Supplementary Figure 1
Supplementary Figure 2

TOTAL SCORE
(min 3/8 needed)

0/8

HOMOLOGY
0/2

- NT based:
  - Blastn
  - Metaphinder2 ANI > 10% (length > 2500)
  0/1

- NR based:
  - DIAMOND
  - CAT
  0/1

GENOME STRUCTURE
0/2

- DeepVirFinder > 0.95
  0/1

FUNCTIONAL
0/2

- EggNOG/Pfam:
  - viral specific genes
  0/1

- pVOG/gene ratio > 0.6
  0/1

VIRSORTER
0/2

- categories 1,2
  2

- categories 3,4,5,6
  1
Supplementary Figure 3

(a) Distribution of contig lengths.
(b) Correlation between contig length and completeness of phage contigs.
(c) Distribution of number of phage contigs.
(d) Scatter plot showing the relationship between the completeness of phage contigs and the log10-transformed reads count, with linear regression lines indicating a significant correlation ($R = 0.85, p < 2.2e^{-16}$) for (b) and ($R = 0.77, p < 2.2e^{-16}$) for (d).
Supplementary Figure 4

a

b

C

| Relative abundance | Days after birth |
|--------------------|-----------------|
| 0.00               | 0               |
| 0.02               | 100             |
| 0.04               | 100             |
| 0.06               | 100             |

Supplementary Figure 4 a, b, and c show various diagrams and data representations.
Supplementary Figure 5
Supplementary Figure 6

[Diagram showing the distribution of various viral families over days after birth for different subjects (S003 to S011). Each subject has a separate graphical representation with viral families shaded in different colors: Small Circular ssDNA (light green), Disease-associated Mamal-Inflicting Viruses (light blue), Plant-Infecting Viruses (dark blue), and Fungus-Infecting Viruses (brown). The x-axis represents days after birth ranging from 0 to 400.]
Supplementary Figure 7
Supplementary Figure 8
Supplementary Figure 9
Supplementary Figure 10
Supplementary Figure 11

Pearson’s r = 0.55

Pearson’s r = 0.78
Supplementary Figure 12

Mann-Whitney U test:

p-value < 2.2 e-16
R2 = 0.84
Supplementary Figure 14
SUPPLEMENTARY FIGURES

Supplementary Figure 1: Overview of the collected and selected samples per infant.

Grey dots (●): All samples collected by the parents of the enrolled infants
Blue diamonds (●): Samples selected for the study of the longitudinal dynamics at predefined timepoints with no clinical signs (n = 143)
Green diamonds (●): Additional ad hoc selected samples at specific external events (n = 161)
Black filled triangles (▲): Three vaccination events in every infant
Black open triangles (▼): Day care entrance
Red triangles (▼): External events around which extra samples were selected.

Abbreviations: fever (F), diarrhoea (D), vomit (V), antibiotics (AB), probiotics (PB)

Supplementary Figure 2: Phage identification scoring scheme.

To identify bacteriophages, a new scoring scheme was developed. At four different levels, contigs were scored. First, homology-based classification was performed on nucleotide level (using BLASTn hit with e-value < 1e-10 and MetaPhinder2ref ANI > 10% only if length >2500 nt) and at protein level (DIAMOND56 with ‘sensitive’ option and CAT57). Secondly, a score on two was given based on the genome structure of the contig defined at kmer-level (DeepVirFinder58 score > 0.95 and p < 0.01) and a gene-to-length ratio of above 1.6 per kb. Third, at functional level, the presence of virus-specific hallmark genes (Supplementary Table 1) was scored, as well as a pVOG/gene ratio above 0.6. Finally, all contigs were scored based on their VirSorter59 category. In total, a minimum score of three of eight was needed to identify a contig as bacteriophage unless the VirSorter Category was one or two.
Supplementary Figure 3: CheckV results.

(a) Histogram showing the distribution of the lengths of the identified phage contigs. (b) Scatterplot showing that the lengths of the phage contigs are significantly correlated to the completeness of these contigs measured by CheckV (Pearson correlation coefficient). (c) Histogram showing the distribution of the read counts (log10 transformed) of the phage contigs identified. (d) Scatterplot showing that the read counts (log10 transformed) of the phage contigs are significantly correlated to the completeness of these contigs measured by CheckV (Pearson correlation coefficient).

Supplementary Figure 4: Overview of the two bacteriophage contigs shared by all eight BaBel infants.

(a-b) Visualisation of the functions of annotated genes (performed by Cenote-Taker2). (c) Relative abundance of the only two phage contigs that are shared by all the eight BaBel infants (based on rarefied reads). Horizontal lines indicate the diet: breastmilk (dark blue), formula milk (light blue), vegetables (dark green) and fruit (light green).

Supplementary Figure 5: Alpha diversity of the BIG phages over the time.

Observed Richness (top) and Shannon diversity (bottom) are shown for the BIG phages over time, coloured per infant and shown for all the samples (n = 304, Loess smoothing with span equals 0.25).

Supplementary Figure 6: Overview of the detection of eukaryotic viruses.

Dense colouring of the dots indicates the sample is positive for the eukaryotic viral family shown on the y-axis. Vertical black lines indicate vaccination time points and vertical red
lines indicate entry into day care. Horizontal lines indicate the diet: breastmilk (dark blue), formula milk (light blue), vegetables (dark green) and fruit (light green).

Supplementary Figure 7: Rarefied number of reads per genus of the different disease-associated mammal infecting viral (DaMiV) genera detected in the BaBel infants over their first year of life (>100 reads shown).

Vertical black lines indicate vaccination time points and vertical red lines indicate entrance of day care. Horizontal lines indicate the diet: breastmilk (dark blue), formula milk (light blue), vegetables (dark green) and fruit (light green).

Supplementary Figure 8: Accumulation of infections by disease-associated mammal infecting viruses (DaMiVs).

Same as Figure 3b, but here the individual profiles per infant are shown. Vertical black lines indicate vaccination time points and vertical red lines indicate entrance of day care. Horizontal lines indicate the diet: breastmilk (dark blue), formula milk (light blue), vegetables (dark green) and fruit (light green).

Supplementary Figure 9: Richness of the Anelloviridae contigs, individual profiles.

Same as Figure 4c, but here the individual profiles per infant are shown. Vertical black lines indicate vaccination time points and vertical red lines indicate entrance of day care. Horizontal lines indicate the diet: breastmilk (dark blue), formula milk (light blue), vegetables (dark green) and fruit (light green).

Supplementary Figure 10: Accumulation of unique Anelloviridae contigs.
Same as Figure 4c, but here the individual profiles per infant are shown. Vertical black lines indicate vaccination time points and vertical red lines indicate entrance of day care. Horizontal lines indicate the diet: breastmilk (dark blue), formula milk (light blue), vegetables (dark green) and fruit (light green).

**Supplementary Figure 11: Abundances of the two detected CrAssphages and their predicted hosts.**

The cooccurrence profiles of the two predicted CrAss-viruses (bottom) and the predicted hosts (top) are shown over time.

**Supplementary Figure 12: Plant 18S reads and the link with diet.**

(top) Number of plant 18S reads (phylum Phragmoplastophyta) detected in the BaBel infants over time. The green vertical line indicates the time of weaning. (bottom) Plant 18S reads are very strongly correlated with weaning (Mann-Whitney, p-value < 2.2e-16, R² = 0.84).

**Supplementary Figure 13: Relative abundances of fungal ASVs found in the eight BaBel infants**

Relative abundances of the fungal ASVs, calculated by dividing the number of reads corresponding to the ASV by the total number of sequenced reads of that sample, are shown. Blue bars indicate antibiotics treatment, red bars correspond to treatment with *S. Boulardii*, green bars indicate fungal infections (S004: thrush and S007: fungal infection labia) and the yellow bars to timepoints with miconazole treatment. Only relative abundances above 0.01 are shown.

**Supplementary Figure 14: Number of Cryptosporidium reads found in infant S009.**
The number of *Cryptosporidium* reads found in every sample of S009 is shown over time. The red bar indicates the start of symptoms (diarrhoea, vomiting and fever). The infant remained sick for three days, however *Cryptosporidium* reads were detected up to 23 days after the first symptoms.
Supplementary Table 1

Description of the infants enrolled in the BaBel infants (n = 8) and description of the analysed samples (n = 304). Abbreviations: F(emale)/M(ale), Secretor(se+)/Non-Secretor(se−), Lewis:negative(Le−)/Lewis B(LeB)/LewisA. Sample indentifier, Infant identifier, age of sample collection (days after birth), included as a predefined timepoint when the infant was healthy (LDA, Yes/No), Consistency (1=Aqueous, 2=Soft, 3=Solid), Food Category (Only Breast milk, No solid food, Solid food), drug category (according the ATC classification system), disease category (according the ICD-10 classification), Holiday (Yes/No), Family stay (Yes/No), Attending DayCare (Yes/No)

Supplementary Table 2

Characteristics of the BaBel Infant Gut (BIG) phages found to be present in 50% or more of the BaBel infants

Supplementary Table 3

Pairwise comparisons using Fisher’s exact test on the lifestyle per two-months age bin

Supplementary Table 4

The first DaMiVs detected in the BaBel infants.

Supplementary Table 5

Overview of the infections by disease-associated mammal-infection viral genera (DaMiVs) detected in the BaBel infants and their association with enteric signs (associated=yes, if the signs are detected within seven days after the detection of the infection)

Supplementary Table 6

Paired Wilcoxon tests to compare number of infections and the infection rate before and after specific events (such as changes in diet (stop of breastfeeding, start with formula feeding or start with solid foods) or starting in day care).

Supplementary Table 7

Paired Wilcoxon tests to compare number of accumulated unique Anellovirus contigs and the accumulated unique Anellovirus contigs rate before and after specific events (such as changes in diet (stop of breastfeeding, start with formula feeding or start with solid foods) or starting in day care).

Supplementary Table 8

Overall abundances of bacterial families in the 16S rRNA library and how many times a BIG phage is predicted to have this family as a host

Supplementary Table 9

Transkingdom interactions found between plant viruses and 18S rRNA ASVs

Supplementary Table 10

Fungal ASVs and reads distribution

Supplementary Table 11

Viral Hallmark gene keywords used for identification of bacteriophages
Supplementary Table 12

Lysogeny specific genes identified by cenote-taker2