



Breast Milk Virome and Bacterial Microbiome Resilience in Kenyan Women Living with HIV

Rabia Maqsood,^a Joshua B. Reus,^a Lily I. Wu,^a LaRinda A. Holland,^a Ruth Nduati,^b Dorothy Mbori-Ngacha,^c Elizabeth Maleche-Obimbo,^b Emily R. Begnel,^d  Soren Gantt,^e Ednah Ojee,^b Dalton Wamalwa,^b Grace John-Stewart,^d Jennifer Slyker,^{d,f} Dara A. Lehman,^{d,g}  Efreem S. Lim^{a,h}

^aCenter for Fundamental and Applied Microbiomics, Biodesign Institute, Arizona State University, Tempe, Arizona, USA

^bDepartment of Paediatrics and Child Health, University of Nairobi, Kenyatta National Hospital, Nairobi, Kenya

^cUnited Nations Children's Fund (UNICEF), Nairobi, Kenya

^dDepartment of Global Health, University of Washington, Seattle, Washington, USA

^eDépartement de Microbiologie, Infectiologie et Immunologie, Université de Montréal, Centre de Recherche du CHU Ste-Justine, Montréal, Québec, Canada

^fDepartment of Epidemiology, University of Washington, Seattle, Washington, USA

^gDivision of Human Biology, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA

^hSchool of Life Sciences, Arizona State University, Tempe, Arizona, USA

ABSTRACT Breast milk is nutritionally and immunologically beneficial in early life but is also a potential source of infection. Little is known about breast milk microbiota of women living with HIV (WLHIV), the impact of severe immunosuppression, and the contribution to mortality of HIV-exposed infants. Here, we performed metagenomic sequencing to characterize the bacterial microbiome and DNA virome of breast milk samples at 1 month postpartum from Kenyan WLHIV who were not receiving combination antiretroviral therapy (cART), 23 women with CD4 counts of <250 and 30 women with CD4 of >500; and additionally, 19 WLHIV with infants that lived and 26 WLHIV with infants that died during the first 2 years of life were included. We found that breast milk bacterial microbiomes in this study population were highly diverse but shared a core community composed of the *Streptococcaceae*, *Staphylococcaceae*, *Moraxellaceae*, and *Eubacteriaceae* families. The breast milk virome was dominated by human cytomegalovirus (CMV) and included the bacteriophage families *Myoviridae*, *Siphoviridae*, and *Podoviridae*. Bacterial microbiome and virome profiles and diversity were not significantly altered by HIV immunosuppression, as defined by a CD4 of <250. CMV viral load was not associated with maternal CD4 counts or infant mortality. In conclusion, we show that the core bacterial and viral communities are resilient in breast milk despite immunosuppression in WLHIV.

IMPORTANCE Breastfeeding plays an important role in seeding the infant gut microbiome and mammary health. Although most studies focus on the diverse breast milk bacterial communities, little is known about the viral communities harbored in breast milk. We performed the first breast milk virome study of an HIV population. In this study cohort of Kenyan women living with HIV from the pre-antiretroviral therapy era, we found that breast milk harbors a core bacterial microbiome and a virome dominated by human cytomegalovirus. The virome and bacterial microbiome were not substantially altered by immunosuppression or associated with infant mortality. Together, these findings indicate resilience of the microbial community in breast milk compartmentalization. These findings advance our fundamental understanding of the breast milk core microbiome and virome interactions in the context of HIV disease.


KEYWORDS breast milk, CMV, HIV, microbiome, virome

Citation Maqsood R, Reus JB, Wu LI, Holland LA, Nduati R, Mbori-Ngacha D, Maleche-Obimbo E, Begnel ER, Gantt S, Ojee E, Wamalwa D, John-Stewart G, Slyker J, Lehman DA, Lim ES. 2021. Breast milk virome and bacterial microbiome resilience in Kenyan women living with HIV. *mSystems* 6:e01079-20. <https://doi.org/10.1128/mSystems.01079-20>.

Editor Emily K. Cope, Northern Arizona University

Copyright © 2021 Maqsood et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Efreem S. Lim, Efreem.Lim@asu.edu.

 First virome study of pre-ART women living with HIV finds remarkable microbiome resilience despite immunosuppression.

Received 16 October 2020

Accepted 23 February 2021

Published 16 March 2021

Breastfeeding confers many benefits to an infant, including optimal nutrition, protective immunity through antibodies, and seeding of the gut microbiota (1–3). Breast milk (BM) harbors a diverse community of bacteria (bacterial microbiome) that contributes to the establishment of healthy infant gut flora (1, 4–10). The small number of studies of the human bacterial microbiome in breast milk conducted to date have revealed common “core” bacterial taxa, comprising an array of bacterial families, including *Streptococcaceae* and *Staphylococcaceae*, consistently detected in the breast milk of women across multiple geographic regions (11–14). For example, even though women in rural Ethiopia had abundant *Rhizobium* in their breast milk, they still shared a high abundance of core taxa with women from other countries, such as Spain and Kenya (11). In addition, it has been shown that the bacterial taxa in breast milk and infant stool are highly correlated in early life (9), and thus alterations in the breast milk microbiome are likely to impact the establishment of a healthy infant gut microbiome.

Relatively little is known about the community of viruses (virome) in breast milk. A study of healthy U.S. women found that the majority of the viruses detected in breast milk were bacteriophages from the *Myoviridae*, *Siphoviridae*, and *Podoviridae* families, with very few eukaryotic viruses reported (15). While a relatively small number of pathogenic viruses are known to be transmitted through human milk (including HIV, cytomegalovirus [CMV], and human T lymphotropic virus type 1 [HTLV-1]) (16–18), these and other viruses may impact both the infant bacterial microbiome and virome through inflammation and immune modulation of the breast milk or infant gut, with subsequent relevance for child outcomes. Studies from women in Italy and the United States found evidence of mother-to-infant virome transmission, as breast milk and stool viromes from mother-infant pairs shared significant homology of bacteriophages (15, 19). Alterations in the breast milk virome may thus influence the establishment of both the infant virome and bacterial microbiome early in life, which could have long-term health consequences (20–24).

Emerging studies indicate that HIV disease has profound effects on microbiota. Immunodeficiency in progressive HIV infections is associated with alterations in the enteric bacterial microbiome and virome (25). For example, immunosuppression among individuals living with HIV (CD4 <200) has been observed to be discriminately associated with bacterial families such as *Enterobacteriaceae* and significantly more *Adenoviridae* in their virome (25). Although one study suggests that women living with HIV (WLHIV) and HIV-uninfected women have similar breast milk bacterial microbiomes (14), the effects of immunosuppression or antiretroviral therapy (ART) on the breast milk bacterial microbiome and virome are not well understood. Given the scarcity of data, it is important to characterize the breast milk bacterial microbiome and virome, particularly in populations with high infant mortality, such as sub-Saharan Africa and populations living with HIV. Based on the low risk of HIV transmission among virally suppressed women, and the nutritional and immunologic benefits of breast milk, current World Health Organization (WHO) guidance recommends women living with HIV (WLHIV) in low-resource settings exclusively breastfeed their infants for 6 months and continue to breastfeed for 1 year or longer as they receive combination antiretroviral therapy (cART) (26). Hence, in the current era of HIV treatment and prevention, HIV-exposed uninfected (HEU) infants will continue to be exposed to a myriad of viruses (other than HIV) in maternal milk. It remains unclear whether maternal HIV infection elicits alterations in the breast milk virome and microbiome or whether changes in the breast milk virome impact infant health outcomes. We used next-generation sequencing to test the hypothesis that severe immunosuppression alters the breast milk DNA virome and bacterial microbiome in WLHIV and that breast milk virome diversity may be associated with mortality in HIV-infected infants.

RESULTS

Comparison of bacterial microbiome and virome diversity between women with high and low CD4 counts. We performed bacterial microbiome 16S rRNA gene sequencing and DNA virome metagenomic sequencing on breast milk samples collected from Kenyan WLHIV at 1 month postpartum as follows: 30 women with CD4

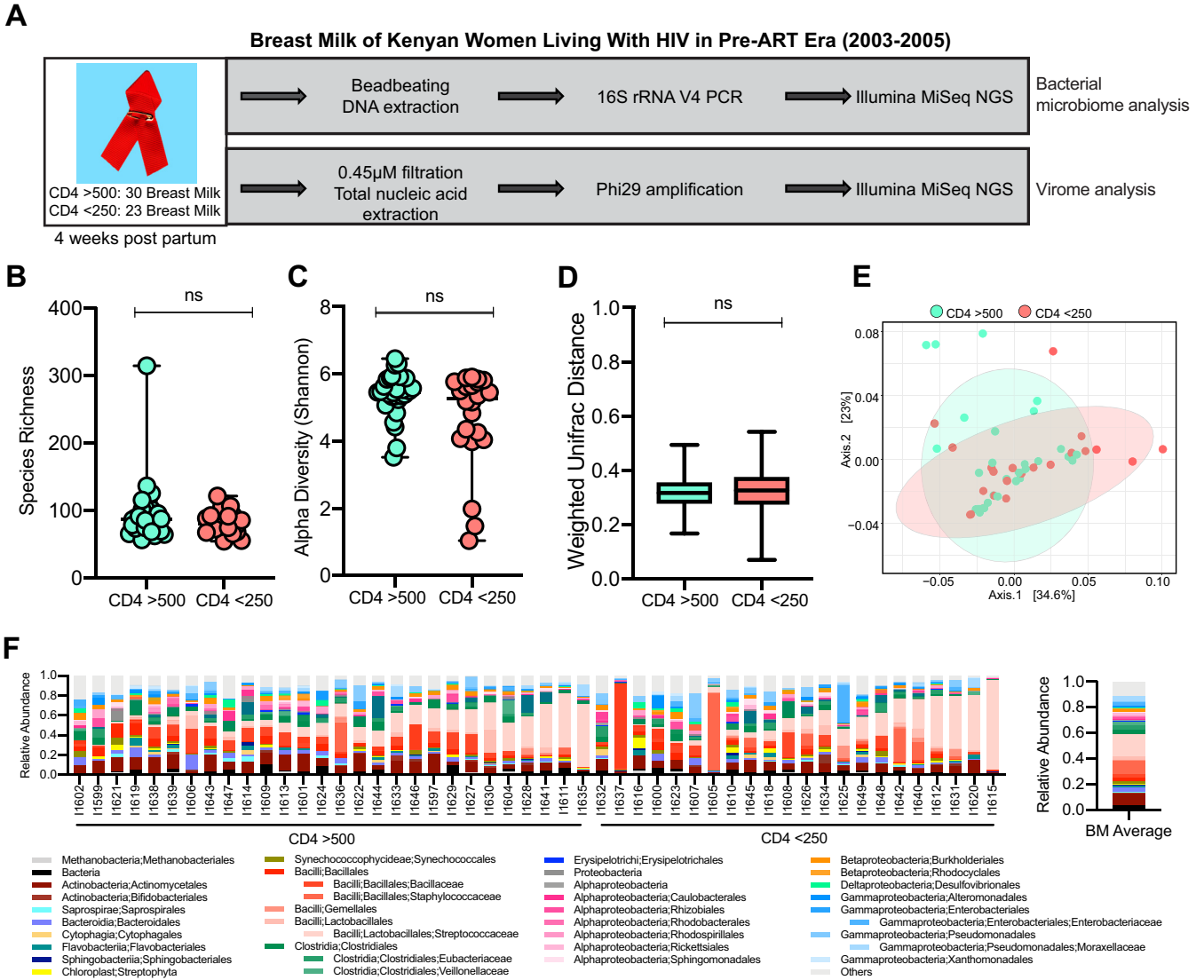


FIG 1 Bacterial microbiome analysis of BM samples from WLHIV by immunosuppression. (A) Overview of study design. (B) Richness of bacterial ASVs in BM samples (median with range). Statistical significance assessed by Mann-Whitney test. (C) Alpha diversity of bacterial ASVs in BM samples (median with range). Statistical significance assessed by Mann-Whitney test. (D) Weighted UniFrac pairwise comparison between samples (median with range). Statistical significance assessed by Mann-Whitney test. (E) PCoA plot of weighted UniFrac distances. Colors represent CD4 count groups. (F) Individual and average relative abundance of bacteria at order and family level in BM samples.

counts of >500 cells/ mm^3 and 23 with CD4 counts of <250 cells/ mm^3 (Fig. 1A, Table 1). As was the standard of care at that time of the original study (2003 to 2005), mothers were given short-course zidovudine during their last trimester to reduce the risk of mother-to-child transmission but received no other cART during follow-up. As expected, the median plasma and breast milk HIV RNA levels were significantly lower in women with CD4 counts of >500 than counts of <250 (plasma, 31,783 versus 455,400 copies/ml [$P < 0.0001$]; BM, 142.5 versus 5,380 copies/ml [$P < 0.0001$]; Fig. S1A and B).

To determine whether immunosuppression due to more severe HIV infection alters the breast milk bacterial microbiome, we compared breast milk samples with sufficient sequencing depth (see Materials and Methods) between women with CD4 counts of >500 ($n = 28$) to those with <250 ($n = 22$). The median bacterial microbiome sequence depth between the two groups was not significantly different (20,450 versus 21,875 average reads/sample for CD4 of >500 versus <250 , respectively; $P = 0.59$). Women with high and low CD4 counts had similar bacterial amplicon sequence variant (ASV)

TABLE 1 Characteristics of Kenyan WLHIV CD4 cohort

Maternal characteristics ^a	Data for samples		P value
	CD4 > 500 (n = 30)	CD4 < 250 (n = 23)	
Maternal age (yr)			
Mean (SD)	23.90 (4.04)	27.48 (5.2)	
Median (IQR)	23 (6.3)	26 (8)	0.012
CD4 counts			
Mean (SD)	832.6 (291.3)	158 (62.2)	
Median (IQR)	782 (368)	170 (102)	<0.0001
CD8 counts			
Mean (SD)	1,263 (498.2)	1,220 (576.8)	
Median (IQR)	1,187 (714.7)	1,066 (758)	0.67
BM HIV viral load (copies/ml)			
Mean (SD)	2,024 (5,856)	30,135 (56,151)	
Median (IQR)	142.5 (266.3)	5,380 (36,930)	<0.0001
Plasma HIV viral load (copies/ml)			
Mean (SD)	157,518 (448,113)	1,136,291 (1,720,360)	
Median (IQR)	31,783 (101,940)	455,400 (1,086,650)	<0.0001

^aIQR, interquartile range.

richness (median, 86.50 versus 80 for CD4 of >500 and <250, respectively; $P=0.65$), Shannon alpha diversity (median, 5.47 versus 5.27 for CD4 of >500 and <250; $P=0.13$), and weighted UniFrac distances (median, 0.32 versus 0.33 for CD4 of >500 and <250; $P=0.17$; Fig. 1B to D). In addition, the weighted principal-coordinate analysis (PCoA) did not show clustering by group (Fig. 1E). While the unweighted UniFrac distance was significantly higher in breast milk from women with CD4 of >500 than CD4 of <250 (median, 0.67 versus 0.64; $P < 0.0001$; Fig. S1C), the unweighted UniFrac PCoA plot showed four distinct clusters, all of which had samples from both CD4 groups (Fig. S1D). Both groups had high abundance of *Streptococcaceae* (16.6% in CD4 of >500 and 20% in CD4 of <250), *Staphylococcaceae* (6.8% and 14.2%), and *Actinomycetales* (10.6% and 7.7%) (Fig. 1F). Potentially discriminant bacterial taxa by CD4 group were rare and very low in abundance (Fig. S1E to G).

We next compared the DNA virome in breast milk samples with sufficient sequencing depth (see Materials and Methods) between women with CD4 counts of >500 ($n=26$) and CD4 of <250 ($n=22$). The median virome sequence depth between the two groups was not significantly different (544,318 versus 465,067 average reads/sample for CD4 of >500 versus <250, respectively; $P=0.32$). Breast milk virome richness was similar between the women with high and low CD4 counts (median, 38 versus 40 for CD4 of >500 versus <250, respectively; $P=0.95$; Fig. 2A). Shannon alpha diversity of the breast milk virome was significantly higher in women with CD4 of >500 than CD4 of <250 (median, 1.77 versus 0.93; $P=0.037$; Fig. 2B). The beta diversities were not significantly different between the women with high and low CD4 counts (median, 0.42 versus 0.41, respectively; $P=0.38$; Fig. 2C). PCoA analysis of virome sequence reads showed the two groups were overlapping without any distinct clusters, which was confirmed by PCoA generated by virome contig analysis (Fig. 2D, Fig. S1H). Cytomegalovirus (CMV) dominated the breast milk virome in both groups, and the average relative abundance of CMV sequence reads was higher in women with a CD4 count of <250 versus >500 (70.6% versus 52.9%; $P=0.05$). CMV sequencing read counts were strongly correlated with CMV genome copies from qPCR measurements (Pearson $r=0.89$, $P < 0.0001$) (Fig. 2F), validating that CMV abundance by next-generation sequencing (NGS) was reflective of actual viral loads in the breast milk (i.e., not the result of amplification bias). However, CMV viral loads, as measured by qPCR, were not significantly different between CD4 groups (Fig. S1I). Removing the CMV sequences from the virome data set eliminated

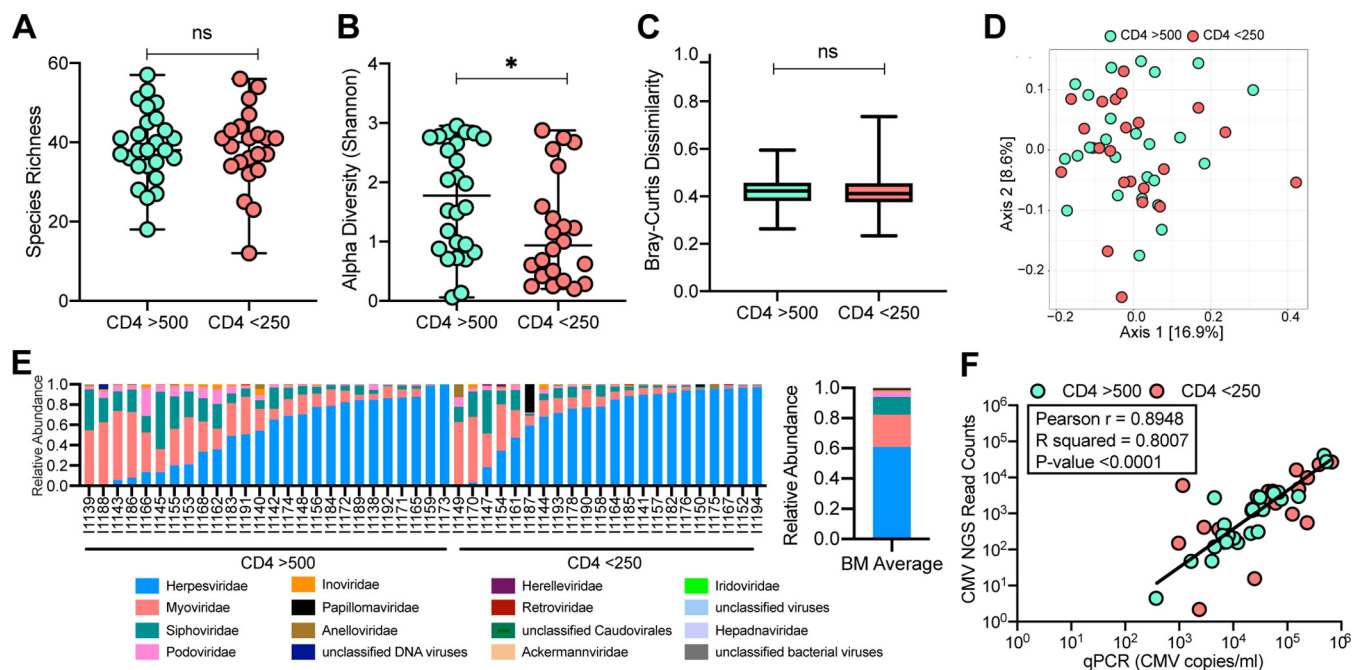


FIG 2 Virome analysis of BM samples from WLHIV by immunosuppression. (A) Richness of viral species in BM samples (median with range). Statistical significance assessed by Mann-Whitney test. (B) Alpha diversity of viral species in BM samples (median with range). Statistical significance assessed by Mann-Whitney test. (C) Unweighted Bray-Curtis distance box plot of BM samples (median with range). Statistical significance assessed by Mann-Whitney test. (D) Unweighted Bray-Curtis PCoA. Colors represent CD4 counts group. (E) Individual and average relative abundance of viruses at family level in BM samples. (F) Plot of CMV sequence reads (\log_{10}) versus CMV qPCR copy numbers (\log_{10}). Pearson correlation and linear regression line fit to data.

the difference in alpha diversity (median, 2.77 versus 2.83; $P=0.98$; Fig. S1J). With the exception of a higher relative abundance of CMV, the other potentially discriminant viruses by CD4 group were very low in abundance (Fig. S1L to N).

Characteristics of the core breast milk bacterial microbiome and virome in Kenyan women living with HIV. We found shared features of the bacterial microbiome across the WLHIV in our Kenyan cohort. An average of 49 bacterial families were detected in each sample, the following seven of which had an average relative abundance of $>2.5\%$ of sequence reads: *Streptococcaceae* (average relative abundance, 18.8%), *Staphylococcaceae* (10.1%), *Moraxellaceae* (4.1%), *Eubacteriaceae* (3.6%), *Veillonellaceae* (3.0%), *Bacillaceae* (2.8%), and *Enterobacteriaceae* (2.5%; Fig. 1F). While the relative abundance of each bacterial family differed between women, six bacterial families were present in $>80\%$ of breast milk samples in our cohort, supporting the presence of a shared core breast milk bacterial microbiome as follows: *Streptococcaceae* (present in 96% of breast milk samples), *Staphylococcaceae* (96%), *Moraxellaceae* (82%), *Eubacteriaceae* (96%), *Bacillaceae* (88%), and *Enterobacteriaceae* (86%).

The breast milk DNA virome was dominated by a high average relative abundance of *Herpesviridae* (60.1%) that was almost exclusively CMV (99.99%), which was found in 98% of breast milk samples in this cohort. Phylogenetic analysis of the assembled CMV genomes demonstrated diverse isolates, arguing against contamination (Fig. S2). Bacteriophages also dominated the virome with a high average relative abundance of *Myoviridae* (21.2%), *Siphoviridae* (11.8%), and *Podoviridae* (3.5%) (Fig. 2E).

Breast milk bacterial microbiome community states and associated bacteriophages.

Bacterial microbiome communities with similar composition profiles, also called community states, can indicate functional groups as in previous microbiome studies of different sample types (27–29). Breast milk microbiome community profiles were examined by performing Euclidean clustering on the bacterial microbiome data set (at the order taxonomic level). This analysis identified three distinct clusters indicative of bacterial community states (Fig. 3). The largest cluster had similar proportions of *Lactobacillales*, *Bacillales*, *Clostridiales*, *Actinomycetales*, and *Pseudomonadales* across the samples (Fig. 3, yellow cluster), the midsized cluster included seven samples

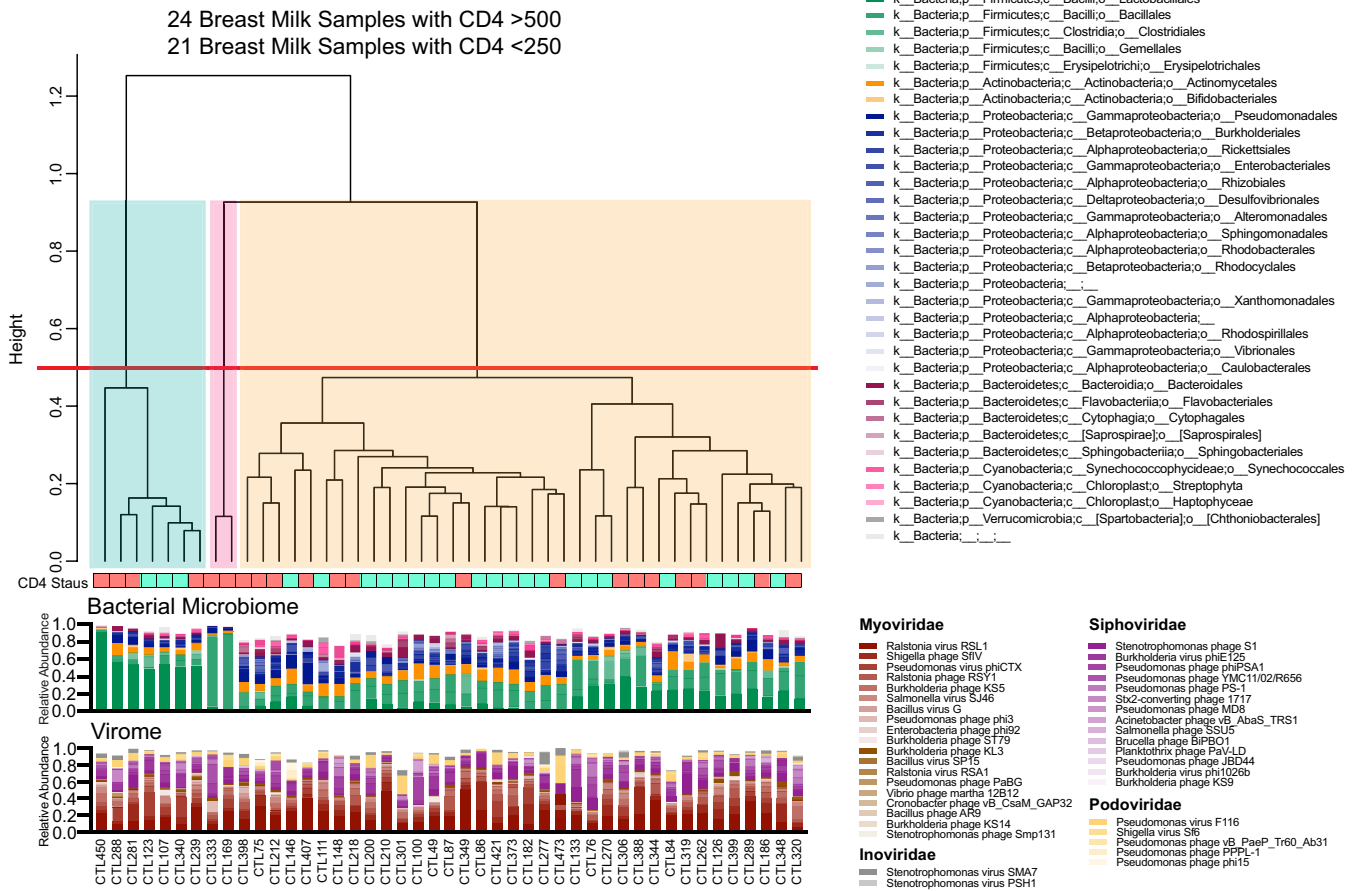


FIG 3 Breast milk microbiome community states. Relative abundance of bacterial at order level, clustered using Euclidean distance. Cutoff of 0.5 applied to separate clusters. Relative abundance of the most frequent phages for the corresponding samples plotted.

dominated by *Lactobacillales* (blue cluster), and the smallest cluster had only two samples which were both dominated by *Bacillales* (pink cluster).

As bacteriophages were found in similar frequency in the breast milk samples but differed in relative abundance between women, it was hypothesized that individual breast milk DNA virome profiles and bacterial microbiome profiles might influence each other. A number of bacteriophages were found across community states at high abundance, including *Ralstonia* virus RSL1 (average relative abundance, 20.5%), *Shigella* phage SfIV (8.5%), *Pseudomonas* virus phiCTX (5.8%), *Stenotrophomonas* phage S1 (5.8%), *Ralstonia* phage RSY1 (5.7%), and *Burkholderia* virus phiE125 (5.7%), while other phages were present at lower relative abundance of below 5% (Fig. 3, Fig. S3A). However, bacteriophage community state clusters did not correspond to the bacterial community states (Fig. S3B). To validate this, we also assembled and identified bacteriophage contigs using VirSorter. Using these assembled bacteriophage contig sequences, we confirmed that the bacterial microbiome community states did not have distinct bacteriophage profiles (Fig. S3C). We did not find any significant differences in viral loads (plasma HIV, breast milk HIV, or breast milk CMV) between the community states. Taken together, this indicates that breast milk harbors distinct bacterial microbiome community states that are not associated with specific bacteriophage profiles.

Breast milk virome and infant mortality. We next sought to evaluate whether the breast milk DNA virome might be associated with infant health outcomes, defined as HIV-infected infants that lived or died during the first 2 years of life. A second set of breast milk samples at 1 month postpartum was sequenced (Table 2). The women in this substudy were transmitting mothers whose infants acquired HIV, and they did not overlap the CD4 substudy above. A total of 45 breast milk samples were sequenced, of

TABLE 2 Characteristics of Kenyan WLHIV infant mortality cohort

Infant status ^a	Data for infants that:		P value
	Lived (n = 19)	Died (n = 26)	
Maternal age (yr)			
Mean (SD)	24.5 (4.2)	24.6 (4.06)	
Median (IQR)	25 (5)	24.5 (6.3)	0.90
CD4 counts			
Mean (SD)	585.3 (300)	496.8 (185.4)	
Median (IQR)	555 (331)	506 (256.8)	>0.48
CD8 counts			
Mean (SD)	1,168 (729)	1,257 (729)	
Median (IQR)	1,058 (835)	1,167 (672.5)	0.32
BM HIV viral load (copies/ml)			
Mean (SD)	8,742 (15,198)	7,482 (12,780)	
Median (IQR)	880 (6,925)	2,010 (4,844)	0.26
Plasma HIV viral load (copies/ml)			
Mean (SD)	494,702 (1,165,814)	5,046,281 (17,813,193)	
Median (IQR)	65,920 (502,940)	258,570 (660,773)	0.039

^aIQR, interquartile range.

which 9 were omitted due to limited sequencing depth, resulting in 14 WLHIV with infants that lived and 22 WLHIV with infants that died. There was insufficient specimen remaining to also conduct 16S bacterial microbiome analyses on these samples. Maternal plasma HIV-1 viral load was significantly higher in the group with infants that died, but not by breast milk HIV-1 viral load (Fig. S4A and B). As in the prior sample set selected for maternal CD4 count (Fig. 2E), the breast milk virome in this sample set selected for infant mortality was also dominated by CMV. However, no significant difference was found in average relative abundance of CMV (97.6% lived and 89.7% died, respectively; $P=0.58$) or CMV DNA level by quantitative PCR (qPCR) between women whose infants lived versus died (median, 29,218 versus 18,302 copies/ml; $P=0.96$) (Fig. 4A, Fig. S4C). CMV detected by NGS was correlated with CMV viral load (Fig. 4B). There was no significant difference in the virome species richness (Fig. 4C) or alpha diversity (Fig. 4D) between infant mortality groups. Virome beta diversity was significantly higher in the mothers of infants who died than the mothers of those who lived (0.56 lived versus 0.61 died, $P=0.002$) (Fig. 4E and F). We found no significant difference in beta diversity when we omit anellovirus sequences (Fig. S4D and E). Although this suggests that the differences may be driven in part by anelloviruses, there was no significant difference in anellovirus levels by qPCR between infant outcome groups (Fig. S4F).

DISCUSSION

We investigated the breast milk DNA virome and bacterial microbiome of Kenyan WLHIV that were not on cART and found that breast milk harbors a core bacterial microbiome and a virome dominated by CMV. Notably, this study is unique, as samples were collected during a time when cART was not the standard of care in Kenya, which allowed us to compare women with extreme differences in CD4 count. We found similar bacterial microbiome and DNA virome profiles and diversity across women with high versus low CD4 count and no specific bacteria or viruses that clearly differed based on CD4. Our findings indicate that the breast milk microbiota is highly resilient and is not substantially altered by severe immunosuppression in Kenyan WLHIV.

Previous studies found similar breast milk bacterial microbiomes comparing WLHIV receiving cART and women without HIV infection (14). Taken together with our data from women not on cART, this suggests that the breast milk bacterial microbiome and

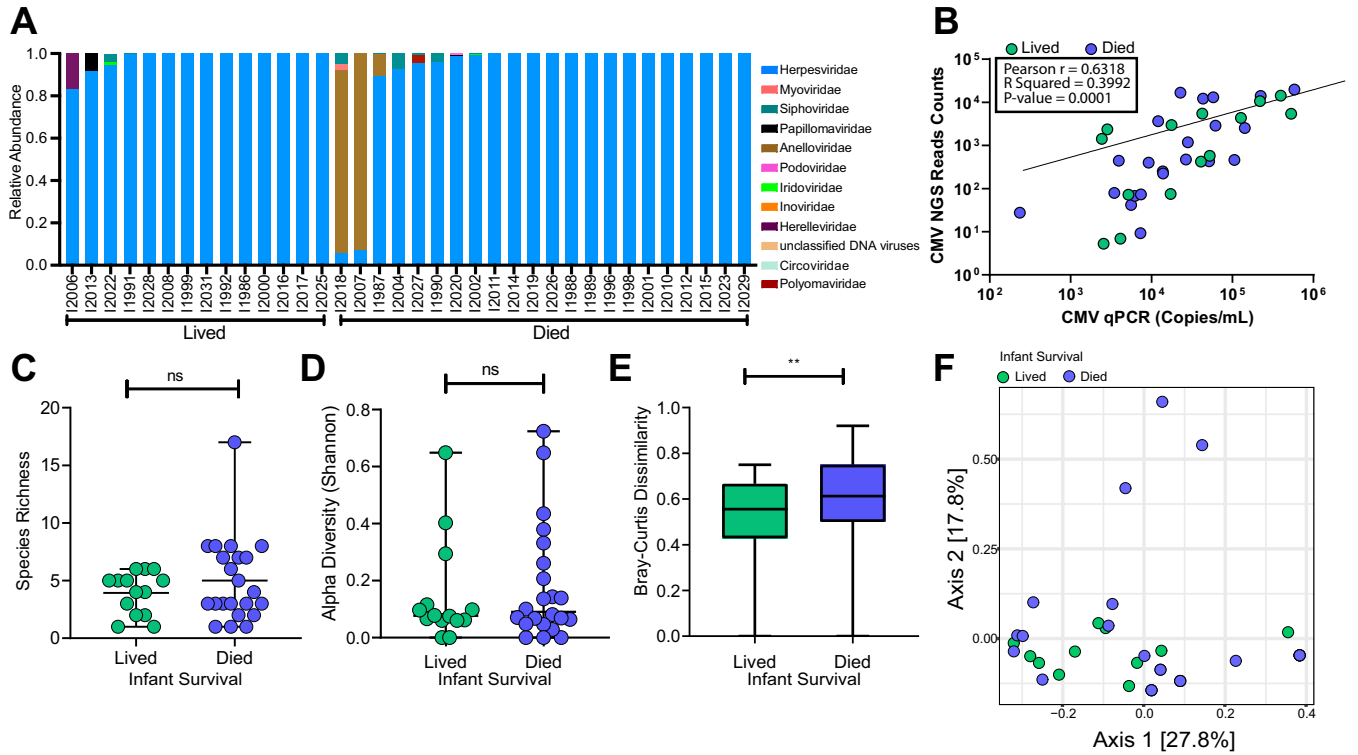


FIG 4 Virome analysis of BM samples from WLHIV by infant mortality. (A) Family-level relative abundance of breast milk virome. (B) Pearson correlation of CMV qPCR copy number and CMV NGS read counts normalized to read depth. (C) Species richness comparing breast milk samples grouped by infant survival (median with range). Statistical significance assessed by Mann-Whitney test. (D) Breast milk virome alpha diversity grouped by infant survival (median with range). Statistical significance assessed by Mann-Whitney test. (E) Box plot comparison of beta diversity (unweighted Bray-Curtis dissimilarity) (median with range). Statistical significance assessed by Mann-Whitney test. (F) PCoA of unweighted Bray-Curtis dissimilarity.

DNA virome are remarkably resilient to the effects of HIV infection. This resilience is notable, given the effects of immunosuppression that have been observed in the gut and other sites (30, 31). While alterations in the gut microbiome have previously been associated with progressive HIV disease (25, 32), the breast milk bacterial and virome profiles appear to be relatively protected from change. This difference may be due to differing effects of HIV replication at the two sites. In the first weeks of HIV infection, CD4⁺ T cells are dramatically depleted from the gut lamina propria, which is accompanied by profound changes to gut mucosal permeability and resident microbial communities; these changes are only partially reversed by long-term cART. In contrast, breast milk is a continuously replenished and concentrated mixture of innate immune factors originating from the blood or mammary tissues, including antibodies, human milk oligosaccharides, alpha-defensins, cytokines, chemokines, lactoferrin, and others (33). It appears that even in severe immunosuppression, this compartmentalized immune system remains sufficiently intact to preserve the microbiome community structures.

CMV was found in 98% of the samples and dominated the breast milk virome in women with both high and low CD4 counts, with differences in the relative abundance and alpha diversity due to increased CMV in immunosuppressed women. This is consistent with previous studies by our group and others that have reported nearly universal detection of CMV DNA in the breast milk of women with HIV and an inverse association between CD4 counts and breast milk CMV levels (34, 35). Because CMV was recovered from nearly all women in this study, we cannot determine the impact of CMV replication or CMV immune responses on the breast milk bacterial microbiome or virome in our study. However, CMV replication induces host inflammation that could affect reactivation of latent viruses or eubiosis (36), and further research is needed to understand whether it is a causal determinant in the core bacterial microbiome or virome. Whether

HIV infection alters the breast milk microbiome and virome will need to be addressed in future studies by comparing WLHIV to uninfected women. Changes in anelloviruses (small circular single-stranded DNA [ssDNA] viruses) detected in the blood and enteric viromes have been implicated with immunosuppression (25, 37, 38). We found that anellovirus was at low abundance in breast milk and was not significantly different between women with CD4 of >500 compared to CD4 of <250 ($P = 0.35$) (Fig. S1K). Similarly, anellovirus levels were not significantly different in infant mortality groups (Fig. S4D). Although virome beta diversity was significantly higher in the mothers of infants who died than those who lived, this was likely driven by three samples with high anellovirus abundance, which when removed, eliminated the differences (Fig. S4E to F).

Given evidence that the breast milk virome contributes to establishment of the infant gut virome (15), we hypothesized that the breast milk virome could affect the infant virome through transmission of pathogenic viruses, transmission of nonpathogenic but potentially inflammatory or immunomodulating viruses, or disruption of protective breast milk immunity for which there is inadequate maternal immune protection from ART-naïve women (39). While there was higher beta diversity in the breast milk virome of women whose infants died, we did not find differences in alpha diversity or richness and did not identify any discriminating taxa between the two groups. To reduce confounding, we only included mothers of children that became HIV-infected during follow-up, and thus the results might differ in populations of HIV-exposed but uninfected infants which were not included here. Because viruses and commensal bacteria in breast milk contribute to seeding the infant gut (9), innate immune responses in the infant and/or specific breast milk microbiota may influence the risk of HIV transmission, as has been observed with the vaginal microbiome (40, 41). However, there were only 15 infants in the original cohort that were HIV infected through breast milk transmission, not all of which had maternal breast milk samples available, and thus we were not able to examine the association between breast milk virome and HIV transmission. Because these specimens were not prospectively collected for microbiome studies and had been used for other virology and immunology studies, our sample size was limited, and we did not have adequate specimen volumes remaining to examine the association between breast milk bacterial microbiome and infant mortality. In addition, our study design was cross-sectional and limited to analysis of maternal samples at 1 month postpartum. Longitudinal studies are needed which adjust for confounding variables such as preterm birth, age of infant death, and malnutrition and that include infant samples to address whether maternal HIV affects maternal-infant transmission.

In summary, this study advances our fundamental understanding of the breast milk core microbiome and virome in the context of HIV disease in Kenya and demonstrates remarkable resilience of the microbiome to the effects of severe HIV-induced immunosuppression. More research is needed to identify the determinants of breast milk virome diversity and composition and the potential impact of CMV on microbiome and virome diversity.

MATERIALS AND METHODS

Study population. Breast milk samples were collected from women enrolled in the CTLs and Prevention of HIV-1 Transmission Study, in Nairobi, Kenya, between 1999 and 2002 (42). As described previously, women enrolled during pregnancy and were followed together with their infants for 2 years postpartum. As was the standard of care at the time of the study, mothers were given short-course zidovudine during their last trimester to reduce the risk of mother-to-child transmission but received no other cART during follow-up. In addition, exclusive breastfeeding was recommended for most WLHIV in sub-Saharan Africa because replacement feeding was not feasible, affordable, or safe for most women (43–45), and all women in this analysis breastfed their infants. The substudy presented here describes two cross-sectional sample sets and analyses of women from this cohort. In the first substudy, we selected all participants who had a breast milk sample available from 1 month postpartum and had CD4 counts of >500 cells/mm³ ($n = 30$) or <250 cells/mm³ ($n = 23$). In the second substudy, we selected available breast milk samples, 1 month postpartum, from mothers of HIV-infected infants that died ($n = 26$) or survived ($n = 19$) until age 2 years. These sample sizes resulted in 80% power to detect a 0.78 and 0.85 difference in Shannon diversity index in the CD4 and infant mortality substudies, respectively. This study

was approved by the University of Nairobi/Kenyatta National Hospital Ethics and Research Committee, the University of Washington, The Fred Hutchinson Cancer Research Center, and Arizona State University Institutional Review Boards; all women provided written informed consent for study participation and for storage and use of their specimens in future studies of HIV.

Virome and bacterial microbiome 16S rRNA sequencing. Breast milk samples were vortexed briefly, centrifuged at 13,500 rpm for 10 min, and separated into pellet (bacterial microbiome analysis) and supernatant (virome analysis). From the pellet, DNA extraction was performed using a DNeasy PowerSoil kit (Qiagen). From the supernatant fraction, total nucleic acid extraction was performed using the eMAG instrument (bioMérieux) according to the manufacturer's guidelines. Negative controls consisting of phosphate-buffered saline (PBS) spiked with *Enterobacteria* phage λ were extracted in parallel to the breast milk samples to assess contamination during extraction. Additionally, controls of PBS spiked with lambdavirus DNA were used to assess cross-sample contamination during amplification and sequencing. Bacterial and virome sequencing were performed as previously described (46, 47). Briefly, PCR was performed on the DNA from the DNeasy eluate using primers for the V4 region (F515/R806) for bacterial microbiome sequencing, and total nucleic acid from the eMAG eluate was amplified with GenomiPhi V2 (GE Healthcare) for virome sequencing. Libraries were sequenced on the Illumina MiSeq platform (version 2, 2 \times 250) at the ASU Biodesign Institute Genomics Core facility.

Bacterial microbiome 16S rRNA analysis. Illumina sequencing reads (2 \times 250 paired-end reads; average, 29,818 \pm 24,604 reads per sample) were processed through QIIME 2 using DADA2 to obtain denoised sequence reads and amplicon sequence variants (ASV) (48). To remove contaminant ASVs, the Decontam (version 1.2.1) package in R, which uses the "prevalence" method to identify contaminants, was applied at threshold 0.25 (Fig. S5A) (49). Additionally, two ASVs found in high abundance and similar proportions in participant samples and PBS were removed. To control for intersample depth variability, QIIME 2 was used to rarefy data to 8,000 reads (subsampling without replacement) and create a phylogenetic tree and taxonomy using the GreenGenes database (Fig. S5B). Out of 53 women's samples sequenced, 3 samples were omitted due to limited sequencing depth; 28 samples with CD4 counts of >500 and 22 samples with CD4 of <250 were analyzed. We performed core ecological metrics of alpha diversity, unweighted UniFrac distances, and weighted UniFrac distances using QIIME 2. PCoA was plotted using ggplot2 (version 3.3.1) in R (<https://ggplot2.tidyverse.org>; 50).

Virome analysis. Illumina sequencing reads (2 \times 250 paired-end reads; average, 1.40 \pm 0.75 million reads per sample) were quality filtered to remove adapters and low-quality bases using BBTools (version 37.64) (51). High-quality-filtered reads were queried against the viral RefSeq + Neighboring sequences database (downloaded March 2019) using BLASTx to identify viral reads. The taxonomic lineage of viral reads was assigned using the MEGAN (version 6.18.0) naive lowest common ancestor (LCA) algorithm with the parameters min support = 1, min support percentage = 0, and bitscore of 100 (52). Contamination was assessed by an unweighted Bray-Curtis PCoA plot clustering the breast milk samples separately from the PBS samples and Decontam at a threshold of 0.25 (Fig. S5C to E). Samples were removed if they had sequence reads below 168,000 (CD4 analysis) or 50,000 (infant mortality analysis), to which the data were then normalized, or if they had viral reads of <250 . Virus species with less than 1 normalized read were masked to remove low-abundance species. In the CD4 substudy, out of 53 women's samples sequenced, 5 samples were omitted due to quality control (QC) described above, 26 samples with CD4 counts of >500 and 22 samples with CD4 of <250 were analyzed. In the second infant mortality substudy, out of 45 women sequenced, 9 samples were omitted due to QC described above, 22 samples from mothers of HIV-infected infants that died and 14 that survived were analyzed. Ecological analyses (richness, alpha diversity-Shannon Index) were performed using vegan R package (version 2.5-5) (53). The unweighted Bray-Curtis beta diversity distances and PCoA plots were obtained by applying QIIME 2 and R package ggplot2 on a presence-absence species data set.

For validating the reads-based analysis, we built contigs using IDBA-UD with quality-filtered reads for each sample and used minimus 2 to merge any overlapping contigs (54, 55). The contigs of each sample were concatenated and were filtered to remove duplicates (minidentity = 99) and contigs with a length of <800 kb using bbtools. These contigs were then used to identify bacteriophages using VirSorter (parameters -virome, -db 1) (version 1.0.6), and the identified contigs were then used to build a database (54, 56). Quality-filtered reads for each sample were mapped to the VirSorter contig database using the Burrows-Wheeler Aligner MEM algorithm (BWA-MEM; -M -L 95,95) (version 0.7.17-r1188) to validate the cohort trends with PCoA plots (38). To determine the proportion of viral contigs (VirSorter) that were not identifiable (i.e., viral "dark matter"), we queried the viral contigs against the viral RefSeq + Neighboring sequences database using BLASTx (E value, $1E^{-3}$) (Fig. S5F).

qPCR assays. CMV qPCR was performed in a QuantStudio 3 (Applied Biosystems) PCR system in 20- μ l reaction mixtures containing 10 μ l of TaqMan fast universal PCR 2X master mix (Applied Biosystems), 1.8 μ l (10 μ M) of each primer (HCMV-FOR, 5'-TGGGCGAGGACACGAA-3', and HCMV-REV, 5'-TGAGGCTGGGAAGCTGACAT-3'), 0.5 μ l (10 μ M) probe (5' 56-FAM/TGG GCA ACC/ZEN/ACC GCA CTG AGG/3IABkFQ/-3'), 0.9 μ l water, and 5 μ l of extracted total nucleic acid (TNA) (57). The assay was performed with preliminary denaturation for 20 s at 95°C (slope, 4.14°C/s), followed by 40 cycles of denaturation at 95°C for 1 s (slope, 4.14°C/s) and annealing at 60°C for 20 s (slope, 3.17°C/s). A standard curve was generated using synthesized double-stranded DNA gene fragment (IDT gBlocks) of the CMV genome region from genome position 82468 to 82948 in reference to CMV complete genome GenBank accession number [MT044484.1](https://www.ncbi.nlm.nih.gov/nuccore/MT044484.1). Analysis was performed with a

0.04 threshold determined by the standard curve; detection measured by the VIC reporter and NFQ-MGB Quencher.

Anellovirus qPCR was performed in 25- μ l reaction mixtures containing 12.5 μ l of TaqMan fast universal PCR master mix (2X Applied Biosystems), 2.3 μ l (10 μ M) of each primer (ATV-FOR, 5'-GTGCCGIAGGTGAGTTTA-3', and ATV-REV, 5'-AGCCCGGCCAGTCC-3'), 0.7 μ l (10 μ M) probe (5' 56-FAM/TCAAGGGGCAATTCGGGCT/36-TAMSp/-3'), 4.2 μ l water, and 3 μ l of extracted TNA. The assay was performed with preliminary denaturation for 20s at 95°C (slope, 4.14°C/s), followed by 40 cycles of denaturation at 95°C for 1 s (slope, 4.14°C/s) and annealing at 60°C for 20 s (slope, 3.17°C/s). A standard curve was generated using a double-stranded gene fragment cloned an anellovirus genome corresponding to the genome sequence of position 1694 to 1756 in reference to Anellovirus complete genome GenBank accession number [MH649099.1](#). Analysis was performed with a 0.03 threshold as determined by the standard curve, measuring FAM reporter and NFQ-MGB Quencher.

Statistical analysis. Since sample selection criteria were based on women with CD4 counts of >500 or <250 and women with HIV-infected infants that lived or died within 2 years of life, we treated the data as ordinal. Additionally, since the data follow a nonnormal distribution, statistical significance of virome diversity and bacterial 16s diversity measures comparing groups of women based on CD4 or infant mortality was assessed using a nonparametric Mann-Whitney U test (two-tailed). Pearson correlation was performed between CMV sequence read counts from the virome sequencing and the CMV qPCR copy numbers. The Mann-Whitney U test (two-tailed) was used to determine significance between qPCR levels in groups of women.

Cooccurring bacterium and bacteriophage analysis. The Euclidean distance between the rarefied bacterial microbiome sequences at the order level was computed and used to cluster the samples. The relative abundance of each order of bacteria was plotted using the clustering order, and a cutoff of 0.5 was used to determine the breast milk community states. To evaluate trends of the phages in the 16S clusters, the relative abundances of the top phage species (as chosen by average and frequency across breast milk samples) were plotted for each sample to assess virome profiles within community states. Additionally, the community states of viromes were also clustered separately using Euclidean distance on the phage data set. We also corroborated the phage trend by plotting the top VirSorter identified bacteriophage contigs according to community states. We performed a one-way nonparametric analysis of variance (ANOVA) test (Kruskal-Wallis and Dunn's multiple-comparison test) to assess whether community states were associated with viral loads (copies/ml) of plasma HIV, breast milk HIV, or breast milk CMV. For cooccurrence analyses, 45 WLHIV had data for both virome and bacterial microbiome, 24 samples with CD4 counts of >500 and 21 samples with CD4 of <250.

Phylogenetic analysis. Contigs were assembled from quality-filtered sequencing reads with metaSPAdes (version 3.14.0) for each individual breast milk sample specimen (58). CMV contigs were extracted by mapping against the human herpesvirus 5 strain Merlin (GenBank Accession number [NC_006273](#)) using BWA. The CMV UL55, UL144, US27, and US28 genes were annotated and concatenated in Geneious Prime (version 2020.0.5) (59). Multiple sequence alignment was performed with CMV sequences and 11 reference genomes (NCBI GenBank accession numbers [FJ616285](#), [GQ221975](#), [KJ872542](#), [KT959235](#), [KJ361946](#), [KJ361959](#), [JX512198](#), [KU550090](#), [KR534213](#), [HQ380895](#), and [NC_006273](#)) using MAFFT (default parameters) (60). The sequence alignment was trimmed using trimAl (default parameters) (61). Statistical selection of best-fit models was determined using jModelTest 2 (g 4 -i -f -BIC -a) (62). Maximum likelihood (ML) phylogenetic trees were constructed using PhyML (K80 +I +G, ti/tv 3.8561) (63). Support for ML trees was assessed by 1,000 nonparametric bootstraps.

Data availability. Sequence data have been deposited in the NCBI under the BioProject number [PRJNA635738](#). The code script to reproduce the analysis and plots is available in the supplemental material (CD4_Data_Script).

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

TEXT S1, DOCX file, 0.01 MB.

FIG S1, PDF file, 0.7 MB.

FIG S2, PDF file, 0.3 MB.

FIG S3, PDF file, 1 MB.

FIG S4, PDF file, 0.2 MB.

FIG S5, PDF file, 0.6 MB.

ACKNOWLEDGMENTS

We thank the study participants and their families.

This research was funded by NIH grant R01 HD092311 awarded to D.A.L. and E.S.L.

We declare no competing interests.

Conceptualization: D.A.L., E.S.L.; data curation: R.M., J.B.R.; formal analysis: R.M., J.B.R.; funding acquisition: D.A.L., E.S.L.; investigation: R.M., J.B.R., L.I.W., L.A.H., E.R.B., S.G., E.O., D.W., G.J.-S., J.S., D.A.L., E.S.L.; project administration: L.I.W.; resources: R.N., D.M.-N., E.M.-O., D.W., G.J.-S.; supervision: E.S.L.; validation: L.I.W.; visualization: R.M.,

J.B.R.; writing-original draft: R.M., J.B.R., E.S.L.; writing-review and editing: R.M., J.B.R., R.N., D.M.-N., E.M.-O., E.R.B., S.G., E.O., D.W., G.J.-S., J.S., D.A.L., E.S.L. All authors reviewed and approved the final manuscript.

REFERENCES

- Ho NT, Li F, Lee-Sarwar KA, Tun HM, Brown BP, Pannaraj PS, Bender JM, Azad MB, Thompson AL, Weiss ST, Azcarate-Peril MA, Litonjua AA, Kozyrskij AL, Jaspan HB, Aldrovandi GM, Kuhn L. 2018. Meta-analysis of effects of exclusive breastfeeding on infant gut microbiota across populations. *Nat Commun* 9:4169. <https://doi.org/10.1038/s41467-018-06473-x>.
- Mueller NT, Bakacs E, Combellick J, Grigoryan Z, Dominguez-Bello MG. 2015. The infant microbiome development: mom matters. *Trends Mol Med* 21:109–117. <https://doi.org/10.1016/j.molmed.2014.12.002>.
- Le Huerou-Luron I, Blat S, Boudry G. 2010. Breast- v. formula-feeding: impacts on the digestive tract and immediate and long-term health effects. *Nutr Res Rev* 23:23–36. <https://doi.org/10.1017/S0954422410000065>.
- Henrick BM, Yao XD, Nasser L, Roozrogousheh A, Rosenthal KL. 2017. Breastfeeding behaviors and the innate immune system of human milk: working together to protect infants against inflammation, HIV-1, and other infections. *Front Immunol* 8:1631. <https://doi.org/10.3389/fimmu.2017.01631>.
- Hanson LA, Korotkova M, Telemo E. 2003. Breast-feeding, infant formulas, and the immune system. *Ann Allergy Asthma Immunol* 90:59–63. [https://doi.org/10.1016/s1081-1206\(10\)61662-6](https://doi.org/10.1016/s1081-1206(10)61662-6).
- Brandtzaeg P. 2003. Mucosal immunity: integration between mother and the breast-fed infant. *Vaccine* 21:3382–3388. [https://doi.org/10.1016/s0264-410x\(03\)00338-4](https://doi.org/10.1016/s0264-410x(03)00338-4).
- Jost T, Lacroix C, Braegger CP, Rochat F, Chassard C. 2014. Vertical mother-neonate transfer of maternal gut bacteria via breastfeeding. *Environ Microbiol* 16:2891–2904. <https://doi.org/10.1111/1462-2920.12238>.
- Martin V, Maldonado-Barragan A, Moles L, Rodriguez-Banos M, Campo RD, Fernandez L, Rodriguez JM, Jimenez E. 2012. Sharing of bacterial strains between breast milk and infant feces. *J Hum Lact* 28:36–44. <https://doi.org/10.1177/0890334411424729>.
- Pannaraj PS, Li F, Cerini C, Bender JM, Yang S, Rollie A, Adisetiyo H, Zabih S, Lincez PJ, Bittinger K, Bailey A, Bushman FD, Sleasman JW, Aldrovandi GM. 2017. Association between breast milk bacterial communities and establishment and development of the infant gut microbiome. *JAMA Pediatr* 171:647–654. <https://doi.org/10.1001/jamapediatrics.2017.0378>.
- Lim ES, Zhou Y, Zhao G, Bauer IK, Droit L, Ndao IM, Warner BB, Tarr PI, Wang D, Holtz LR. 2015. Early life dynamics of the human gut virome and bacterial microbiome in infants. *Nat Med* 21:1228–1234. <https://doi.org/10.1038/nm.3950>.
- Lackey KA, Williams JE, Meehan CL, Zachek JA, Benda ED, Price WJ, Foster JA, Sellen DW, Kamau-Mbuthia EV, Kamundia EW, Mbugua S, Moore SE, Prentice AK, Dg Kvist LJ, Otoo GE, Garcia-Carral C, Jimenez E, Ruiz L, Rodriguez JM, Pareja RG, Bode L, McGuire MA, McGuire MK. 2019. What's normal? Microbiomes in human milk and infant feces are related to each other but vary geographically: the INSPIRE study. *Front Nutr* 6:45. <https://doi.org/10.3389/fnut.2019.00045>.
- Hunt KM, Foster JA, Forney LJ, Schutte UM, Beck DL, Abdo Z, Fox LK, Williams JE, McGuire MK, McGuire MA. 2011. Characterization of the diversity and temporal stability of bacterial communities in human milk. *PLoS One* 6:e21313. <https://doi.org/10.1371/journal.pone.0021313>.
- Browne PD, Aparicio M, Alba C, Hechler C, Beijers R, Rodriguez JM, Fernandez L, de Weerth C. 2019. Human milk microbiome and maternal postnatal psychosocial distress. *Front Microbiol* 10:2333. <https://doi.org/10.3389/fmicb.2019.02333>.
- Bender JM, Li F, Martelly S, Byrt E, Rouzier V, Leo M, Tobin N, Pannaraj PS, Adisetiyo H, Rollie A, Santiskulvong C, Wang S, Autran C, Bode L, Fitzgerald D, Kuhn L, Aldrovandi GM. 2016. Maternal HIV infection influences the microbiome of HIV-uninfected infants. *Sci Transl Med* 8:349ra100. <https://doi.org/10.1126/scitranslmed.aaf5103>.
- Pannaraj PS, Ly M, Cerini C, Saavedra M, Aldrovandi GM, Saboory AA, Johnson KM, Pride DT. 2018. Shared and distinct features of human milk and infant stool viromes. *Front Microbiol* 9:1162. <https://doi.org/10.3389/fmicb.2018.01162>.
- Slyker JA, Richardson B, Chung MH, Atkinson C, Asbjornsdottir KH, Lehman DA, Boeckh M, Emery V, Kiarie J, John-Stewart G. 2017. Maternal highly active antiretroviral therapy reduces vertical cytomegalovirus transmission but does not reduce breast milk cytomegalovirus levels. *AIDS Res Hum Retroviruses* 33:332–338. <https://doi.org/10.1089/AID.2016.0121>.
- Richardson BA, John-Stewart G, Atkinson C, Nduati R, Asbjornsdottir K, Boeckh M, Overbaugh J, Emery V, Slyker JA. 2016. Vertical cytomegalovirus transmission from HIV-infected women randomized to formula-feed or breastfeed their infants. *J Infect Dis* 213:992–998. <https://doi.org/10.1093/infdis/jiv515>.
- Durand AJ, Goga AE, Waitt C, Gessain A, Taylor GP, Rollins N, Abrams EJ, Lyall EH, de Perre PV. 2019. Transmission of CMV, HTLV-1, and HIV through breastmilk. *Lancet Child Adolesc Health* 3:264–273. [https://doi.org/10.1016/S2352-4642\(19\)30024-0](https://doi.org/10.1016/S2352-4642(19)30024-0).
- Duranti S, Lugli GA, Mancabelli L, Armanini F, Turroni F, James K, Ferretti P, Gorfer V, Ferrario C, Milani C, Mangifesta M, Anzalone R, Zolfo M, Viappiani A, Pasolli E, Bariletti I, Canto R, Clementi R, Cologna M, Crifo T, Cusumano G, Fedi S, Gottardi S, Innamorati C, Mase C, Postai D, Savoia D, Soffiati M, Tateo S, Pedrotti A, Segata N, van Sinderen D, Ventura M. 2017. Maternal inheritance of bifidobacterial communities and bifidophages in infants through vertical transmission. *Microbiome* 5:66. <https://doi.org/10.1186/s40168-017-0282-6>.
- Schwartz A, Jacobi M, Frick JS, Richter M, Rusch K, Kohler H. 2010. Microbiota in pediatric inflammatory bowel disease. *J Pediatr* 157:240–244e1. <https://doi.org/10.1016/j.jpeds.2010.02.046>.
- Gevers D, Kugathasan S, Denson LA, Vazquez-Baeza Y, Van Treuren W, Ren B, Schwager E, Knights D, Song SJ, Yassour M, Morgan XC, Kostic AD, Luo C, Gonzalez A, McDonald D, Haberman Y, Walters T, Baker S, Rosh J, Stephens M, Heyman M, Markowitz J, Baldassano R, Griffiths A, Sylvester F, Mack D, Kim S, Crandall W, Hyams J, Huttenhower C, Knight R, Xavier RJ. 2014. The treatment-naïve microbiome in new-onset Crohn's disease. *Cell Host Microbe* 15:382–392. <https://doi.org/10.1016/j.chom.2014.02.005>.
- Kostic AD, Gevers D, Siljander H, Vatanen T, Hyotylainen T, Hamalainen AM, Peet A, Tillmann V, Poho P, Mattila I, Lahdesmaki H, Franzosa EA, Vaarala O, de Goffau M, Harmsen H, Ilonen J, Virtanen SM, Clish CB, Oresic M, Huttenhower C, Knip M, Group DS, Xavier RJ, DIABIMMUNE Study Group. 2015. The dynamics of the human infant gut microbiome in development and in progression toward type 1 diabetes. *Cell Host Microbe* 17:260–273. <https://doi.org/10.1016/j.chom.2015.01.001>.
- White RA, Bjornholt JV, Baird DD, Midtvedt T, Harris JR, Pagano M, Hide W, Rudi K, Moen B, Iszatt N, Peddada SD, Eggesbo M. 2013. Novel developmental analyses identify longitudinal patterns of early gut microbiota that affect infant growth. *PLoS Comput Biol* 9:e1003042. <https://doi.org/10.1371/journal.pcbi.1003042>.
- Rautava S. 2016. Early microbial contact, the breast milk microbiome and child health. *J Dev Orig Health Dis* 7:5–14. <https://doi.org/10.1017/S2040174415001233>.
- Monaco CL, Gootenberg DB, Zhao G, Handley SA, Ghebremichael MS, Lim ES, Lankowski A, Baldrige MT, Wilen CB, Flagg M, Norman JM, Keller BC, Luevano JM, Wang D, Boum Y, Martin JN, Hunt PW, Bangsberg DR, Siedner MJ, Kwon DS, Virgin HW. 2016. Altered virome and bacterial microbiome in human immunodeficiency virus-associated acquired immunodeficiency syndrome. *Cell Host Microbe* 19:311–322. <https://doi.org/10.1016/j.chom.2016.02.011>.
- World Health Organization. 2020. Infant feeding for the prevention of mother-to-child transmission of HIV. https://www.who.int/elena/titles/hiv_infant_feeding/en/.
- Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SS, McCulle SL, Karlebach S, Gorle R, Russell J, Tacket CO, Brotman RM, Davis CC, Ault K, Peralta L, Forney LJ. 2011. Vaginal microbiome of reproductive-age women. *Proc Natl Acad Sci U S A* 108:4680–4687. <https://doi.org/10.1073/pnas.1002611107>.
- Gottschick C, Deng ZL, Vital M, Masur C, Abels C, Pieper DH, Wagner-Dobler I. 2017. The urinary microbiota of men and women and its changes in women during bacterial vaginosis and antibiotic treatment. *Microbiome* 5:99. <https://doi.org/10.1186/s40168-017-0305-3>.
- Pearce MM, Zilliox MJ, Rosenfeld AB, Thomas-White KJ, Richter HE, Nager CW, Visco AG, Nygaard IE, Barber MD, Schaffer J, Moalli P, Sung VW, Smith AL, Rogers R, Nolen TL, Wallace D, Meikle SF, Gai X, Wolfe AJ, Brubaker L, Pelvic Floor Disorders Network. 2015. The female urinary microbiome in

- urgency urinary incontinence. *Am J Obstet Gynecol* 213:347.e1–347.e11. <https://doi.org/10.1016/j.ajog.2015.07.009>.
30. Kordy K, Gaufrin T, Mwangi M, Li F, Cerini C, Lee DJ, Adisetiyo H, Woodward C, Pannaraj PS, Tobin NH, Aldrovandi GM. 2020. Contributions to human breast milk microbiome and enteromammary transfer of *Bifidobacterium breve*. *PLoS One* 15:e0219633. <https://doi.org/10.1371/journal.pone.0219633>.
 31. Becquart P, Hocini H, Garin B, Sepou A, Kazatchkine MD, Belec L. 1999. Compartmentalization of the IgG immune response to HIV-1 in breast milk. *AIDS* 13:1323–1331. <https://doi.org/10.1097/00002030-199907300-00008>.
 32. Lozupone CA, Li M, Campbell TB, Flores SC, Linderman D, Gebert MJ, Knight R, Fontenot AP, Palmer BE. 2013. Alterations in the gut microbiota associated with HIV-1 infection. *Cell Host Microbe* 14:329–339. <https://doi.org/10.1016/j.chom.2013.08.006>.
 33. Cacho NT, Lawrence RM. 2017. Innate immunity and breast milk. *Front Immunol* 8:584. <https://doi.org/10.3389/fimmu.2017.00584>.
 34. Slyker J, Farquhar C, Atkinson C, Asbjornsdottir K, Roxby A, Drake A, Kiarie J, Wald A, Boeckh M, Richardson B, Odem-Davis K, John-Stewart G, Emery V. 2014. Compartmentalized cytomegalovirus replication and transmission in the setting of maternal HIV-1 infection. *Clin Infect Dis* 58:564–572. <https://doi.org/10.1093/cid/cit727>.
 35. Viljoen J, Tuailon E, Nagot N, Danaviah S, Peries M, Padayachee P, Foulongne V, Bland R, Rollins N, Newell ML, van de Perre P. 2015. Cytomegalovirus, and possibly Epstein-Barr virus, shedding in breast milk is associated with HIV-1 transmission by breastfeeding. *AIDS* 29:145–153. <https://doi.org/10.1097/QAD.0000000000000527>.
 36. Santos RC, Hirao LA, Weber MG, Mendez-Lagares G, Chang WLW, Jiang G, Deere JD, Sparger EE, Roberts J, Barry PA, Hartigan-O'Connor DJ, Dandekar S. 2018. Subclinical cytomegalovirus infection is associated with altered host immunity, gut microbiota, and vaccine responses. *J Virol* 92:e00167-18. <https://doi.org/10.1128/JVI.00167-18>.
 37. Young JC, Chehoud C, Bittinger K, Bailey A, Diamond JM, Cantu E, Haas AR, Abbas A, Frye L, Christie JD, Bushman FD, Collman RG. 2015. Viral metagenomics reveal blooms of anelloviruses in the respiratory tract of lung transplant recipients. *Am J Transplant* 15:200–209. <https://doi.org/10.1111/ajt.13031>.
 38. Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv:1303.3997*. <https://arxiv.org/abs/1303.3997>.
 39. Bosire R, Farquhar C, Nduati R, Broliden K, Luchters S, Van de Perre P, De Vincenzi I, Merkel M, Wachuka V, Mbori-Ngacha D, John-Stewart G, Lohman-Payne B, Reilly M. 2018. Higher transplacental pathogen-specific antibody transfer among pregnant women randomized to triple antiretroviral treatment versus short course zidovudine. *Pediatr Infect Dis J* 37:246–252. <https://doi.org/10.1097/INF.0000000000001749>.
 40. Nahui Palomino RA, Vanpouille C, Laghi L, Parolin C, Melikov K, Backlund P, Vitali B, Margolis L. 2019. Extracellular vesicles from symbiotic vaginal lactobacilli inhibit HIV-1 infection of human tissues. *Nat Commun* 10:5656. <https://doi.org/10.1038/s41467-019-13468-9>.
 41. Abdool Karim SS, Baxter C, Passmore JS, McKinnon LR, Williams BL. 2019. The genital tract and rectal microbiomes: their role in HIV susceptibility and prevention in women. *J Int AIDS Soc* 22:e25300. <https://doi.org/10.1002/jia2.25300>.
 42. John-Stewart GC, Mbori-Ngacha D, Payne BL, Farquhar C, Richardson BA, Emery S, Otieno P, Obimbo E, Dong T, Slyker J, Nduati R, Overbaugh J, Rowland-Jones S. 2009. HIV-1-specific cytotoxic T lymphocytes and breast milk HIV-1 transmission. *J Infect Dis* 199:889–898. <https://doi.org/10.1086/597120>.
 43. Ministry of Health, National AIDS & STI Control Program. 2018. Guidelines on use antiretroviral drugs for treating and preventing HIV infection in Kenya, 2018 ed. NASCOP, Nairobi, Kenya.
 44. World Health Organization. 2015. Guideline on when to start antiretroviral therapy and on pre-exposure prophylaxis for HIV. WHO, Geneva, Switzerland.
 45. World Health Organization. 2010. Antiretroviral drugs for treating pregnant women for their own health and to prevent HIV infection in their infants. In Antiretroviral drugs for treating pregnant women and preventing HIV infection in infants: recommendations for a public health approach. WHO, Geneva, Switzerland.
 46. Maqsood R, Rodgers R, Rodriguez C, Handley SA, Ndao IM, Tarr PI, Warner BB, Lim ES, Holtz LR. 2019. Discordant transmission of bacteria and viruses from mothers to babies at birth. *Microbiome* 7:156. <https://doi.org/10.1186/s40168-019-0766-7>.
 47. Toigo G, Situlin R, Tamaro G, Del Bianco A, Giuliani V, Dardi F, Vianello S, Toffoletto P, Faccini L, Guarnieri G. 1989. Effect of intravenous supplementation of a new essential amino acid formulation in hemodialysis patients. *Kidney Int Suppl* 27:S278–S281.
 48. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ, Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope EK, Da Silva R, Diener C, Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M, Fouquier J, Gauglitz JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley GA, Janssen S, Jarmusch AK, Jiang L, Kaehler BD, Kang KB, Keefe CR, Keim P, Kelley ST, Knights D, et al. 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* 37:852–857. <https://doi.org/10.1038/s41587-019-0209-9>.
 49. Davis NM, Proctor DM, Holmes SP, Relman DA, Callahan BJ. 2018. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* 6:226. <https://doi.org/10.1186/s40168-018-0605-2>.
 50. Wickham H. 2016. ggplot2: elegant graphics for data analysis. Springer-Verlag, New York.
 51. Bushnell B. *BBMap*. sourceforge.net/projects/bbmap/.
 52. Huson DH, Beier S, Flade I, Gorska A, El-Hadidi M, Mitra S, Ruscheweyh HJ, Tappu R. 2016. MEGAN community edition: interactive exploration and analysis of large-scale microbiome sequencing data. *PLoS Comput Biol* 12:e1004957. <https://doi.org/10.1371/journal.pcbi.1004957>.
 53. Simpson L, Solymos P, Henry H, Stevens MHH, Szoecs E, Wagner H. 2019. vegan: community ecology package. <https://CRAN.R-project.org/package=vegan>.
 54. Peng Y, Leung HC, Yiu SM, Chin FY. 2012. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* 28:1420–1428. <https://doi.org/10.1093/bioinformatics/bts174>.
 55. Sommer DD, Delcher AL, Salzberg SL, Pop M. 2007. Minimus: a fast, lightweight genome assembler. *BMC Bioinformatics* 8:64. <https://doi.org/10.1186/1471-2105-8-64>.
 56. Roux S, Enault F, Hurwitz BL, Sullivan MB. 2015. VirSorter: mining viral signal from microbial genomic data. *PeerJ* 3:e985. <https://doi.org/10.7717/peerj.985>.
 57. Boeckh M, Huang M, Ferrenberg J, Stevens-Ayers T, Stensland L, Nichols WG, Corey L. 2004. Optimization of quantitative detection of cytomegalovirus DNA in plasma by real-time PCR. *J Clin Microbiol* 42:1142–1148. <https://doi.org/10.1128/JCM.42.3.1142-1148.2004>.
 58. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
 59. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28:1647–1649. <https://doi.org/10.1093/bioinformatics/bts199>.
 60. Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* 30:3059–3066. <https://doi.org/10.1093/nar/gkf436>.
 61. Capella-Gutiérrez S, Silla-Martinez JM, Gabaldon T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25:1972–1973. <https://doi.org/10.1093/bioinformatics/btp348>.
 62. Darriba D, Taboada GL, Doallo R, Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nat Methods* 9:772. <https://doi.org/10.1038/nmeth.2109>.
 63. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 59:307–321. <https://doi.org/10.1093/sysbio/syq010>.