

Metagenomic characterization of viruses in the serum of children with newly diagnosed cancer

Gustaf Leijonhufvud^{a,b}, Tatiany Aparecida Teixeira Soratto^{c,d}, Gabriel Machado Matos^c, Amanj Bajalan^g, Claudia Eichler-Jonsson^f, Britt Gustafsson^{b,e}, Gordana Bogdanovic^f, Tobias Allander^{f,g}, Gustaf Ljungman^{a,b,1}, Björn Andersson^{c,*,1}

^a Department of Women's and Children's Health, Uppsala University, SE-75237 Uppsala, Sweden

^b Department of Pediatric Hematology/Oncology, Children's University Hospital, SE-75185 Uppsala Sweden

^c Department of Cell and Molecular Biology, Karolinska Institutet, SE-171 77 Stockholm, Sweden

^d Department of Microbiology, Immunology and Parasitology, Federal University of Santa Catarina, Florianópolis, SC, Brazil

^e Department of Clinical Science, Intervention and Technology, CLINTEC, Karolinska Institutet, S-141 86 Stockholm, Sweden

^f Department of Clinical Microbiology, Karolinska University Hospital, SE-171 76 Stockholm, Sweden

^g Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, SE-171 77 Stockholm, Sweden

ARTICLE INFO

Keywords:

Virus metagenome
Pediatric cancer
Anellovirus
GBV-C
Virus infection
Cancer treatment

ABSTRACT

Background and Objectives: A large cohort of pediatric patients with various forms of childhood cancer was investigated for the presence of viruses using metagenomics. A total of 476 patient samples, collected between 1989 and 2018, were analyzed, representing various pediatric oncological diagnoses and a control group of non-malignant diagnoses.

Study design: The study was carried out using metagenomic sequencing of serum samples. Viruses were identified and analyzed using bioinformatics methods, followed by Polymerase chain reaction (PCR) confirmation

Results: The results indicate that a wide range of viruses can be detected in the bloodstream of children with newly diagnosed cancer. Nine viral genomes were identified: Human Pegivirus (HPgV), Hepatitis C virus, Parvovirus 1, Rhinovirus C, Human papillomavirus 116, Human polyomavirus 10, Parvovirus B19, and different variants of Torque Teno Virus (TTV). In this study, a previously unknown virus was found belonging to the Iflaviridae family in the order Picornavirales. HPGV was significantly more common in patients with leukemia compared to other conditions.

Conclusions: These results highlight the abundance of systemic virus infections in children, and the value of metagenomic sequencing for hypothesis forming regarding the associations between virus infections and cancer.

1. Background and objectives

Viruses are important pathogens in humans, and viral infections play a role in defining individual immune systems [1].

Metagenomic sequencing can be used to study viruses more effectively at a genomic level. The increase in the capacity of Next-Generation Sequencing (NGS) and improved bioinformatics methods have made it possible to characterize microbial communities using

shotgun sequencing [2].

Analyzing the viral load in children with cancer is interesting for several reasons. Many malignancies result in immunosuppression, and if some symptoms may be attributable to virus infections, this has relevance for treatment. New insights regarding the involvement of viruses in cancer may also emerge. Viral infections are the leading cause of approximately 10 % of the global cancer burden [3]. Also, certain types of cancer cells may be sensitive to infection by a particular virus [4].

Abbreviations: PCR, Polymerase chain reaction; RT-PCR, Reverse transcriptase polymerase chain reaction; TTV, Torque teno virus; ALL, Acute lymphatic leukemia; HPGV, GB virus C; HBV, hepatitis B virus.

* Corresponding author: Björn Andersson, Department of Cell and Molecular Biology, Karolinska Institutet, Biomedicum C9, 17177 Stockholm, Sweden, +46-70-355 3987

E-mail address: bjorn.andersson@ki.se (B. Andersson).

Shared last authorship

<https://doi.org/10.1016/j.jcv.2024.105736>

Received 13 May 2024; Received in revised form 3 October 2024;

Available online 6 October 2024

1386-6532/© 2024 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Hematological cancers are of particular interest, as infections that affect immunologically untrained cells seem to increase the risk of developing a leukemic condition [5].

This study uses a cohort of pediatric patients, sampled at diagnosis, with various forms of childhood cancer. Serum samples have been analyzed using shotgun metagenomic sequencing for the detection and characterization of viruses. The results have revealed a multitude of viruses present in these samples, including rarely detected viral species. The results also provide support for the previously suggested association between Human Pegivirus (HPgV) and lymphatic malignancies.

2. Study design

A total of 476 serum samples, collected between 1989 and 2018, from patients at the Department of Pediatric Hematology and Oncology, Uppsala Children's University Hospital, were analyzed. The sampled children suffered from various pediatric oncological diagnoses and a group of non-malignant hematological diagnoses was also included for comparison (Table 1). For metagenomic sequencing, the samples were divided into 23 pools based on diagnosis and age. The number of patients per pool ranged from 8 to 47.

The brain tumor group included common childhood brain tumors such as: astrocytoma, medulloblastoma, pilocytic astrocytoma, ependymoma, craniopharyngioma, and pons glioma. The group of solid tumors mostly consisted of patients diagnosed with Wilm's tumor, Ewing's sarcoma, osteosarcoma, teratoma, neuroblastoma, and germ cell tumor. The histiocytosis group included Langerhans cell histiocytosis, hemophagocytic lymphohistiocytosis, and familial erythrophagocytic lymphohistiocytosis. The non-malignant diagnoses group included patients initially suspected to have malignant disease. Anemias (other than aplastic anemia) were grouped together in one pool, and included iron deficiency anemia, but also Diamond-Blackfan and Fanconi anemias. The rest of the non-malignant group mostly consisted of patients investigated for various types of lymphadenopathy, but also hemangioma, hereditary spherocytosis, hereditary thrombocytopenia, and benign teratoma.

2.1. Nucleic acid extraction

21 µL to 50 µL of each sample were pooled, aiming for a total volume of 1 mL. PBS was added up to 1 mL final volume if necessary. Total nucleic acids were extracted by MagNA Pure 96 (Roche). Total nucleic

Table 1

The table shows how the 23 pools were distributed.

Samples/pool	Number of patient per group	Samples/pool	Number of patient per group
Pre B-ALL (0–2 y. o.)	22	Solid tumors (0–2 y.o.)	17
Pre B-ALL (3–5 y. o.)	22	Solid tumors (3–5 y.o.)	25
Pre B-ALL (3–5 y. o.)	22	Solid tumors (6–18 y.o.)	28
Pre B-ALL (6–18 y. o.)	22	Solid tumors (6–18 y.o.)	28
Pre B-ALL (6–18 y. o.)	22	CNS tumors (0–5 y.o.)	15
Pre B-ALL (6–18 y. o.)	22	CNS tumors (6–18 y.o.)	23
Acute myeloid leukemia	31	CNS tumors (6–18 y.o.)	22
T-ALL	15	Aplastic anemia	10
Non-Hodgkin Lymphoma	22	Other anemias	10
Hodgkin's lymphoma	18	Benign diagnoses (0–2 y.o.)	12
Histiocytoses	8	Benign diagnoses (3–5 y.o.)	13
		Benign diagnoses (6–18 y.o.)	47

acid elution was done in 100 µL 60 mM Tris–HCl buffer. For individual samples, 50 µL of serum, diluted with 150 µL of PBS), were extracted by MagNA Pure 96 (Roche) and eluted in 50 µL Tris–HCl buffer.

2.2. cDNA library

Amplified cDNA was obtained using QIAseq FX Single Cell RNA Library Kit (QIAGEN) according to the manufacturer's instructions, followed by purification with AGENCOURT® AMPure®XP PCR purification kit (Beckman Coulter) quantification with a Qubit Fluorometer.

2.3. DNA library

DNA whole-genome amplification was performed using the Illustra GenomiPhi V2 DNA Amplification Kit (GE Healthcare). Amplified DNA was purified with AGENCOURT® AMPure®XP PCR purification kit (Beckman Coulter) or Illustra MicroSpin G-50 Columns (GE Healthcare) and quantified as above.

2.4. Shotgun sequencing and bioinformatics analysis

Illumina sequencing libraries were produced using the ThruPLEX DNA-seq kit (Rubicon Genomics) followed by Illumina MiSeq and Novaseq sequencing at Science for Life Laboratory, Stockholm, Sweden.

Trimming was carried out using Cutadapt (v. 2.8)[6] and Trim Galore (v. 0.6.5–1). Human reads were removed by mapping to the Human genome build GRCh 38 (GCA_000001405.15) using Bowtie 2 (v. 2.4.1) [7].

The reads were analyzed and assembled using an in-house pipeline built using Nextflow (v. 19.07.0.5106) called "Discovery2" (<https://github.com/Amanj1/Discovery2>). In brief, the reads were assembled using MEGAHIT (v. 1.2.8) [8]. The reads and contigs were classified using Kraken2 (v. 2.0.8-beta)[9] and Diamond Blastx (v. 2.0.6), which aligns protein query sequences to the "nr" (Non-redundant protein sequences) database. The NCBI Taxonomic database and in-house scripts were used for taxonomic classification.

2.5. PCR verification assays

Virus-specific PCR assays were used to validate viral findings. PCR was conducted on individual samples in the pool positive for the virus by metagenomics, except for HPgV which was analysed by nested PCR for all samples. When possible, diagnostic assays used at the Dept of Clinical Microbiology were used. Primers are listed in Table 4. One-step RT-PCR reactions were performed using SuperScript™ III One-Step RT-PCR System with Platinum™ TaqHigh Fidelity DNA Polymerase (Invitrogen) and PCR reactions were performed with AmpliTaq Gold™ DNA Polymerase with Buffer II and MgCl₂ (Applied Biosystems) following manufacturer recommendations. PCR testing for HPgV was essentially performed as described in [10], and for TTV as described in [11]. Amplification reactions for Papillomavirus type 116, Polyomavirus 10, and Parvovirus B19 were carried out essentially as described above. The PCR products were analyzed by electrophoresis in 1,5 % agarose gels.

Parechovirus was detected by an in-house diagnostic real-time PCR assay composed of 5 µL of sample, 1x TaqPath™ One-Step RT-PCR Master Mix kit (Life Technologies), 1 µM of each forward and reverse primers and 0,2 µM of FRET-based MGB probe conjugated with fluorochrome FAM, which enables detection of the PCR products in the ABI7500 FAST Real Time PCR System (Applied Biosystems). Rhinovirus was detected by Allplex respiratory panel 3 (Seegene, Seoul, South Korea).

3. Results

Shotgun metagenomic sequencing for detection of viruses was

performed on pooled pediatric cancer samples. A total of 476 patient samples were analyzed (Table 1). The samples were divided into 23 pools based on diagnosis and age and the number of samples per pool ranged from 8 to 47.

The read-based analysis identified a total of 83,286,813 reads assigned to 14 virus families (Fig. 2 and Fig. 3). The dominant finding was, as expected for serum samples, different types of anelloviruses. Phage genomes and sequences present in the non template control were excluded from further analysis.

Eight genomes of viruses known to infect humans could be fully or partially assembled. These included 7 genomes classified at the species level (HPgV, Hepatitis C virus, Parechovirus 1, Rhinovirus C, Human papillomavirus 116, Human polyomavirus 10, and Parvovirus B19), and different anelloviruses. 54 complete anellovirus genomes, and multiple partial genomes, were assembled, but are here simply reported as Torque Teno viruses (TTV) (Table 3). In addition, a previously unknown genome belonging to the Iflaviridae family was identified.

HPgV reads were detected in 14 pools (Table 2) with the highest read number observed in the lymphoma group (274,543 reads).

A total of 87,574 reads from Hepatitis C virus (HCV) were detected in one sample pool: Solid malignant tumors, 0–2 years. Review of medical records confirmed that one of the samples in this group came from a child with a known HCV infection.

A total of 1365 reads of Parechovirus 1 were detected in the pool of solid malignant tumor patients (0–2 years). One sample, derived from an 11-month-old child who underwent surgery for sacral teratoma, was PCR-positive. The patient was febrile before and during surgery with no other obvious explanation.

A limited number of Rhinovirus C reads (12 reads) was detected in the B-ALL 0–2 years pool. The real-time PCR identified one positive sample. The patient was reported to have upper respiratory symptoms at the time of cancer diagnosis.

A complete human papillomavirus genome, classified as human Papillomavirus 116, was assembled from 20,111 reads from the B-ALL (3–5 years) pool. A single serum sample from a 3-year-old child with B-ALL was identified as the source using PCR.

A complete Human polyomavirus 10 genome was assembled from 1518 reads from the CNS tumor (0–5 years) pool. A serum sample from a 1-year-old child with pontine glioma was identified as the source using a specific PCR [12].

A total of 16,246 Parvovirus B19 reads were detected in the “other

benign diagnoses” (6–18 years) pool. Two individual samples were confirmed to be positive for parvovirus B19 by PCR. Both patients suffered from fever and lymphadenopathy, and one of the patients also suffered from severe anemia (Hemoglobin 50 g/L). The latter patient was diagnosed with hereditary spherocytosis. In both patients the parvovirus infection was the likely cause of the acute symptoms.

Six contigs corresponding to a previously unknown genome belonging to the Iflaviridae family were identified in the pool Solid malignant tumors, 0–2 years. Individual samples were tested by nested PCR (Table 4) and a single serum sample was positive. Gap closure produced an 8464 nucleotide contig (Supplementary figure 1). A phylogenetic analysis revealed a previously unknown species belonging to the *Iflavirus* genus. Since known iflaviruses have mainly been found in insects, the source and natural host of this virus is unclear. PCR screening of 1000 blood donor sera and 10 species of locally captured mosquitos (kindly provided by Dr. Tobias Lilja, Swedish Veterinary Agency), did not result in additional positive samples. We have tentatively named this virus Uppsala iflavirus.

TTV was abundant in all pools (20,862,935 reads). A clear pattern was observed, where levels of TTV in blood were significantly higher in the 3–5 year age group (Fig. 1). This appears to be the case regardless of the underlying diagnosis.

HPgV was identified in most pools. A higher number of reads was observed in five pools (Table 2). The presence of HPgV in individual samples was evaluated by PCR (Table 3). Of a total of 476 individuals, 37 tested positive for HPgV (Table 2). Patients with leukemia-related disease (B-ALL, T-ALL, Acute myeloid leukemia and Non-Hodgkin lymphoma) were compared with all other diagnoses (non-leukemia group). In the leukemia group 22/200 (11.0 %) of individuals were positive for HPgV. In the non-leukemia group 15/276 (5.4 %) were positive. The higher frequency of HPgV infection among leukemia patients was significant by chi-squared test ($p < 0.05$). The highest proportions of HPgV positive patients were seen in B-ALL of various ages (12.1 %) and non-Hodgkin lymphoma (13.7 %). In contrast, none of the Hodgkin's lymphoma patients were positive for HPgV.

A comparison of only lymphatic malignancies, i.e. ALL and Non-Hodgkin lymphoma, with the other patients resulted in a smaller difference that did not reach statistical significance ($p = 0.061$).

4. Discussion

Metagenomic sequencing is a powerful tool for unbiased characterization of the viral flora of clinical samples and sample collections [13]. Advantages of this approach include the possibility of finding all viruses, including novel strains and viruses that are normally not tested for by regular diagnostic assays.

HPgV is a single-stranded RNA virus belonging to the genus Pegivirus in the Flaviviridae family, that spreads via infected blood, vertical transmission from mother to child and sexual contact [14].

The viremia is in most cases cleared, but can in some cases persist. Presence of antibodies to the envelope protein E2 has been associated with viral clearance and is believed to protect against reinfection [15, 16]. The results from this study indicate that children diagnosed with leukaemia have a significantly higher prevalence of HPgV viremia at diagnosis than healthy children and children with other types of cancers. The prevalence of HPgV viremia among healthy adult blood donors in the United States is 1–2 % [17]. Studies of healthy children in Australia and Denmark showed prevalences of HPgV RNA of 1.3 % and 1.4 %, respectively [18,19]. The proportions found in our study are higher, particularly in leukemia and lymphoma patients.

Several large recent studies of adult patients suggested an association between HPgV and non-Hodgkin lymphoma, including a cohort study showing that viremia could be detected years prior to lymphoma diagnosis [20]. Our results support these earlier findings and indicate a similar association also with childhood leukemia.

Single-stranded DNA viruses dominate the blood virome, and

Table 2
HPgV results from cDNA sequencing and PCR verification assays.

Samples/pools	Coverage (x)	Reads classified by Kraken	PCR results positive samples
Pre B-ALL (0–2 y.o.)	2.93	154	5
Pre B-ALL (3–5 y.o.)	652.6	29,753	6
Pre B-ALL (6–18 y.o.)	998.33	47,034	3
Pre B-ALL (6–18 y.o.)	115.63	5221	2
Acute myeloid leukemia	123.45	7180	3
Non-Hodgkin Lymphoma	6526.16	274,543	3
Solid tumors (0–2 y.o.)	1430.71	72,625	1
Solid tumors (3–5 y.o.)	1.81	85	1
Solid tumors (6–18 y.o.)	5208.72	166,982	2
CNS tumors (6–18 y.o.)	8.99	434	1
Aplastic anemia	2973.8	132,443	1
Other anemias	604.12	24,108	1
Histiocytoses	42.3	1533	1
Benign diagnoses (0–2 y.o.)	169.84	6303	2
Benign diagnoses (6–18 y.o.)	13.52	515	3

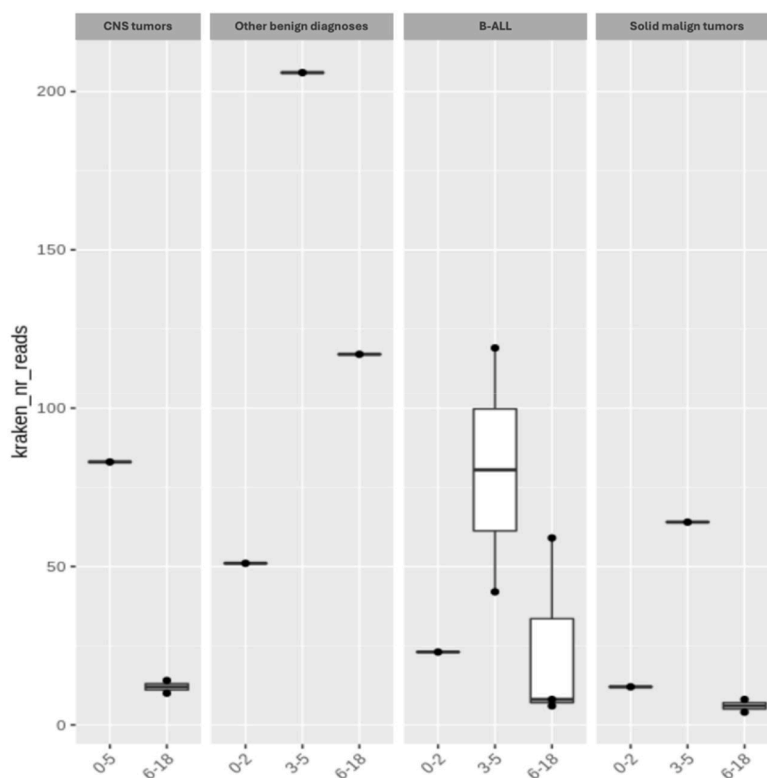


Fig. 1. TTV levels as determined by read numbers were highest in the 3–5 year age group, according to underlying diagnosis.

anelloviruses account for two-thirds of the total virome [21]. A rapid increase in TTV prevalence has been observed during the first year of life, with a continued increase up to 4 years of age [22]. TTV has not been linked to any human disease, but increased levels of TTV have been observed, for example, in septicemias [23], untreated solid tumors [24] and after bone marrow transplantation [25]. We did not observe a difference in TTV read numbers between the children diagnosed with non-malignant diseases or malignant diseases, but TTV read numbers were clearly higher in the 3–5 year age group, coinciding with the incidence peak for acute lymphoblastic leukemia at 2–5 years of age. However, quantitative analysis of anelloviruses is difficult and could be method- and primer-dependent [26]. Further metagenomics studies of this virus are needed.

Other, less prevalent virus findings include parvovirus B19, parechovirus, rhinovirus C, and polyomavirus 10. These viruses have been detected in blood samples previously [27,28,29] and were in some cases the likely causing symptoms of acute infection at the time of sampling.

In two cases of non-malignant disease, the likely cause of the symptoms was a parvovirus B19 infection which was overlooked both by the primary physician and the oncologist, but detected in this study. In an additional case, we identified the likely cause of a febrile episode (parechovirus) in a tumor surgery patient. These cases illustrate how this unbiased approach is a diagnostic strength of metagenomic sequencing.

HPV116 was detected in serum of a patient with pre-B ALL. The substantial amount and stable and reproducible detection makes contamination from skin unlikely.

A previously unknown picornavirales genome belonging to the Iflaviridae family was unexpectedly identified in a sample from a solid tumor patient. The genome was repeatedly detected in a single sample, which ruled out reagent contamination. Iflaviruses are not well-studied, but most infect insects. However, a few clades of the related dicistroviruses have been repeatedly detected in human serum samples, indicating that they could possibly replicate in humans [30,31,32]. The novel virus genome could not be detected by nested PCR in a limited study of local mosquitoes and blood donor sera. The origin and host

species of this virus genome remains unknown, but data are insufficient for ruling out that the finding represents an infection.

The previous discovery of various oncoviruses has had great importance for prevention and treatment of cancer. We expect that future research will generate new insights into the origins of childhood cancer and effective ways to prevent and treat it. In this study, we have analyzed a material collected over several decades and found that systemic virus infections are prevalent in children with newly diagnosed cancer, and that many of these infectious agents are not well studied. There is a need for further studies about the nature and consequences of many newly identified viruses in humans.

Ethical approval

The project is based on analysis of human clinical samples. Informed consent was obtained from the legal guardians of the patients sampled. The study was approved by the Regional ethics boards in Uppsala, 2018/258, and Stockholm, Dnr 02–212, 02–422, 04–836/4, 2008/1846–31.

Role of funder/sponsor (if any)

The funders had no role in the design and conduct of the study.

Data statement

The sequence data was deposited in the BioSample database at NCBI under bioproject ID PRJNA1036630.

CRediT authorship contribution statement

Gustaf Leijonhufvud: Writing – review & editing, Writing – original draft, Visualization, Validation, Resources, Methodology, Investigation, Formal analysis, Data curation. **Tatiany Aparecida Teixeira Soratto:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis. **Gabriel**

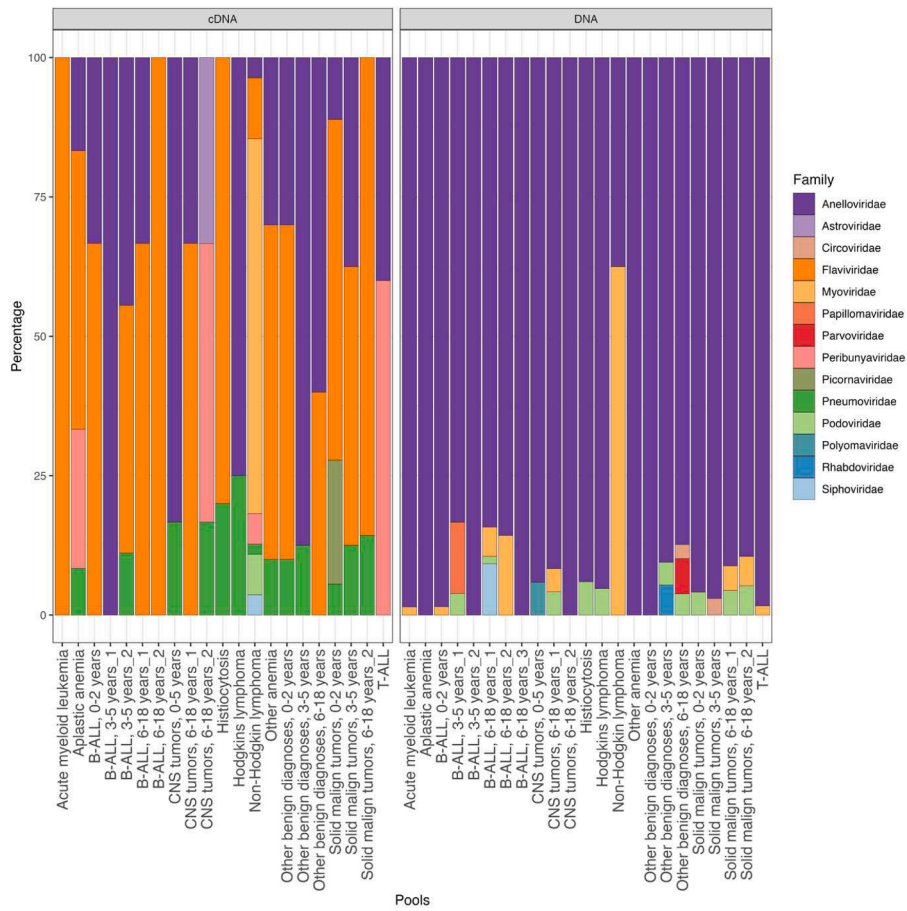


Fig. 2. Viral family diversity per pool of pediatric samples (unconfirmed raw data by Kraken2 classification). The number of reads per family in each pool was normalized and is presented as percentage.

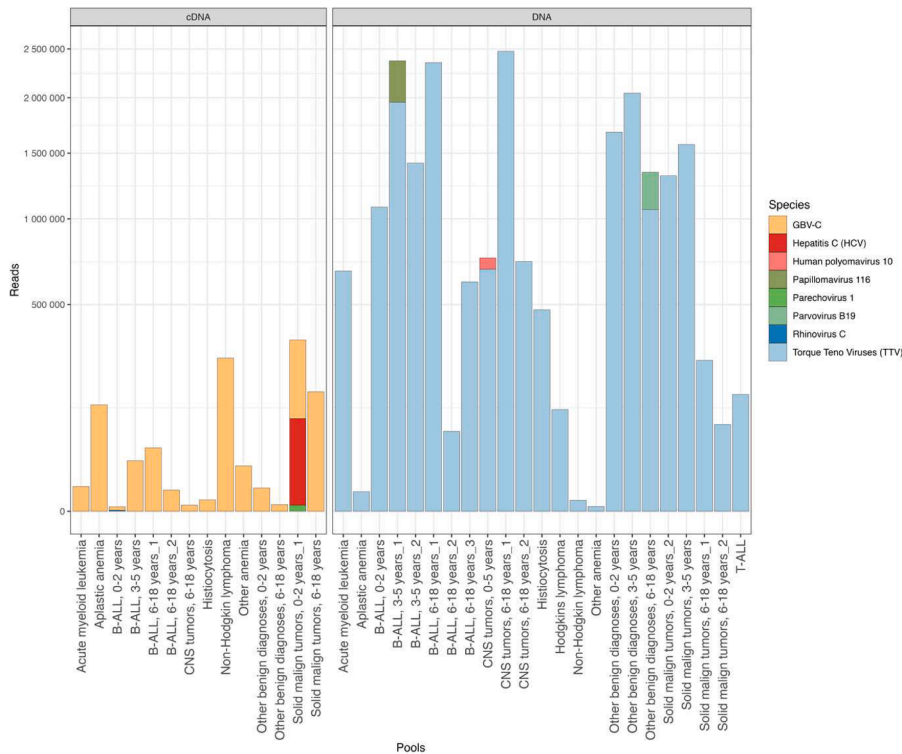


Fig. 3. Number of reads mapped to the confirmed viral genomes assembled from the samples.

Table 3
Results from PCR verification assays of virus-positive samples.

Family	Genus	Virus name	Virus-positive groups by PCR	Number of positive patient samples per group
Flaviviridae	Pegivirus	HPGV/HPgV-1	15/23 groups positive	See table 2
	Hepacivirus	Hepatitis C-virus	Pre-B-ALL (3–5 y.o.)	1
Picornaviridae	Enterovirus	Rhinovirus C	Pre-B-ALL (0–2 y.o.)	1
	Parechovirus	Parechovirus 1	Solid tumors (0–2 y.o. groups)	1(0–2 y.o.)*
Iflaviridae	Iflavirus	Uppsala iflavirus	Solid tumors (0–2 y.o.)	1
Papillomaviridae	Gammmapapillomavirus	Papillomavirus type 116	Pre-B-ALL (3.5 y.o.)	1
Polyomaviridae	Deltapolyomavirus	Polyomavirus 10	CNS tumors (0–5 y.o.)	1
Parvoviridae	Erythrovirus	Parvovirus B19	Benign diagnose (6–18 y.o.)	2
Circoviridae	Anellovirus	Torque teno virus (TTV)	Only tested for one pool Pre-B-ALL (3–5 y.o.)	13

* Individual samples in the group 6–18 years old were not tested due to high ct. value (40,8) in the group pool.

Table 4
Oligonucleotides used for PCR verification assays.

Virus	Primer ^a /Probe	Sequence (5'-3')	Product size (bp)
Parechovirus	Parecho_F547	CAC TAG TTG TAA GGC CCA CGA	70
	Parecho_R590	TCA GAT CCA YAG TGT CDC TTG TTA CCT	
	Parecho_P570	6-FAM - GAT GCC CAG AAG GTA - MGBNFQ	
Uppsala Iflavirus	2276 nedre F1	AAT GTG CCC AGA GCT GAA A	392
	2276 nedre R1	TCT AAC GCT CCT TTC TTG ACT ATA C	
	2276 nedre F3	GTT TAG GCA TCC AGA AGC AAT TTA	228
	2276 nedre R3	TCA TGA TAG GCC TTG TGT TCC	
Papillomavirus type 116	Papillo-F1	TCC ATG TTG GCA GTG GAT TTA	375
	Papillo-R1	CTT GGT TCT CCT TGT TCT CCT C	
	Papillo-F2	GGC ATA AAG TTG AAG GTG CTG	167
	Papillo-R2	GCT AGT GAC AGA GGC AGA AAT	
Parvovirus B19	Parvo-F	ACA AGC CTG GGC AAG TTA GC	75
	Parvo-R	ACA AGC CTG GGC AAG TTA GC	
Polyomavirus 10	HPyV10-4589. F	Gustafsson et al., 2013.	133
	HPyV10-4759. R		
Torque Teno Virus (TTV)	AMTS-F	Maggi et al., 2003	63
	AMTS-R AMTPTU-probe		
Pegivirus (HPGV)	GBV-F1	Sullivan et al., 2011.	299
	GBV-R1		
	GBV-F2		251
	GBV-R2		

^a F, sense primer; R, antisense primer.

Machado Matos: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation. **Amanj Bajalan:** Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis. **Claudia Eichler-Jonsson:** Writing – review & editing, Validation, Investigation, Data curation. **Britt Gustafsson:** Writing – review & editing, Writing – original draft, Supervision, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Gordana Bogdanovic:** Writing – review & editing, Writing – original draft, Project administration, Formal analysis, Data curation, Conceptualization. **Tobias Allander:** Writing – review & editing, Writing – original draft, Validation, Supervision, Resources, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Gustaf Ljungman:** Writing –

review & editing, Writing – original draft, Supervision, Resources, Project administration, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Björn Andersson:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Project administration, Supervision, Visualization, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors have no conflicts of interest to disclose.

Acknowledgments

The authors acknowledge support from the National Genomics Infrastructure in Stockholm funded by Science for Life Laboratory. The authors have no conflict of interest to declare.

Funding/Support

This study was funded by the Knut and Alice Wallenberg Foundation, The Swedish Childhood Cancer Fund, and the Swedish Research Council.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.jcv.2024.105736](https://doi.org/10.1016/j.jcv.2024.105736).

References

- [1] H.W. Virgin, E.J. Wherry, R. Ahmed, Redefining chronic viral infection, *Cell* 138 (1) (2009) 30–50, <https://doi.org/10.1016/j.cell.2009.06.036>.
- [2] M.R. Capobianchi, E. Giombini, G. Rozera, Next-generation sequencing technology in clinical virology, *Clin. Microbiol. Infect.* 19 (1) (2013) 15–22, <https://doi.org/10.1111/1469-0691.12056>.
- [3] M. Plummer, C. de Martel, J. Vignat, J. Ferlay, F. Bray, S. Franceschi, Global burden of cancers attributable to infections in 2012: A synthetic analysis, *Lancet Glob. Health* 4 (9) (2016) e609–e616, [https://doi.org/10.1016/S2214-109X\(16\)30143-7](https://doi.org/10.1016/S2214-109X(16)30143-7).
- [4] J.T. Schiller, D.R. Lowy, An introduction to virus infections and human cancer, *Recent Results Cancer Res.* 217 (2021) 1–11, https://doi.org/10.1007/978-3-030-57362-1_1.
- [5] M. Greaves, V. Cazzaniga, A.M. Ford, Can we prevent childhood Leukaemia? *Leukemia* 35 (5) (2021) 1258–1264, <https://doi.org/10.1038/s41375-021-01211-7>.
- [6] M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads, *EMBnet.J* 17 (1) (2011) 10, <https://doi.org/10.14806/ej.17.1.200>.
- [7] B. Langmead, S.L. Salzberg, Fast gapped-read alignment with Bowtie 2, *Nat. Methods* 9 (4) (2012) 357–359, <https://doi.org/10.1038/nmeth.1923>. Published 2012 Mar 4.
- [8] D. Li, C.M. Liu, R. Luo, K. Sadakane, T.W. Lam, Megahit: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph, *Bioinform.* 31 (10) (2015) 1674–1676, <https://doi.org/10.1093/bioinformatics/btv033>.
- [9] D.E. Wood, J. Lu, B. Langmead, Improved metagenomic analysis with Kraken 2, *Genome Biol.* 20 (1) (2019) 257, <https://doi.org/10.1186/s13059-019-1891-0>. Published 2019 Nov 28.

- [10] P.F. Sullivan, T. Allander, F. Lysholm, et al., An unbiased metagenomic search for infectious agents using monozygotic twins discordant for chronic fatigue, *BMC Microbiol.* 11 (2011) 2, <https://doi.org/10.1186/1471-2180-11-2>. Published 2011 Jan 2.
- [11] F. Maggi, M. Pifferi, C. Fornai, et al., TT Virus in the nasal secretions of children with acute respiratory diseases: Relations to viremia and disease severity, *J. Virol.* 77 (4) (2003) 2418–2425, <https://doi.org/10.1128/jvi.77.4.2418-2425.2003>.
- [12] B. Gustafsson, R. Jp, P. Priftakis, G. Giraud, T. Ramqvist, T. Dalianis, Human polyomaviruses were not detected in cerebrospinal fluid of patients with neurological complications after hematopoietic stem cell transplantation, *Future Virol* 8 (8) (2013) 809–814, <https://doi.org/10.2217/fvl.13.55>.
- [13] K.M. Wylie, G.M. Weinstock, G.A. Storch, Emerging view of the human virome, *Transl. Res.* 160 (4) (2012) 283–290, <https://doi.org/10.1016/j.trsl.2012.03.006>.
- [14] W.B. Supapol, R.S. Remis, J. Raboud, et al., Mother-to-child transmission of GB Virus C in a cohort of women coinfecting with GB Virus C and HIV in Bangkok, Thailand, *J. Infect. Dis.* 200 (2) (2009) 227–235, <https://doi.org/10.1086/599793>.
- [15] R.A. Gutierrez, G.J. Dawson, M.F. Knigge, et al., Seroprevalence of GB virus C and persistence of RNA and antibody, *J. Med. Virol.* 53 (2) (1997) 167–173, [https://doi.org/10.1002/\(sici\)1096-9071\(199710\)53:2<167::aid-jmv10>3.0.co;2-g](https://doi.org/10.1002/(sici)1096-9071(199710)53:2<167::aid-jmv10>3.0.co;2-g).
- [16] D.L. Thomas, D. Vlahov, H.J. Alter, et al., Association of antibody to GB virus C (hepatitis G virus) with viral clearance and protection from reinfection, *J. Infect. Dis.* 177 (3) (1998) 539–542, <https://doi.org/10.1086/514245>.
- [17] W.K. Roth, D. Waschk, S. Marx, et al., Prevalence of hepatitis G virus and its strain variant, the GB agent, in blood donations and their transmission to recipients, *Transfus. (Paris)* 37 (6) (1997) 651–656, <https://doi.org/10.1046/j.1537-2995.1997.37697335162.x>.
- [18] D.J. Siebert, D.S. Bowden, S.L. Tracy, W. Hardikar, Prevalence of hepatitis g/HPGV in a healthy paediatric population, *J. Paediatr. Child Health* 38 (4) (2002) 423, <https://doi.org/10.1046/j.1440-1754.2002.00029.x>.
- [19] P.B. Christensen, N. Fisker, L.H. Mygind, et al., GB Virus C epidemiology in denmark: Different routes of transmission in children and low- and high-risk adults, *J. Med. Virol.* 70 (1) (2003) 156–162, <https://doi.org/10.1002/jmv.10359>.
- [20] A. Famà, M.C. Larson, B.K. Link, et al., Human pegivirus infection and lymphoma risk: A systematic review and meta-analysis, *Clin. Infect. Dis.* 71 (5) (2019) 1221–1228, <https://doi.org/10.1093/cid/ciz940>.
- [21] I. De Vlaminc, K.K. Khush, C. Strehl, et al., Temporal response of the human virome to immunosuppression and antiviral therapy, *Cell* 155 (5) (2013) 1178–1187, <https://doi.org/10.1016/j.cell.2013.10.034>.
- [22] E. Väisänen, I. Kuisma, M. Mäkinen, et al., Torque teno virus primary infection kinetics in early childhood, *Viruses*. 14 (6) (2022) 1277, <https://doi.org/10.3390/v14061277>. Published 2022 Jun 11.
- [23] A.H. Walton, J.T. Muenzer, D. Rasche, et al., Reactivation of multiple viruses in patients with sepsis, *PLoS One* 9 (2) (2014) e98819, <https://doi.org/10.1371/journal.pone.0098819>. Published 2014 Jun 11.
- [24] S. Zhong, W. Yeo, M.W. Tang, et al., Gross elevation of TT virus genome load in the peripheral blood mononuclear cells of cancer patients, *Ann. N. Y. Acad. Sci.* 945 (2001) 84–92, <https://doi.org/10.1111/j.1749-6632.2001.tb03868.x>.
- [25] A. Pradier, S. Masouridi-Levrat, C. Bosshard, et al., Torque teno virus as a potential biomarker for complications and survival after allogeneic hematopoietic stem cell transplantation, *Front. Immunol.* 11 (2020) 998, <https://doi.org/10.3389/fimmu.2020.00998>. Published 2020 May 27.
- [26] G. Leijonhufvud, A. Bajalan, T.A. Teixeira Soratto, et al., Better detection of Torque teno virus in children with leukemia by metagenomic sequencing than by quantitative PCR, *J. Med. Virol.* 94 (2) (2022) 634–641, <https://doi.org/10.1002/jmv.27409>.
- [27] S.C.M. De Crom, J.W.A. Rossen, A.M. Van Furth, C.C. Obihara, Enterovirus and parechovirus infection in children: A brief overview, *Eur. J. Pediatr.* 175 (8) (2016) 1023–1029, <https://doi.org/10.1007/s00431-016-2725-7>.
- [28] N. Fuji, A. Suzuki, S. Lupisan, et al., Detection of human rhinovirus c viral genome in blood among children with severe respiratory infections in the Philippines, *PLoS One* 6 (11) (2011) e27247, <https://doi.org/10.1371/journal.pone.0027247>.
- [29] S. Kamminga, E. van der Meijden, C. de Brouwer, M. Feltkamp, H. Zaaijer, Prevalence of DNA of fourteen human polyomaviruses determined in blood donors, *Transfus. (Paris)* 59 (12) (2019) 3689–3697, <https://doi.org/10.1111/trf.15557>.
- [30] T.G. Phan, J. Del Valle Mendoza, M. Sadeghi, E. Altan, X. Deng, E. Delwart, Sera of Peruvians with fever of unknown origins include viral nucleic acids from non-vertebrate hosts, *Virus. Genes.* 54 (1) (2018) 33–40, <https://doi.org/10.1007/s11262-017-1514-3>.
- [31] S. Cordey, F. Laubscher, M.A. Hartley, et al., Detection of dicistroviruses RNA in blood of febrile Tanzanian children, *Emerg. Micro. Infect* 8 (1) (2019) 613–623, <https://doi.org/10.1080/22221751.2019.1603791>.
- [32] J.U. Oguzie, B.A. Petros, P.E. Oluniyi, et al., Metagenomic surveillance uncovers diverse and novel viral taxa in febrile patients from Nigeria, *Nat. Commun.* 14 (1) (2023) 4693, <https://doi.org/10.1038/s41467-023-40247-4>. Published 2023 Aug 4.