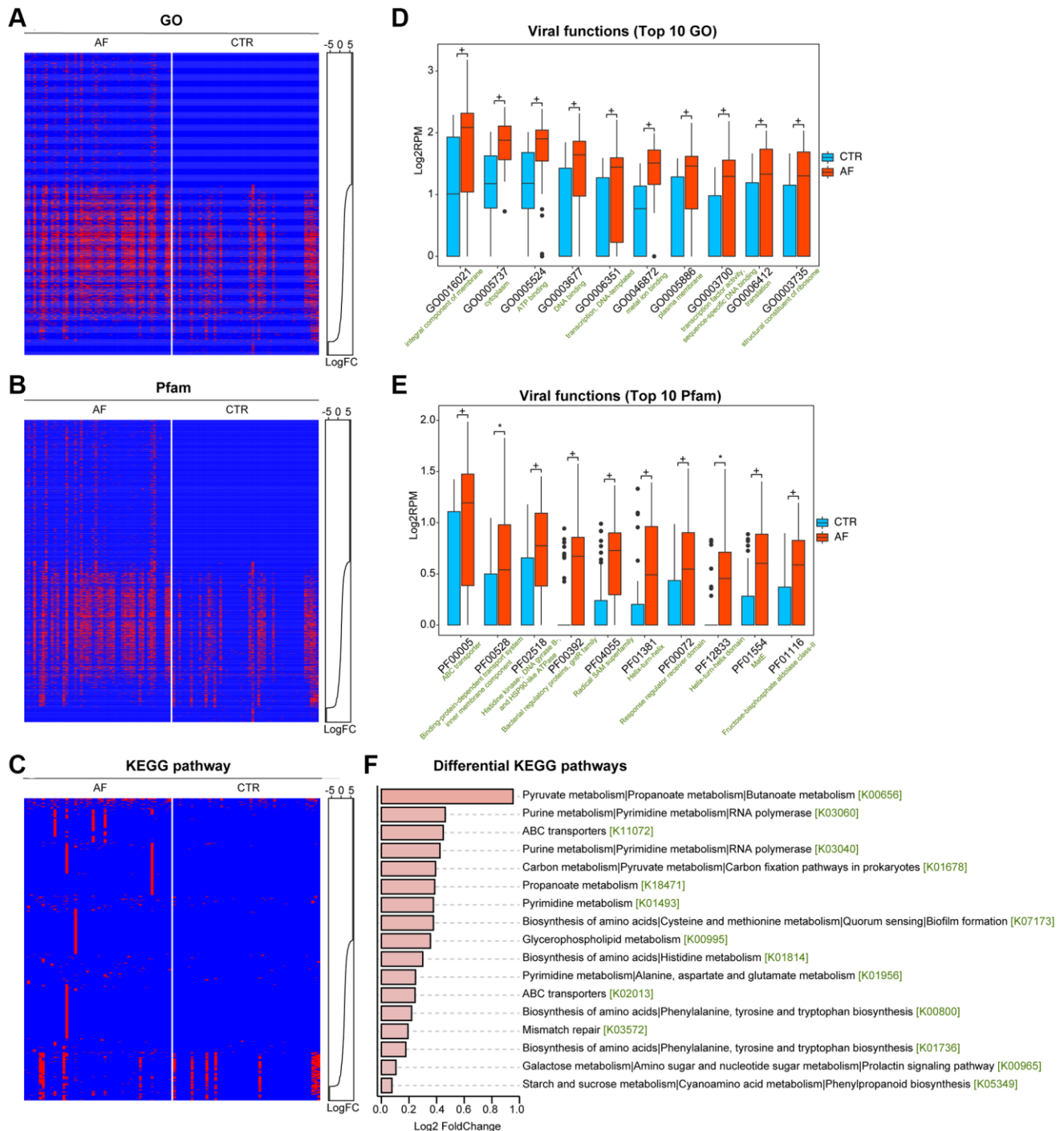


**Figure 2. Altered gut viral composition in AF.** The Manhattan plot shows an overview of differentially enriched viral taxa between non-AF and AF individuals (A). Box plots of differentially enriched viral families (B) between non-AF control and AF individuals. Boxes are interquartile ranges; lines denote medians; circles are outliers. The scatter plot shows absolute values of Log<sub>2</sub> (fold change of CTR/AF), while dots colored in purple and green denote enriched viruses in AF and controls, respectively. + denotes adjusted  $p$  ( $q$ ) value < 0.01, \* $q$  < 0.05, and NS indicates no significant difference based on DESseq analysis.



were driven by both depleted and newly acquired ones; meanwhile, these correlations were lower in AF cases versus control patients (Supplementary Figure 3B). Notably, the associations of the viruses *Synechococcus phage S-SM1* and *Cronobacter phage CR5* with

bacterial species were very tight in control individuals but markedly dampened in AF cases. The above findings suggested a changed viral-bacterial relationship in AF, with viruses and bacterial organisms becoming intertwined and more specialized, further revealing the



**Figure 3. Functional alterations of the gut viral signals in AF.** Presence-absence heat map of categorized viral functions in non-AF and AF individuals, including in Gene Ontology (GO, **A**), the Pfam protein family database (**B**), and the KEGG pathway (**C**). The abundance distributions were graphed with line charts, and abundance levels were presented as Log2 reads per million mapped reads (RPM). The box plots show the top 10 differential viral functions (**D** for GO and **E** for Pfam). Boxes represent interquartile ranges; lines denote medians, and circles are outliers. + denotes adjusted  $p$  ( $q$ ) value  $< 0.01$ , \* $q < 0.05$ , and NS indicates no significant difference based on limma analysis. (**F**) Differential abundance in the discriminative KEGG pathways in AF compared to the control group.

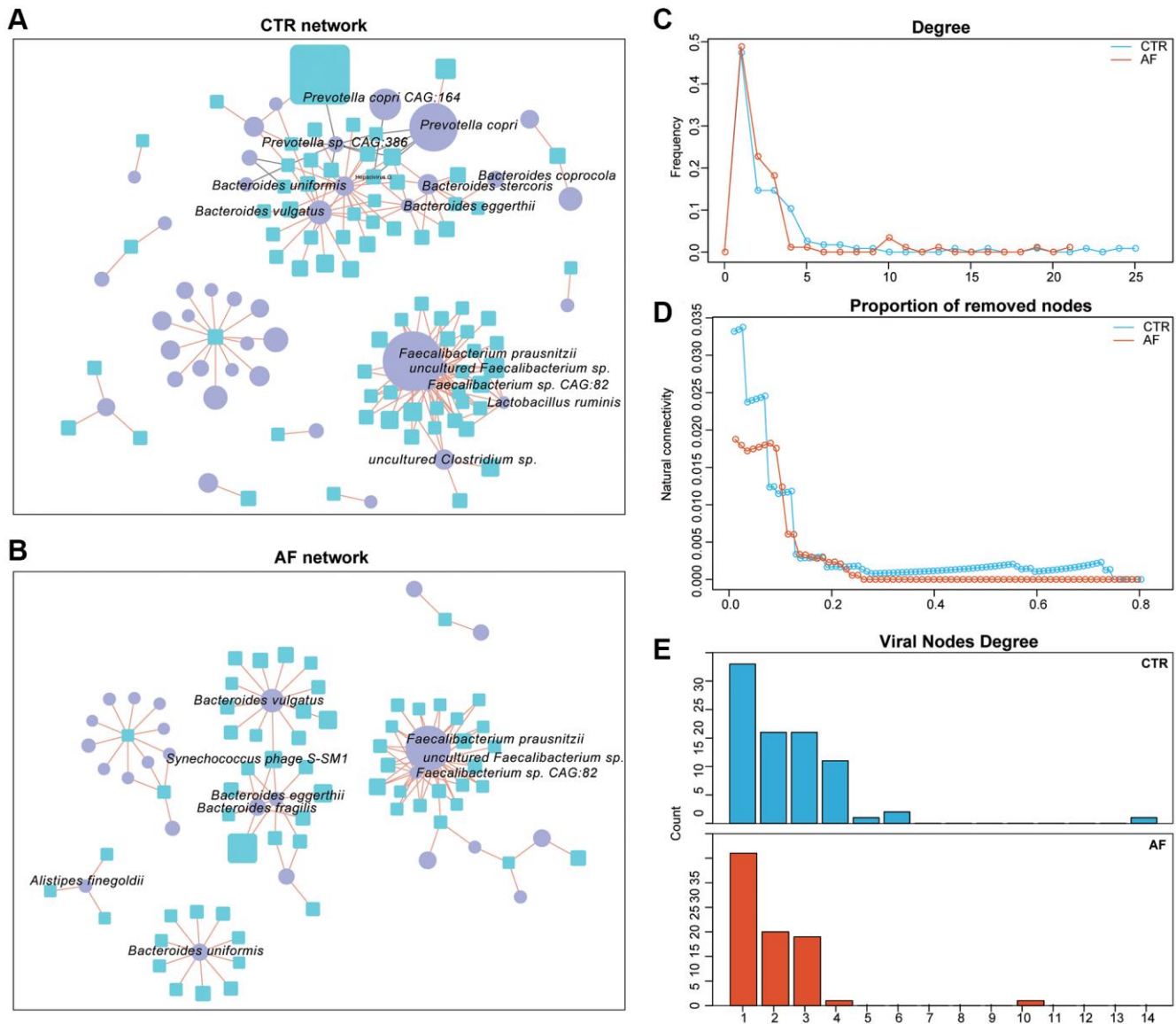
significance of inter-kingdom equilibrium among gut viruses and bacterial organisms in human health.

### Association of the gut viral signals with the risk of AF recurrence

Given the significance of the risk of AF recurrence after radiofrequency catheter ablation, whether the gut virome could help predict recurrent AF (RAF) was examined to identify individuals who could highly benefit from catheter ablation. In this work, we

constructed a gut virome-dependent signature for assessing the risk of AF recurrence.

First, taxa with the highest predictive values in AF recurrence were selected by LASSO analysis. Totally 21 viruses of all candidates (124 viruses with differences between the non-RAF and RAF groups with Wilcoxon rank-sum test  $q < 0.05$ ) retained statistical significance, showing non-zero coefficients in 40 AF cases (Figure 5A, 5B). Then, a risk score (viral score based on a linear combination of the 21 viral taxa-based markers)



**Figure 4. Virus-bacteria co-occurrence networks.** Virus-bacteria co-occurrence networks for non-AF controls (A) and AF patients (B). The nodes are viruses and bacterial species. Purple ellipses and blue rectangles represent bacterial species and viruses, respectively. Edges represent positive (orange) and negative (grey) associations of viruses with bacterial organisms at the species level. Nodes denote distinct bacterial organisms or viruses, with their sizes being proportional to mean relative abundances. The cutoff Spearman correlation coefficient and the adjusted  $p$ -value were  $|0.7|$  and  $0.05$ , respectively. Degrees of distribution of virus-bacteria co-occurrence networks (C). Comparison of the alteration levels of the virus-bacteria co-occurrence networks based on the rates of removed nodes (D). Degrees of viral nodes in virus-bacteria co-occurrence networks (E).

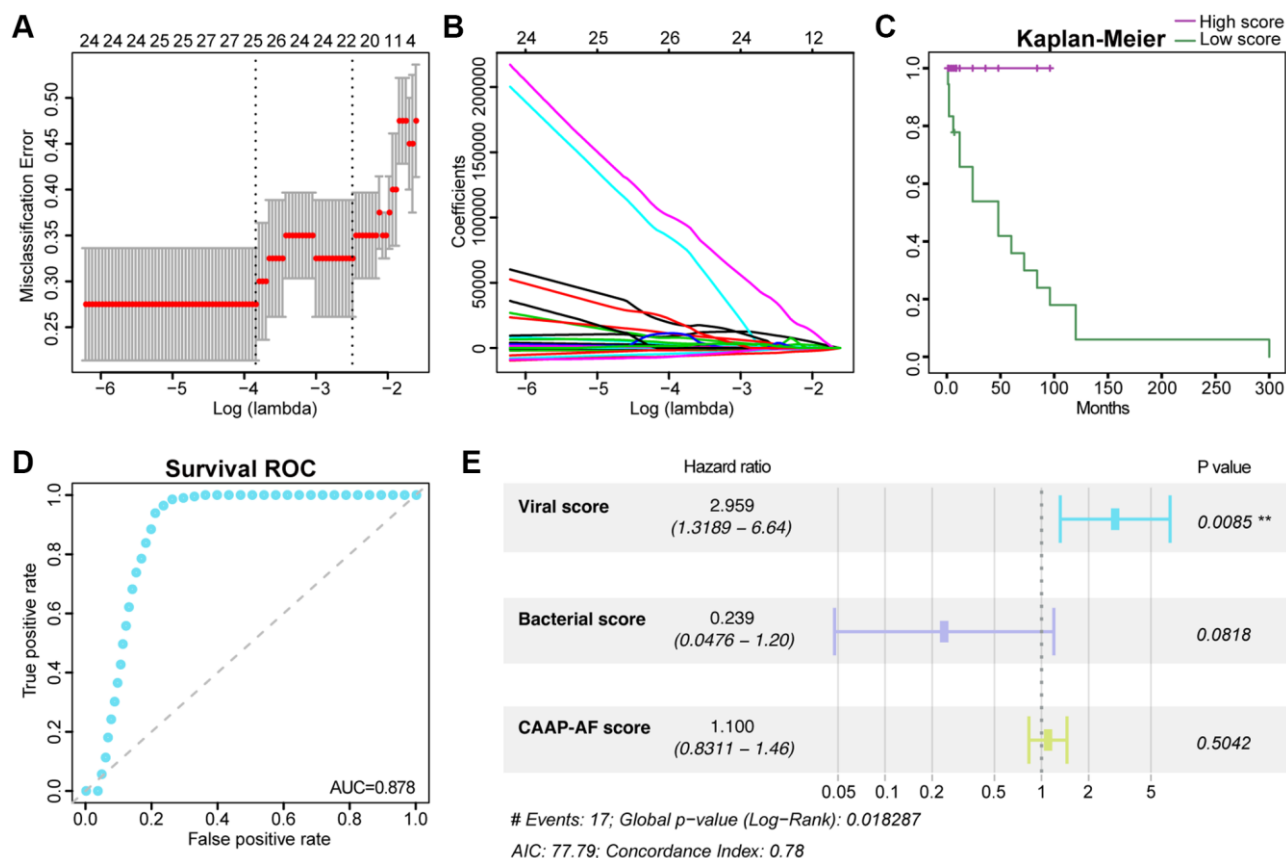
was obtained. The value of  $-0.5727$  was used as a cut-off for the viral score to determine the high- and low-risk groups of RAF cases by the nearest neighbor estimation (NNE) method. Kaplan-Meier curves showed that the high-risk group had an increased risk of RAF compared with low-risk individuals ( $P = 0.002$ , Figure 5C). The AUC of the time-dependent ROC curve for the viral score was 0.741 for sinus rhythm maintained on an average follow-up of  $15.6 \pm 12.57$  months (Figure 5D). Finally, we found that the viral score (HR = 2.959, 95% CI: 1.3189–6.64,  $p = 0.0085$ ) independently predicted prognosis upon adjustment for potential confounders, e.g., the bacterial score which was built in our previous study [15] (HR = 0.239, 95% CI: 0.0476–1.20,  $p = 0.0818$ ) and a clinical scoring system called CAAP-AF score (HR = 1.100, 95% CI: 0.8311–1.46,  $p = 0.5042$ ) (Figure 5E, Supplementary Table 5).

## DISCUSSION

Based on metagenomics data, the current work described the first and an in-depth human gut viral

signature in AF patients. As demonstrated above, AF cases had an altered gut viral signal, characterized by increased viral diversity and changed viral composition. In addition, marked functional alterations in the gut virus were detected in AF, indicating essential roles for viruses in virus-host interactions. Furthermore, a virus-based strategy was utilized to build a predictive model for RAF after ablation. Moreover, we provide initial support regarding the virus as a vital part of the microbiome ecological network.

The human gut contains a vast array of viruses that are highly diverse and stable [16]. Disordered gut virome has been revealed in several human diseases [17], including colorectal cancer [18, 19], ulcerative colitis [20], nonalcoholic fatty liver disease [21], carcinogenesis [22], diabetes mellitus [23] and hypertension [24]. There are few studies examining the relationship between gut viruses and diseases, and the substantial role of gut viruses in human health is unknown. Viruses might regulate the compositions of the microbial community, therefore affecting the



**Figure 5. Association of the gut viral signals with the risk of recurrent AF.** The tuning index (lambda) was identified utilizing the LASSO. Then, receiver operating characteristic (ROC) curve analysis was performed, and the AUC was plotted against log (lambda) (A). Dotted vertical lines represent the optimal values based on the minimum criteria and one standard error of the minimum criteria (1-SE criteria). Coefficients versus log (lambda) are shown (B). Kaplan-Meier curve analysis of overall survival based on the viral score (C). Survival ROC curves for the viral score (D). Forest plot of multivariate Cox regression analysis of virome-related risk groups and baseline patient covariates (E). Hazard ratios and 95% CIs are shown. Abbreviations: C-index: concordance index; CI: confidence interval.

functions of the microbiota and disease progression [25]. Bacteriophages change bacterial community composition, promoting colonization by driver bacterial species. Then, passenger bacterial organisms initiate biofilm formation, which increases with phage-induced dispersal. Epithelial cell transformation, tight junction disruption, and bacterial infiltration also occur. Bacteria step into a dysbiotic microenvironment and may thrive off specific metabolites, which are disease-related and participates in the progression of the disease [26–28].

Viral richness across the human life stages was assessed in healthy, western individuals. Across the human lifespan in the human gut virome database, the highest overall viral richness was observed in infants and adults and there were significant increases between children and adults, and significant decreases between adults and elderly individuals [29]. Notably, the tendency of the viral diversity, which increases with aging and decreases between adults and elderly individuals, is similar to the age-dependent patterns revealed in western individuals. The role of aberrant gut viruses in the pathogenesis of age-related AF remains for further study.

Previous reports have suggested that viral markers may help diagnose pathologies and predict therapeutic responses [30, 31]. In the current study, multiple viruses had elevated levels and significant between-group variations, including *Synechococcus phage S-SM1*, *Cronobacter phage CR5*, and *Staphylococcus phage SPbeta-like*. The above viruses are firstly linked to AF in this report, and their functions in the gut microbiota remain largely undefined. The random forest model constructed based on viral species may provide some clues for the screening and detection of occult AF, which refers to AF with no typical clinical symptoms and no clear history detected by ECG monitoring. Occult AF accounts for 15%–40% of the total AF and is also a significant cause of cryptic stroke [32]. Therefore, it is of great clinical significance to carry out effective screening. The gut virome might be of specific value for the screening and stratification of the high-risk population for occult AF. In addition, altered gut virome independently predicted AF recurrence. Catheter ablation is broadly utilized in clinical AF, with high efficacy. However, high post-ablation recurrence demands the development of tools capable of improving the selection of individuals who would potentially benefit from ablation. Although no clear association of viruses with AF has been reported, we examined viruses as potential markers for detecting asymptomatic latent AF or building a predictive model for RAF.

Application of the gut virome mainly targets the modulation of the gut bacteriome, where lytic

bacteriophages could reshape the structure of the gut microbiome [33]. With growing metagenomic evidence, the engineering of viral genomes could become a patient-specific therapy as a “final destination” [34]. For example, *Caudovirales* and *Microviridae* may be applied for future phage-specific therapies in Crohn's disease, and inhibiting bacteriophage infection of the respective bacterial hosts might also curve the natural course of inflammatory bowel diseases [35].

Finally, a large gap remains between human intestinal virome and AF. The current study is not typical virome research but a study mining viral signal from the vastly available bulk gut metagenomic sequencing data, which is accompanied by the limitation of omission about the viruses which need to be enriched during sampling purification. A standard virome study based on the deep shotgun metagenomic sequencing of virus-like particles-derived DNA should be performed to explore the role of the viral community in disease development. Further studies such as fecal virome transplantation trying to examine whether the AF phenotype is transferrable by altered gut virome and the subsequent effects on pathological changes related to atrial remodeling are still needed.

Overall, we described the altered gut viral profile in AF patients based on metagenomic data, which could be summarized as increased viral diversity, disturbed viral composition and functions, and disordered linkage between virus and bacteria. In addition, viruses were shown to have a superior predictive value for AF recurrence after ablation.

## MATERIALS AND METHODS

### Description of the study population of the metagenomic data

The whole-metagenome sequencing data analyzed in the current study were acquired from our previously published human non-valvular AF study (PRJEB28284) [6], which examined 50 non-AF control and 50 non-valvular AF cases in northern China. Fifty patients with nonvalvular AF were consecutively enrolled in the Heart center of Beijing Chaoyang Hospital since March 2016, and 50 individuals as matched controls were enrolled from the Kailuan cohort who received biennial medical examination in Kailuan General Hospital. Individuals with a history of heart failure, coronary heart disease, structural heart disease, comorbidities (inflammatory bowel diseases, irritable bowel syndrome, autoimmune diseases, liver diseases, renal diseases, or cancer), or use of antibiotics or probiotics in the past one month were excluded.



According to AF history and cardiogram manifestations [1], the 50 AF patients were assigned to the paroxysmal (PAF,  $n = 30$ ) and persistent (psAF,  $n = 20$ ) groups. Subsequently, the psAF cases were grouped into the <12-month (Pers<12 m,  $n = 12$ ) and >12-month (Pers>12 m) groups, reflecting cases with psAF durations shorter and longer than one year, respectively. In addition, 40 AF patients underwent radiofrequency catheter ablation during the hospital stay. AF recurrence (RAF) occurred in 17 AF cases after a follow-up of  $15.6 \pm 12.57$  months [36].

As previously reported [6–8, 36], AF patients were older, with more females than males, with a higher incidence of type 2 diabetes mellitus (T2DM), lower total cholesterol serum levels, and a higher incidence of medications compared to the non-AF CTR group. The baseline clinical characteristics among the different AF types were similar (PAF vs. psAF; Pers< 12 m vs. Pers> 12 m; non-RAF vs. RAF), with no remarkable difference in terms of age, sex, gender, body mass index, hypertension, T2DM, fasting blood glucose, serum creatinine, or alanine aminotransferase. The information about the treatment history of AF patients has been provided in Supplementary Table 6, including renin-angiotensin system inhibitors, amiodarone, statin, metformin, and radiofrequency ablation history. The study had approval from the ethics committees of Beijing Chaoyang and Kailuan General Hospitals. Each patient provided signed informed consent at the time of enrolment.

### **Assembly of metagenomic data, taxonomic assignment, and abundance profiling**

MEGAHIT v1.1.3 was applied to assemble using presets meta-large, with a k-mer ranging from 27 to 127, step 10. Only contigs >1,000 bp were retained for subsequent assessment for high predictive accuracy for viral sequences [37]. Viral sequences from these contigs were predicted and annotated based on the Kraken virus database (default parameters) and the taxonomies of various viruses [38]. The viral sequences were utilized for constructing a viral database to calculate viral composition in 100 fecal specimens. Contig coverage (Reads Per Kilobase per Million mapped reads, RPKMs) was determined with Bowtie2 (default parameters), normalizing to contig length and the number of mapped reads in a given specimen [39]. Average RPKM for a given virus was determined in every specimen for relative abundance calculation.

### **Rarefaction curve and diversity analyses**

Rarefaction analysis based on relative abundance levels of viral contigs and viruses was conducted to evaluate

whether the sequencing data were sufficiently deep and the current sample size was optimal for the present analysis. The patient population underwent 100-time random sampling with replacement, and common viral contigs/viruses and viral contigs/viruses joint in sample pairs were assessed, and R v2.15.3 (vegan package) was utilized for plotting. To determine taxonomic diversity, the total number of viruses,  $\alpha$  (within-individual) diversity indexes such as Simpson index, Shannon index, Pielou index, and Chao 1 richness, and  $\beta$  (between-individual) diversity using nonmetric dimensional scaling (NMDS) according to Bray-Curtis differences at the species level were calculated.

### **Virome function analysis**

Virome functions were classified by annotating all viral-contig derived reads with HUMANN2 v0.9.4 (default parameters). Gene ontology (GO) terms, Pfam protein family identities, and KEGG orthology data were utilized to predict functions, and abundance levels were determined in RPK (reads per kilobase) [20].

### **Virus and bacteria co-occurrence network analysis**

To determine the associations of bacterial and viral communities in non-AF controls and AF patients, the top 100 viruses and top 100 bacterial species based on their relative abundances were selected, and virus-bacteria linkages were further examined. Spearman correlation analysis was carried out to determine these associations, using the “cor” function in R and correcting for multiple testing by the Benjamini-Hochberg method. The cutoffs for the correlation coefficient and adjusted  $p$ -value were 0.7 and 0.05, respectively. The igraph package (v1.1.2) in R was used to quantitate distribution, node betweenness, and the network’s natural connectivity. Entropy for every network was obtained based on node degree using the entropy package (v1.2.1) in R. Finally, associations with statistical significance were imported into Cytoscape for visualizing virus-bacteria bipartite networks, in which nodes and edges represent viral and bacterial entities, and positive and negative associations of bacteria with viruses, respectively [24].

### **Design and validation of a predictive model for RAF**

The least absolute shrinkage and selection operator (LASSO) algorithm was utilized to select indexes with the highest predictive efficiencies based on significant taxa between the non-RAF and RAF groups. Viral scores were obtained for all cases by a linear combination of the retained taxa weighted by their respective coefficients. Internal validation was carried out based on a method reported in our previous study [36].

## Data availability

Raw data was available at European Nucleotide Archive via PRJEB28384; the datasets generated and analyzed in this study are available from the corresponding author upon reasonable request.

## AUTHOR CONTRIBUTIONS

Kun Zuo and Jing Li designed the experiments, interpreted the results, and wrote the manuscript. Chen Fang recruited and collected the clinical details from the participants. Jiuchang Zhong directed the project. Li Xu and Xinchun Yang conceived the study, directed the project, and revised the manuscript. All authors read and approved the final manuscript.

## CONFLICTS OF INTERESTS

The authors declare no conflicts of interest related to this study.

## ETHICAL STATEMENT AND CONSENT

The study had approval from the ethics committees of Beijing Chaoyang and Kailuan General Hospitals. Each patient provided signed informed consent at the time of enrolment.

## FUNDING

The present study was funded by the National Natural Science Foundation of China (82100334, 81970271, 81670214, 81500383 and 81870308).

## REFERENCES

1. Hindricks G, Potpara T, Dagres N, Arbelo E, Bax JJ, Blomström-Lundqvist C, Boriani G, Castella M, Dan GA, Dilaveris PE, Fauchier L, Filippatos G, Kalman JM, et al, and ESC Scientific Document Group. 2020 ESC Guidelines for the diagnosis and management of atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic Surgery (EACTS): The Task Force for the diagnosis and management of atrial fibrillation of the European Society of Cardiology (ESC) Developed with the special contribution of the European Heart Rhythm Association (EHRA) of the ESC. *Eur Heart J.* 2021; 42:373–498.  
<https://doi.org/10.1093/eurheartj/ehaa612>  
PMID:32860505
2. Jin L, Shi X, Yang J, Zhao Y, Xue L, Xu L, Cai J. Gut microbes in cardiovascular diseases and their potential therapeutic applications. *Protein Cell.* 2021; 12:346–59.  
<https://doi.org/10.1007/s13238-020-00785-9>  
PMID:32989686
3. Li J, Zhao F, Wang Y, Chen J, Tao J, Tian G, Wu S, Liu W, Cui Q, Geng B, Zhang W, Weldon R, Auguste K, et al. Gut microbiota dysbiosis contributes to the development of hypertension. *Microbiome.* 2017; 5:14.  
<https://doi.org/10.1186/s40168-016-0222-x>  
PMID:28143587
4. Cui X, Ye L, Li J, Jin L, Wang W, Li S, Bao M, Wu S, Li L, Geng B, Zhou X, Zhang J, Cai J. Metagenomic and metabolomic analyses unveil dysbiosis of gut microbiota in chronic heart failure patients. *Sci Rep.* 2018; 8:635.  
<https://doi.org/10.1038/s41598-017-18756-2>  
PMID:29330424
5. Jonsson AL, Bäckhed F. Role of gut microbiota in atherosclerosis. *Nat Rev Cardiol.* 2017; 14:79–87.  
<https://doi.org/10.1038/nrcardio.2016.183>  
PMID:27905479
6. Zuo K, Li J, Li K, Hu C, Gao Y, Chen M, Hu R, Liu Y, Chi H, Wang H, Qin Y, Liu X, Li S, et al. Disordered gut microbiota and alterations in metabolic patterns are associated with atrial fibrillation. *Gigascience.* 2019; 8:giz058.  
<https://doi.org/10.1093/gigascience/giz058>  
PMID:31149718
7. Zuo K, Li J, Wang P, Liu Y, Liu Z, Yin X, Liu X, Yang X. Duration of Persistent Atrial Fibrillation Is Associated with Alterations in Human Gut Microbiota and Metabolic Phenotypes. *mSystems.* 2019; 4:e00422-19.  
<https://doi.org/10.1128/mSystems.00422-19>  
PMID:31822598
8. Zuo K, Yin X, Li K, Zhang J, Wang P, Jiao J, Liu Z, Liu X, Liu J, Li J, Yang X. Different Types of Atrial Fibrillation Share Patterns of Gut Microbiota Dysbiosis. *mSphere.* 2020; 5:e00071-20.  
<https://doi.org/10.1128/mSphere.00071-20>  
PMID:32188747
9. Laliberté F, Moore Y, Dea K, LaMori JC, Mody SH, Jones JL, Arledge MD, Damaraju CV, Schein JR, Lefebvre P. Gastrointestinal comorbidities associated with atrial fibrillation. *Springerplus.* 2014; 3:603.  
<https://doi.org/10.1186/2193-1801-3-603>  
PMID:25392776
10. Rowan-Nash AD, Korry BJ, Mylonakis E, Belenky P. Cross-Domain and Viral Interactions in the Microbiome. *Microbiol Mol Biol Rev.* 2019; 83:e00044-18.  
<https://doi.org/10.1128/MMBR.00044-18>  
PMID:30626617

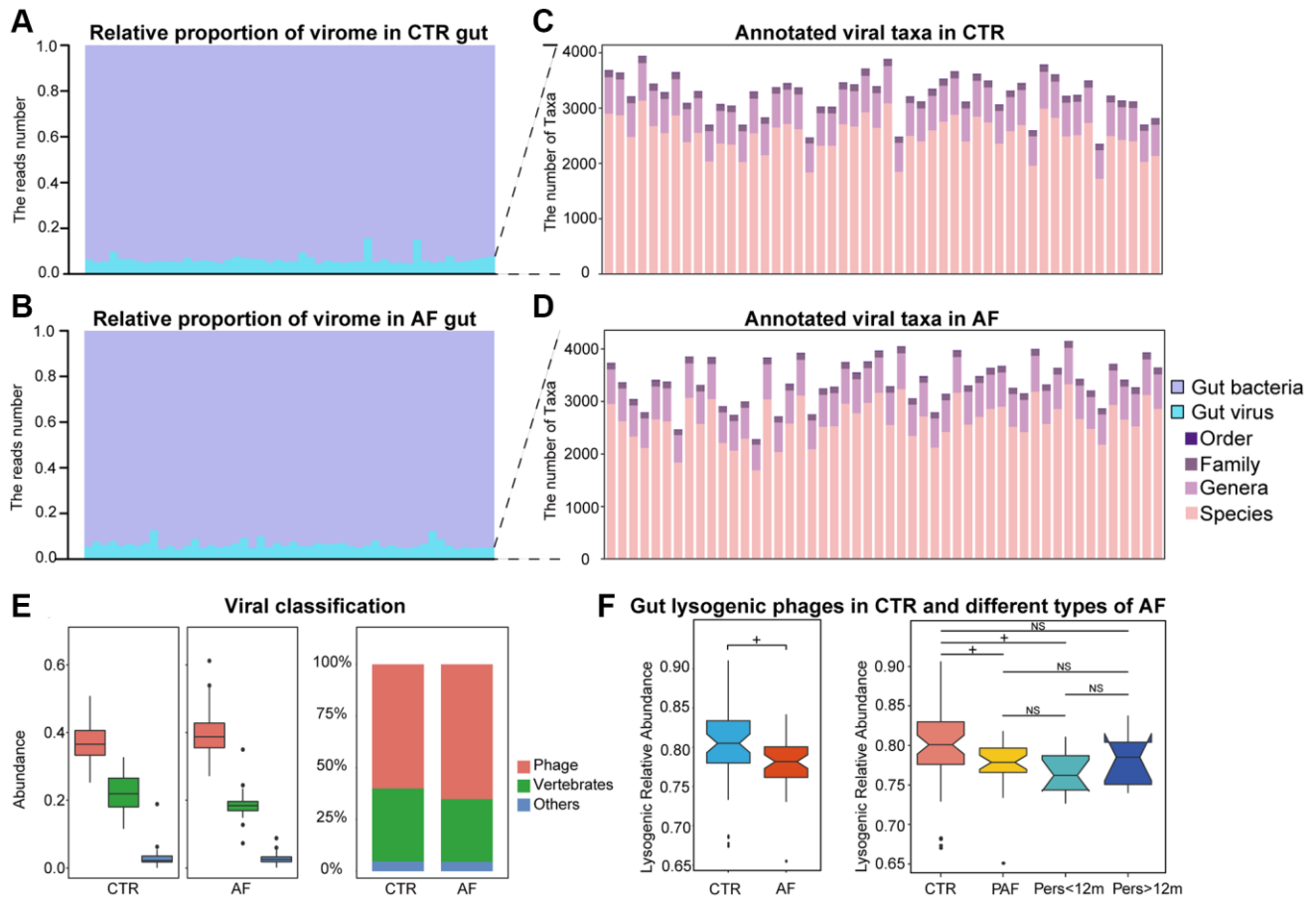
11. Clooney AG, Sutton TDS, Shkoporov AN, Holohan RK, Daly KM, O'Regan O, Ryan FJ, Draper LA, Plevy SE, Ross RP, Hill C. Whole-Virome Analysis Sheds Light on Viral Dark Matter in Inflammatory Bowel Disease. *Cell Host Microbe*. 2019; 26:764–78.e5.  
<https://doi.org/10.1016/j.chom.2019.10.009>  
PMID:[31757768](https://pubmed.ncbi.nlm.nih.gov/31757768/)
12. Zheng T, Li J, Ni Y, Kang K, Misiakou MA, Imamovic L, Chow BKC, Rode AA, Bytzer P, Sommer M, Panagiotou G. Mining, analyzing, and integrating viral signals from metagenomic data. *Microbiome*. 2019; 7:42.  
<https://doi.org/10.1186/s40168-019-0657-y>  
PMID:[30890181](https://pubmed.ncbi.nlm.nih.gov/30890181/)
13. Bikel S, Valdez-Lara A, Cornejo-Granados F, Rico K, Canizales-Quinteros S, Soberón X, Del Pozo-Yauner L, Ochoa-Leyva A. Combining metagenomics, metatranscriptomics and viromics to explore novel microbial interactions: towards a systems-level understanding of human microbiome. *Comput Struct Biotechnol J*. 2015; 13:390–401.  
<https://doi.org/10.1016/j.csbj.2015.06.001>  
PMID:[26137199](https://pubmed.ncbi.nlm.nih.gov/26137199/)
14. Minot S, Sinha R, Chen J, Li H, Keilbaugh SA, Wu GD, Lewis JD, Bushman FD. The human gut virome: inter-individual variation and dynamic response to diet. *Genome Res*. 2011; 21:1616–25.  
<https://doi.org/10.1101/gr.122705.111>  
PMID:[21880779](https://pubmed.ncbi.nlm.nih.gov/21880779/)
15. Winkle RA, Jarman JW, Mead RH, Engel G, Kong MH, Fleming W, Patrawala RA. Predicting atrial fibrillation ablation outcome: The CAAP-AF score. *Heart Rhythm*. 2016; 13:2119–25.  
<https://doi.org/10.1016/j.hrthm.2016.07.018>  
PMID:[27435586](https://pubmed.ncbi.nlm.nih.gov/27435586/)
16. Shkoporov AN, Clooney AG, Sutton TDS, Ryan FJ, Daly KM, Nolan JA, McDonnell SA, Khokhlova EV, Draper LA, Forde A, Guerin E, Velayudhan V, Ross RP, Hill C. The Human Gut Virome Is Highly Diverse, Stable, and Individual Specific. *Cell Host Microbe*. 2019; 26:527–41.e5.  
<https://doi.org/10.1016/j.chom.2019.09.009>  
PMID:[31600503](https://pubmed.ncbi.nlm.nih.gov/31600503/)
17. Mukhopadhyay I, Segal JP, Carding SR, Hart AL, Hold GL. The gut virome: the 'missing link' between gut bacteria and host immunity? *Therap Adv Gastroenterol*. 2019; 12:1756284819836620.  
<https://doi.org/10.1177/1756284819836620>  
PMID:[30936943](https://pubmed.ncbi.nlm.nih.gov/30936943/)
18. Hannigan GD, Duhaime MB, Ruffin MT 4th, Koumpouras CC, Schloss PD. Diagnostic Potential and Interactive Dynamics of the Colorectal Cancer Virome. *mBio*. 2018; 9:e02248-18.  
<https://doi.org/10.1128/mBio.02248-18>  
PMID:[30459201](https://pubmed.ncbi.nlm.nih.gov/30459201/)
19. Nakatsu G, Zhou H, Wu WKK, Wong SH, Coker OO, Dai Z, Li X, Szeto CH, Sugimura N, Lam TY, Yu AC, Wang X, Chen Z, et al. Alterations in Enteric Virome Are Associated With Colorectal Cancer and Survival Outcomes. *Gastroenterology*. 2018; 155:529–41.e5.  
<https://doi.org/10.1053/j.gastro.2018.04.018>  
PMID:[29689266](https://pubmed.ncbi.nlm.nih.gov/29689266/)
20. Zuo T, Lu XJ, Zhang Y, Cheung CP, Lam S, Zhang F, Tang W, Ching JYL, Zhao R, Chan PKS, Sung JJY, Yu J, Chan FKL, et al. Gut mucosal virome alterations in ulcerative colitis. *Gut*. 2019; 68:1169–79.  
<https://doi.org/10.1136/gutjnl-2018-318131>  
PMID:[30842211](https://pubmed.ncbi.nlm.nih.gov/30842211/)
21. Lang S, Demir M, Martin A, Jiang L, Zhang X, Duan Y, Gao B, Wisplinghoff H, Kasper P, Roderburg C, Tacke F, Steffen HM, Goeser T, et al. Intestinal Virome Signature Associated With Severity of Nonalcoholic Fatty Liver Disease. *Gastroenterology*. 2020; 159:1839–52.  
<https://doi.org/10.1053/j.gastro.2020.07.005>  
PMID:[32652145](https://pubmed.ncbi.nlm.nih.gov/32652145/)
22. Emler C, Ruffin M, Lamendella R. Enteric Virome and Carcinogenesis in the Gut. *Dig Dis Sci*. 2020; 65:852–64.  
<https://doi.org/10.1007/s10620-020-06126-4>  
PMID:[32060814](https://pubmed.ncbi.nlm.nih.gov/32060814/)
23. Ma Y, You X, Mai G, Tokuyasu T, Liu C. A human gut phage catalog correlates the gut phageome with type 2 diabetes. *Microbiome*. 2018; 6:24.  
<https://doi.org/10.1186/s40168-018-0410-y>  
PMID:[29391057](https://pubmed.ncbi.nlm.nih.gov/29391057/)
24. Han M, Yang P, Zhong C, Ning K. The Human Gut Virome in Hypertension. *Front Microbiol*. 2018; 9:3150.  
<https://doi.org/10.3389/fmicb.2018.03150>  
PMID:[30619215](https://pubmed.ncbi.nlm.nih.gov/30619215/)
25. Wahida A, Tang F, Barr JJ. Rethinking phage-bacteria-eukaryotic relationships and their influence on human health. *Cell Host Microbe*. 2021; 29:681–8.  
<https://doi.org/10.1016/j.chom.2021.02.007>  
PMID:[33735620](https://pubmed.ncbi.nlm.nih.gov/33735620/)
26. Beller L, Matthijssens J. What is (not) known about the dynamics of the human gut virome in health and disease. *Curr Opin Virol*. 2019; 37:52–7.  
<https://doi.org/10.1016/j.coviro.2019.05.013>  
PMID:[31255903](https://pubmed.ncbi.nlm.nih.gov/31255903/)
27. Lourenço M, Chaffringeon L, Lamy-Besnier Q, Pédrón T, Campagne P, Eberl C, Bérard M, Stecher B, Debarbieux L, De Sordi L. The Spatial Heterogeneity of the Gut Limits Predation and Fosters Coexistence of

- Bacteria and Bacteriophages. *Cell Host Microbe*. 2020; 28:390–401.e5.  
<https://doi.org/10.1016/j.chom.2020.06.002>  
PMID:32615090
28. Hsu BB, Gibson TE, Yeliseyev V, Liu Q, Lyon L, Bry L, Silver PA, Gerber GK. Dynamic Modulation of the Gut Microbiota and Metabolome by Bacteriophages in a Mouse Model. *Cell Host Microbe*. 2019; 25:803–14.e5.  
<https://doi.org/10.1016/j.chom.2019.05.001>  
PMID:31175044
29. Gregory AC, Zablocki O, Zayed AA, Howell A, Bolduc B, Sullivan MB. The Gut Virome Database Reveals Age-Dependent Patterns of Virome Diversity in the Human Gut. *Cell Host Microbe*. 2020; 28:724–40.e8.  
<https://doi.org/10.1016/j.chom.2020.08.003>  
PMID:32841606
30. Ludwig JA, Weinstein JN. Biomarkers in cancer staging, prognosis and treatment selection. *Nat Rev Cancer*. 2005; 5:845–56.  
<https://doi.org/10.1038/nrc1739>  
PMID:16239904
31. Mangalea MR, Paez-Espino D, Kieft K, Chatterjee A, Chriswell ME, Seifert JA, Feser ML, Demoruelle MK, Sakatos A, Anantharaman K, Deane KD, Kuhn KA, Holers VM, Duerkop BA. Individuals at risk for rheumatoid arthritis harbor differential intestinal bacteriophage communities with distinct metabolic potential. *Cell Host Microbe*. 2021; 29:726–39.e5.  
<https://doi.org/10.1016/j.chom.2021.03.020>  
PMID:33957082
32. Siontis KC, Gersh BJ, Killian JM, Noseworthy PA, McCabe P, Weston SA, Roger VL, Chamberlain AM. Typical, atypical, and asymptomatic presentations of new-onset atrial fibrillation in the community: Characteristics and prognostic implications. *Heart Rhythm*. 2016; 13:1418–24.  
<https://doi.org/10.1016/j.hrthm.2016.03.003>  
PMID:26961300
33. Kabwe M, Dashper S, Bachrach G, Tucci J. Bacteriophage manipulation of the microbiome associated with tumour microenvironments-can this improve cancer therapeutic response? *FEMS Microbiol Rev*. 2021; 45:fuab017.  
<https://doi.org/10.1093/femsre/fuab017>  
PMID:33765142
34. Grubb DS, Wrigley SD, Freedman KE, Wei Y, Vazquez AR, Trotter RE, Wallace TC, Johnson SA, Weir TL. PHAGE-2 Study: Supplemental Bacteriophages Extend *Bifidobacterium animalis* subsp. *lactis* BLO4 Benefits on Gut Health and Microbiota in Healthy Adults. *Nutrients*. 2020; 12:2474.  
<https://doi.org/10.3390/nu12082474>  
PMID:32824480
35. Scarpellini E, Ianiro G, Attili F, Bassanelli C, De Santis A, Gasbarrini A. The human gut microbiota and virome: Potential therapeutic implications. *Dig Liver Dis*. 2015; 47:1007–12.  
<https://doi.org/10.1016/j.dld.2015.07.008>  
PMID:26257129
36. Li J, Zuo K, Zhang J, Hu C, Wang P, Jiao J, Liu Z, Yin X, Liu X, Li K, Yang X. Shifts in gut microbiome and metabolome are associated with risk of recurrent atrial fibrillation. *J Cell Mol Med*. 2020; 24:13356–69.  
<https://doi.org/10.1111/jcmm.15959>  
PMID:33058365
37. Li D, Luo R, Liu CM, Leung CM, Ting HF, Sadakane K, Yamashita H, Lam TW. MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods*. 2016; 102:3–11.  
<https://doi.org/10.1016/j.ymeth.2016.02.020>  
PMID:27012178
38. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol*. 2014; 15:R46.  
<https://doi.org/10.1186/gb-2014-15-3-r46>  
PMID:24580807
39. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012; 9:357–9.  
<https://doi.org/10.1038/nmeth.1923>  
PMID:22388286

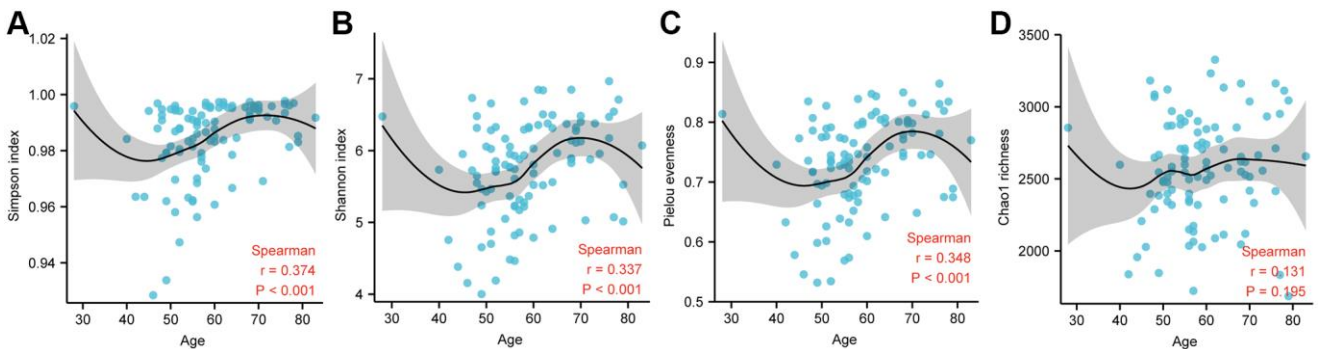


SUPPLEMENTARY MATERIALS

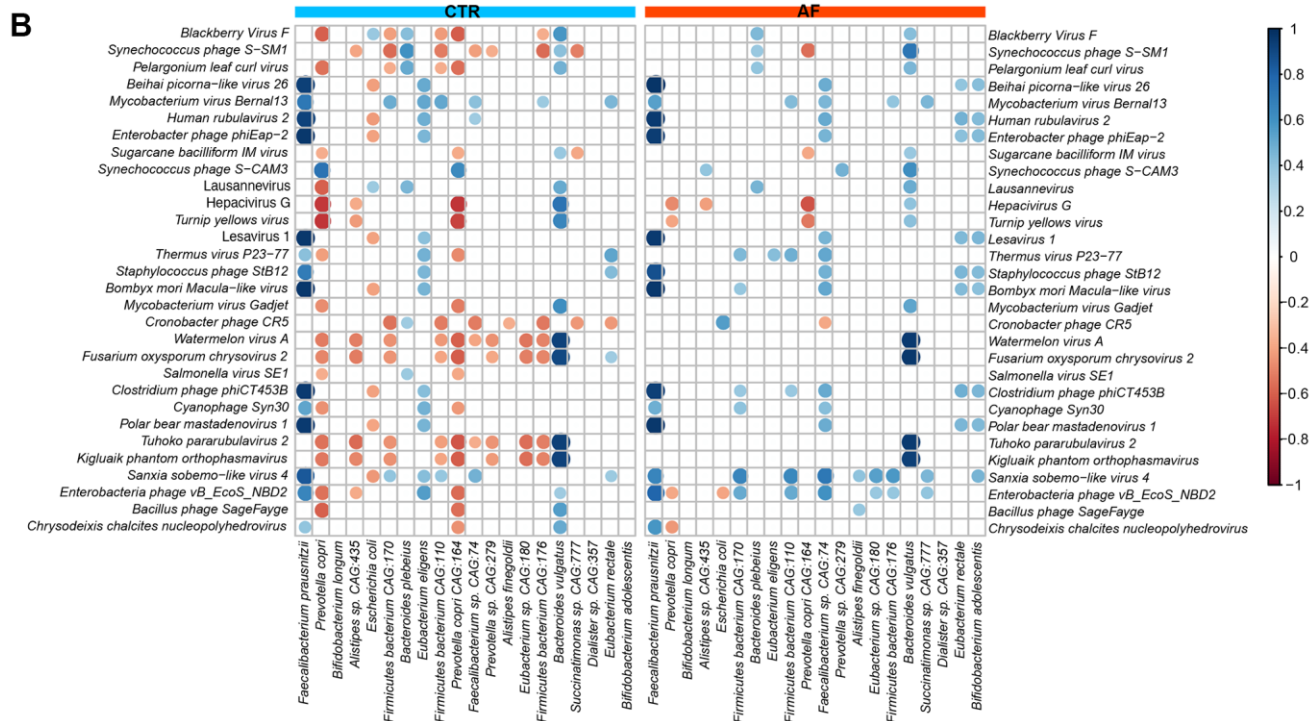
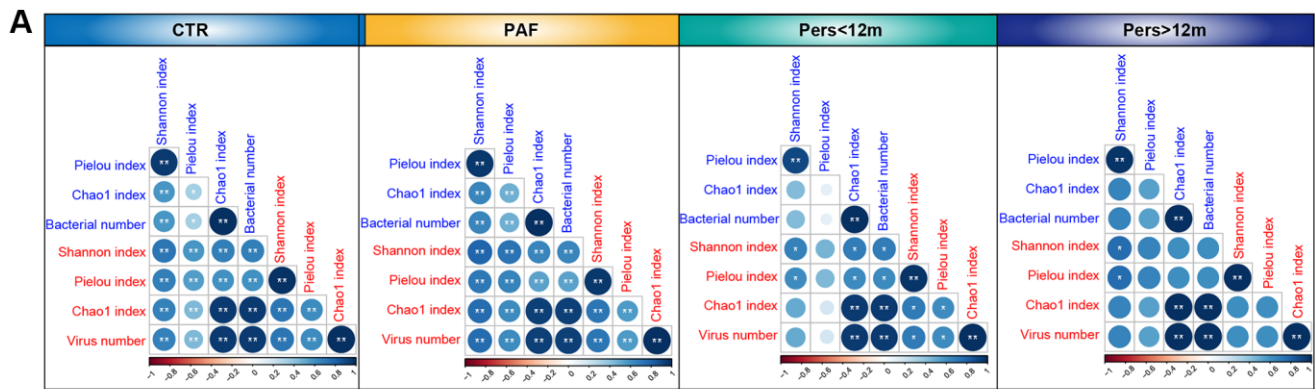
Supplementary Figures



**Supplementary Figure 1. Taxonomic profile of the gut viral signals in AF.** Relative proportions of bacterial and viral read numbers in the gut of controls (A) and AF patients (B). Taxonomic landscapes of the numbers of viral taxa at the order, family, genera, and species levels in controls (C) and AF patients (D). The box plot shows the abundance of vertebrate-infecting viruses and phages (E) and lysogenic phages (F). Boxes are interquartile ranges; lines denote medians, and circles are outliers. \*indicates  $p < 0.05$ , and NS indicates no significant difference based on the  $t$ -test.



**Supplementary Figure 2.** Spearman correlation between age and viral diversity including (A) Simpson index, (B) Shannon index, (C) Pielou evenness, and (D) Chao1 richness.



**Supplementary Figure 3. Trans-kingdom linkage between gut virus and bacteria in AF.** (A) Correlations between the  $\alpha$  diversity indexes of intestinal bacteria and viruses in non-AF individuals, PAF, Pers<12 m and Pers>12 m. Spearman's correlation coefficients were determined for all pairs. \* $P < 0.05$  and \*\* $P < 0.01$  indicate statistical significance. The circle's size and intensity reflect the correlation's magnitude, the darker and greater the circle, the tighter the correlation. (B) Associations of the most common 30 viruses with the most common 20 bacterial species in non-AF controls and AF patients. Spearman's correlation coefficients were determined for all pairs. Only correlations with statistical significance are shown. Blue and orange circles denote positive and inverse correlations, respectively. The circle's size and intensity reflect the correlation's magnitude, the darker and greater the circle, the tighter the correlation.

## Supplementary Tables

Please browse Full Text version to see the data of Supplementary Tables 2, 3, 4 and 6.

**Supplementary Table 1. The association between potential confounders and viral diversity.**

Spearman correlation analyses								
	Shannon index		Pielou evenness		Chao1 richness		Simpson index	
	Correlation	<i>p</i> value	Correlation	<i>p</i> value	Correlation	<i>p</i> value	Correlation	<i>p</i> value
AF	0.499	<0.001	0.522	<0.001	0.168	0.095	0.567	<0.001
Age	0.337	0.001	0.348	<0.001	0.131	0.195	0.374	<0.001
Gender	0.004	0.972	0.036	0.726	-0.153	0.130	0.040	0.691
BMI	0.048	0.638	0.032	0.752	0.053	0.605	0.032	0.751
TC	-0.215	0.032	-0.229	0.022	-0.107	0.289	-0.248	0.013
T2DM	0.035	0.728	0.058	0.569	-0.102	0.311	0.091	0.370
HTN	0.018	0.858	0.024	0.810	-0.041	0.683	0.006	0.956
Statin	-0.034	0.740	-0.035	0.727	-0.042	0.675	0.018	0.861
Metformin	0.089	0.379	0.124	0.219	-0.086	0.395	0.140	0.165
RASI	0.312	0.002	0.304	0.002	0.145	0.150	0.296	0.003
Amiodarone	0.382	<0.001	0.359	<0.001	0.379	<0.001	0.358	<0.001

Multivariate Linear Regression								
	Shannon index		Chao richness		Pielou evenness		Simpson	
	<i>Beta</i>	<i>p</i> value	<i>Beta</i>	<i>p</i> value	<i>Beta</i>	<i>p</i> value	<i>Beta</i>	<i>p</i> value
AF	0.386	0.003	0.064	0.631	0.419	0.001	0.380	0.004
Age	0.111	0.324	0.112	0.353	0.108	0.339	0.175	0.131
Gender	-0.092	0.333	-0.171	0.094	-0.074	0.432	-0.088	0.363
BMI	-0.051	0.587	0.053	0.600	-0.067	0.476	-0.090	0.351
TC	-0.064	0.532	-0.126	0.252	-0.049	0.632	0.009	0.931
T2DM	-0.170	0.192	-0.100	0.473	-0.171	0.187	-0.095	0.473
HTN	-0.078	0.443	-0.168	0.124	-0.058	0.566	-0.185	0.078
RASI	0.125	0.267	0.007	0.951	0.136	0.223	0.133	0.246
Statin	-0.149	0.121	-0.114	0.266	-0.147	0.126	-0.067	0.492
Amiodarone	0.199	0.050	0.377	0.001	0.153	0.127	0.091	0.378
Metformin	0.103	0.423	-0.068	0.621	0.127	0.321	0.095	0.469

Abbreviations: AF: atrial fibrillation; BMI: body mass index; HTN: hypertension; DM: diabetes mellitus; TC: total cholesterol; HTN: hypertension; RASI: renin-angiotensin system inhibitor.

**Supplementary Table 2. DESeq analysis of taxa.**

**Supplementary Table 3. Limma analysis of function (GO, Pfam, and KEGG pathway).**

**Supplementary Table 4. Virus-bacteria co-occurrence network.**

**Supplementary Table 5. Viral, bacterial, and CAAP-AF score for each sample.**

<b>ID</b>	<b>Viral score</b>	<b>Bacteria score</b>	<b>CAAP-AF score</b>
no_reAF1	-1.643436274	-0.438057175	8
no_reAF2	-0.572739537	-0.559589758	4
no_reAF3	-1.112909752	-0.41292474	5
no_reAF4	-1.720926254	-0.480547348	4
no_reAF5	-1.663225355	-0.629424774	3
no_reAF6	-1.207326045	-0.457982056	2
no_reAF7	-0.794490746	-0.403373535	7
no_reAF8	-1.449086837	-0.924431659	4
no_reAF9	-0.690488681	-0.372348362	3
no_reAF10	-1.590535439	-0.83456788	5
no_reAF11	-1.454642971	-1.018637523	1
no_reAF12	-1.38143236	-0.522411896	3
no_reAF13	-1.17920171	-0.784698581	6
no_reAF14	-0.745314246	-0.941740217	7
no_reAF15	-1.446141423	-0.846711054	3
no_reAF16	-2.376555056	-0.406248875	4
no_reAF17	-1.028053231	-0.121126375	0
no_reAF18	-0.853312857	-1.094805179	2
no_reAF19	-1.086303952	-0.405425016	1
no_reAF20	-1.390471095	-0.474613524	1
no_reAF21	-1.001423967	-0.838099047	4
no_reAF22	-1.534434775	-0.913203781	3
no_reAF23	-0.39917529	-0.515052	1
reAF1	0.857412367	0.130452354	5
reAF2	0.662074237	-0.114841987	2
reAF3	0.788297736	-0.430826956	4
reAF4	0.808764009	-0.072169315	1
reAF5	0.089692379	0.094527915	4
reAF6	1.711957888	0.744633894	8
reAF7	1.333083101	-0.166743886	7
reAF8	0.539843254	-0.285436126	7
reAF9	0.67832355	0.1271324	2
reAF10	1.638005186	0.942875207	4
reAF11	1.182311213	-0.386605779	4



reAF12	0.194674374	-0.160757943	4
reAF13	1.581301599	-0.215528759	6
reAF14	0.785256346	0.821312269	5
reAF15	0.2598357	0.680733825	6
reAF16	0.638192391	-0.413989751	7
reAF17	0.067054588	0.611781801	8

Abbreviations: viral score = (-0.4191 \*(Intercept)) + (-151.9898 \*Lepidopteran hudovirus) + (-824.8187 \*Acanthocystis turfacea chlorella virus 1) + (-32.1303 \*Synechococcus phage S-RIM8) + (-36.6109660550033 \*Agrotis segetum nucleopolyhedrovirus B) + (-1001.4687 \*Mokola lyssavirus) + (1046.1101 \*Rhodococcus phage REQ1) + (564.2624 \*Chicken anemia virus) + (1061.2306 \*Escherichia virus HX01) + (1216.3961 \*Guajara orthobunyavirus) + (-879.1352 \*Escherichia virus ADB2) + (-80.5973 \*Azospirillum phage Cd) + (-3751.0734 \*Helicoverpa armigera granulovirus) + (1722.6372 \*Salmonella phage SPN3UB) + (73.0350 \*Listeria virus P100) + (112.2706 \*Beihai zhaovirus-like virus 1) + (9441.7402 \*Maize rayado fino virus) + (34369.6655 \*Ribgrass mosaic virus) + (568.4422 \*Wuhan millipede virus 3) + (2657.2776 \*Wuhan insect virus 14) + (3857.4019 \*Narcissus mosaic virus) + (1303.9541 \*Paracoccus phage vB\_PmaS-R3).

**Supplementary Table 6. Medication history.**