Supplementary methods

Human beta-2 Microglobulin (B2M) gene expression

Successful RNA extraction was confirmed by the detection of beta-2-microglobulin (B2M) mRNA using RT-qPCR. Here, 5 µL RNA eluate were used as input in the Invitrogen Superscript III one step RT-PCR system with Platinum Taq Polymerase (Thermo Fisher Scientific) including 0.4 µM of each primer (B2M forward primer: 5'-GAGTATGCCTGCCGTGTG-3', B2M reverse primer: 5'-AATCCAAATGCGGCATCT-3', 0,28 µM probe (B2M Probe FAM-CCTCCATGATGCTGCTTACATGTCTC-TAMRA; PMID: 17026756), 1 µg of nonacetylated bovine serum albumin (Roche), 1 µL enzyme mix, and additional 0.8 mM MgSO4 (Thermo Fisher Scientific). Thermal cycling was performed at 50 °C for 15 minutes, followed by 95 °C for 3 minutes and then 45 cycles of 95 °C for 15 seconds, 52 °C for 20 seconds.

High throughput-sequencing

Up to 47.2ng RNA were used for library preparation using the KAPA RNA HyperPrep Kit (Roche, Mannheim, Germany) according to the manufacturer's protocol. Briefly, for native sequencing, RNA was fragmented for 5 min at 85°C and amplified for 10-12 cycles. Resulting libraries were equimolar pooled and paired-end sequenced on an Illumina HiSeq4000 (2x75 base pairs (bp)). For rRNA (and globin) depletion, the QIAseq FastSelect -rRNA and Globin HMR Kit (Qiagen) was applied following the recommended temperature cycling protocol after fragmentation for 6 min at 85°C. Library amplification was performed for 14 cycles. Resulting libraries were equimolar pooled and paired-end sequenced on an Illumina NextSeq (2x75bp).

HTS data analyses

In-house computational pipeline

For virus detection, we applied a computational pipeline that consists of a collection of shell scripts that invoke third-party and custom in-house programs. Overall execution is coordinated by a Slurm Pipeline Python package (T. C. Jones, Slurm Pipeline, available at https://github.com/acorg/slurm-pipeline). The subsequent steps are as follows:

The adapter sequences are trimmed from sequencing reads using AdapterRemoval (version 2.3.2) [9]. The reads are combined using flash (version 1.2.11) [7] and mapped against a series of databases using bwa (version 0.7.17-r1188) [6] with all matching reads removed from further pipeline processing using samtools (version 1.12) [4].

The databases are:

i) GRCh38.p7 Homo sapiens (March 2016) genome
ii) GRCh38 (August 2019) homo sapiens RNA
iii) Homo sapiens mitochondrial DNA (July 2020)
iv) Ribosomal RNA (July 2020)
v) Long non-coding RNA (July 2020)

Duplicate sequences are removed, using the filter-fasta.py command from the dark-matter Python tools (version 4.0.58) (T. C. Jones, B. Mühlemann, T. Veith, S. Mathias, U. Gieraths, Dark Matter, available at https://github.com/acorg/dark-matter). Reads are matched at the amino acid level against protein databases using DIAMOND (version 2.1.7) [3].

The databases are:

i) National Center for Biotechnology Information (NCBI) viral reference sequences (downloaded June 2023) [2].
ii) Protein sequences from a collection of insect virus genomes [5].

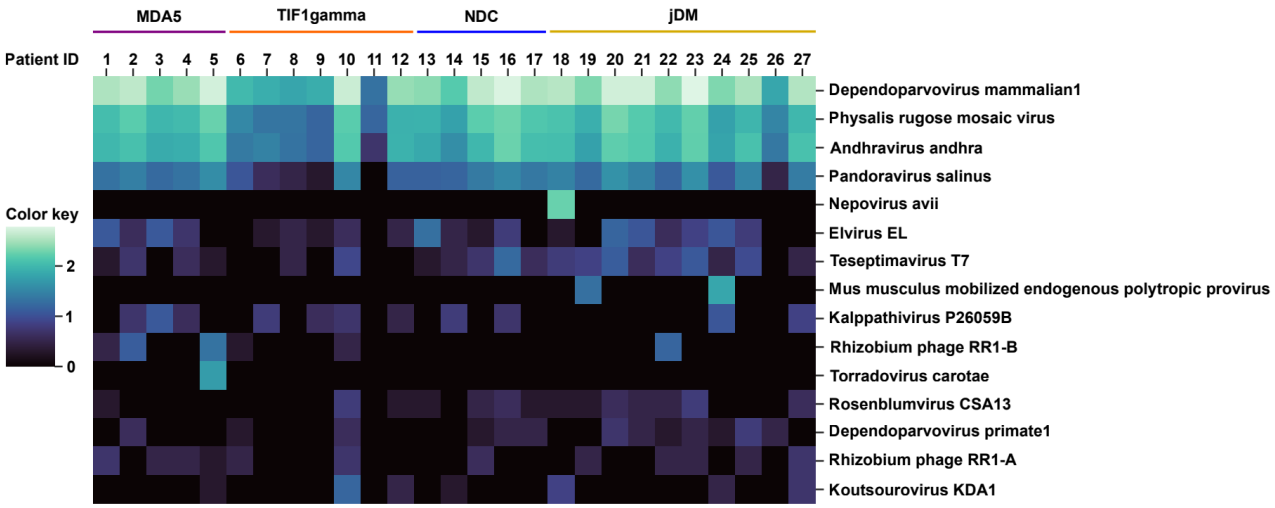
The protein databases are built using an in-house process using code from the dark-matter package (T. C. Jones, B. Mühlemann, T. Veith, S. Mathias, U. Gieraths, Dark Matter, available at https://github.com/acorg/dark-matter). Protein match results are combined into virus level information, and this is converted into an HTML summary using code from dark-matter (T. C. Jones, B. Mühlemann, T. Veith, S. Mathias, U. Gieraths, Dark Matter, available at https://github.com/acorg/dark-matter). The summary incorporates taxonomy information from an sqlite3 database built using in-house Python (T. C. Jones, Taxonomy database, available at https://github.com/acorg/ncbi-taxonomy-database) based on the NCBI Taxonomy database (National Center for Biotechnology Information, NCBI Taxonomy Database, available at https://www.ncbi.nlm.nih.gov/taxonomy). Pipeline results were checked independently by two trained virologists. Potential viral hits were manually reviewed using Geneious Prime v2022.0.1 (Biomatters, Auckland, New Zealand, https://www.geneious.com).

Taxonomic analyses of sequencing reads using Kraken2

The metagenomic analysis was done at the nucleotide level using Kraken2 (version 2.1.3) [10]. Here, all sequencing reads were aligned against the 'standard' (revision data 10/9/2023) database provided by the authors of Kraken2 at https://benlangmead.github.io/aws-indexes/k2. This database includes, in particular, human DNA and viral sequences. The analysis was done in paired read mode and using a confidence of 0.05.

The results were visualized using Krona [8] and Pavian [1] and again checked independently by two trained virologists.

Supplementary figure



Supplementary figure 1: Heatmap (Log10) of Kraken2/Pavian hits for the 15 viral taxa with most assigned reads after rRNA depletion by Kraken2/Pavian. Samples are highlighted according to their respective cohort (MDA5, TIF-1gamma (TIF1γ); non-diseased controls (NDC), and jDM). Virus names are given on the right.

Supplementary references

1. Breitwieser FP, Salzberg SL (2020) Pavian: interactive analysis of metagenomics data for microbiome studies and pathogen identification. Bioinformatics 36: 1303-1304. DOI: https://doi.org/10.1093/bioinformatics/btz715

2. Brister JR, Ako-Adjei D, Bao Y, Blinkova O (2015) NCBI viral genomes resource. Nucleic Acids Res 43: D571-577. DOI: https://doi.org/10.1093/nar/gku1207

3. Buchfink B, Reuter K, Drost HG (2021) Sensitive protein alignments at tree-of-life scale using DIAMOND. Nat Methods 18: 366-368. DOI: https://doi.org/10.1038/s41592-021-01101-x

4. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, Li H (2021) Twelve years of SAMtools and BCFtools. Gigascience 10:. DOI: https://doi.org/10.1093/gigascience/giab008

5. Kafer S, Paraskevopoulou S, Zirkel F, Wieseke N, Donath A, Petersen M, Jones TC, Liu S, Zhou X, Middendorf Met al (2019) Re-assessing the diversity of negative strand RNA viruses in insects. PLoS Pathog 15: e1008224. DOI: https://doi.org/10.1371/journal.ppat.1008224

6. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25: 1754-1760. DOI: https://doi.org/10.1093/bioinformatics/btp324

7. Magoc T, Salzberg SL (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics 27: 2957-2963. DOI: https://doi.org/10.1093/bioinformatics/btr507

8. Ondov BD, Bergman NH, Phillippy AM (2011) Interactive metagenomic visualization in a Web browser. BMC Bioinformatics 12: 385. DOI: https://doi.org/10.1186/1471-2105-12-385

9. Schubert M, Lindgreen S, Orlando L (2016) AdapterRemoval v2: rapid adapter trimming, identification, and read merging. BMC Res Notes 9: 88. DOI: https://doi.org/10.1186/s13104-016-1900-2

10. Wood DE, Lu J, Langmead B (2019) Improved metagenomic analysis with Kraken 2. Genome Biol 20: 257. DOI: https://doi.org/10.1186/s13059-019-1891-0

Supplementary table 1

Supplementary table 1: Beta-2-microglobulin (B2M) mRNA ct-values and concentrations for all patients.

Patient ID	B2M ct-value	B2M cp/µL
1	26.61	2.90E3
2	35.16	4.91E0
3	28.60	6.58E2
4	24.69	1.22E4
5	27.85	1.15E3
6	24.06	1.95E4
7	23.82	2.33E4
8	21.13	1.74E5
9	21.08	1.80E5
10	28.34	7.98E2
11	19.68	5.12E5
12	33.23	2.08E1
13	26.26	3.78E3
14	27.91	1.10E3
15	36.03	2.57E0
16	32.55	3.45E1
17	30.05	2.22E2
18	27.46	1.54E3
19	23.08	4.04E4
20	36.59	1.70E0
21	28.93	5.15E2
22	27.27	1.78E3
23	25.93	4.81E3
24	27.50	1.49E3
25	26.55	3.04E3
26	22.25	7.50E4
27	23.55	2.85E4

Supplementary table 2

Supplementary table 2: Proportion of human, bacterial, viral and unclassified reads for native sequencing and after the removal of rRNA as classified by the kraken2.

Patient ID	native sequencing					after rRNA removal				
	% human	% archaea	% bacterial	% viral	% unclassfied	% human	% archaea	% bacterial	% viral	% unclassified
1	94	0.01	1	0.00005	5	91	0.0004	1	0.006	8
2	98	0.01	1	0.0001	0.3	76	0.001	6	0.01	18
3	98	0.009	1	0.00005	0.2	91	0.0005	1	0.004	8
4	98	0.007	1	0.00003	0.7	93	0.0003	0.9	0.004	6
5	98	0.01	1	0.001	0.4	80	0.0007	3	0.02	17
6	98	0.008	1	0.00004	0.2	97	0.0002	0.3	0.001	3
7	98	0.008	1	0.00003	1	98	0.00009	0.2	0.001	2
8	97	0.008	1	0.00003	2	99	0.00009	0.1	0.0007	1
9	97	0.008	1	0.0001	2	98	0.00006	0.1	0.0008	2
10	97	0.009	1	-	2	88	0.0007	2	0.009	11
11	96	0.01	1	0.0001	2	99	0.00004	0.06	0.0003	0.8
12	98	0.009	1	0.00001	0.4	92	0.0003	1	0.005	7
13	97	0.008	1	0.00003	2	96	0.00006	0.6	0.003	4
14	98	0.01	2	0.00006	0.2	95	0.0003	0.5	0.002	4
15	98	0.01	1	0.00008	0.3	74	0.0007	4	0.02	22
16	96	0.009	1	0.00006	3	86	0.0005	2	0.01	12
17	95	0.01	1	0.0001	4	85	0.0003	1	0.009	14
18	98	0.008	1	0.00001	0.4	79	0.001	3	0.02	18
19	99	0.008	1	0.00005	0.3	94	0.0002	0.6	0.003	6
20	98	0.009	1	0.0001	0.2	48	0.006	6	0.03	46
21	98	0.01	2	0.0005	0.3	66	0.0005	5	0.03	29
22	59	0.006	2	0.0003	39	66	0.0005	3	0.02	30
23	98	0.008	1	0.00005	1	79	0.0007	2	0.02	18
24	98	0.008	1	0.00009	0.4	83	0.002	2	0.01	15
25	98	0.01	1	-	0.5	83	0.0003	2	0.01	16
26	94	0.01	4	0.002	2	99	0.00003	0.2	0.0009	1
27	98	0.01	1	0.00004	0.4	88	0.0002	1	0.007	10