# Four complete genomes of human parvovirus B19 from amniotic fluid specimens.

Enagnon Kazali Alidjinou, Lander de Coninck, Jill Swinnen, Mouna Lazrek, Didier Hober, Jelle Matthijnssens

Microbiology Resource Announcements

| Clinical Microbiology | Announcement

# Four complete genomes of human parvovirus B19 from amniotic fluid specimens

**Enagnon Kazali Alidjinou,[1,2] Lander De Coninck,[2] Jill Swinnen,[2] Mouna Lazrek,[1] Didier Hober,[1] Jelle Matthijnssens[2]**

**AUTHOR AFFILIATIONS** See affiliation list on p. 2.

**ABSTRACT** We report the sequences of four complete genomes of parvovirus B19, extracted from human amniotic fluid specimens collected from pregnant women with abnormal ultrasound features in France. The genome sequences are 5,596 nucleotides long and include long terminal repeats. Several amino acid substitutions were observed in nonstructural protein (NS1).

**KEYWORDS** parvovirus B19, genomes, amniotic fluid

Parvovirus B19 belongs to the species *Erythroparvovirus primate 1*, in the genus *Erythroparvovirus* in the family *Parvoviridae* (1). Clinically, B19 is associated with a wide range of diseases, such as "erythema infectiosum" in children, acute polyarthritis in adults, and aplastic crisis in sickle cell disease patients. Primary infection in pregnant women can cause hydrops fetalis in the developing second-trimester fetus. The diagnosis of congenital infection is usually confirmed by testing amniotic fluid samples (AFS) (2). The B19 genome is a single-stranded DNA (ssDNA) molecule of 5,596 nucleotides (nt) with long (383 nt) terminal repeats (TRs). Three genotypes (G) are currently described, including the widely circulating G1 and two others—G2 and G3—which diverge in genome nucleotide sequence by ~10% (1, 3). Sequences of several isolates from human samples are available. However, no sequence detected from AFS has been reported.

Five AFS (P1–P5) collected between 2017 and 2019 at the University Hospital of Lille, France, through amniocentesis from pregnant women presenting with abnormal fetal ultrasound features tested positive for B19 by a qPCR assay (AltoStar Parvovirus B19 PCR Kit, Altona Diagnostics). The viral loads were 6.4, 8.1, 2.8, 8.6, and 8.1 Log copies/mL for P1–P5, respectively. Samples were stored at –80°C and later used for sequencing.

We performed the previously described NetoVIR protocol for viral particle purification (4, 5). Briefly, 200 µL of AFS were enriched for virus-like particles (centrifugation, filtration, and nuclease treatment) and submitted to nucleic acid extraction using the QIAamp Viral RNA Mini Kit (Qiagen). Then, cDNA synthesis and random PCR amplification were done using the whole transcriptome amplification kit 2 (Sigma Aldrich). Finally, library preparation using the Nextera XT DNA kit (Illumina) was performed, followed by 2 × 150 bp paired-end sequencing on a NovaSeq 6000 platform.

Paired reads from each sample were analyzed using ViPER, a bioinformatic pipeline designed to trim and assemble paired-end Illumina reads and classify the resulting contigs (6).

Thereafter, trimmed reads were mapped to the B19 reference genome (B19-J35, AY386330) using bwa-mem2. Mapping results were obtained with samtools coverage (7), and fasta consensus sequences were generated using samtools mpileup and ivar consensus (8, 9). All tools were run with default parameters unless otherwise specified.

The reference mapping yielded full genome coverage (5596 nt) for four samples (LP1, LP2, LP4, and LP5) (see Table 1). All the sequences shared a nucleotide identity higher

**Editor** Simon Roux, DOE Joint Genome Institute, Berkeley, California, USA

Address correspondence to Jelle Matthijnssens, jelle.matthijnssens@kuleuven.be.

The authors declare no conflict of interest.

**Received** 3 July 2023
**Accepted** 5 August 2023
**Published** 15 September 2023

**TABLE 1** Mapping results and analysis of coding regions compared to reference strain B19-J35 (AY386330)

| Specimens | | | LP1 | LP2 | LP4 | LP5 |
|---|---|---|---|---|---|---|
| Nucleoditic sequences (mapping results) | Number of trimmed reads aligned to the reference | | 3,454,461 | 400,427 | 47,893,753 | 3,968,351 |
| | Covered bases (%)[a] | | 100 | 100 | 100 | 100 |
| | Mean depth of coverage | | 77,100 | 7,530 | 1,060,210 | 87,800 |
| | Mean baseQ in covered region | | 36.1 | 36.3 | 36 | 35.9 |
| Proteins | Nonstructural protein (NS1) (672 nucleotides) | Coverage (%) | 100 | 100 | 100 | 100 |
| | | Concordance (%) | 99.3 | 99 | 99.3 | 99.8 |
| | | AA[c] changes | L8I, L57F, A71V, F111L, S190G, T505P, F554S | L8I, L57F, A71V, F111L, S190G, S195P, D280N, T505P, F554S | L57F, A71V, F111L, S190G, S344T, T505P, F554S | C17S, L57F, A71V, F111L, E114G, G159A, T163N, I181M |
| | 7.5 kDa protein (75 nucleotides) | Coverage (%) | 100 | 100 | 100 | 100 |
| | | Concordance (%) | 100 | 98.7 | 100 | 100 |
| | | AA changes | None | M1T[b] | None | None |
| | Capsid protein 1 (782 nucleotides) | Coverage (%) | 100 | 100 | 100 | 100 |
| | | Concordance (%) | 99.9 | 99.9 | 99.9 | 99.3 |
| | | AA changes | E14K | E14K | E14K | D12N, E14K, V30L, S98N, D107N, A260T, N533S |
| | Protein X (82 nucleotides) | Coverage (%) | 100 | 100 | 100 | 100 |
| | | Concordance (%) | 100 | 100 | 100 | 99.4 |
| | | AA changes | None | None | None | A15T |
| | Capsid protein 2 (555 nucleotides) | Coverage (%) | 100 | 100 | 100 | 100 |
| | | Concordance (%) | 100 | 100 | 100 | 99.7 |
| | | AA changes | None | None | None | A33T, N306S |
| | 11 kDa protein (95 nucleotides) | Coverage (%) | 100 | 100 | 100 | 100 |
| | | Concordance (%) | 100 | 99.2 | 100 | 98.9 |
| | | AA changes | None | T61K | None | M1T[b], I54V |

[a]Few nucleotidic positions in the terminal repeats (59, 15, and 19 in LP1, LP4, and LP5, respectively) were covered by less than 10 reads and were, therefore, removed (external nucleotides) or labeled as "N" (internal nucleotides) in the consensus sequences. A G insertion was observed at nt 567 in the LP5 5′ TR.
[b]A mutation in the start codon (resulting in the M1T amino acid change according to the reference) was observed for the 7.5 kDa protein in LP2 and the 11 kDa protein in LP5, and new coding DNA sequences (CDS) were defined using the NCBI ORFfinder.
[c]AA: amino acid.

than 98% with the reference genome, suggesting that they belong to genotype 1. The GC content ranged between 43.6% and 43.9%.

The coding DNA sequences shared 98.7–100% identity with the reference. CDS features were defined by pairwise alignment with the reference. The highest number of amino acid substitutions was observed in nonstructural protein 1 (NS1) for all isolates (7, 9, 7, and 8 for LP1, LP2, LP4, and P5, respectively). Regarding the structural proteins, only the capsid protein 1 substitution E14K was observed in the LP1, LP2, and LP4 genomes, while LP5 displayed 7 and 3 mutations in the capsid proteins 1 and 2, respectively.

## ACKNOWLEDGMENTS

We thank the University Hospital of Lille for providing a scholarship to support E.K.A.'s stay at the Rega Institute, KU Leuven.

## AUTHOR AFFILIATIONS

[1]Univ Lille, CHU de Lille, Laboratoire de Virologie ULR 3610, Lille, France
[2]KU Leuven, Department of Microbiology, Immunology and Transplantation, Rega Institute, Laboratory of Viral Metagenomics, Leuven, Belgium

## AUTHOR ORCIDs

Enagnon Kazali Alidjinou http://orcid.org/0000-0002-1106-9826
Jelle Matthijnssens http://orcid.org/0000-0003-1188-9733

## AUTHOR CONTRIBUTIONS

Enagnon Kazali Alidjinou, Conceptualization, Formal analysis, Investigation, Methodology, Writing – original draft | Lander De Coninck, Formal analysis, Methodology | Jill Swinnen, Methodology | Mouna Lazrek, Investigation | Didier Hober, Investigation | Jelle Matthijnssens, Conceptualization, Formal analysis, Validation, Writing – review and editing

## DATA AVAILABILITY

All raw reads have been submitted to NCBI's SRA at PRJNA979336. BioSample accessions are: LP1: SAMN35578145, LP2: SAMN35578146, LP4: SAMN35578147, LP5: SAMN35578148. The consensus sequences have been deposited in GenBank under the accession nos. OR138119, OR138120, OR138121, OR138122.

## ETHICAL APPROVAL

Informed consent was obtained from the pregnant women. The study was approved by the institutional data protection authority of CHU Lille under the number DEC22-299.

## REFERENCES

1. Cotmore SF, Agbandje-McKenna M, Canuti M, Chiorini JA, Eis-Hubinger A-M, Hughes J, Mietzsch M, Modha S, Ogliastro M, Pénzes JJ, Pintel DJ, Qiu J, Soderlund-Venermo M, Tattersall P, Tijssen P, ICTV Report Consortium. 2019. ICTV virus taxonomy profile: *Parvoviridae*. J Gen Virol 100:367–368. https://doi.org/10.1099/jgv.0.001212
2. Young NS, Brown KE. 2004. Parvovirus B19. N Engl J Med 350:586–597. https://doi.org/10.1056/NEJMra030840
3. Servant A, Laperche S, Lallemand F, Marinho V, De Saint Maur G, Meritet JF, Garbarg-Chenon A. 2002. Genetic diversity within human erythroviruses: identification of three genotypes. J Virol 76:9124–9134. https://doi.org/10.1128/jvi.76.18.9124-9134.2002
4. Conceição-Neto N, Zeller M, Lefrère H, De Bruyn P, Beller L, Deboutte W, Yinda CK, Lavigne R, Maes P, Van Ranst M, Heylen E, Matthijnssens J. 2015. Modular approach to customise sample preparation procedures for viral metagenomics: a reproducible protocol for virome analysis. Sci Rep 5:16532. https://doi.org/10.1038/srep16532
5. Conceição-Neto N, Yinda KC, Van Ranst M, Matthijnssens J. 2018. Netovir: modular approach to customize sample preparation procedures for viral metagenomics. Methods Mol Biol 1838:85–95. https://doi.org/10.1007/978-1-4939-8682-8_7
6. ViPER. Available from: https://github.com/Matthijnssenslab/ViPER
7. Samtools coverage. 2023 Available from: http://www.htslib.org/doc/samtools-coverage.html
8. Samtools mpileup. 2023 Available from: http://www.htslib.org/doc/samtools-mpileup.html
9. Ivar consensus. Available from: https://andersen-lab.github.io/ivar/html/manualpage.html