


RESEARCH

Open Access



Metagenomic sequencing reveals time, host, and body compartment-specific viral dynamics after lung transplantation

Stefanie Widder^{1,2*} , Irene Görzer³, Benjamin Friedel^{3,4}, Nina Rahimi¹, Stefan Schwarz⁵, Peter Jaksch⁵, Sylvia Knapp¹ and Elisabeth Puchhammer-Stöckl³

Abstract

Background: The virome of lung transplant recipients (LTRs) under immunosuppressive therapy is dominated by non-pathogenic *Anelloviridae* and further includes several pathogenic viruses such as Herpesviruses or respiratory viruses. It is unclear whether the donor-derived virome in the transplanted lung influences recipient virome dynamics in other body compartments and if so, to which degree. Likewise, it is unknown whether dependencies exist among virus populations that mutually shape viral loads and kinetics.

Results: To address these questions, we characterized viral communities in airways and plasma of 49 LTRs and analyzed their abundance patterns in a data modeling approach. We found distinct viral clusters that were specific for body compartments and displayed independent dynamics. These clusters robustly gathered specific viral species across the patient cohort. In the lung, viral cluster abundance associated with time after transplantation and we detected mutual exclusion of viral species within the same human host. In plasma, viral cluster dynamics were associated with the indication for transplantation lacking significant short-time changes. Interestingly, pathogenic viruses in the plasma co-occurred specifically with *Alpha torque virus* genogroup 4 and *Gamma torque virus* strains suggesting shared functional or ecological requirements.

Conclusions: In summary, the detailed analysis of virome dynamics after lung transplantation revealed host, body compartment, and time-specific dependency patterns among viruses. Furthermore, our results suggested genetic adaptation to the host microenvironment at the level of the virome and support the hypothesis of functional complementarity between *Anellovirus* groups and other persistent viruses.

Keywords: Virus metagenomics, Network analysis, Viral community clusters, Lung transplantation, Data modeling

Introduction

Within the last years, it was shown that a multitude of different eukaryotic viruses colonize the human host [1, 2]. This eukaryotic human virome may vary over time, and differs between healthy persons and specific

patient populations as well as between different human body sites as blood or lung [3, 4]. In immunosuppressed patients, as solid organ transplant recipients, who undergo lifelong drug induced decrease of different immune responses, the level of distinct virus populations is massively increased [4–6]. While in a number of pathogenic viruses, as herpesviruses or polyomaviruses, replication under these circumstances may cause disease, the most prominent eukaryotic virus populations found in clinical material of transplant

*Correspondence: stefanie.widder@meduniwien.ac.at

¹ Research Laboratory of Infection Biology, Department of Medicine I, Medical University of Vienna, Vienna, Austria

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



© The Author(s) 2022, corrected publication 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

recipients belong to the non-pathogenic *Anelloviridae*, consisting of *Alpha-*, *Beta-*, and *Gamma-torque viruses* [4, 5, 7, 8]. Within the last years, it turned out that among these, especially the group of *Alpha torque viruses* (TTV) are strongly elevated in patients under immunosuppressive therapy after solid organ transplantation [9, 10]. The overall total level of TTV, which comprise 5 genogroups (TTV1-5), and numerous TTV strains, was shown meanwhile to serve as a marker for the level of immunosuppression especially after lung or kidney transplantation [9, 11, 12]. Previous data in lung transplant recipients have provided evidence that there is not only a multitude of different TTV strains detected within a single patient at a given time point in BAL or blood, but that TTV populations also change substantially over time within a single patient [4].

It is known that a variety of different viruses may co-exist within the same human compartment [5]. While it is well established that the human bacteria microbiome engages in a plethora of interactions, it is unknown whether there exist causal dependencies between specific virus populations and whether the replication of individual virus populations is associated with mutual reduction or increase of other virus populations in the same microenvironment at a given time point or over time. Due to high genetic variability in *Anelloviridae*, it has been proposed that some *Torque teno viruses* may require complementary strains to gather all functionalities for successful infection and immune evasion [13]. This may generate possible liabilities between pathogenic and non-pathogenic viruses as well as the relation between different *Anellovirus* populations.

Especially in transplant recipients, potential virome cross-talk may have impact on the development of clinical complications, including microbial infections or transplant rejection, which limits transplant survival [11, 14]. In addition, as the total TTV load is increasingly used to measure the immunosuppressive net state after transplantation [15], knowledge about the distribution and potential interactions of different TTV genogroups within this patient setting needs to be further elucidated.

In the present study, we therefore investigated the quantitative relations between different virus populations in the virome of lung transplant recipients (LTRs), as well as their association with demographic or clinical parameters. We analyzed viral co-occurrence patterns and revealed robust clustering among specific virus groups in two different body compartments, i.e., lung and plasma. Importantly, we also studied cluster differences at specific time points after

lung transplantation to characterize viral adaptation dynamics after solid organ transplantation.

Methods

Patient collective

In this exploratory study, samples from 49 patients were included, who all received a lung transplant at the Medical University of Vienna between 2012 and 2018 and from whom clinical specimens were available for virome analyses. In detail, we received 96 bronchoalveolar lavage (BALF) samples from 41 patients, as well as 45 plasma samples from 22 patients of which 14 patients provided both types of samples. Demographic and clinical data were recorded and all patients were followed up for at least 1 year. Patient demographics and samples are described in detail in SI Table 1A and B, respectively. All patients received induction therapy with 30 mg Alemtuzumab (Berlex), and maintenance therapy with tacrolimus (Prograf) and corticosteroids (Prednisolone). Additionally, all patients received antibiotic prophylaxis for two weeks, anti-fungal prophylaxis with aerosolized amphotericin B for 3 months, and HCMV-prophylaxis consisted of four doses of intravenous HCMV hyperimmunoglobuline during the first month and Val/ganciclovir administered for at least 3 months post-transplantation. The study was approved by the local Ethics Committee at the Medical University of Vienna (EK-number 1880/2016 and EK-number 1710/2014).

Data generation

Data used for analysis comprise newly generated virome data derived from 68 BALF and 13 plasma samples from 34 lung transplant patients as well as previously generated virome data of 15 LTRs (Segura-Wang et al., 2018).

Virus enrichment

Viral enrichment of BALF and plasma samples was done as described previously [4]. In short, 1 ml BALF aliquot was centrifuged for 5 min at 1200 g. Both, BALF and plasma volumes were adjusted to 1.2 ml with PBS and filtered through 0.45 µm filters (Sartorius). To remove cell-free DNA and RNA, the supernatant was treated with RNase A (QIAGEN) and TURBO DNase (Ambion) to final concentrations of 0.77 mg/ml and 20 Units/ml, respectively, incubated for 1 h at 37 °C, and deactivated by Protease (QIAGEN) treatment for 30 min at 37 °C.

Nucleic acid extraction

Total nucleic acids were extracted using 1 ml of the filtered and nuclease treated material, with the automated NucliSENS EasyMAG system (BioMérieux).

Metagenomic sequencing

For each extracted sample, RNA and DNA were amplified by random primer amplification in separate reactions and pooled before sequencing. RNA was reverse transcribed using a random octamer linked to an anchor sequence ATCGTCGTCGTAGGCTGC TCNNNNNNNN 15–18. Amplification reactions used 5 μ l of extracted eluate and 5 μ M random primers and were incubated for 5 min at 65 °C. Then 20 μ l total volume reactions were set up as described in [16]. The resulting cDNA was denatured for 2 min at 94 °C, cooled 5 min at 10 °C, and used for second-strand synthesis. Similarly, the DNA from the extracted eluate was randomly amplified by using 5 μ l of template. The resulting DNA derived from both, RNA and DNA templates, was subsequently amplified in separate reactions by PCR using 4 μ l of input DNA from the previous step, 0.25 μ M dNTP mix (Promega), 2 mM MgCl₂, 1x AmpliTaq Gold Buffer II (Applied Biosystems), 0.05 U/ μ l AmpliTaq Gold Polymerase (Applied Biosystems), and 1 μ M of an anchor specific primer for a total volume of 50 μ l. The products were quantified by Qubit dsDNA BR Assay (Invitrogen) and diluted to 0.4 ng/ μ l. Equal amounts of the amplified and diluted cDNA and DNA were pooled at 0.4 ng/ μ l, and used as input for library preparation using the NexteraXT Library Preparation kit (Illumina). Samples were sequenced on an Illumina MiSeq platform, single-read or paired-read, 150bp and 250bp, respectively.

Viral bioinformatics

Bioinformatic processing of the metagenomic sequencing data was performed with CLC Genomics Workbench 12.0 (QIAGEN Bioinformatics). Raw sequencing reads were trimmed including filtering for length \geq 70 bp and removing low quality reads with an average Phred score below 30. Subsequently, the sequencing data were cleaned from non-viral reads by sequential mapping to a human reference genome (GRCh38), and to sets of bacterial and bovine reference genomes. To identify sequences of viral origin the remaining reads were aligned to a set of viral reference genomes by NCBI BLASTn [17] integrated into CLC Genomics Workbench. The used viral reference database was originally derived from the NCBI GenBank nuccore database [16, 18] which was cleaned from unverified isolates of TTV and thus consisted of 49,069 complete viral genomes including all 29 TTV genotypes officially recognized by ICTV [19] and verified TTV isolates according to NCBI GenBank [18]. BLAST search parameters included a minimum word size of 20, an E value \leq 0.001 and filtering genetic regions with low entropy. For each sequenced read, the BLAST hit with the lowest E value was used to identify viral species. The

identified sequences were further filtered to an E value \leq 1E-15 and a species was reported provided that at least five reads were found in the filtered data set of a sample. TTV reads were further categorized to the respective *Alpha torque virus* genogroup according to their phylogenetic relationship, which was evaluated as described before [13].

Data preparation

Sequencing counts were normalized against total metagenome counts to reduce compositional effects of viral species in subsequent analyses. The fold excess of total metagenome counts vs. total viral counts per sample amounted to 5367.2 on average. Subsequently, viral counts were normalized against the respective genome sizes according to [20] and technical replicates were averaged. To ensure robust abundance signals for downstream analyses, we denoised the data set and removed organisms with a presence of \leq 10% in plasma and BALF samples, respectively and an average abundance per sample \leq 0.01 according to [21]. The resulting read abundances were 4th root transformed to reduce data skewness for further analyses as in [22].

Site-specific adaptation of lung and plasma virome after transplantation, exclusion dynamics

To clarify whether virome composition and abundance in the allograft and plasma converge over time, we evaluated the similarity of 56 matched sequencing samples from both body parts. In detail, first samples were categorized by time of collection after surgery using 5–12 weeks, 13–34 weeks, 35–95 weeks intervals. Then we calculated centroids from samples derived from both sides individually and compared its Bray-Curtis distance to all samples within and among sites. Centroids were calculated using R package *usedist* [23] and *vegan* [24]. Subsequently, all within site distances and between site distances were grouped and similarity of distributions was tested using Wilcoxon test statistics.

Virome predictors and variable hierarchy

We investigated driving variables and their hierarchy in the data using independent and nested permutational multivariate analysis of variation (PERMANOVA). We tested sample variance with the categorical covariates patient and time after transplantation. Time was log-transformed to study proportional changes after transplantation and categorized into intervals (2–4 weeks, 5–12 weeks, 13–34 weeks, 35–95 weeks) that entered as factor into the statistical models. Sequencing data were 4th root transformed as in [22] and PERMANOVA

was performed on sample Bray-Curtis distances using `adonis2` from R package `vegan` [24]. Variable hierarchy was determined by iteratively nesting either covariate.

Clusters of viral co-occurrence

We performed co-occurrence analyses to identify robust patterns among viral organisms using `SparCC` [25]. To correct for biases introduced by individual patients, we pursued a leave-one-out (LOO) strategy by iteratively removing samples derived from every patient resulting in 41/22 correlation matrices for BALF and plasma, respectively.

To identify the optimal number of clusters and to overcome technical limitations imposed by sparse and signed correlation matrices, we performed partitioning around medoids (PAM) of the full correlation matrices in parallel with graph clustering (blockmodels) of only strong correlations ($\rho > |0.2|$, $p < 0.05$) on the core virome present in all graphs. The data were clustered with varying $k = 3-10$ and the resulting partitioning quality was assessed by complementary quality measures (Dunn Index, Dunn2 Index, Separation Index, Widest Gap, Pearson Gamma). Subsequently, the mean value for each measure was assessed across all 41/22 matrices and the most supported cluster number was identified by majority vote throughout these measures, for BALF and plasma independently. Partitioning and quality check was performed in R using the packages `clusters`, `signet`, `fpc` [26–28].

To assign core organisms to their final clusters, we harmonized the information on cluster association α obtained from `pam` partitions and `blockmodels` according to

$$\alpha(ij) = \sum_0^N c_{ij}/N$$

with c_{ij} being instances where organisms i and j are classified to the same cluster within one partitioned set and N representing the total number of partitioned sets. PAM was performed on the resulting pairwise distance $d(ij) = 1 - \alpha(ij)$.

Correlation networks

To correct for potential patient bias, we generated 41 BAL/22 plasma networks from the co-occurrence analyses described above using $\rho > |0.2|$ and $p < 0.05$ as cut-off for including edges into the graph. To integrate these into a single BALF and plasma graph, we first identified organism present in all respective graphs (core virome) and added edges, if present in $> 50\%$ of the graphs (majority vote). The average weight of the individual edges across all graphs is indicated in the final graph.

Ordination of virome and association testing with genogroup

For the core virome, a single distance matrix for BALF and plasma was inferred by averaging over the LOO distance matrices derived from correlation analyses. Ordination was performed on averaged virome distances using `PHATE` [29] to ensure proper denoising and improved embedding of high dimensional correlation data. Variance testing of distances, genogroups, and interactions thereof was performed with `PERMANOVA` [24]. To test associations between viral genogroups and partitioned clusters, a χ^2 test was performed on the contingency table of genotype frequency in clusters using R package `psych` [30]. The association among genogroups within clusters was assessed using Pearson correlation of the contingency table and Benjamini-Hochberg correction for multiple testing using `corrplot` [31].

Linear mixed effect models for assessing drivers of virome structure

We analyzed the relationship of transplant recipient virome at the level of composition and richness with time and patient characteristics. We tested whether individual viral clusters can be significantly predicted by fitting a mixed effect intercept model (LMM) with following standard structure

$$y = X\alpha + Z\beta + \varepsilon$$

where y is a vector of the response variable, X is the vector of fixed effect predictor, Z is the matrix of random effects, and α , β , ε represent vectors of fitted regression coefficients.

The response variable was constructed as sum of 4th root transformed reads associated with the same cluster per sample or alternatively, as sum of different organism per cluster per sample normalized relative to maximum value achieved. For plasma-derived clusters, we used patient identity as fixed effect and $\log(\text{time})$ after transplant as random effect. For BAL-derived clusters, both patient ID or $\log(\text{time})$ were tested as fixed effects independently, with the respective other variable used as random effects. Subsequently, statistical significance of the dependence was determined by ANOVA testing on model F-statistic (type III analysis, Satterthwaite's method) using packages `lmerTest` [32] and `core package stats`. For visualization of effect sizes, model fits, and dynamics over time, we used `effects` [33], `ggplot2` and `ggpubr` [34], respectively.

For detecting driver organism in time and for TX indication, denoised and `clr` transformed read data were used as response variable and $\log(\text{time})/\text{indication}$ categories were used as fixed effects; patient and batch ID were included as random effects.

Spatial exclusion patterns in the lung

To quantify exclusion of particular viral organism in the same BALF sample, we compared the 4th root transformed abundances of viruses in time-associated lung clusters and defined the degree of segregation between the community clusters. The viral abundances were summed by cluster and their ratios stratified into 5 exclusion categories, these are presence of organisms of only one cluster (100% exclusion), presence of organisms of both clusters: abundance ratio $\leq 1:4$ (75% exclusion), $\leq 1:2$ (50% exclusion), $\leq 3:4$ (25% exclusion), and comparable abundance levels (no exclusion). Sample numbers according to exclusion categories are presented ($n = 56$).

Association of patient groups and clinical meta-information

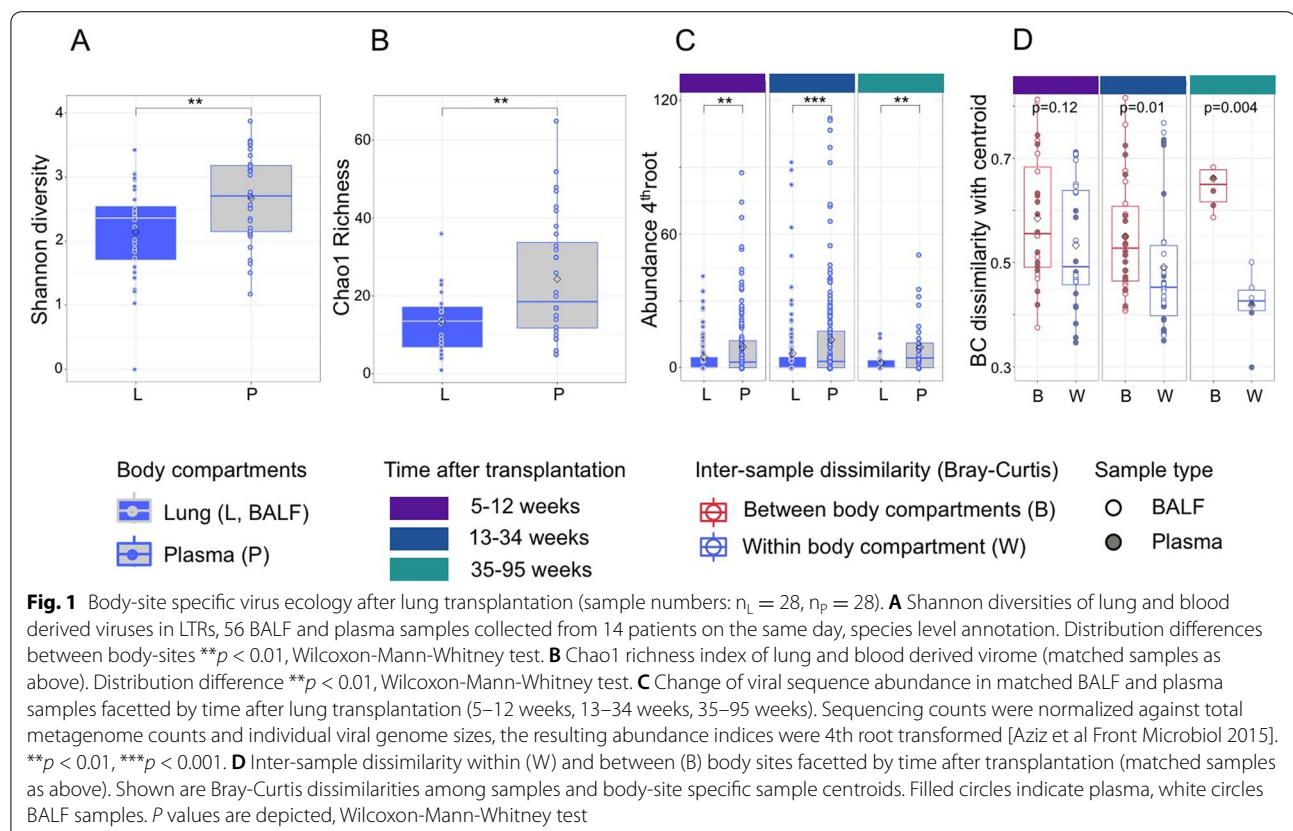
For LMM models, where a significant association of viral clusters with patient identity was detected, patients were split into 2 subgroups based on their estimated regression parameters α (fixed effects). In detail, patient groups were formed from regression parameters with different sign (+/-). Subsequently, we tested association of these patient groups with clinical meta-information, i.e., recipient TX indication, donor CMV status, acute rejection

episodes, induction therapy, immuno-suppressive treatment categories, and recipient sex. Contingency tables were inferred, categories represented with a single instance were removed, and χ^2 test was performed to assess association and statistical significance.

Results

Virome structure in lung and blood develops independently after lung transplantation

To better understand how the recipient virome rearranges after lung transplantation and to determine relevant drivers in this process, we analyzed the variation of viral composition and abundance in matched samples; these were 28 BALF and 28 plasma samples collected at the same day from 14 patients based on metagenomic sequencing. The virome was annotated to species level and as it was strongly dominated by eukaryotic viruses, only these were included in the analyses. Overall, *Anelloviridae* dominated the virome composition (*Anelloviridae:non-Anelloviridae* = 16:1) with *Alpha torque viruses* (TTV) belonging to genogroup 3 being most abundant in 70% of the samples. We observed that blood and lung-derived viromes differed significantly (Fig. 1). Shannon diversity and Chao1 richness both indicated a significantly higher taxonomic diversity in plasma



samples compared to BALF (Fig. 1A, B). Similarly, the viral abundance at different, matched time intervals after TX (5–12 weeks, 13–34 weeks, and 35–95 weeks) was significantly higher in plasma than in BALF at all time points tested (Fig. 1C).

We further asked whether the site-specific differences adjusted over time, i.e., whether the viromes of blood and alveolar compartment became more similar with time after transplantation. To this end, we determined the centroid composition of all plasma and BALF samples within three defined time periods independently, and tested the temporal changes in Bray-Curtis dissimilarity of samples to centroid, both within and between body sites (Fig. 1D). The results showed that differences between plasma and airway virome persisted and even increased over the first year after transplantation, whereas samples from the same body site became more alike over time, suggesting the development of a sampling site specific viral community.

To test whether these site-specific virus variations can be explained by temporal changes after transplantation (time-dependency), by patient identity, or by a combination of both, we performed permutational multivariate analysis of variance (PERMANOVA). We included all 141 study samples to increase the statistical power, i.e., 45 plasma, 96 BALF samples (detailed data description in SI Table 1). Again, lung-derived and blood-derived viromes yielded different results (SI Table 2A). While both variables yielded statistically significant results if tested independently, we identified hierarchical effects by nesting the variables iteratively. In this limited data, patient identity accounted significantly for variance at both sites (plasma: $R^2 = 0.67$, $p < 0.001$; lung: $R^2 = 0.63$, $p < 0.001$), while time was significantly associated with lung ($R^2 = 0.09$, $p < 0.002$) but not plasma ($R^2 = 0.14$, $p = 0.2$) virome variance. These trends persisted also if the number of lung samples was reduced to the number of plasma samples ($n = 45$, 1000 random subsamples, SI Table 2B).

Together, our analyses suggested that the lung virome underwent detectable changes with time after transplantation. For both body sites, patient-specific factors robustly impacted sample variance.

Co-occurrence analysis reveals viral clusters that delimit genetic groups

Next, we analyzed the co-occurrence patterns of viral species after transplantation. Due to the dynamical and compositional differences among body sites, we conducted an independent analysis on 45 plasma and 96 BALF samples, respectively, including all time points of collection. To correct for potential biases due to individual patients, we iteratively removed individual patients from the analysis (leave-one-out stratification).

Interestingly, we observed strong exclusive viral communities characterized by positive association within and negative associations between groups of virus species in BALF (Fig. 2A) and plasma (Fig. 3A). To better discriminate the correlation patterns, we performed a cluster analysis that accommodated the particular properties of co-occurrence matrices; these are positive and negative (signed) association values, and sparsity due to weak correlations. We performed cluster analysis on blockmodels of sparse and signed adjacency matrices and combined the outcome with results from standard partitioning around medoids (PAM) clustering on quantitative distance matrices. We found five strong community clusters in the lung and three in plasma (PERMANOVA $p < 0.001$ in both sites) (Fig. 2B, Fig. 3B), whose composition was strongly defined by different *Anelloviruses*. We performed χ^2 testing to examine the association of individual virus species with clusters (Fig. 2C). In the lung, TTV species belonging to genogroup 1 (TTV1), genogroup 3 (TTV3), and genogroup 4 (TTV4) dominated their respective clusters and pathogenic viruses like Coronavirus, Cytomegalovirus, and other human Herpesviruses co-occurred with *Gamma torque viruses* (GTV) in a separate cluster ($\chi^2 = 185.2$, $p < 0.0001$).

In the blood compartment, the association of pathogenic viruses with GTV persisted, but was extended by co-occurrence of TTV4 in the same cluster (Fig. 3C). The other two clusters determined in plasma presented mixed communities comprising TTV1, 3, 5, and *Beta torque viruses*, respectively ($\chi^2 = 49.7$, $p < 0.0001$). We confirmed the associations among *Anellovirus* subpopulations and *non-Anelloviruses* in parallel by Pearson correlation for BALF samples (SI Fig. 1A) and plasma samples (SI Fig. 1B). These results highlight a distinct and significant association between specific *Anellovirus* subpopulations and between these and pathogenic viruses in the different compartments.

Viruses display specific temporal dynamics in the lung

To dissect in detail the viral adaptation in the lung compartment over time after transplantation, we analyzed the changes in abundance of the viral sequences in the five virus clusters found in the lung over the first 8 months post transplantation. Therefore, we built linear mixed effect models (LMM) on the 96 BALF samples, to identify which of the five clusters significantly associated with log-transformed continuous time after transplantation, while introducing patient identity as random effects to the model (Fig. 4A, SI Fig. 2A–C). We discovered that only clusters 1, 2, and 4 displayed a significant increase in viral abundance over time (cluster nomenclature as depicted in Fig. 2). These clusters were rich in species belonging especially to TTV genogroup

3 or 4 (LMM, $p < 0.05$). On the level of single virus strains, SEN34, HD21, HD22, 19, 27, 29, L1984, and JT05F displayed a significant time-associated increase (LMM, $p < 0.05$, Fig. 4B). The results highlight that virome adaptation did not proceed uniformly within the *Anelloviruses* and perturbation of the microenvironment in the lung compartment after transplantation seemed to benefit replication of certain virus strains over others.

We further asked whether these time-associated TTV strains increased simultaneously or whether increase of one strain prohibited the emergence of others. We found that only those strains associated with the same cluster increased simultaneously, while strains from different clusters were not observed together at high levels. In 80% (45 out of 56) BALF samples, the dominant viruses were more than twice as abundant in the sample as viruses confined to the inferior cluster.

Recipients' underlying disease imprints long-term plasma virome structure

Based on the initial finding that virome variation in plasma was affected not by time but by patient identity, we sought to identify clinical parameters that may explain differences in recipient virome variations in an exploratory manner. To this end, we modeled viral cluster abundance in 45 plasma samples with linear mixed effect models using patient ID as explanatory variable and time after transplantation entered as random effects. Only plasma cluster 3 showed a significant linear relationship with the underlying patient identity (LMM, $p < 0.05$). Analysis of the model parameter estimates revealed that these split the patient collective that have viral species belonging to cluster 3 (dominated by TTV1 and BTV) into two groups with high or low abundances, respectively (Fig. 5A, C). In a next step, we asked, which patient-related factors might be associated with the grouping. We tested group association with clinical categories including indication for transplantation, survival, and prevalence of chronic lung allograft dysfunction (CLAD) performing χ^2 tests. In our limited data set, we found that solely

the underlying disease significantly specified the patient groups ($p = 0.024$, $p = 0.004$, respectively, Fig. 5B). As such, patients originally diagnosed with chronic obstructive pulmonary disease (COPD), fibrosis, or idiopathic pulmonary fibrosis (IPF) presented a different plasma virome than patients with CF. As lung recipients with CF tended to be younger than recipients with different diagnosis, we assessed whether patient age could explain patient grouping in a similar manner (SI Fig. 3). The resulting association between age and patient grouping shows a positive trend $\chi^2 = 5.949$, but misses statistical significance ($p = 0.051$). While an explorative association analysis like this one cannot reinforce causal conclusions nor enable to distinguish drivers in a set of covariates like CF and age, our results underline the importance of host factors for the long-term dynamics of the human virome.

Discussion

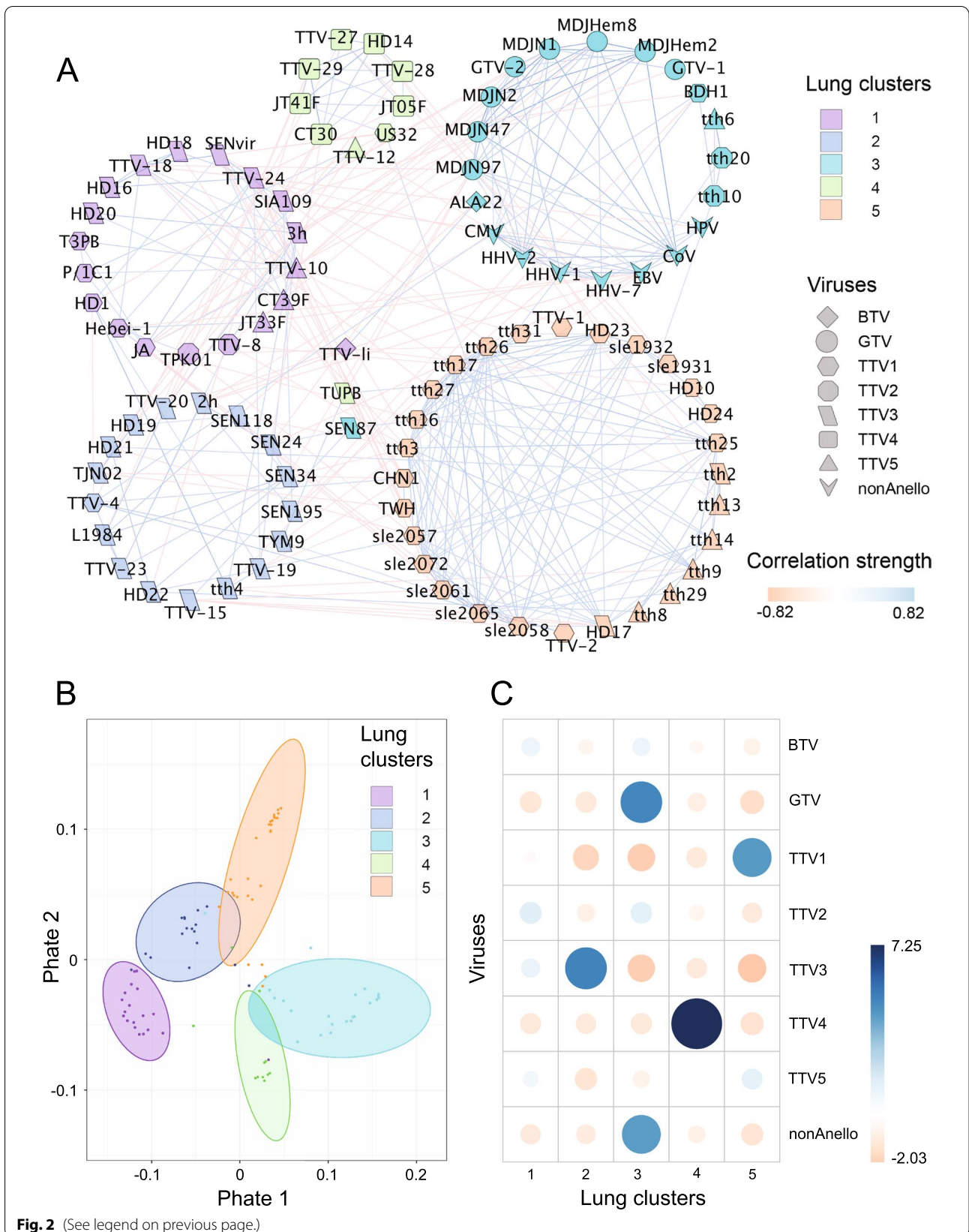
The human virome is intimately related with the health status of its host. Viral populations colonize the neonate shortly after birth and develop a body site specific signature later on [35]. While the adult virome is relatively stable in individual body compartments of a person over time [36], the transplantation of an organ from a donor to a recipient organism leads to the introduction of an unrelated viral population to a host. It is unclear how such a perturbation affects viral population dynamics in the allograft and in other body compartments of the recipient.

The present study on a cohort of LTRs analyzing eukaryotic viruses in 141 plasma and BALF samples revealed compartment-specific viromes that display separate dynamics.

The analysis of inter-sample distances over time demonstrated that these differences between recipient blood and transplanted lung-derived virome not only persisted but even increased over the first year after transplantation. This modeling approach is in line with previous clinical observations in LTRs that detected the transfer of *Anelloviruses* from donor to recipient within the lung [37], but showed no clear imprint of the donor *Anellovirus* population on the recipient blood virome after kidney transplantation [38]. Thus, the data suggest that the *Anellovirus*

(See figure on next page.)

Fig. 2 *Anelloviruses* with same genogroup tend to form community clusters in the lung (Sample number: $n_L = 96$). **A** Co-occurrence network of viruses in the lung. Nodes represent core viral species shaped according to TTV genogroups (TTV1-5), *Beta-* (BTV) and *Gamma torque virus* (GTV) or *non-Anelloviruses* (nonAnello), edges depict correlation (red) and exclusion (blue). Networks were calculated iteratively removing every patient (leave-one-out strategy) to correct potential patient bias. The fused graph depicts viruses present in all subnetworks (core virome), edges present in > 50% of subnetworks, edge weight as average over the respective values in the subnetworks. **B** 2-dimensional PHATE ordination of distances between core viral strains in the lung virome. Depicted clusters were inferred from partitioning around medoids of correlation matrix combined with block models of graph adjacency matrices. $p < 0.001$, PERMANOVA. **C** Association of *Anelloviruses*, genogroups and pathogenic viruses with viral clusters in the lung. BTV *Beta torque virus*, GTV *Gamma torque virus*, TTV1-5 *Alpha torque virus* genogroups, nonAnello *non-Anelloviridae*, Pearson χ^2 test, association strength color-coded, circle size depicts normalized strength, $\chi^2 = 185$, $p < 0.0001$



dominated viromes of different hosts do not significantly interchange after solid organ transplantation.

Under healthy conditions, host antiviral immune responses exert control over persisting viruses by suppressing viral replication. Immunosuppressive therapy after transplantation releases this balance resulting in increased virus loads in LTRs, with *Anelloviridae* comprising a dominant fraction thereof [5]. A main question of the present study was to clarify whether epidemiologic or ecologic dependencies existed among viruses and viral strains that could become visible during such viral re-activations after transplantation. An important finding was the significant co-exclusion observed in the blood compartment between pathogenic viruses on one and *Alpha-* and *Beta torque virus* populations on the other hand. Earlier work showed similar epidemiological features of *Alpha-* and *Beta torque virus* strains in the context of acute respiratory diseases [39]. One possible explanation for the observed clustering and the co-exclusion with pathogenic viruses may be that we observe the result of viral adaptation to its microenvironment [40, 41]. Replication of pathogenic viruses such as *Herpesviruses* or *Hepatitis C Virus* leads to immune responses which elicit substantial cytokine expression [42]. These high cytokine levels in the environment may indirectly cause also a rapid limitation of *Anellovirus* replication affecting the majority of *Alpha-* and *Beta torque viruses*. In fact, it was described that therapeutic administration of alpha interferon due to specific diseases caused a secondary decrease of *Alpha-* and *Beta torque virus* levels [43, 44]. *Anelloviruses* themselves do most likely not induce high cytokine levels in the human host.

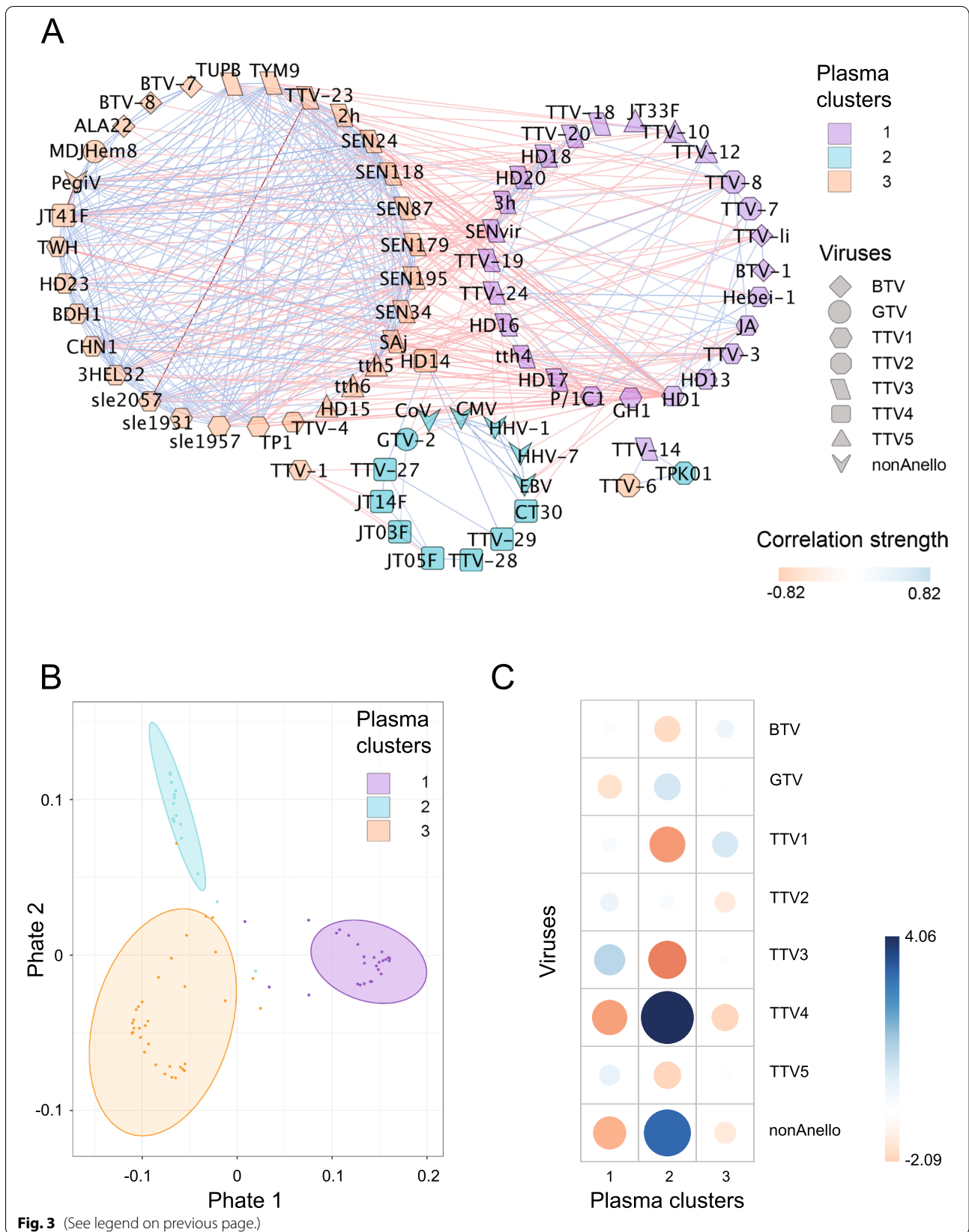
Interestingly, the pathogenic viruses in the studied plasma samples formed a strong synergistic cluster with GTVs and a specific subgroup of TTV, namely genogroup 4. Therefore, the question arose, whether TTV genogroup 4 is different in some way to other TTV genogroups. Borkosky and colleagues [45] discovered a functional relation between *Herpes* and *Anelloviridae* and showed that co-infection with *EBV* activated the replication of genogroup 4 isolate

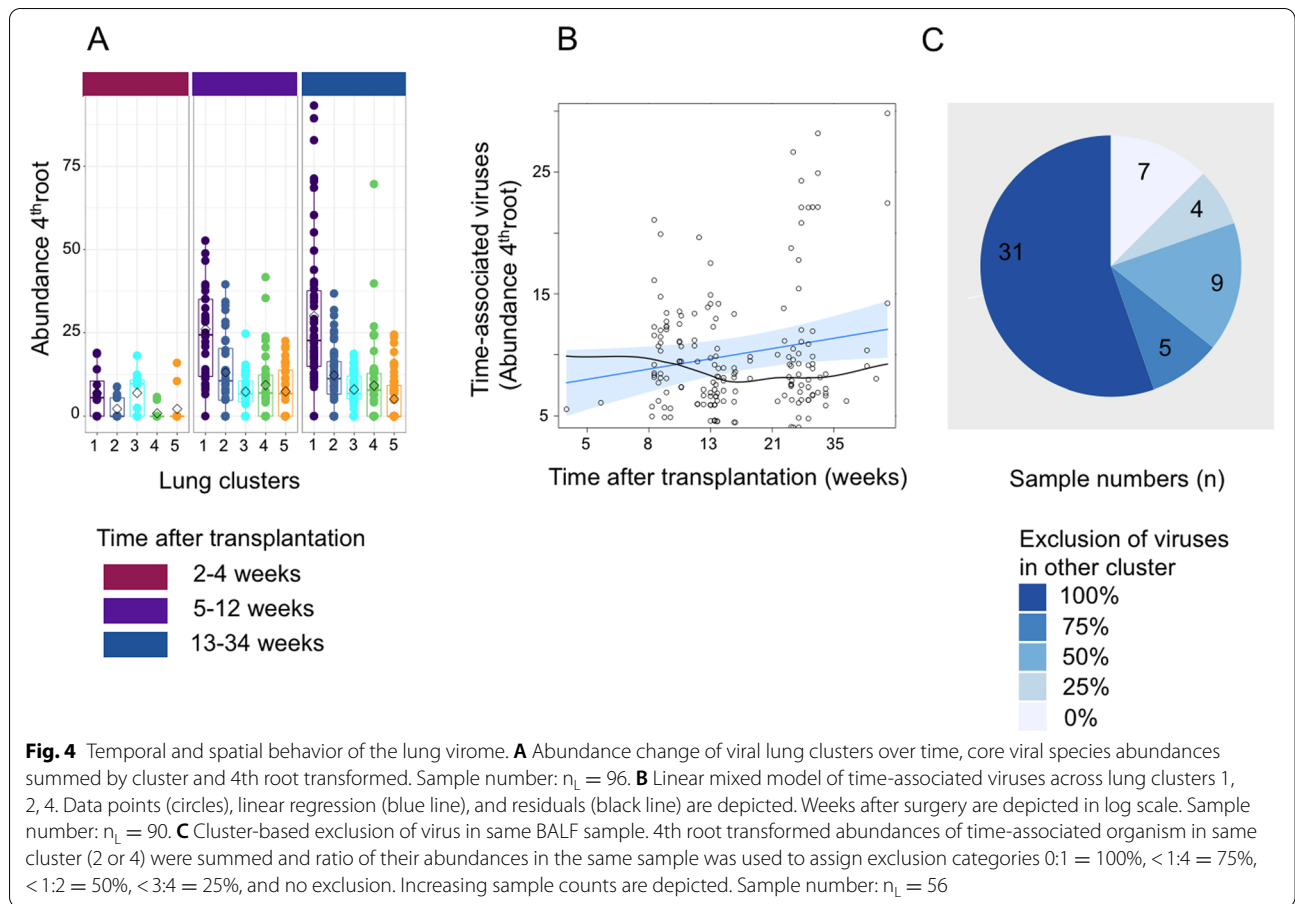
TTV-HD14 in vitro. Moreover, there are limited previous data, suggesting that the TTV4 genogroup may eventually be more pathogenic than other *Anelloviruses* [46]. The group of Maggi found increased TTV4 loads in children with acute respiratory diseases developing complications such as severe bronchopneumonia [39]. Whether these effects are evidence for a direct pathogenic effect is unclear, but fit the current findings that point toward a special status of TTV4 compared to other *Alpha torque viruses*. While *Anelloviruses* are in general seen as commensal, non-pathogenic viruses, the high proportion of *Anellovirus* co-infections may hide the fact that on overall one or the other member of this group may have some pathogenic potential or benefit from pathogenic virus infections. It has been shown that TTV and several *Herpes viruses* target similar blood cells in the human host, specifically monocytes and lymphocytes [2, 47, 48]. Thus, there may exist overlapping ecological niches of TTV and pathogenic viruses, which are possibly different between individual *Anellovirus* populations. This could also explain the clustering between pathogenic viruses and GTVs. While GTVs were found frequently in nasopharyngeal secretions of children with acute respiratory disease and also in the context of lung transplantation [49], it is supposed that they are innocent bystanders more than pathogenic agents [50].

In the lung, the level of viral replication was significantly associated with time after transplantation. Five viral clusters were identified with a more pronounced co-occurrence pattern among *Anellovirus* genogroups than in plasma. At the level of the population, it has been proposed that only few TTV strains possess all functional components to persist in the host, therefore co-infection is frequently observed to gather the required components [13, 46]. The higher number and definition of clusters in the BALF samples compared to plasma may be caused by a temporal instability or re-activation of the donor-derived lung virome. While the precise mechanisms remain obscure, the robust clustering of specific virus groups suggests that viruses

(See figure on next page.)

Fig. 3 Virus community clusters in plasma (sample number: $n_p = 45$). **A** Co-occurrence network of viral abundances in patient plasma. Nodes represent core viral species shaped according to TTV genogroups (TTV1-5), *Beta-* (BTV) and *Gamma torque virus* (GTV) or non-*Anelloviruses* (nonAnello), edges depict correlation (red), and exclusion (blue). Networks were calculated iteratively removing every patient (leave-one-out strategy) to correct potential patient bias. The fused graph depicts viruses present in all subnetworks (core virome), edges present in > 50% of subnetworks, edge weight as average over the respective values in the subnetworks. **B** 2-dimensional PHATE ordination of distances between core viral strains in the plasma virome. Depicted clusters were inferred from partitioning around medoids of correlation matrix combined with block models of graph adjacency matrices. $p < 0.001$, PERMANOVA. **C** Association of *Anellovirus* genogroups and pathogenic viruses with viral clusters in blood. BTV *Beta torque virus*, GTV *Gamma torque virus*, TTV1-4 *Alpha torque virus* genogroups, nonTTV pathogenic viruses. Pearson χ^2 test, association strength color-coded, circle size depicts normalized strength, $\chi^2 = 50$, $p < 0.0001$

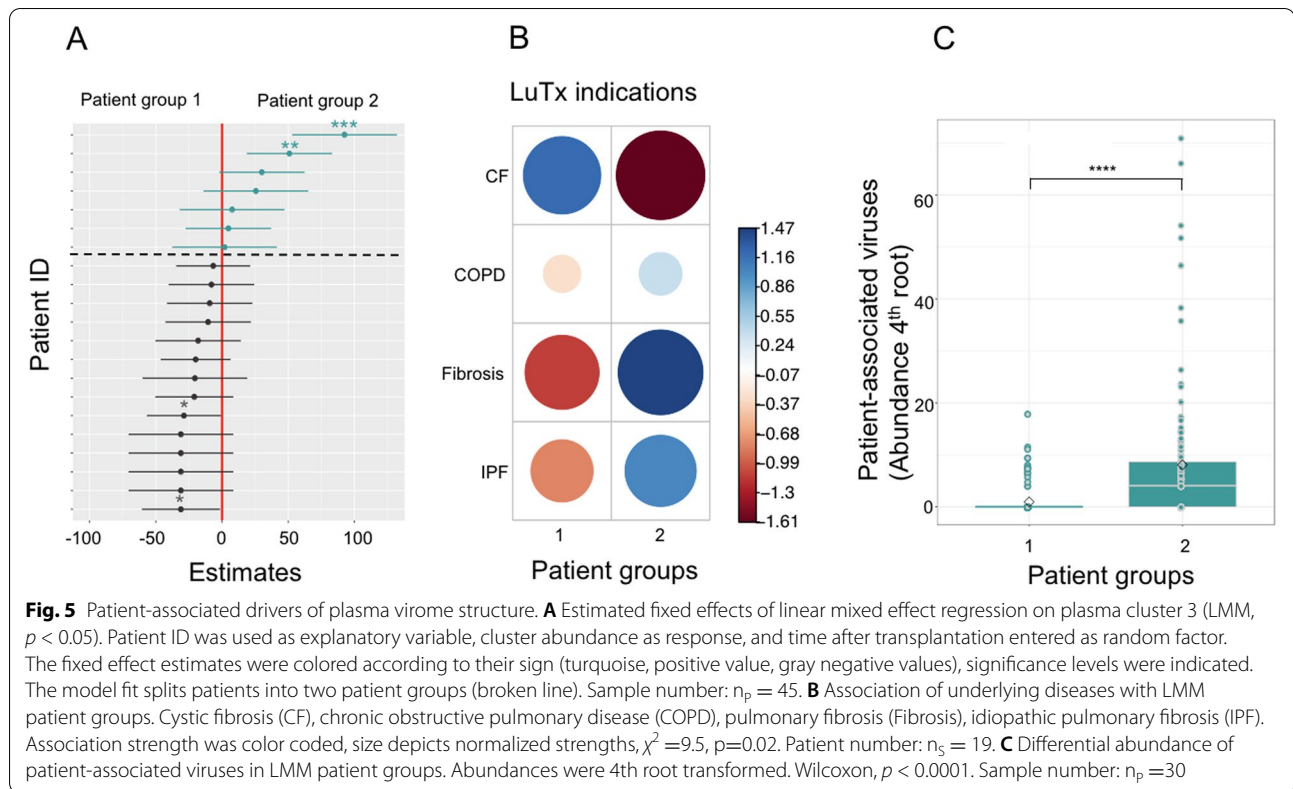




within one cluster display similar activity in response to the local environment, which is distinct from viruses in antagonistic clusters. Moreover, the perturbation of the microenvironment after transplantation seems to benefit replication of certain virus strains, which at the same time preclude the increase of others in a predictable way.

Regarding the more pronounced time dynamics in the airways, the modeling presented in this retrospective study accurately reflects the characteristics of the studied data. However, the direct comparison between body compartments at early time points and potential temporal dynamics in plasma are limited due to lack of early, matched plasma samples in the collection and resulting sample size differences. Therefore, it cannot be excluded that also the plasma virome exhibits time dynamics after lung transplantation. We recommend to investigate the time behavior of the post-transplant plasma virome dynamics in an increased sample collection.

An important clinical question is whether the abundance of specific TTV strains or clusters can be linked to specific clinical complications after transplantation. We tested potential correlations with patient survival, the prevalence of CLAD development, and CMV state of donor and recipient. From these analyses, no significant drivers could be identified, but confirmation by a much larger patient collective would be required. However, we observed significant association between specific clusters in the blood compartment of LTRs and clinical background. Interestingly, patients with CF developed a significantly higher level of the TTV3 cluster than patients with COPD or fibrosis. Of note, the CF patient group tended to be younger than other LTRs, which could imply a potential role of age-dependent virome patterns in the lung [51]. Data correlation could not distinguish between these co-varying predictors; therefore, further validation in an extended patient cohort, with specific investigation on the underlying reasons is needed.



Conclusions

Our study on a limited data set provides further evidence for mutual dependencies among different *Anelloviridae* and other opportunistic viruses that result in specific virome signatures that can be linked to host parameters, body site and time after lung transplantation.

Abbreviations

LTR: Lung transplant recipients; COPD: Chronic obstructive pulmonary disease; IPF: Idiopathic pulmonary fibrosis; CLAD: Chronic lung allograft dysfunction; TTV: *Alpha torque teno virus*; 1-5 TTV1-5: *Alpha torque teno virus* genogroup; BTV: *Beta torque teno virus*; BTV: *Gamma torque teno virus*; BALF: Bronchoalveolar lavage fluid; n_p : Number of plasma samples; n_l : Number of BALF (lung) samples; n_s : Number of subjects.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40168-022-01244-9>.

Additional file 1: SI Figure 1. Association of *Anelloviruses* and non-*Anelloviruses* in lung and blood compartments. **SI Figure 2.** Temporal changes of community cluster abundances show significant time-association in the lung. **SI Figure 3.** Association of age with patient grouping. **SI Table 1.** Description of patient demographics (A) and samples (B) used in this study. **SI Table 2.** Data structure and variable hierarchy.

Acknowledgements

We gratefully acknowledge the many individuals who provided the samples used in this study.

Authors' contributions

SW provided conceptualization, investigation, data analysis, methodology, visualization, manuscript writing, and funding acquisition. IG conducted methodology, investigation, and manuscript writing. BF contributed methodology, bioinformatics, and data curation. NR and SS provided data and metadata; PJ contributed patient samples. SK contributed conceptualization, supervision, and writing. EPS provided conceptualization, samples, resources, supervision, and writing. All authors read and approved the final manuscript.

Authors' information

Not applicable.

Funding

SW was supported by the Austrian Science Fund (FWF) Elise Richter V585-B31. The study was supported by the Center of Virology of the Medical University Vienna.

Availability of data and materials

Metagenomic sequencing data as well as meta information were deposited at NCBI's Sequence Read Archive under the BioProject accession number PRJNA764588 with individual sample accessions SAMN21511896-SAMN21512078. R code supporting the findings of this study was made available as supplementary material.

Declarations

Ethics approval and consent to participate

The study was approved by the local Ethics Committee at the Medical University of Vienna (EK-number 1880/2016 and EK-number 1710/2014).

Consent for publication

All authors concur to submit and publish this report to BMC Microbiome.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Research Laboratory of Infection Biology, Department of Medicine I, Medical University of Vienna, Vienna, Austria. ²Konrad Lorenz Institute for Evolution and Cognition Research, Klosterneuburg, Austria. ³Center of Virology, Medical University Vienna, Vienna, Austria. ⁴Department for Internal Medicine, Diabetology, Endocrinology, Diakonissenkrankenhaus, ViDia Kliniken, Karlsruhe, Germany. ⁵Division of Thoracic Surgery, Department of Surgery, Medical University of Vienna, Vienna, Austria.

Received: 15 July 2021 Accepted: 10 February 2022

Published online: 23 April 2022

References

- Delwart E. A roadmap to the human virome. *PLoS Pathog.* 2013;9(2):e1003146. <https://doi.org/10.1371/journal.ppat.1003146>.
- Wylie KM, Weinstock GM, Storch GA. Emerging view of the human virome. *Transl Res.* 2012;160(4):283–90. <https://doi.org/10.1016/j.trsl.2012.03.006>.
- Mitchell AB, Mourad B, Morgan LC, Oliver BGG, Glanville AR. Transplanting the pulmonary virome: dynamics of transient populations. *J Heart Lung Transplant.* 2018;37(9):1111–8. <https://doi.org/10.1016/j.healun.2018.06.002>.
- Segura-Wang M, Gorzer I, Jaksch P, Puchhammer-Stockl E. Temporal dynamics of the lung and plasma viromes in lung transplant recipients. *PLoS One.* 2018;13(7):e0200428. <https://doi.org/10.1371/journal.pone.0200428>.
- De Vlaminck I, Khush KK, Strehl C, Kohli B, Luikart H, Neff NF, et al. Temporal response of the human virome to immunosuppression and antiviral therapy. *Cell.* 2013;155(5):1178–87. <https://doi.org/10.1016/j.cell.2013.10.034>.
- Zanella MC, Cordey S, Kaiser L. Beyond *Cytomegalovirus* and *Epstein-Barr virus*: a review of viruses composing the blood virome of solid organ transplant and hematopoietic stem cell transplant recipients. *Clin Microbiol Rev.* 2020;33(4). <https://doi.org/10.1128/CMR.00027-20>.
- Focosi D, Antonelli G, Pistello M, Maggi F. *Torque teno virus*: the human virome from bench to bedside. *Clin Microbiol Infect.* 2016;22(7):589–93. <https://doi.org/10.1016/j.cmi.2016.04.007>.
- Young JC, Chehoud C, Bittinger K, Bailey A, Diamond JM, Cantu E, et al. Viral metagenomics reveal blooms of *Anelloviruses* in the respiratory tract of lung transplant recipients. *Am J Transplant.* 2015;15(1):200–9. <https://doi.org/10.1111/ajt.13031>.
- Gorzer I, Haloschan M, Jaksch P, Klepetko W, Puchhammer-Stockl E. Plasma DNA levels of *Torque teno virus* and immunosuppression after lung transplantation. *J Heart Lung Transplant.* 2014;33(3):320–3. <https://doi.org/10.1016/j.healun.2013.12.007>.
- Maggi F, Focosi D, Statzu M, Bianco G, Costa C, Macera L, et al. Early post-transplant *Torque teno virus* viremia predicts *Cytomegalovirus* reactivations in solid organ transplant recipients. *Sci Rep.* 2018;8(1):15490. <https://doi.org/10.1038/s41598-018-33909-7>.
- Jaksch P, Kundi M, Gorzer I, Murakozy G, Lambers C, Benazzo A, et al. *Torque teno virus* as a novel biomarker targeting the efficacy of immunosuppression after lung transplantation. *J Infect Dis.* 2018;218(12):1922–8. <https://doi.org/10.1093/infdis/jiy452>.
- Strassl R, Schiemann M, Doberer K, Gorzer I, Puchhammer-Stockl E, Eskandary F, et al. Quantification of *Torque teno virus* viremia as a prospective biomarker for infectious disease in kidney allograft recipients. *J Infect Dis.* 2018;218(8):1191–9. <https://doi.org/10.1093/infdis/jiy306>.
- Spandole S, Cimponeriu D, Berca LM, Mihaescu G. Human *Anelloviruses*: an update of molecular, epidemiological and clinical aspects. *Arch Virol.* 2015;160(4):893–908. <https://doi.org/10.1007/s00705-015-2363-9>.
- Gorzer I, Jaksch P, Strassl R, Klepetko W, Puchhammer-Stockl E. Association between plasma *Torque teno virus* level and chronic lung allograft dysfunction after lung transplantation. *J Heart Lung Transplant.* 2017;36(3):366–8. <https://doi.org/10.1016/j.healun.2016.10.011>.
- Mitchell AB, Glanville AR. Kinetics of TTV-DNA plasma load: a global measure of immune suppression? *Transplantation.* 2019;103(4):660–1. <https://doi.org/10.1097/TP.0000000000002437>.
- Lewandowska DW, Zagordi O, Geissberger FD, Kufner V, Schmutz S, Boni J, et al. Optimization and validation of sample preparation for metagenomic sequencing of viruses in clinical samples. *Microbiome.* 2017;5(1):94. <https://doi.org/10.1186/s40168-017-0317-z>.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol.* 1990;215(3):403–10. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2).
- Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. *Nucleic Acids Res.* 2016;44(D1):D67–72. <https://doi.org/10.1093/nar/gkv1276>.
- Biagini P, Bendinelli M, Hino S, Kakkola L, Mankertz A, Niel C, et al. *Anelloviridae*. In: King AE, Adams M, Carstens E, Lefkowitz E, editors. *Virus taxonomy: Ninth report of the international committee on taxonomy of viruses*; Elsevier Academic Press; 2012.
- Aziz RK, Dwivedi B, Akhter S, Breitbart M, Edwards RA. Multidimensional metrics for estimating phage abundance, distribution, gene density, and sequence coverage in metagenomes. *Front Microbiol.* 2015;6:381. <https://doi.org/10.3389/fmicb.2015.00381>.
- Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, et al. Enterotypes of the human gut microbiome. *Nature.* 2011;473(7346):174–80. <https://doi.org/10.1038/nature09944>.
- Anderson MJ. Permutational multivariate analysis of variance (PERMANOVA): Wiley StatsRef: Statistics Reference Online; 2017. <https://doi.org/10.1002/9781118445112.stat07841>.
- Bittinger K. *usedist*: Distance matrix utilities. R package version 040. 2020; <https://cran.r-project.org/package=usedist>.
- Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlenn D, et al. *vegan*: Community ecology package. R package version 25-6 2019; <https://cran.r-project.org/package=vegan>.
- Friedman J, Alm EJ. Inferring correlation networks from genomic survey data. *PLoS Comput Biol.* 2012;8(9):e1002687. <https://doi.org/10.1371/journal.pcbi.1002687>.
- Maechler M, Rousseeuw P, Struyf A, Hubert M, Hornik K. *cluster*: Cluster analysis basics and extensions. R package version 212. 2021; <https://cran.r-project.org/package=cluster>.
- Gouy A, Daub JT, Excoffier L. Detecting gene subnetworks under selection in biological pathways. *Nucleic Acids Res.* 2017;45(16):e149. <https://doi.org/10.1093/nar/gkx626>.
- Hennig C. *fpc*: Flexible procedures for clustering. R package version 2.2.9. 2020; <https://cran.r-project.org/web/packages/fpc>.
- Moon KR, van Dijk D, Wang Z, Gigante S, Burkhardt DB, Chen WS, et al. Visualizing structure and transitions in high-dimensional biological data. *Nat Biotechnol.* 2019;37(12):1482–92. <https://doi.org/10.1038/s41587-019-0336-3>.
- Revelle W. *psych*: Procedures for psychological, psychometric, and personality research. R package version 1912. 2019; <https://cran.r-project.org/package=psych>.
- Wei T, Simko V. R package *corrplot*: Visualization of a correlation matrix (Version 0.84). 2021. <https://github.com/taiyun/corrplot>.
- Kuznetsova A, Brockhoff PB, Christensen RH. *lmerTest*: tests in linear mixed effects models. *J Stat Softw.* 2017;82(13):1–26.
- Fox J, Weisberg S. *An R companion to applied regression*, 3rd Edition. Thousand Oaks; 2019. <https://socialsciences.mcmaster.ca/jfox/Books/Companion/index.html>
- Wickham H. *ggplot2*: elegant graphics for data analysis. New York: Springer-Verlag; 2016. <https://ggplot2.tidyverse.org/>
- Bushman F, Liang G. Assembly of the virome in newborn human infants. *Curr Opin Virol.* 2021;48:17–22. <https://doi.org/10.1016/j.coviro.2021.03.004>.
- Liang G, Bushman FD. The human virome: Assembly, composition and host interactions. *Nat Rev Microbiol.* 2021. <https://doi.org/10.1038/s41579-021-00536-5>.
- Abbas AA, Diamond JM, Chehoud C, Chang B, Kotzin JJ, Young JC, et al. The perioperative lung transplant virome: *Torque teno viruses* are elevated in donor lungs and show divergent dynamics in primary graft dysfunction. *Am J Transplant.* 2017;17(5):1313–24. <https://doi.org/10.1111/ajt.14076>.
- Kulifaj D, Tilloy V, Scaon E, Guerin E, Essig M, Pichon N, et al. Viral metagenomics analysis of kidney donors and recipients: *Torque teno virus* genotyping and prevalence. *J Med Virol.* 2020. <https://doi.org/10.1002/jmv.26298>.
- Maggi F, Pifferi M, Fornai C, Andreoli E, Tempestini E, Vatteroni M, et al. TT virus in the nasal secretions of children with acute respiratory diseases:

- relations to viremia and disease severity. *J Virol.* 2003;77(4):2418–25. <https://doi.org/10.1128/jvi.77.4.2418-2425.2003>.
40. Jelcic I, Hotz-Wagenblatt A, Hunziker A, Zur Hausen H, de Villiers EM. Isolation of multiple TT virus genotypes from spleen biopsy tissue from a Hodgkin's disease patient: genome reorganization and diversity in the hypervariable region. *J Virol.* 2004;78(14):7498–507. <https://doi.org/10.1128/JVI.78.14.7498-7507.2004>.
 41. Khudyakov YE, Cong ME, Nichols B, Reed D, Dou XG, Viazov SO, et al. Sequence heterogeneity of TT virus and closely related viruses. *J Virol.* 2000;74(7):2990–3000. <https://doi.org/10.1128/jvi.74.7.2990-3000.2000>.
 42. Schultze JL, Aschenbrenner AC. COVID-19 and the human innate immune system. *Cell.* 2021;184(7):1671–92. <https://doi.org/10.1016/j.cell.2021.02.029>.
 43. Maggi F, Pistello M, Vatteroni M, Presciuttini S, Marchi S, Isola P, et al. Dynamics of persistent TT virus infection, as determined in patients treated with alpha interferon for concomitant hepatitis C virus infection. *J Virol.* 2001;75(24):11999–2004. <https://doi.org/10.1128/JVI.75.24.11999-12004.2001>.
 44. Moen EM, Sagedal S, Bjoro K, Degre M, Opstad PK, Grinde B. Effect of immune modulation on TT virus (TTV) and TTV-like-mini-virus (TLMV) viremia. *J Med Virol.* 2003;70(1):177–82. <https://doi.org/10.1002/jmv.10356>.
 45. Borkosky SS, Whitley C, Kopp-Schneider A, zur Hausen H, de Villiers EM. *Epstein-Barr virus stimulates Torque teno virus replication: a possible relationship to multiple sclerosis.* *PLoS One.* 2012;7(2):e32160. <https://doi.org/10.1371/journal.pone.0032160>.
 46. Maggi F, Andreoli E, Lanini L, Fornai C, Vatteroni M, Pistello M, et al. Relationships between total plasma load of *torquetenovirus* (TTV) and TTV genogroups carried. *J Clin Microbiol.* 2005;43(9):4807–10. <https://doi.org/10.1128/JCM.43.9.4807-4810.2005>.
 47. Okamoto H, Takahashi M, Kato N, Fukuda M, Tawara A, Fukuda S, et al. Sequestration of TT virus of restricted genotypes in peripheral blood mononuclear cells. *J Virol.* 2000;74(21):10236–9. <https://doi.org/10.1128/jvi.74.21.10236-10239.2000>.
 48. Mori Y, Yamanishi K, Arvin A, Campadelli-Fiume G, Mocarski E, Moore PS, et al. HHV-6A, 6B, and 7: pathogenesis, host response, and clinical disease. In: *Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis.* Cambridge: Cambridge University Press; 2007. Chapter 46.
 49. Das S, Bernasconi E, Koutsokera A, Wurlod DA, Tripathi V, Bonilla-Rosso G, et al. A prevalent and culturable microbiota links ecological balance to clinical stability of the human lung after transplantation. *Nat Commun.* 2021;12(1):2126. <https://doi.org/10.1038/s41467-021-22344-4>.
 50. Burian Z, Szabo H, Szekeley G, Gyurkovits K, Pankovics P, Farkas T, et al. Detection and follow-up of *Torque teno midi virus* ("small anelloviruses") in nasopharyngeal aspirates and three other human body fluids in children. *Arch Virol.* 2011;156(9):1537–41. <https://doi.org/10.1007/s00705-011-1021-0>.
 51. 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(5):724–40.e8. <https://doi.org/10.1016/j.chom.2020.08.003>.

Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

