



# SCIENTIFIC REPORTS



OPEN

## Diagnosis and characterization of canine distemper virus through sequencing by MinION nanopore technology

Alessia Peserico<sup>1</sup>, Maurilia Marcacci<sup>1</sup>, Daniela Malatesta<sup>1</sup>, Marco Di Domenico <sup>1</sup>, Annamaria Pratelli <sup>2</sup>, Iolanda Mangone<sup>1</sup>, Nicola D'Alterio<sup>1</sup>, Federica Pizzurro<sup>1</sup>, Francesco Cirone<sup>2</sup>, Guendalina Zaccaria<sup>1</sup>, Cesare Cammà<sup>1</sup> & Alessio Lorusso<sup>1</sup>

Prompt identification of the causative pathogen of an infectious disease is essential for the choice of treatment or preventive measures. In this perspective, nucleic acids purified from the brain tissue of a dog succumbed after severe neurological signs were processed with the MinION (Oxford Nanopore Technologies, Oxford UK) sequencing technology. Canine distemper virus (CDV) sequence reads were detected. Subsequently, a specific molecular test and immunohistochemistry were used to confirm the presence of CDV RNA and antigen, respectively, in tissues. This study supports the use of the NGS in veterinary clinical practice with potential advantages in terms of rapidity and broad-range of molecular diagnosis.

Rapid identification of the etiologic agent of an infectious disease is essential for setting up treatment and preventive measures. In general, pathogen identification is performed by direct diagnostic tests which normally include amplification of target nucleic acids by PCR-based assays. Although these approaches are highly specific and often validated, they suffer a number of limitations, including the difficulties of testing for the plethora of rare pathogens that might be expected to cause a given pathology and their inability to identify new or unexpected pathogens, eventually originating from cross-species jumps. Therefore, the availability of other more rapid, broad-range techniques has become more and more important in the milieu of laboratory diagnosis of infectious diseases.

In this context one of these novel and intriguing technologies is certainly represented by next generation sequencing (NGS), even more when performed for direct identification of pathogens from clinical samples. Genome sequencing directly from biological samples can provide insights into how viruses transmit and spread, and it has been successfully applied to both virus discovery and diagnostics<sup>1–3</sup>. Recently the release of the MinION (Oxford Nanopore Technologies, Oxford, UK), a novel portable real-time NGS sequencer, enables the application of sequencing for rapid diagnosis with inexpensive sample preparation even in low-throughput laboratories<sup>4–10</sup>.

Canine distemper (CD) is a highly contagious disease which involves mainly young dogs and other susceptible carnivores with high morbidity and mortality rates. CD infection is characterized by lympho-, neuro- and epithelio-tropism resulting in systemic infection with severe clinical signs and death<sup>11,12</sup>.

The causative agent is canine distemper virus (CDV), a member of the genus *Morbillivirus* of the family *Paramyxoviridae*. CDV is an enveloped, negative-sense, single-stranded RNA virus. Like other paramyxoviruses, the virus contains six structural proteins, termed nucleocapsid (N), phospho (P), large (L), matrix (M), hemagglutinin (H) and fusion (F) protein, and two accessory non-structural proteins (C and V) that were found as extra-transcriptional units within the P gene<sup>13</sup>. Circulating CDV strains are grouped in geographical lineages based on the genetic correlation with the H protein encoding gene. Phylogenetic and molecular evolutionary analyses of CDV have revealed that mutations affecting the binding site of the H protein for virus entry receptors are associated with the occurrence of disease emergence in novel host species<sup>14</sup>. In Italy, CDV has been detected in several occasions with different viral strains belonging to separate lineages circulating in the field<sup>15–22</sup>. Generally,

<sup>1</sup>National Reference Center for Whole Genome Sequencing of microbial pathogens: database and bioinformatic analysis, Istituto Zooprofilattico Sperimentale dell'Abruzzo e del Molise, Campo Boario, 64100, Teramo, Italy.

<sup>2</sup>Dipartimento di Medicina Veterinaria, strada provinciale per Casamassima Km 3, 70100, Valenzano, Bari, Italy. Correspondence and requests for materials should be addressed to A.L. (email: [a.lorusso@izs.it](mailto:a.lorusso@izs.it))

Brain tissue	
Active channels	460
Reads number	88,357
Mean reads length	1,011
Reads length N50	1,033
Quality filtered reads (>Q7)	87,221 (98.7%) 88Mb
Total bases	89,313,926

Brain tissue	
Viral reads /Classified reads	1.2%
CDV reads/Total viral reads	70%

Virus (CDV) isolate	
Active channels	453
Reads number	131,930
Mean reads length	961
Reads length N50	1,021
Quality filtered reads (>Q7)	129,354 (98%) 124.6Mb
Total bases	129,779,778

Virus (CDV) isolate	
Viral reads /Classified reads	4%
CDV reads/Total viral reads	100%

**Figure 1.** Statistics of Nanopore sequencing data from the brain tissue (A,B) and the CDV isolate (C,D). These data were obtained by using NanoPlot.

in Italy CDV circulation in domestic dogs is under control as for the availability of valid vaccines on the market. However, occasional outbreaks occur, and these outbreaks are generally related with the illegal trade of infected dogs from Eastern Europe<sup>17,18,21,23</sup>. The last large CDV outbreak in Italy was caused by a strain belonging to the Arctic lineage (prototype CDV2784/2013). It occurred in 2013 involving primarily the Abruzzi region, an area hosting several natural parks with significant animal biodiversity. In this outbreak, distemper caused overt disease in unvaccinated domestic and shepherd dogs, Apennine wolves (*Canis lupus*) and other wild carnivores living in natural areas of the Abruzzi and neighboring regions, including Molise. CDV circulation was still evidenced in the following years in wild and domestic animals<sup>22</sup>. Direct diagnosis of CDV infection in animals showing neurological signs is primarily performed by molecular assays detecting the viral genome in the cerebrospinal fluid or alternatively in the urine<sup>20,24</sup>.

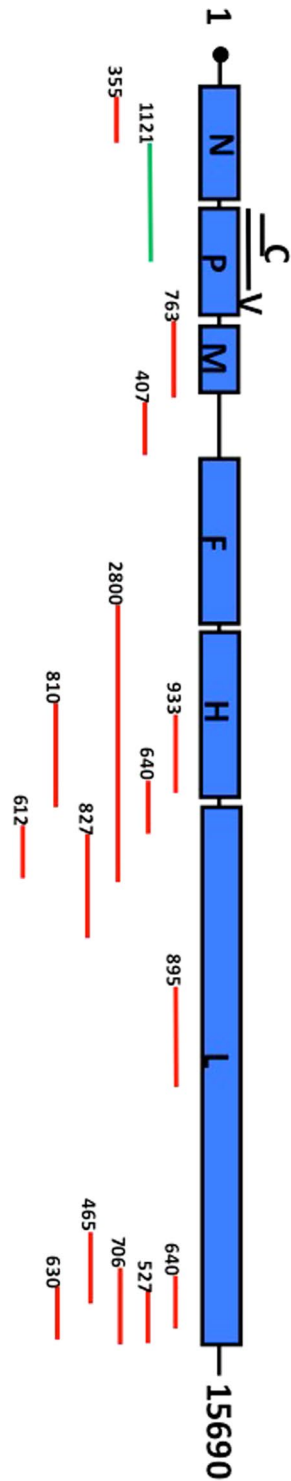
In this manuscript, we described the identification and characterization of a CDV strain from the brain tissue of a dog by using the MinION sequencing Nanopore technology.

## Results

**MinION real time sequencing from the brain tissue.** Data were produced by 460 active channels of the flow cell. A total number of 88.357 reads (89.313.926 bp) were obtained from the cDNA sequencing run. By assessing a quality cut-off (>Q7), a total number of 87.221 reads of sequencing reads was selected. The average length of the reads was 1.011 nt. Of these 1.886, the 96,4% was classified as eukaryota, the 2,4% as bacteria and the 1,2% as viruses (Fig. 1A,B). A total number of 16 reads, representing the 70% of reads assigned to viruses, were further classified as canine distemper virus. These reads contained sequence information of all protein encoding genes of CDV genome, in particular, H and L sequences (Fig. 2). One read of 2.688 nt encompassed a portion of the F, the whole H and of a portion of the L coding sequence. The first of these reads, which encompassed the N and P gene sequences, (Fig. 2) was retrospectively shown to be acquired within the first 20 minutes of sequencing run, the last after 11 hours. Overall, the 50% of CDV reads was revealed within the first 5 hours of real time sequencing. Conversely, within the run performed with the DNA library, no reads were classified as viruses (data not shown).

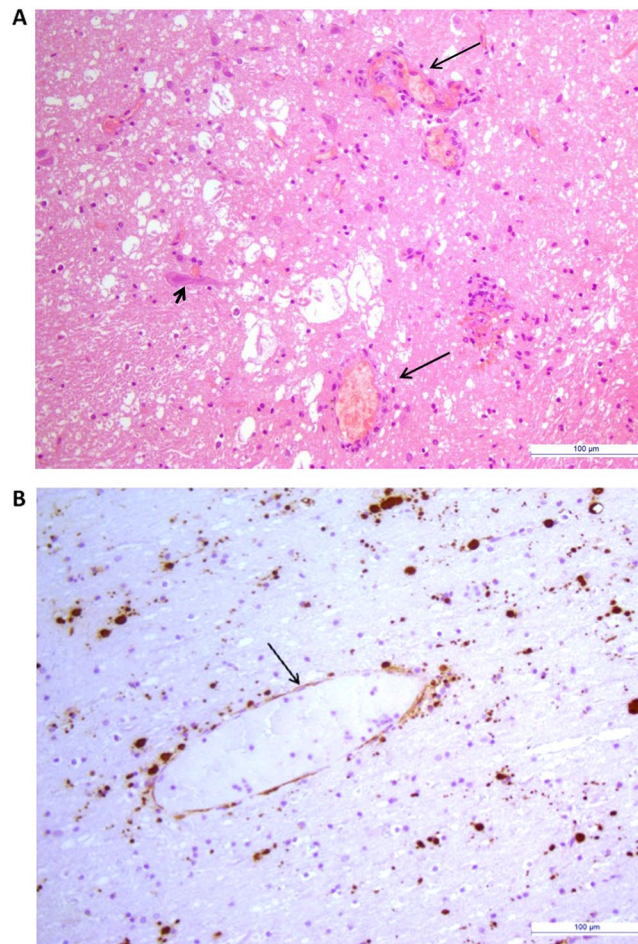
**rtPCR<sub>CDV</sub>, IHC and isolation.** A specific molecular test was performed after the identification of CDV by MinION. The rtPCR<sub>CDV</sub> confirmed the presence of RNA belonging to CDV in both tissues. Threshold cycles ( $C_T$ ) were 24 and 22 from brain and lung, respectively. Histological lesions and visualization of CDV antigen are showed in Fig. 3A,B, respectively. CDV (strain CDV1747/ITL2018) was isolated only from the lung homogenate at the second cell passage onto Vero.DogSLAMtag-cells and confirmed by rtPCR<sub>CDV</sub>.

**Whole genome sequencing of the isolate by MinION.** The whole genome sequence of CDV1747/ITL2018 (cell passage 3 onto Vero.DogSLAMtag-cells) has been obtained by an additional MinION sequencing run. Data were produced using 453 active channels of the flow cell. A total number of 131.930 reads were obtained from the cDNA sequencing run. By assessing a quality cut-off (>Q7), a total number of 129.354 reads (129.779.778 bp) of sequencing reads was selected. The average of the length of the reads was 961 nt. A total number of 5.164 reads, representing the 100% of the virus-assigned reads, were classified as canine distemper virus (Fig. 1C,D). Genomic coverage was 295X with 30X of minimum depth. The consensus sequence of CDV1747/ITL2018 (15.690 nt in length) has been deposited and annotated with the GenBank database (acc. no. MH316137). Raw reads, obtained from the MinION run with nucleic acids purified from the brain tissue, showed a nt identity ranging from 75 to 93% with the deposited consensus sequence. On the other hand, nt identity between the H protein encoding gene sequence of the isolate obtained by MinION and the sequences



**Figure 2.** Graphic representation of CDV genome (15,690 bp) and mapping of raw sequence reads originated from the MinION run with nucleic acid purified from the brain tissue. The green bar indicates the first detected CDV raw read. CDV gene sequences coding for viral proteins are represented by blue boxes. The number next to each bar/read indicates the nt length.

(either from isolate or infected tissues) obtained by Sanger sequencing was 100%. According to the phylogenetic analysis (Fig. 4) of the H protein encoding gene, CDV1747/ITL2018 clusters with strains Th270L and 19.876 identified from dogs in Thailand and USA, respectively. The H protein encoding gene sequence of CDV1747/ITL2018 bears, indeed, the highest sequence identity with strains Th270L (96.29%) and 19876 (96.59%) whereas with reference CDV lineages and vaccine strains, including CDV599/2016 (Europe wildlife lineage)<sup>22</sup> and



**Figure 3.** (A) Brain, acute demyelinating encephalitis with vacuolation of the white matter due to myelin sheath oedema. Acidophilic viral intranuclear inclusion in a shrunken neuron (black arrow head). Perivascular inflammatory changes are minimal with few cells (black arrows). HE staining. Bar = 100 μm. (B) Brain, strong staining of CDV antigen in neurons and in endothelial cells of a blood vessel (black arrow). IHC. Bar = 100 μm.

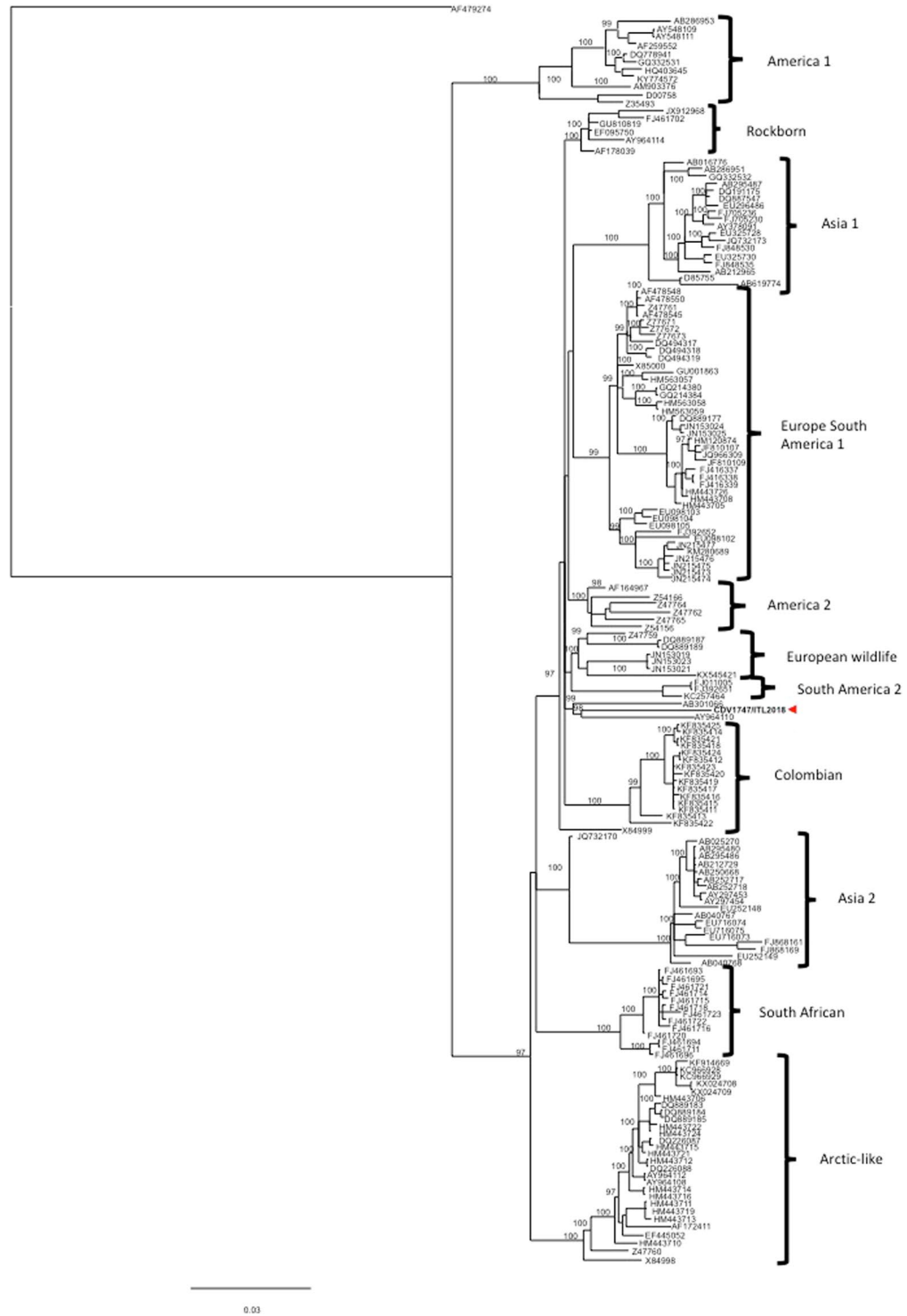
CDV2784/2013 (Arctic lineage)<sup>21,25</sup> which have been recently identified in central and southern Italy, sequence identity ranges from 90.89% to 95.48% (Table 1). Considering also the close temporal relationship between the onset of clinical signs and arrival to Italy, it is tempting to speculate that CDV1747/2018 was likely introduced to Italy from Cuba.

## Discussion

This study was planned and performed to assess the feasibility of the MinION sequencing technology in routine diagnostics. In this occasion, CDV infection was identified in the brain tissue of a dog imported to Italy from Cuba. The dog died after severe neurological infection, which was evidenced two days after the arrival from Cuba.

The MinION, performed with nucleic acids purified from the infected brain tissue, was able to provide indication of CDV infection. A specific molecular assay and IHC were then performed to confirm the presence of CDV in tissues. The same NGS platform was also used to obtain the whole genome sequence of the isolate. The involved strain, CDV1747/ITL2018, has been shown to be different from CDV strains which have been demonstrated to circulate in central/southern Italy in recent years.

One could reasonably argue that the clinical signs were already satisfactory to address the molecular diagnosis of CDV by existing direct assays, including rtPCR<sub>CDV</sub>. Although this aspect is undeniably true, the intention of this study is to promote efforts to validate NGS into routine diagnostics in order to get as soon as possible the identification of the etiological agent of a given disease and potentially opportunistic pathogens as previously demonstrated to occur in the field<sup>26</sup> with a broad-range diagnostic method. Undeniably, our study has some drawbacks. One of them may be represented by the low number of CDV reads originating from the MinION run which may apparently underestimate its use into diagnostic according to the adopted protocol. This could be explained by the fact that we did not perform any viral enrichment. In this regard, novel techniques for viral enrichment from stools, environmental and tissue samples have been recently described for viral metagenomics studies with second-generation sequencing technologies<sup>27–29</sup>. As for today, further efforts are still necessary to



**Figure 4.** Phylogenetic analysis of H protein encoding gene sequences retrieved from GenBank. Accession numbers are provided. Phylogeny carried out by Mr Bayes program implemented within the software package Geneious [version 9.1.8]. Bayesian inference was performed using four “chains” run over one million generations (with the first 2000 trees discarded as “burn-in”) and supplying statistical support with subsampling over 200 replicates. jModelTest was used to identify the most appropriate model of evolution. The identified program settings for all partitions, under the Akaike Information Criteria, included six-character states (General Time Reversible model), a proportion of invariable sites, and a gamma distribution of rate variation across sites (GTR + I + G). H gene sequence of Phocine distemper virus has been used as outgroup. Strain CDV1747/ITL2018 is indicated by a red arrow. Putative CDV lineages are indicated by brackets. Bar indicates the estimated number of nt substitutions per site.

CDV strains	CDV1747/ITL2018	Th270L	19876	Snyder Hill	Rockborn	Yanaka	SNP350/09/ITA	A75/17	CDV599/2016	Argentina 26	39-CO/12	07D111	4L70214	CDV2784/2013
	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt	nt
CDV1747/ITL2018	/	96.29	96.59	90.89	95.48	93.11	93.72	95.29	93.05	93.47	93.23	92.08	93.05	92.44
Th270L	96.29	/	96.09	91.01	95.42	94.23	94.11	95.49	93.81	94.29	94.66	92.35	93.69	93.02
19876	96.59	96.09	/	91.07	95.32	93.93	93.81	95.81	93.69	94.49	94.29	92.71	93.38	92.71
Snyder Hill	90.89	91.01	91.07	/	93.14	91.20	91.38	92.59	90.95	91.74	91.86	90.89	92.05	91.20
Rockborn	95.48	95.42	95.32	93.14	/	96.54	96.84	98.48	96.05	96.66	96.66	94.84	96.11	95.08
Yanaka	93.11	94.23	93.93	91.20	96.54	/	94.54	96.17	94.11	94.35	94.47	93.08	94.29	92.96
SNP350/09/ITA	93.72	94.11	93.81	91.38	96.84	94.54	/	96.24	94.29	94.54	94.74	92.90	94.11	93.26
A75/17	95.29	95.49	95.81	92.59	98.48	96.17	96.24	/	95.63	96.24	96.30	94.35	95.69	94.66
CDV599/2016	93.05	93.81	93.69	90.95	96.05	94.11	94.29	95.63	/	94.17	94.11	92.29	93.62	92.77
Argentina 26	93.47	94.29	94.49	91.74	96.66	94.35	94.54	96.24	94.17	/	94.60	92.65	94.29	92.96
39-CO/12	93.23	94.66	94.29	91.86	96.86	94.47	94.74	96.30	94.11	94.60	/	93.02	94.60	93.14
07D111	92.08	92.35	92.71	90.89	94.84	93.08	92.90	94.35	92.29	92.65	93.02	/	93.62	92.65
4L70214	93.05	93.69	93.38	92.05	96.11	94.29	94.11	95.69	93.62	94.29	94.60	93.62	/	94.29
CDV2784/2013	92.44	93.02	92.71	91.20	95.08	92.96	93.26	94.66	92.77	92.96	93.14	92.65	94.29	/

**Table 1.** Nucleotide (nt) sequence identity (expressed as %) between the H protein gene sequence of CDV1747/ITL2018 and representative strains belonging to existing CDV lineages.

assess protocols for virome preparation for third-generation sequencing platforms. Indeed, although Theuns and colleagues<sup>7</sup> described a viral enrichment protocol from stool samples for MinION sequencing<sup>5</sup>, a specific viral enrichment from tissue samples combined with MinION still needs to be explored.

On the other hand, in this study we performed the SISPA amplification before the sequencing run. SISPA is an established protocol which permitted, through random primer amplification, the identification and characterization of several viruses, including different avian RNA viruses<sup>30</sup>, Schmallenberg virus<sup>31</sup>, Hepatitis C virus from clinical biopsy samples<sup>32</sup>, a novel mink astrovirus<sup>33</sup>, viruses from human stool samples<sup>34</sup>, porcine epidemic diarrhea virus from swine feces<sup>35</sup>, classical and atypical Bluetongue virus serotypes from blood samples<sup>36–39</sup>, feline morbillivirus from urine<sup>40</sup> and viruses responsible for respiratory tract infections<sup>41</sup>. Moreover, recently our group combined successfully SISPA and MinION sequencing for the identification of bovine enterovirus from a cell-culture isolate<sup>6</sup>. By the way, several enrichment strategies from tissues and combination of MinION sequencing are currently under investigation by our team. Furthermore, we did not perform live basecalling and diagnostic indications were provided after the sequencing run was stopped without live acquisition of raw sequence data. The local WIMP analysis, indeed, was completed after 13 hours starting from the load of the flow cell with cDNA. However, the retrospective analysis showed that the first read came out within the first 20 minutes of the sequencing run. Interestingly, when the sample was re-processed by the live pipeline (mimicking a real-time sequencing scenario) the first sequence read was assigned to CDV after 40 minutes from the beginning of the sequencing run. This would suggest that if used as point-of-care device, solid identification of CDV would have been acquired within less than 1 hour starting from the beginning of the sequencing run.

Conversely, what could have happened if rtPCR<sub>CDV</sub> was negative when used as first diagnostic attempt? The only valid strategy for diagnosis would have been the one which includes the testing of purified nucleic acids for all the plethora of pathogens known to cause the clinical signs observed in that dog. Reasonably, this approach would have required a longer period to get, eventually, a valid response. This aspect combined with the high chance of unavailability of specific tests, contributes to the relatively high proportion of cases that remain undiagnosed. Therefore, we attempted to use the MinION sequencing platform to get diagnostic indications. In this case, acquired raw CDV reads were not identical to the deposited consensus sequence. This can be easily explained by the fact that reads originated by the second MinION run, with nucleic acids purified from the viral isolate, were further processed with two bioinformatics, GraphMap and Pilon, tools. These tools fix the errors which occurred during the assembly of sequence reads. Further support to the credibility of the final consensus sequence has been provided by the Sanger sequencing of the H protein encoding gene of the occurring strain, either from the isolate or directly from the tissue specimens. Indeed, Sanger sequencing evidenced 100% of nt identity with the sequence obtained by the second MinION run. Nevertheless, raw reads were indispensable and satisfactory to identify the involved viral species and to get diagnostic indications.

The MinION platform is a unique, highly portable, single molecule real time sequencer currently available. This technology recently emerged to reduce the sequencing costs and to simplify the library preparation procedures. The hallmark of the MinION technology is certainly represented by the fact that this platform can be routinely used to investigate viral and bacterial outbreaks also in low throughput laboratories and, importantly, in field conditions with a real time acquisition of raw data. Further validation experiments are currently ongoing to validate the use of MinION in the daily routine with different biological matrices. Overall, these data demonstrate that MinION is affordable and able to detect CDV quickly and with enough accuracy, supporting its potential for virus surveillance also in field conditions.

## Methods

**Case description.** A male of Mexican hairless dog (Xoloitzcuintli) born on July 4<sup>th</sup>, 2017 was legally introduced from Cuba to Italy (municipality of Campobasso, Molise region, southern Italy) on February 19<sup>th</sup>, 2018. The dog was vaccinated for rabies (Virbac-UK, <https://uk.virbac.com/home/products/vaccines/main/dog-vaccines/canigen-rabies.html>) on November 10<sup>th</sup>, 2017. On February 21<sup>st</sup>, 2018, two days after the arrival in Italy, the dog showed labored breathing and severe neurological signs including disorientation, lack of alertness, convulsive movements of the head and paws (Supplementary Video S1), aimless wandering, hyper-salivation and chewing movements of the jaw. Therapy consisted of administration of fluids, cortisone and antibiotics. However, the dog succumbed few days later. The carcass was sent to IZSAM for necropsy. A complete necropsy examination was accomplished at IZSAM. Lung and brain tissues were employed for diagnosis.

**Virus isolation.** Vero cells stably expressing canine SLAM (Vero. DogSLAMtag<sup>42</sup>) were inoculated with lung and brain homogenates and incubated at 37 °C with 5% CO<sub>2</sub> for a maximum of 5–7 days until cytopathic effect (CPE) was observed. Plates showing no CPE were further passaged up to passage 3.

**Shotgun metagenomics by MinION.** 30 mg of brain were chopped into small pieces with a razor blade, resuspended in 1 ml of Phosphate-buffered saline (PBS) and homogenized with a TissueLyser II (Qiagen, Hilden Germany) for 5 min following manufacturer's guidelines.

The homogenate was centrifuged and 200 µl of the supernatant was used for nucleic acid purification using High Pure Viral Nucleic Acid Kit (Roche, Basel Swiss) and metagenomic analysis by MinION (Oxford Nanopore Technologies, Oxford UK). Nucleic acid elution was divided in two aliquots to perform RNA and DNA sequencing in two different runs as described in a previous study from our group with some modifications<sup>6</sup>.

Likewise, the supernatant of the Vero.DogSLAMtag-cells CDV isolate (passage 3) was collected and processed for RNA purification (High Pure Viral Nucleic Acid Kit, Roche, Basel CH).

RNA samples were processed by means of the Sequence Independent Single Primer Amplification (SISPA) method<sup>43</sup> with some modifications to obtain cDNA<sup>39,40</sup>. FR26RV-N (GCCGGAGCTCTGCAGATATCNNNNNN) and FR20RV (GCCGGAGCTCTGCAGATATC) primers were used for SISPA protocol<sup>43</sup>. Reverse transcriptase and PCR were carried out by using the SuperScript II Reverse Transcriptase (Thermo Fisher Scientific) and Pfu Ultra II Fusion HS DNA Polymerase (Agilent), respectively.

DNA and amplified cDNA were quantified by Qubit dsDNA HS assay (Thermo Fisher Scientific, Waltham, MA) and used for library preparation by Low input genomic DNA by PCR Barcoding (SQK-LWB001, Oxford Nanopore Technologies, Oxford UK) following manufacturer's guidelines. Briefly, DNA ends were end-repaired and dA-tailed using the NEBNext End Repair/dA tailing module (New England Biolabs, Hitchin, UK). PCR adapters were ligated onto the prepared ends using the Blunt/TA Ligase Master Mix (New England Biolabs, Hitchin, UK) and a unique barcode was added by PCR. The amplification program consisted of a first 3 min step at 95 °C followed by 14 cycles with the following conditions: 95 °C for 15 sec, 56 °C for 15 sec and 65 °C for 6 min. The program ended with 1 step at 65 °C for 6 min. Sequencing adapters were added prior library loading on the Flow cell MIN106, R9 version (Oxford Nanopore Technologies). All purification steps were carried out using AMPure XP beads (Agencourt, Beckman Coulter Brea, CA) according the SQK-LWB001 sequencing protocol. For sequencing, the NC\_48hr\_sequencing\_FLO-MIN106\_SQK-LBW001 program was run on MinKNOW Software v.1.4.2. The sequencing run was stopped after 11 hours from starting sequencing.

**MinION data analysis.** Fast5 raw MinION files were processed locally by using the Albacore software (version 2.1.10) to obtain fastq files output with a quality cut-off (>Q7). The local base calling allowed to obtain fastq files within less than 2 hours of processing. "What is in my pot" (WIMP) workflow<sup>43</sup> was then initiated in the EPI2MEAgent Software (version 2.47.537208) for analyzing brain tissue originated-reads produced by Albacore. Briefly, WIMP tool enables the comparison in real time of the raw reads against the NCBI Organism Database; this tool classifies and identifies the species by looking at the k-mers produced. WIMP platform required 10 minutes of analysis for classification.

Reads derived from the CDV isolate sequencing run, after Albacore processing, were further demultiplexed, and trimmed by using Porechop (version 0.2.3). To obtain a draft sequence by mapping, the closest reference sequence was identified performing Blastn search (version 0.5.2). The consensus sequence was obtained using the GraphMap (version 0.4.1<sup>44</sup>) aligner and using Pilon software (version 1.22). Specifically, we used the default parameters for Graphmap and a combination of samtools, bcftools, vcfutils and internal scripts to extract the consensus sequence from the alignment. Then, consensus sequence was refined by Pilon. The final sequence and ORFs were also manually inspected using the Ugene software (version 1.29.0<sup>45</sup>). Quality check analysis of raw reads was performed by using the Nanoplot (version 1.17.2).

**Diagnostic RT-qPCR (rtPCR<sub>CDV</sub>), immunohistochemistry (IHC).** Purified nucleic acids from brain and lung were screened by rtPCR<sub>CDV</sub> with specific primers and probe for the N protein-encoding gene of CDV<sup>46</sup>. The brain sample was fixed in 10% neutral buffered formalin and embedded in paraffin wax. Sections (5 µm) were stained with haematoxylin and eosin (HE). Additional tissue sections were also subjected to immunohistochemistry (IHC) using primary antibody specific for CDV (1:1000, mouse monoclonal antibody (Ab) anti CDVNP; VMRD<sup>TM</sup>, Pulmann WA). Labeling was performed by using a biotinylated goat anti-mouse Ab and subsequently by indirect streptavidin-biotin method (Dako REAL<sup>TM</sup> Detection System, Peroxidase/DAB kit, Santa Clara CA, USA). Antigen retrieval was performed by heat treatment in citrate buffer pH 6.0, 0.01 M at 121 °C for 5 min. Sections were visualized with 3,3 diaminobenzidine solution (DAB, Dako Santa Clara CA, USA), and finally counterstained with Mayer's hematoxylin. CDV-positive lung tissue sections of dogs available *in house* were used as positive controls, whereas the specificity of the immune-labeling was verified with an unrelated Ab.

**CDV isolate and phylogeny.** RNAs purified from the CDV isolate (passage 3) and tissue specimens were used for RT-PCR and Sanger sequencing of the H protein encoding sequence. Amplicons were obtained using RH-3 and RH-4 primers<sup>47</sup> and sequenced by dideoxy Sanger method, performed onto 3130xl Genetic Analyzer (Applied Biosystem, Foster City CA, USA). Sequence reads obtained by Sanger were assembled with DNASTar software package (DNASTar Inc., Madison WI, USA). Representative H protein nucleotide (nt) sequences of canine distemper virus strains were retrieved from GenBank (Supplementary Table S2) and aligned using the Clustal Omega tool from the European Molecular Biology Laboratory (<https://www.ebi.ac.uk/Tools/msa/clustalo/>). Phylogenetic analysis was carried out by Mr Bayes program<sup>48,49</sup> implemented within the software package Geneious (version 9.1.8, Biomatters Ltd., New Zealand). Bayesian inference was performed using four “chains” run over one million generations (with the first 2000 trees discarded as ‘burn-in’) and supplying statistical support with subsampling over 200 replicates. jModelTest<sup>50</sup> was used to identify the most appropriate model of evolution. The identified program settings for all partitions, under the Akaike Information Criteria, included six-character states (General Time Reversible model), a proportion of invariable sites, and a gamma distribution of rate variation across sites (GTR + I + G). H gene sequence of Phocine distemper virus was used as outgroup.

**Ethical Statement.** All methods, including necropsy and sampling, were carried out in accordance with internal guidelines and regulations of IZSAM. All experimental protocols, including handling and processing of infectious tissues, were approved by the Ethical Committee of IZSAM.

## References

- Deamer, D. W. & Branton, D. Characterization of nucleic acids by nanopore analysis. *Acc Chem Res.* **35**, 817–825 (2002).
- Nakane, J. J., Akeson, M. & Marziali, A. Nanopore sensors for nucleic acid analysis. *J Phys Condens Matter.* **15**, R1365–R1393 (2003).
- Healy, K. Nanopore-based single-molecule DNA analysis. *Nanomedicine* **2**, 459–481 (2007).
- Wang, J., Moore, N. E., Deng, Y. M., Eccles, D. A. & Hall, R. J. MinION nanopore sequencing of an influenza genome. *Front. Microbiol.* **6**, 766 (2015).
- Batovska, J., Lynch, S. E., Rodoni, B. C., Sawbridge, T. I. & Cogan, N. O. I. Metagenomic arbovirus detection using MinION nanopore sequencing. *J. Virol. Methods.* **249**, 79–84 (2017).
- Beato, M. S. *et al.* Identification and genetic characterization of bovine enterovirus by combination of two next generation sequencing platforms. *J. Virol. Methods* **260**, 21–25 (2018).
- Theuns, S. *et al.* Nanopore sequencing as a revolutionary diagnostic tool from porcine viral enteric disease complexes identifies porcine kobuvirus as an important enteric virus. *Sci Rep.* **8**(1), 9830 (2018).
- Croville, G. *et al.* Rapid whole-genome based typing and surveillance of avipoxviruses using nanopore sequencing. *J Virol Methods* **261**, 34–39 (2018).
- Vanmechelen, B. *et al.* Identification of a novel species of papillomavirus in giraffe lesions using nanopore sequencing. *Vet Microbiol.* **201**, 26–31 (2017).
- Günther, T. *et al.* Recovery of the first full-length genome sequence of a parapoxvirus directly from a clinical sample. *Sci Rep.* **7**(1), 3734 (2017).
- Lempp, C. *et al.* New aspects of the pathogenesis of canine distemper leukoencephalitis. *Viruses* **6**, 2571–2601 (2014).
- von Messling, V., Milosevic, D. & Cattaneo, R. Tropism illuminated: lymphocyte-based pathways blazed by lethal morbillivirus through the host immune system. *Proc. Natl. Acad. Sci. USA* **101**, 14216–14221 (2004).
- Orvell, C. Structural polypeptides of canine distemper virus. *Arch. Virol.* **66**, 193–206 (1980).
- McCarthy, A. J., Shaw, M. A. & Goodman, S. J. Pathogen evolution and disease emergence in carnivores. *Proc. Biol. Sci.* **274**, 3165–3174 (2007).
- Martella, V. *et al.* Detection and genetic characterization of canine distemper virus (CDV) from free-ranging red foxes in Italy. *Mol Cell Probes.* **16**(1), 77–83 (2002).
- Decaro, N. *et al.* Canine distemper and related diseases: report of a severe outbreak in a kennel. *New Microbiol.* **27**(2), 177–81 (2004).
- Martella, V. *et al.* Heterogeneity within the hemagglutinin genes of canine distemper virus (CDV) strains detected in Italy. *Vet. Microbiol.* **116**, 301–309 (2006).
- Martella, V. *et al.* Genotyping canine distemper virus (CDV) by a hemi-nested multiplex PCR provides a rapid approach for investigation of CDV outbreaks. *Vet Microbiol.* **16**(122(1–2)), 32–42 (2007).
- Monne, I. *et al.* A distinct CDV genotype causing a major epidemic in Alpine wildlife. *Vet Microbiol.* **150**, 63–69 (2011).
- Elia, G. *et al.* Virological and serological findings in dogs with naturally occurring distemper. *J Virol Methods.* **213**, 127–30 (2015).
- Di Sabatino, D. *et al.* Arctic lineage-canine distemper virus as a cause of death in Apennine wolves (*Canis lupus*) in Italy. *PLoS One.* **20**, 9(1), e82356 (2014).
- Di Sabatino, D. Lethal distemper in badgers (*Meles meles*) following epidemic in dogs and wolves. *Infect Genet Evol.* **46**, 130–137 (2016).
- Lorusso, A. & Savini, G. Old diseases for new nightmares: distemper strikes back in Italy. *Vet Ital.* **50**(2), 151–4 (2014).
- Martella, V., Elia, G., Buonavoglia, C. Canine distemper virus. *Vet Clin North Am Small Anim Pract.* **38**(4), 787–97, vii–viii (2008).
- Marcacci, M. *et al.* Whole genome sequence analysis of the arctic-lineage strain responsible for distemper in Italian wolves and dogs through a fast and robust next generation sequencing protocol. *J Virol Methods.* **202**, 64–8 (2014).
- Zaccaria, G. *et al.* Circovirus in domestic and wild carnivores: An important opportunistic agent? *Virology* **490**, 69–74.
- Conceição-Neto, N. *et al.* Modular approach to customise sample preparation procedures for viral metagenomics: a reproducible protocol for virome analysis. *Sci Rep.* **12**(5), 16532 (2015).
- Kumar, A., Murthy, S. & Kapoor, A. Evolution of selective-sequencing approaches for virus discovery and virome analysis. *Virus Res.* **15**(239), 172–179 (2017).
- Kohl, C. *et al.* Protocol for metagenomic virus detection in clinical specimens. *Emerg Infect Dis.* **21**(1), 48–57 (2015).
- Chrzastek, K. *et al.* Use of Sequence-Independent, Single-Primer-Amplification (SISPA) for rapid detection, identification, and characterization of avian RNA viruses. *Virology* **509**, 159–166 (2017).
- Rossee, T. *et al.* DNase SISPA-next generation sequencing confirms Schmallenberg virus in Belgian field samples and identifies genetic variation in Europe. *PLoS One* **7**(7), e41967 (2012).
- Daly, G. M. *et al.* A viral discovery methodology for clinical biopsy samples utilising massively parallel next generation sequencing. *PLoS One* **6**(12), e28879 (2011).
- Blomström, A. L., Widén, F., Hammer, A. S., Belák, S. & Berg, M. Detection of a novel astrovirus in brain tissue of mink suffering from shaking mink syndrome by use of viral metagenomics. *J Clin Microbiol.* **48**(12), 4392–6 (2010).
- Victoria, J. G. *et al.* Metagenomic analyses of viruses in stool samples from children with acute flaccid paralysis. *J Virol.* **83**(9), 4642–51 (2009).



35. Pizzurro, F. *et al.* Outbreak of porcine epidemic diarrhoea virus (PEDV) in Abruzzi region, central-Italy. *Vet Med Sci.* **4**(2), 73–79 (2018).
36. Sghaier, S. *et al.* A novel Bluetongue virus serotype 3 strain in Tunisia, November 2016. *Transbound Emerg Dis.* **64**(3), 709–715 (2017).
37. Savini, G. *et al.* Novel putative Bluetongue virus in healthy goats from Sardinia, Italy. *Infect Genet Evol.* **51**, 108–117 (2017).
38. Lorusso, A. *et al.* Analysis of bluetongue serotype 3 spread in Tunisia and discovery of a novel strain related to the bluetongue virus isolated from a commercial sheep pox vaccine. *Infect Genet Evol.* **59**, 63–71 (2018).
39. Marcacci, M. *et al.* One after the other: A novel Bluetongue virus strain related to Toggenburg virus detected in the Piedmont region (North-western Italy), extends the panel of novel atypical BTV strains. *Transbound Emerg Dis.* **65**, 370–374 (2018).
40. Marcacci, M. *et al.* Genome characterization of feline morbillivirus from Italy. *J Virol Methods.* **234**, 160–163, <https://doi.org/10.1016/j.jviromet.2016.05.002> (2016).
41. Thorburn, F. *et al.* Erratum to “The use of next generation sequencing in the diagnosis and typing of respiratory infections”. *J Clin Virol* **70**, 128 (2015).
42. Seki, F., Ono, N., Yamaguchi, R. & Yanagi, Y. Efficient isolation of wild strains of canine distemper virus in Vero cells expressing canine SLAM (CD150) and their adaptability to marmoset B95a cells. *J Virol.* **77**, 9943–9950 (2003).
43. Juul, S. *et al.* What's in my pot? Real-time species identification on the MinION. *bioRxiv*, <https://doi.org/10.1101/030742> (2015).
44. Sović, I. *et al.* Fast and sensitive mapping of nanopore sequencing reads with GraphMap. *Nat Commun.* **7**, 11307 (2016).
45. Okonechnikov, K., Golosova, O. & Fursov, M. The UGENE team. Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics* **28**, 1166–1167, <https://doi.org/10.1093/bioinformatics/bts091> (2012).
46. Elia, G. *et al.* Detection of canine distemper virus in dogs by real-time RT-PCR. *J Virol Methods.* **136**, 171–176 (2006).
47. Mochizuki, M., Hashimoto, M., Hagiwara, S., Yoshida, Y. & Ishiguro, S. Genotypes of canine distemper virus determined by analysis of the hemagglutinin genes of recent isolates from dogs in Japan. *Journal of Clinical Microbiology* **32**, 2936–2942 (1999).
48. Huelsenbeck, J. P. & Ronquist, F. MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* **17**, 754–755 (2001).
49. Ronquist, F. & Huelsenbeck, J. P. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572–1574 (2003).
50. Posada, D. jModelTest: phylogenetic model averaging. *Mol Biol Evol.* **25**, 1253–1256 (2008).

## Acknowledgements

The authors acknowledge Dr. Dionisio Ianniruberto (Dipartimento Unico Regionale della Prevenzione, Campobasso-Italy). Funding were provided by the Italian Ministry of Health (MSRCTE06/17, Ricerca Corrente 2017 “Nuovi flussi diagnostici in Sanità animale: dalla NGS alla banca antigeni”, recipient Alessio Lorusso). Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the IZSAM.

## Author Contributions

C.C., N.D.A. and A.L. supervised the study, A.L. and A.P.<sup>1</sup> wrote the manuscript. A.P.<sup>1</sup>, M.M. and M.D.D. performed the NGS experiments. I.M. performed the bioinformatics analyses. F.P. and G.Z. performed virus isolation and Sanger sequencing. D.M. performed histology and immunohistochemistry. A.P.<sup>2</sup> and F.C. performed phylogeny. All authors read and approved the final version of the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-37497-4>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019