



Evaluating methods for Avian avulavirus-1 whole genome sequencing

Saar Tal^{a,1}, Meirav Ben Izhak^{b,1}, Chaim Wachtel^{c,1}, Anat Wiseman^{a,g,1}, Tzipi Braun^{b,1}, Elinor Yechezkel^a, Einav Golan^a, Ruth Hadas^d, Adi Turjeman^e, Caroline Banet-Noach^f, Michal Bronstein^e, Avishai Lublin^d, Elyakum Berman^a, Ziv Raviv^g, Michael Pirak^a, Eyal Klement^g, Yoram Louzoun^{h,*}

^a The Poultry Health Laboratories, The Egg and Poultry Board, Israel

^b Life Science faculty and Gonda Brain Research Center, Bar Ilan University, Israel

^c HyLabs NGS Services, Israel

^d Kimron Veterinary Institute, Division of Avian Diseases, Israel

^e The Center for Genomic Technologies, The Alexander Silberman Institute of Life Sciences, The Hebrew University of Jerusalem, Israel

^f Phibro Animal Health Corp. Ltd, Israel

^g Koret School of Veterinary Medicine, The Robert H. Smith Faculty of Agriculture, Food and Environment, Hebrew University of Jerusalem, Israel

^h Department of Mathematics and Gonda Brain Research Center, Bar Ilan University, Israel

ARTICLE INFO

Keywords:

Newcastle disease virus
AAvV-1
NGS
WGS
Mutations
Sequencing

ABSTRACT

Background: Avian avulavirus-1 (AAvV-1, previously Newcastle Disease Virus) is responsible for poultry and wild birds' disease outbreaks. Numerous whole genome sequencing methods were reported for this virus. These methods included cloning, specific primers amplification, shotgun PCR approaches, Sequence Independent Single Primer Amplification and next generation sequencing platform kits.

Methods: Three methods were used to sequence 173 Israeli Avian avulavirus-1 field isolates and one vaccine strain (VH). The sequencing was performed on Proton and Ion Torrent Personal Genome Machine and to a lesser extent, Illumina MiSeq and NextSeq sequencers. Target specific primers (SP) and Sequence Independent Single Primer Amplification (SISPA) products sequenced via the Ion torrent sequencer had a high error rate and truncated genomes. All the next generation sequencing platform sequencing kits generated high sequence accuracy and near-complete genomic size.

Results: A high level of mutations was observed in the intergenic regions between the avian avulavirus-1 genes. Within genes, multiple regions are more mutated than the Fusion region currently used for typing.

Conclusions: Our findings suggest that the whole genome sequencing by the Ion torrent sequencing kit is sufficient. However, when higher fidelity is desired, the Illumina NextSeq and Proton torrent sequencing kits were found to be preferable.

1. Introduction

Newcastle Disease Virus, currently named Avian avulavirus 1 (Amarasinghe et al., 2017) is the causative agent of Newcastle disease (ND). The disease affects over 200 species of birds including wild birds, pigeons and poultry Newcastle disease outbreaks can have a devastating effect on flocks with mortality approaching 100%. AAvV-1 was first described in 1926 in Newcastle-on-Tyne England (hence the name) and on the island of Java (Alexander, 2001).

AAvV-1 is a member of the order *Mononegavirales*, family *Paramyxoviridae*, genus *Avulaviruse* (Afonso et al., 2016; Amarasinghe et al., 2017; De Battisti et al., 2013). The virus has an approximately 15.2 kb ssRNA genome with varying lengths (Shi et al., 2011). The viral genome is composed of six genes expressing six proteins: nucleoprotein [NP], matrix, phosphoprotein [P], [M], fusion [F], hemagglutinin-neuraminidase [HN], and RNA polymerase [L]. AAvV-1 has only one serotype but has varying degrees of virulence (Gogoi et al., 2015). The F protein is cleaved between amino acid 111–117. The AAvV-1 virulence

Abbreviations: AAvV-1, Newcastle Disease virus; SP, Specific primers; NGS, Next Generation Sequencing; WGS, Whole Genome Sequence or sequencing; SISPA, Sequence Independent Single Primer Amplification; PGM, Personal Genome Machine; NCBI, National Center for Biotechnology Information; nt, nucleotide

* Corresponding author at: Math Department and Gonda Brain Research Center, Bar Ilan University, Ramat Gan 52900, Israel.

E-mail address: louzouy@math.biu.ac.il (Y. Louzoun).

¹ Equal contributors.

<https://doi.org/10.1016/j.gene.2019.100004>

Received 28 November 2018; Accepted 3 January 2019

Available online 09 January 2019

2590-1583/ © 2019 Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

is correlated with the frequency of basic amino acids in the F protein proteolytic cleavage site (Gogoi et al., 2015). Accordingly, AAVV-1 has five virulence categories or pathotypes. These virulence pathotypes are according to severity, starting with the highly virulent to the non-virulent: Velogenic neurotropic, Velogenic viscerotropic, Mesogenic, Lentogenic and asymptomatic (Ganar et al., 2014).

The AAVV-1 genome was previously sequenced using several methods, including: cloning (De Leeuw and Peeters, 1999), sequencing portions of the viral genome (Barbezange and Jestin, 2003, 2005), or sequencing the full genome using target specific primers (Absalón et al., 2014; Meng et al., 2012; Munir et al., 2011). Prior to whole genome sequencing kits, most publications used target specific primer amplification of the viral genome. This method utilizes AAVV-1-universal primers that can be applied to any AAVV-1 subtype as well as a rapid amplification of cDNA ends (RACE) to obtain the viral genome 5' and 3' end sequences (Chenchik et al., 1996; Li et al., 2005; Schaefer, 1995; Shi et al., 2011; Shi and Jarvis, 2006). Pyrosequencing and shotgun PCR methods were also successful in producing a whole genome sequence or parts of the whole genome (Campana et al., 2014; Ma et al., 2014). SISPA, a shotgun PCR approach, was also shown to successfully sequence the AAVV-1 genome (Allander et al., 2001; Chrzastek et al., 2017; Djikeng et al., 2008; Karlsson et al., 2013; Van Borm et al., 2013). The AAVV-1 whole genome was successfully sequenced using the Illumina NGS (Dimitrov et al., 2017; Forrester et al., 2013; Frey et al., 2014; Satharasinghe et al., 2016) as well as via the Ion Torrent sequencing kits (Jakhesara et al., 2016; Karlsson et al., 2013).

In this study we evaluated the performances of three different methods for AAVV-1 whole genome sequencing during AAVV-1 epizootic research. The three methods used were the target specific primer (SP) approach followed by SISPA and finally the whole genome sequencing platforms: Ion Torrent PGM and Proton Torrent. Few samples were sequenced using the Illumina NextSeq and MiSeq sequencing kits for comparison.

2. Materials and methods

2.1. Viral isolation

Five tracheal or cloacal swabs were pooled in a tube containing 2-3 ml PBS solution (Hylabs, Cat. BP 532/100S). The sample was incubated for 1 h, at room temperature. A 250 µl aliquot of the sample was added to a 2 ml Eppendorf containing 1.5 ml of PBS and 250 µl of X10 antibiotics solution for injection (HyLabs, Cat. TT359). The solution was incubated for 1 h at room temperature before injection to 9–11 days old embryonated chicken eggs (ECE's). A total of 200 µl of the prepared solution was injected per egg. Eggs were examined for seven days post injection using an egg lamp. Eggs containing dead embryos were opened and the allantoic fluid was collected and stored at -20 °C.

2.2. Viral RNA purification

AAVV-1 RNA purification was based on previously published methods (Dimitrov et al., 2014; Jakhesara et al., 2016; Hussain et al., 1989). Briefly, a total of 6 ml thawed allantoic fluid were added to 4 ml PBS. The samples were centrifuged in the Beckman Avanti J-E centrifuge at 10,000g for 15 min at 4 °C to remove tissues and cells. The clear supernatant was transferred into clear ultracentrifuge tubes (Beckman #344058) and was supplemented with 25 ml of PBS to a final volume of 35 ml. The samples were centrifuged in the Beckman Optima XE-90K using the swing rotor (SW 28) at 70,000g for 3 h at 4 °C. Following centrifugation, the supernatant was decanted and the pellet containing the virus was resuspended in 110 µl PBS (final volume 170 µl). The viral suspension was transferred to an Eppendorf tube containing 30 µl of DNase-RNase buffer. The DNase-RNase buffer was comprised of 10 units of DNaseI and 4.5 units of RNaseA suspended in x1 DNaseI buffer (NEB Cat. M0303L and Sigma Cat. R4875,

Table 1

AAVV-1 universal reverse transcription primers (AAVV-1uniRT) and amplification primers (AAVV-1final).

Final primers	Sequence	From	To	Product
AAVV-1uni RT_1	gactttgYgcaRgcgatgatg	2427	2461	
AAVV-1uni RT_2	caaatgcYtctccYcaggtDgcYaagatac	4207	4236	
AAVV-1uni RT_3	aaccaatgaRgctgtRcaYgaRgtcac	5027	5053	
AAVV-1uni RT_4	acYcgRccaggtagtrtccc	7757	7776	
AAVV-1uni RT_5	gcRgagcaYcagatYatcctacc	8409	8431	
AAVV-1uni RT_6	gYcagaagctRtggacRatgatctc	10,546	10,570	
AAVV-1uni RT_7	gaYggRtcacaccaRcttgc	12,975	12,994	
AAVV-1uni RT_8	tatgcMtgMgaggRgatgatg	14,223	14,243	
AAVV-1final_IF	gccatgactgcRtatgagac	656	675	3281
AAVV-1final_IR	tcYacStccacatYaataatgtac	3915	3937	
AAVV-1final_IIIF	agYctgtaYaactYgcRctcaatgtcac	3891	3947	3533
AAVV-1final_IIR	ccctccRtaRactgRgaacc	7353	7424	
AAVV-1final_IIIF	gcYtgYatgtaYtcaagac	5652	5671	2851
AAVV-1final_IIIR	agYggtagccccRgtYaattcc	8482	8503	
AAVV-1final_IVF	ctMtaYtactggaaatRacYgggctacc	8472	8500	3442
AAVV-1final_IVR	tctatattgtYggRagatgRaacc	11,890	11,914	
AAVV-1final_VF	aacacYgtaatgtcYtgtgc	10,905	10,924	3791
AAVV-1final_VR	ctgYttRagtgatgYctct	14,677	14,696	

respectively). The viral suspension was incubated at 37 °C for 2 h to decrease host DNA and RNA. RNA purification was performed using the Trizol LS kit (Thermo Fisher, Cat 10296028) according to the manufacturer's instructions. Viral RNA was stored at -80 °C.

2.3. Sequence specific primer amplification of AAVV-1 RNA

A total of 260 AAVV-1 full genome sequences from NCBI and one Israeli AAVV-1 vaccine (VH) sequence were used (sequence courtesy of Dr. Banet-Noach, Phibro Israel). The sequences were aligned using Clustal Omega v1.1.0 (Sievers et al., 2011). The resulting reverse transcription primers and PCR primers (see Table 1) were designed to amplify 3.5 kb overlapping amplicons of the AAVV-1 genome.

Optimized reverse transcription was performed using Agilent's Accuscript high fidelity Reverse Transcriptase (Cat. 600089) with the following modifications. A total of 100 ng RNA template was supplemented with 2 µl of accucript buffer (×10), 2 µl of 10 µM the reverse transcription primers mix, 0.8 µl of dNTP's 100 mM and DEPC treated DDW to a final volume of 10 µl. Pre-boil was performed at 65 °C for 15 min and the tubes were allowed to cool slowly to room temperature, according to the manufacturer's instructions. The pre-boiled mix was supplemented with 3 µl of DEPC DDW, 2 µl DTT, 3 µl of 20% DMSO solution, 1 µl of Betain 5 M solution and 1 µl of Accucript reverse transcriptase enzyme. The reverse transcription was performed at 37 °C for 2 h followed by a heat inactivation step at 70 °C for 15 min.

The Q5 high fidelity PCR mix (NEB, Cat. M0493) was optimized to efficiently amplify up to 6 kb of the viral genome. To this end, 2 µl of the reverse transcription reaction were used as template supplemented with 2 µl of 10 µM forward and reverse primers, 2 µl of 10 mM dNTP's, 4 µl of NEB Q5 buffer (×5), 0.4 µl Q5 enzyme, 3 µl of 20% DMSO solution, 3 µl of Betain 5 M solution and 3.6 µl DEPC DDW (20 µl final reaction volume). The PCR reaction was performed in the S1000 thermal cycler (Biorad). The reaction conditions were as follows: 95 °C for 3 min, 35 cycles of 95 °C - 20 s, 35 °C - 30 s, 72 °C - 45 s followed by a final step of 72 °C - 10 min. The PCR results were examined in a 1.5% agarose gel using Ethidium bromide prior to sequencing. The PCR reactions were purified using Stratec's PCR and enzymatic reaction purification kit (MSB PCRapace kit, Statec, Cat. 1,020,220,300) according to the manufacturer's instructions.

2.4. SISPA

The SISPA protocol was performed according to the literature with the modifications listed below (Djikeng et al., 2008). Briefly, 100 ng of

purified RNA were reverse transcribed using Agilent Accuscript reverse transcriptase using 1 μ l of 10 mM FR26RV-N primer (5' GCCGGAGCT CTGCAGATATC NNNNNN 3'), 2 μ l Agilent RT buffer, 0.8 μ l of 100 mM dNTP's and DEPC DDW were added to a final volume of 13 μ l. The pre-boil was performed at 65 °C for 15 min and the samples were allowed to cool to room temperature. The reverse transcription reaction supplemented contained 2 μ l DTT, 3 μ l DMSO 20%, 1 μ l Betain 5 M and 1 μ l Agilent accuscript reverse transcriptase, reaching a final reaction volume of 20 μ l. The reaction mix was incubated for 1 h at 37 °C followed by heat inactivation at 75 °C for 15 min. The reverse transcription reaction was supplemented with Klenow fragment reaction mix. The Klenow reaction contained 5 μ l NEB buffer 2 (X10), 1 μ l FR26RV-N primer, 1 μ l (5 units) of Klenow fragment (NEB Cat. M0210S), 1 μ l of 10 mM dNTP's and DEPC DDW to a mix volume of 30 μ l. The Klenow reaction was incubated at 37 °C for 1 h followed by heat inactivation at 70 °C for 10 min. The reaction was purified using Strattec's PCR and enzymatic reaction purification kit, as before, having a final elution volume of 20 μ l.

The PCR reaction was performed using the Q5 high fidelity polymerase kit, as listed above, using 2 μ l of the RT-Klenow purified reaction as template and 2 μ l of 10 mM primer FR26RV (5' GCCGGAGCTC TGCAGATATC 3'), 2 μ l dNTP's, 4 μ l Q5 buffer X10, 0.4 μ l Q5 enzyme, 3 μ l Betain 5 M, 3 μ l DMSO 20% and 3.6 DEPC DDW to a final reaction volume of 20 μ l. The reaction conditions were as follows: 95 °C-5 min, 40 cycles of 95 °C-30 s, 55 °C-35 s, 72 °C-7 s and a final elongation step of 72 °C-5 min. The PCR results were examined in a 1.5% agarose gel using Ethidium bromide prior to sequencing.

2.5. Sequencing platforms

The Specific Primers amplified DNA, SISPA amplified DNA and purified RNA samples were sent for sequencing using the Ion Torrent, Ion Proton (HyLabs) and Illumina NextSeq and MiSeq platforms (The Center for Genomic Technologies, the Hebrew university of Jerusalem).

2.5.1. SISPA and SP

The amplified samples were fragmented by sonication using a Bioruptor sonicator (Diagenode). The fragmented DNA was then used to prepare barcoded libraries using the The Ion Xpress™ DNA library prep kit (Life Technologies) according to manufacturer's instructions. Sequencing was performed using the Ion Torrent PGM (Bar-Ilan university).

2.5.2. Ion torrent PGM sequencing kit

RNA libraries were prepared from purified viral RNA using the Ion res™ RNA library prep kit (Life Technologies) and the resulting bar-coded DNA libraries were then sequenced on the Ion Torrent Proton according to the manufacturer's instructions. Sequencing was performed on the Ion and Proton Torrent PGM (HyLabs).

2.5.3. Illumina sequencing

Samples were sequenced using the Illumina MiSeq and NextSeq. The cDNA synthesis, sample denaturation, loading and Illumina kits were used according to the manufacturer's instructions. The samples were pair ended (251 \times 2 and 38 \times 2 for MiSeq and NextSeq, respectively) using the Illumina TruSeq sample preparation kit. The samples were sequenced either using the MiSeq reagent kit V2 (500 cycles) or the NextSeq 500 high output kit V2 (75 cycle).

2.6. Genome assembly

As a part of the quality control, nucleotides from the edges were trimmed and reads with poor quality scores were removed. Short reads were discarded. The remaining reads were used to de novo reconstruct the consensus sequence using Trinity v2.0.6 (Grabherr et al., 2011). Reads were mapped to the viral consensus sequence using bowtie2

v2.2.4 (Langmead and Salzberg, 2012). The resulting sam files were used to determine the quality of the sequence. Positions that had coverage of less than 5 reads were considered undetermined and written as N.

2.7. Accession numbers

The sequences of the AAVV-1 isolates included in the study were submitted to GenBank under the accession numbers MH371022-MH371102, MH377246-MH377325. Accession numbers are detailed in Supplementary Table S1.

3. Results

A total of 174 AAVV-1 samples were sequenced (See Appendix A and Supplementary Information). Of these, 173 were field isolates (10 wild bird; 31 backyard birds; 132 commercial poultry) and the Israeli AAVV-1 vaccine (VH) were sequenced using different platforms. The viruses were sequenced using three different methods including the specific primers (34), SISPA (57) and 96 samples via NGS sequencing kits (8 IonTorrent, 84 Proton Torrent, 2 Illumina MiSeq and 2 Illumina NextSeq). Several samples were sequenced using multiple methods. Except for the VH, all these methods utilized enriched viral RNA that was purified from the allantoic fluid (Jakhesara et al., 2016). The VH genome was purified from a commercial vaccination tablet.

For the template-specific primers method, AAVV-1 universal primers were designed using the entire Pubmed database AAVV-1 sequences. Eight forward AAVV-1 universal primers were used for the reverse transcription reaction step and five universal primer pairs amplified 3.5 kb overlapping segments of the 15.2 kb viral cDNA. The resulting PCR products were purified prior to library preparation and sequenced via the Ion Torrent PGM platform kits followed by genome assembly. Preliminary tests using the lentogenic VH vaccine RNA as a template generated a contig of 14,000 nt's. The resulting VH contig covered ~94% of the viral genome (680–14,680 nt's). Following this test, 34 AAVV-1 vologenic field isolates were sequenced using this method. However, in all the field isolates the first amplicon (656–3937) was missing, resulting in a partial contig covering 72% of the viral genome (3890 nt–14,680 nt, see Fig. 1). The first amplicon PCR failure presumably occurred due to a poly-C region in the AAVV-1 genome located in a ~100 bp region near 1700 nt (1656–1761). The vologenic AAVV-1 field isolates have a poly-C region containing 3–4 repeats of 5–6 tandem Poly-Cytosine while the VH sequence had a smaller poly-C region containing only 3 repeats of 3–4 tandem repeats (data not shown). This region has a highly consistent drop in viral genome coverage in all the vologenic samples sequenced, regardless of the amplification method or NGS platform used (see Figs. 1 and 2). This is likely due to G-quadruplex (Chambers et al., 2015).

The SISPA method was developed to sequence novel RNA viruses (Froussard, 1992). It has been shown to be efficient and to successfully sequence the AAVV-1 genome as well as other RNA viruses (Allander et al., 2001; Djikeng et al., 2008; Djikeng and Spiro, 2009). The standard SISPA primers were examined in silico to verify no match to the AAVV-1 known sequences, thus avoiding technical issues (Karlsson et al., 2013). The SISPA method resulted in a contig coverage of over 99% of the viral genome. Additional 56 AAVV-1 isolates were sequenced using the SISPA method via the Ion Torrent sequencer. This resulted in an average coverage of 98–99% of the viral genomes (See Figs. 1 and 2).

The Ion Torrent, Illumina MiSeq and NextSeq RNA library preparation and sequencing kits were also examined using the VH vaccine and Israeli isolates. Purified VH RNA sequenced using the Ion Torrent sequencing kit resulted in 99.98% of the viral genome with the small exception of 5–7 mismatches compared to the known sequence. These mismatches were found only at the ends of the viral genome (at the sense 5' end). The VH RNA was successfully sequenced using Illumina

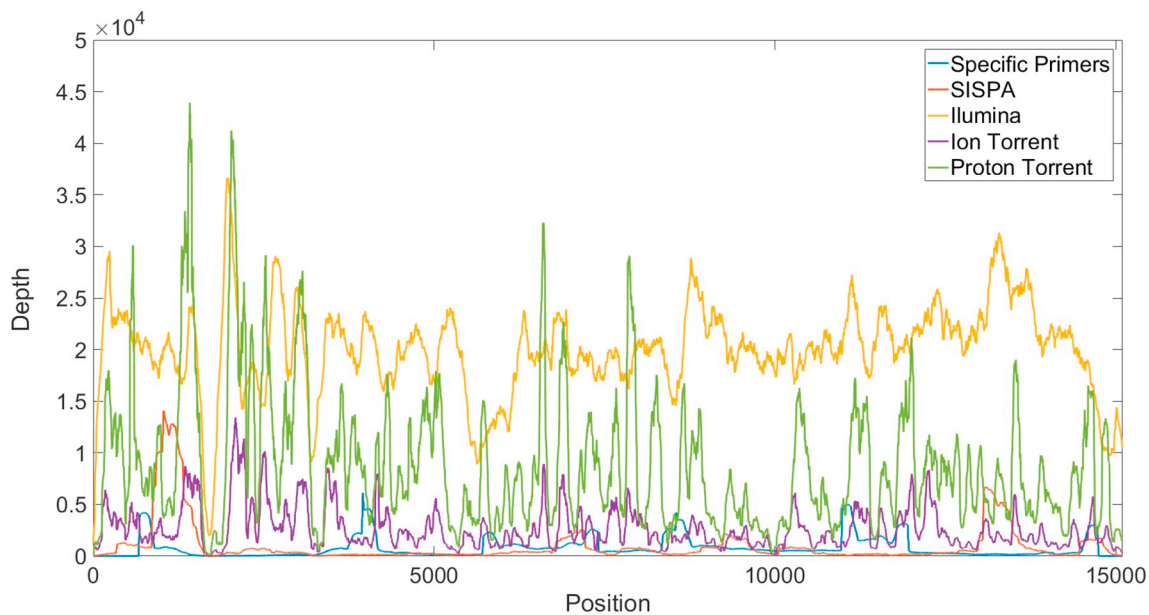


Fig. 1. Average coverage of each position in the AAV-1 viral genome sequences, per method. The viruses were sequenced using three different methods including the specific primers (34), SISPA (57) and 96 samples via NGS sequencing kits (8 IonTorrent, 84 Proton Torrent, 4 Illumina). In all the methods a drop in the coverage near nucleotide 1700 is seen, possibly due to the G-quadruplexes. The Illumina and Proton Torrent have the highest coverage. The lowest coverage was obtained by the specific primers and SISPA methods (sequenced via the Ion Torrent).

NextSeq reaching an average depth of 80,000 reads/nt, and a perfect 100% match to the known VH sequence. A total of 6 AAV-1 field isolates were sequenced using the Ion Torrent sequencing kit followed by a second set of 82 isolates sequenced using the ProtonTorrent sequencing kit with a final average cover of > 99% of the viral genome (see Figs. 1,2).

All three sequencing platforms and their kits (IonTorrent, ProtonTorrent and Illumina) had similar nucleotide profile with few N reads or gaps. The SP and SISPA methods had higher N reads compared to the NGS sequencer kit average (8.9 and 10.3 respectively) (see Fig. 3).

Given the high sequencing accuracy and coverage of the NGS

methods, we addressed the mutational load at each site within a host (sample). The high coverage allows us to detect not only the main consensus sequence, but also the quasi-species of the virus genomic population. A high mutation load was seen in SISPA and the specific primers sequences all along the viral genome, suggesting that these observed mutations are mainly sequencing errors. However, in the NGS based methods, mutations are limited to specific regions. The high mutation and deletion rate is observed in all sequencing methods near the 3' and 5' ends of the viral genome is possibly due to the folding of the viral genome ends leading to misreading. The mutation hotspots in the viral genome mutations are mainly between the genes, regardless of the method or NGS sequencer used (see Fig. 4a and b). Note that the

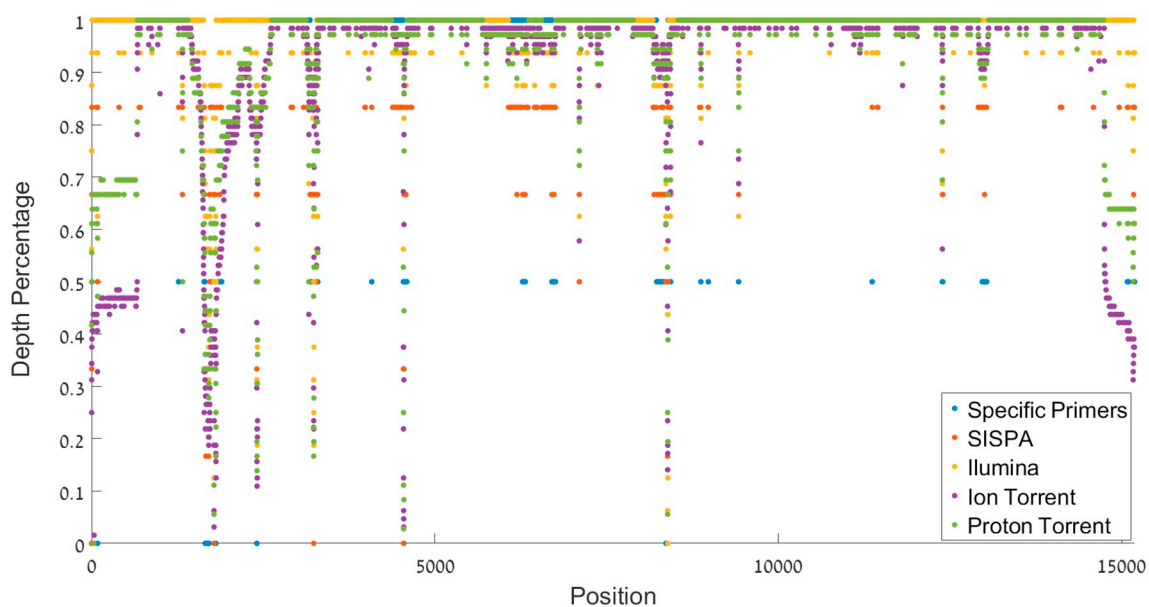


Fig. 2. Fraction of covered samples in each position along the sequence per method as measured by the fraction of samples that have at least five reads for a given position (otherwise, the position is assigned an N). The coverage drop at position 1,700 is clearly observed, as well as the coverage drop in the 5' and 3' ends of the viral genome. Intra-genic regions also have a high error rate.

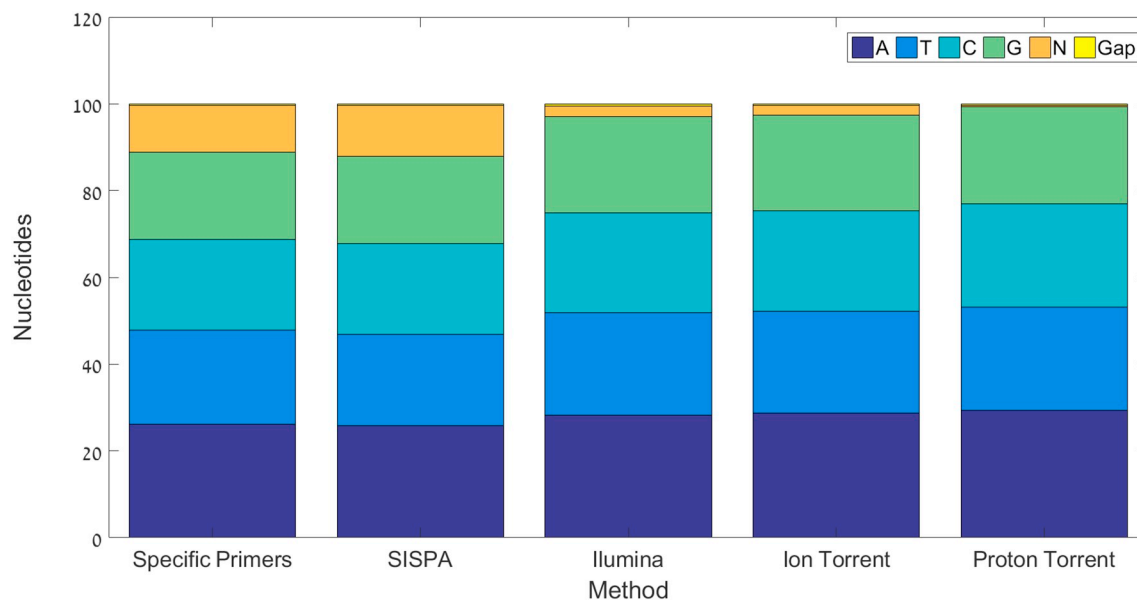


Fig. 3. The mean nucleotide percentage, N reads and average gap per sequencing method. The graph indicates a similarity in the nucleotide consistency between the sequencing methods. The SP method has an 8.9 and SISPA has a 10.3 times higher N reads compared to the NGS sequencer kit average. The Proton torrent had the least N reads, approximately 10 times lower than the other NGS sequencing kits. All the sequencing kits (Illumina, IonTorrent and Proton Torrent) had the lowest N frequency.

average mutation rate with a sample in the F gene typically used to classify the virus into sub-types is lower than in most other genes.

To further estimate the variability, we computed the fraction of NS mutations. While, there is no practically no difference between sequencing methods, as expected intergenic regions have significantly higher NS fractions compared with genic regions (Data not shown).

4. Discussion

Brute force NGS sequencing approach (i.e. total DNA or RNA purification and sequencing) has its own advantages (Li et al., 2016). It is mostly used for sequencing a small number of samples at low coverage. However, for sequencing of tens to hundreds of samples it is more cost-effective to remove the host genomic DNA or RNA, as much as possible (Dimitrov et al., 2017; Thomson et al., 2016; Tyler et al., 2016). This can be done using the physical properties of the virion (Jakhesara et al., 2016; Hussain et al., 1989) or by using purification kits post DNA or RNA purifications (Dimitrov et al., 2014, 2017). The latter involves a considerable loss of target genomic material due to the purification steps losses. During this study we opted for the former approach for viral RNA purification (Temmam et al., 2015). This method resulted in 90–95% viral RNA sequences (data not shown). The low host DNA and RNA enabled loading the NGS platforms with a large pool of tagged cDNA samples while still obtaining high coverage. For the Ion Torrent and Proton Torrent up to 12 and 16 samples were loaded per sequencing run, respectively.

Several methods and sequencing platforms were reported to amplify the AAVV-1 genome. The specific primer approach was mostly used for AAVV-1 whole genome sequencing (Absalón et al., 2014; Meng et al., 2012; Munir et al., 2011; Wei et al., 2008). However, the partial sequence obtained, the need for the RACE method to amplify the viral genome ends, the sequence inaccuracies and the amplification termination near position 1700 nt's of the field isolate genomes all suggest that such a method is not appropriate for whole genome sequencing of AAVV-1 (Chenchik et al., 1996; Li et al., 2005; Liu and Gorovskiy, 1994; Schaefer, 1995; Shi and Jarvis, 2006).

The SISPA method was also shown to reliably sequence the AAVV-1 genome (Allander et al., 2001; Chrzastek et al., 2017; Djikeng et al., 2008; Karlsson et al., 2013). The sequencing of other viruses using this

method has failed, possibly due to technical issues (Karlsson et al., 2013). In our hands, SISPA successfully sequenced the AAVV-1 genome via the Ion Torrent PGM. Similarly to the specific primers method, it was found to have high error rates resulting in a considerably higher mutation rate compared to the sequencing kits of the NGS platforms. The SISPA was usually unable to sequence 100-200 bp from the 5' and 3' ends of the viral genome. This problem could possibly be resolved using a higher coverage, at the cost of running fewer samples per chip.

Our findings concur with the reports of PCR amplification bias in NGS sequencing which were recently reported in several articles (Chrzastek et al., 2017; Dimitrov et al., 2017; Jones et al., 2015; Pinard et al., 2006; Thomson et al., 2016; Tyler et al., 2016; Van Borm et al., 2016). While PCR amplified samples can be used for sequencing low titer samples, the bias involved should be considered and avoided when possible (Aigrain et al., 2016; Chrzastek et al., 2017; Dimitrov et al., 2017; Huggett et al., 2015; Thomson et al., 2016; Tyler et al., 2016; Van Borm et al., 2016). In our hands, the Illumina NextSeq, ProtonTorrent and Ion Torrent sequencing kits generated nearly full viral genome (98–99% for the Ion and proton torrent systems and 100% for the Illumina NextSeq) with few mismatches at the 3' and 5' ends. From our experience, it is advisable to opt for a high coverage sequence (a minimum coverage of 5000, preferably 50,000) to cover the 1700 nt position (the poly-C region) as well as the 7000 nt region. We therefore recommend using the Proton Torrent or the Illumina Nextseq sequencing kits for whole genome sequencing of AAVV-1 isolates.

Given the high fidelity of the NGS based sequencing, it is possible to study beyond the main variant. The quasi-species cloud (surrounding each main variant) is highly mutated in the intergenic regions, and in the 3' and 5'. We have found a mutation rate of 30 mutations/year in the viral genomes examined (data not shown). Such mutations are expected, since their effect on the viral fitness may be limited. Considering the viral reverse transcriptase mutation rate, we expected a higher mutation in the viral genome (Menéndez-Arias, 2002; Menéndez-Arias et al., 2017). The AAVV-1 found mutation rate suggests there is a natural selection process. This process is possibly affecting the expected mutation rate and limiting the regions that can be affected (Fan et al., 2017; Miller et al., 2009).

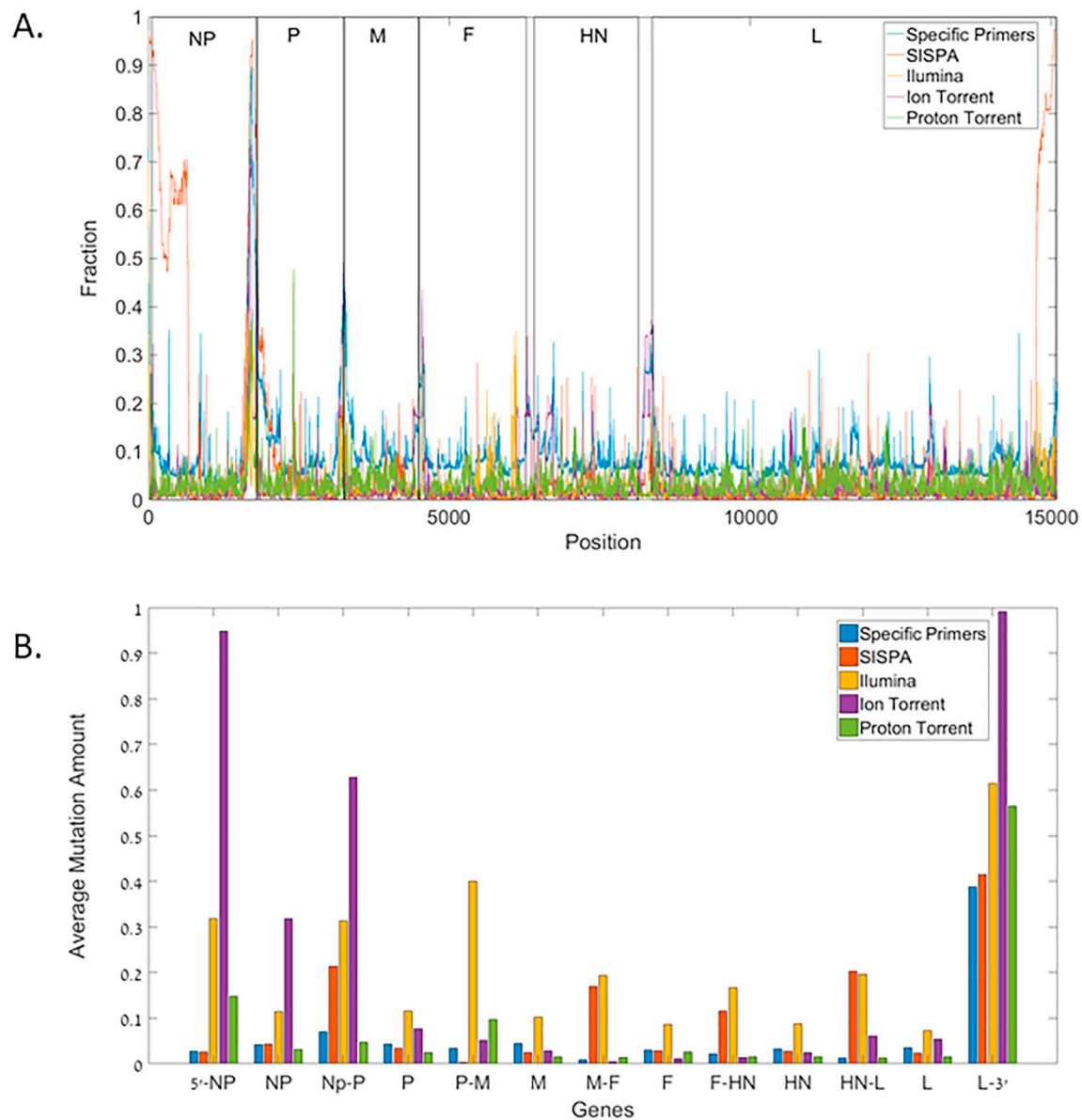


Fig. 4. A: Mutation frequency per position in the sequence. The rectangles represent the gene locations along the sequence. The SISPA and Specific primer methods were unable to reproduce the sequence near the 3' and 5' ends (Qiu et al., 2014). 4b. The average mutation fraction in each segment of the sequence for each method. It is notable that the segments between the genes, and in the 5' and 3' ends are the ones with higher percentage of mutations.

5. Conclusions

We have evaluated different methods of AAV-1 full genome sequencing. The main conclusions are a clear recommendation to purify the viral genome prior to sequencing. PCR amplifications of the viral genomes, either using shotgun PCR (SISPA in this case) or via target specific primers have generated poor results compared to NGS platform kits. Sequencing kits of the Illumina NextSeq and Proton torrent PGM have successfully resulted in 99–100% of the viral genome at a high coverage. The high coverage obtained via NGS methods shows a map of mutations in the AAV-1 genome which are mainly in intergenic regions.

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.gene.2019.100004>.

Availability of data and material

The datasets used and/or analysed during the current study are

available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the Egg and Poultry Board [Grant No. 22-2222-14].

Authors' contributions

Tal Saar, Ben Izhak Meirav, Wachtel Chaim and Wiseman Anat and Broun Tzipi have made substantial contribution to the experimental design and acquisition of data.

Yechezkel Elinor, Golan Einav, Hadash Ruth, Adi Turjeman, Banet Noach Caroline, Michal Bronstein and Avishai Lublin have all have

made substantial contributions to acquisition of data, methodology and resources.

Berman Elyakum, Raviv Ziv, Pirak Michael, Klement Eyal and Louzoun Yoram have made substantial contributions to analysis and interpretation of data, as well as funding acquisition.

Tal Saar, Wiseman Anat and Louzoun Yoram have been involved in drafting the original manuscript draft as well as reviewing and editing the final draft.

Pirak Michael, Klement Eyal, Raviv Ziv and Louzoun Yoram have been involved in Project administration.

Appendix A

Table of AAVV-1 isolates and their origin.

Sample number	Source	Specific primers	SISPA	Platform sequencing kits			
				Ion torrent	Proton torrent	MiSeq	NextSeq
1818	EPB	v					
11278	EPB	v					
120420	EPB	v					
173849	EPB	v					
176674	EPB	v					
120147	EPB	v					
120430	EPB	v					
195810	EPB	v		v		v	v
Lasota	Phibro	v		v		v	v
120809	EPB	v					
124026	EPB	v					
141353	EPB	v					
120860	EPB	v					
120654	EPB	v					
120423	EPB	v					
124027	EPB	v		v			
123498	EPB	v					
117699	EPB	v					
120807	EPB	v					
137495	EPB	v					
123526	EPB	v					
121808	EPB	v					
104477	EPB	v		v			
138692	EPB	v					
137592	EPB	v					
122509	EPB	v					
137569	EPB	v					
99188	EPB	v					
114599	EPB	v					
121118	EPB	v					
125877	EPB	v					
118279	EPB	v					
115038	EPB	v					
114112	EPB	v					
141759	EPB		v				
142551	EPB		v				
142626	EPB		v				
142784	EPB		v				
143114	EPB		v				
143696	EPB		v				
144702	EPB		v				
147004	EPB		v				
149165	EPB		v				
152628	EPB		v				
153812	EPB		v				
156916	EPB		v				
159057	EPB		v				
166631	EPB		v				
168757	EPB		v				
169730	EPB		v				
170338	EPB		v				
173294	EPB		v	v			
175571	EPB		v				
176851	EPB		v				
201029	EPB		v				
203145	EPB		v				
203308	EPB		v	v			
203762	EPB		v				
205659	EPB		v				
205770	EPB		v				
206519	EPB		v				
207436	EPB		v				
209108	EPB		v				

217142	EPB	v	
218637	EPB	v	
223030	EPB	v	
223068	EPB	v	
223759	EPB	v	
123168	EPB	v	
123945	EPB	v	
125057	EPB	v	
127508	EPB	v	
127935	EPB	v	
129688	EPB	v	
131986	EPB	v	
133813	EPB	v	
137866	EPB	v	
138629	EPB	v	
139013	EPB	v	
139141	EPB	v	
140266	EPB	v	
140339	EPB	v	
147993	EPB	v	
151754	EPB	v	
208243	EPB	v	
224349	EPB	v	
224903	EPB	v	
225193	EPB	v	
226419	EPB	v	
226501	EPB	v	
28469	EPB		v
30507	EPB		v
34617	EPB		v
36911	EPB		v
37114	EPB		v
40978	EPB		v
76558	EPB		v
87809	EPB		v
88629	EPB		v
95252	EPB		v
95493	EPB		v
97202	EPB		v
116035	EPB	v	
117126	EPB	v	
117244	EPB		v
119228	EPB		v
121082	EPB		v
121551	EPB		v
125051	EPB		v
141932	EPB		v
143483	EPB		v
145002	EPB		v
149777	EPB		v
150791	EPB		v
154728	EPB		v
156566	EPB		v
167788	EPB		v
169208	EPB		v
173633	EPB		v
204094	EPB		v
2041	EPB		v
5620	EPB		v
5869	EPB		v
9273	EPB		v
11365	EPB		v
11543	EPB		v
23536	EPB		v
24317	EPB		v
24639	EPB		v
25318	EPB		v
27571	EPB		v
17222	EPB		v
238796	EPB		v
24153	EPB		v
24993	EPB		v
151347	EPB		v
147406	EPB		v
223397	EPB		v
232809	EPB		v
238651	EPB		v
237660	EPB		v
204642	EPB		v
208252	EPB		v

204895	EPB	v
174621	EPB	v
176533	EPB	v
251453	EPB	v
251612	EPB	v
251747	EPB	v
253817	EPB	v
253685	EPB	v
257758	EPB	v
264746/1	KVI	v
264747/1	KVI	v
264748/1	KVI	v
264750/1	KVI	v
264751/1	KVI	v
264752/1	KVI	v
264753/1	KVI	v
264754/1	KVI	v
264755/1	KVI	v
264756/1	KVI	v
264757/1	KVI	v
264758/1	KVI	v
246332	EPB	v
249275	EPB	v
251542	EPB	v
255913	EPB	v
264220	EPB	v
265941	EPB	v
265559	EPB	v
266018	EPB	v
256987	EPB	v
253633	EPB	v
245632	EPB	v
264759	EPB	v
265752	EPB	v

Note: Specific primers and SISPA samples were sequenced using the IonTorrent. Acronyms: EPB - egg and poultry board. KVI - Kimron Veterinary institute.

References

- Absalón, A.E., Mariano-Matías, A., García, L.J., Morales-Garzón, A., Toscano-Contreras, A., Lucio-Decanini, E., Cortés-Espinosa, D.V., 2014. Complete genome analysis of velogenic Newcastle disease virus reference strain "chimalhuacan": evolution of viral lineages in Mexico. *Virus Genes* 49, 233–236. <https://doi.org/10.1007/s11262-014-1082-8>.
- Afonso, C.L., Amarasinghe, G.K., Bányai, K., Bào, Y., Basler, C.F., Bavari, S., Bejerman, N., Blasdel, K.R., Briand, F.-X., Briese, T., Bukreyev, A., Calisher, C.H., Chandran, K., Chéng, J., Clawson, A.N., Collins, P.L., Dietzgen, R.G., Dolnik, O., Domier, L.L., Dürrwald, R., Dye, J.M., Easton, A.J., Ebihara, H., Farkas, S.L., Freitas-Astúa, J., Formenty, P., Fouchier, R.A.M., Fù, Y., Ghedin, E., Goodin, M.M., Hewson, R., Horie, M., Hyndman, T.H., Jiāng, D., Kitajima, E.W., Kobinger, G.P., Kondo, H., Kurath, G., Lamb, R.A., Lenardon, S., Leroy, E.M., Li, C.-X., Lin, X.-D., Liú, L., Longdon, B., Marton, S., Maisner, A., Mühlberger, E., Netesov, S.V., Nowotny, N., Patterson, J.L., Payne, S.L., Paweska, J.T., Randall, R.E., Rima, B.K., Rota, P., Rubbenstroth, D., Schwemmler, M., Shi, M., Smither, S.J., Stenglein, M.D., Stone, D.M., Takada, A., Terregino, C., Tesh, R.B., Tian, J.-H., Tomonaga, K., Tordo, N., Towner, J.S., Vasilakis, N., Verbeek, M., Volchkov, V.E., Wahl-Jensen, V., Walsh, J.A., Walker, P.J., Wang, D., Wang, L.-F., Wetzl, T., Whitfield, A.E., Xiè, J.T., Yuen, K.-Y., Zhang, Y.-Z., Kuhn, J.H., 2016. Taxonomy of the order *Mononegavirales*: update 2016. *Arch. Virol.* 161, 2351–2360. <https://doi.org/10.1007/s00705-016-2880-1>.
- Aigrain, L., Gu, Y., Quail, M.A., 2016. Quantitation of next generation sequencing library preparation protocol efficiencies using droplet digital PCR assays - a systematic comparison of DNA library preparation kits for Illumina sequencing. *BMC Genomics* 17, 458. <https://doi.org/10.1186/s12864-016-2757-4>.
- Alexander, D.J., 2001. Gordon memorial lecture. Newcastle disease. *Br. Poult. Sci.* 42, 5–22. <https://doi.org/10.1080/713655022>.
- Allander, T., Emerson, S.U., Engle, R.E., Purcell, R.H., Bukh, J., 2001. A virus discovery method incorporating DNase treatment and its application to the identification of two bovine parvovirus species. *Proc. Natl. Acad. Sci. U. S. A.* 98, 11609–11614. <https://doi.org/10.1073/pnas.211424698>.
- Amarasinghe, G.K., Bào, Y., Basler, C.F., Bavari, S., Beer, M., Bejerman, N., Blasdel, K.R., Bochnowski, A., Briese, T., Bukreyev, A., Calisher, C.H., Chandran, K., Collins, P.L., Dietzgen, R.G., Dolnik, O., Dürrwald, R., Dye, J.M., Easton, A.J., Ebihara, H., Fang, Q., Formenty, P., Fouchier, R.A.M., Ghedin, E., Harding, R.M., Hewson, R., Higgins, C.M., Hong, J., Horie, M., James, A.P., Jiāng, D., Kobinger, G.P., Kondo, H., Kurath, G., Lamb, R.A., Lee, B., Leroy, E.M., Li, M., Maisner, A., Mühlberger, E., Netesov, S.V., Nowotny, N., Patterson, J.L., Payne, S.L., Paweska, J.T., Pearson, M.N., Randall, R.E., Revill, P.A., Rima, B.K., Rota, P., Rubbenstroth, D., Schwemmler, M., Smither, S.J., Song, Q., Stone, D.M., Takada, A., Terregino, C., Tesh, R.B., Tomonaga, K., Tordo, N., Towner, J.S., Vasilakis, N., Volchkov, V.E., Wahl-Jensen, V., Walker, P.J., Wang, B., Wang, D., Wang, F., Wang, L.F., Werren, J.H., Whitfield, A.E., Yan, Z., Ye, G., Kuhn, J.H., 2017. Taxonomy of the order *Mononegavirales*: update 2017. *Arch. Virol.* 162, 2493–2504. <https://doi.org/10.1007/s00705-017-3311-7>.
- Barbezange, C., Jestin, V., 2003. Molecular characterisation of three avian paramyxovirus type 1 isolated from pigeons in France. *Virus Genes* 26, 175–183. <https://doi.org/10.1023/A:1023439530750>.
- Barbezange, C., Jestin, V., 2005. Quasispecies nature of an unusual avian paramyxovirus type-1 isolated from pigeons. *Virus Genes* 30, 363–370. <https://doi.org/10.1007/s11262-004-6780-1>.
- Campana, M.G., Robles Garcia, N., Ruhli, F.J., Tuross, N., 2014. False positives complicate ancient pathogen identifications using high-throughput shotgun sequencing. *BMC Res. Notes* 7, 111. <https://doi.org/10.1186/1756-0500-7-111>.
- Chambers, V.S., Marsico, G., Boutell, J.M., Di Antonio, M., Smith, G.P., Balasubramanian, S., 2015. High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nat. Biotechnol.* <https://doi.org/10.1038/nbt.3295>.
- Chenchi, A., Diachenko, L., Moqadam, F., Tarabykin, Y., Lukyanov, S., Siebert, P.D., 1996. Full-length cDNA cloning and determination of mRNA 5' and 3' ends by amplification of adaptor-ligated cDNA. *BioTechniques* 21, 526–534.
- Chrzastek, K., Lee, D. hun, Smith, D., Sharma, P., Suarez, D.L., Pantin-Jackwood, M., Kapczynski, D.R., 2017. Use of sequence-independent, single-primer-amplification (SISPA) for rapid detection, identification, and characterization of avian RNA viruses. *Virology* 509, 159–166. <https://doi.org/10.1016/j.virol.2017.06.019>.
- De Battisti, C., Salomoni, A., Ormelli, S., Monne, I., Capua, I., Cattoli, G., 2013. Rapid pathotyping of Newcastle disease virus by pyrosequencing. *J. Virol. Methods* 188, 13–20. <https://doi.org/10.1016/j.jviromet.2012.11.021>.
- De Leeuw, O., Peeters, B., 1999. Complete nucleotide sequence of Newcastle disease virus: evidence for the existence of a new genus within the subfamily *Paramyxovirinae*. *J. Gen. Virol.* 131–136.
- Dimitrov, K., Clavijo, A., Sneed, L., 2014. RNA extraction for molecular detection of Newcastle disease virus—comparative study of three methods. *Rev. Med. Vet.* 165, 172–175.
- Dimitrov, K.M., Sharma, P., Volkening, J.D., Goraichuk, I.V., Wajid, A., Rehmani, S.F., Basharat, A., Shittu, I., Joannis, T.M., Miller, P.J., Afonso, C.L., 2017. A robust and cost-effective approach to sequence and analyze complete genomes of small RNA viruses. *Virology* 509, 159–166. <https://doi.org/10.1016/j.virol.2017.06.019>.
- Djikeng, A., Halpin, R., Kuzmickas, R., Depasse, J., Feldblyum, J., Sengamalay, N., Afonso, C., Zhang, X., Anderson, N.G., Ghedin, E., Spiro, D.J., 2008. Viral genome sequencing by random priming methods. *BMC Genomics* 9, 5. <https://doi.org/10.1186/1471-2164-9-5>.
- Djikeng, A., Spiro, D.J., 2009. Advancing full length genome sequencing for human RNA viral pathogens. *Futur. Virol.* 4, 47–53. <https://doi.org/10.2217/17460794.4.1.47>.
- Fan, W., Wang, Y., Wang, S., Cheng, Z., Guo, H., Zhao, X., Liu, J., 2017. Virulence in

- Newcastle disease virus: a genotyping and molecular evolution spectrum perspective. Res. Vet. Sci. 111, 49–54. <https://doi.org/10.1016/j.rvsc.2016.12.001>.
- Forrester, N.L., Widen, S.G., Wood, T.G., Travassos Da Rosa, A.P., Ksiązek, T.G., Vasilakis, N., Tesh, R.B., 2013. Identification of a new Newcastle disease virus isolate from Indonesia represents an ancestral lineage of class II genotype XIII. Virus Genes 47, 168–172. <https://doi.org/10.1007/s11262-013-0900-8>.
- Frey, K.G., Herrera-Galeano, J., Redden, C.L., Luu, T.V., Servetas, S.L., Mateczun, A.J., Mokashi, V.P., Bishop-Lilly, K.A., 2014. Comparison of three next-generation sequencing platforms for metagenomic sequencing and identification of pathogens in blood. BMC Genomics 15, 96. <https://doi.org/10.1186/1471-2164-15-96>.
- Froussard, P., 1992. A random-PCR method (rPCR) to construct whole cDNA library from low amounts of RNA. Nucleic Acids Res. 20, 2900. <https://doi.org/10.1093/nar/20.11.2900>.
- Ganar, K., Das, M., Sinha, S., Kumar, S., 2014. Newcastle disease virus: current status and our understanding. Virus Res. 184, 71–81. <https://doi.org/10.1016/j.virusres.2014.02.016>.
- Gogoi, P., Ganar, K., Kumar, S., 2015. Avian paramyxovirus: a brief review. Transbound. Emerg. Dis. 1–15. <https://doi.org/10.1111/tbed.12355>.
- Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Muceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29, 644–652. <https://doi.org/10.1038/nbt.1883>.
- Huggett, J.F., Cowen, S., Foy, C.A., 2015. Considerations for digital PCR as an accurate molecular diagnostic tool. Clin. Chem. <https://doi.org/10.1373/clinchem.2014.221366>.
- Hussain, Zahid, M. K., S. S.T.A., 1989. Purification of NDV.pdf.
- Jakhesara, S.J., Prasad, V.V.S.P., Pal, J.K., Jhala, M.K., Prajapati, K.S., Joshi, C.G., 2016. Pathotypic and sequence characterization of Newcastle disease viruses from vaccinated chickens reveals circulation of genotype II, IV and XIII and in India. Transbound. Emerg. Dis. 63, 523–539. <https://doi.org/10.1111/tbed.12294>.
- Jones, M.B., Highlander, S.K., Anderson, E.L., Li, W., Dayrit, M., Klitgord, N., Fabani, M.M., Seguritan, V., Green, J., Pride, D.T., Yooseph, S., Biggs, W., Nelson, K.E., Venter, J.C., 2015. Library preparation methodology can influence genomic and functional predictions in human microbiome research. Proc. Natl. Acad. Sci. U. S. A. 112 <https://doi.org/10.1073/pnas.1519288112>. 1519288112.
- Karlsson, O.E., Belák, S., Granberg, F., 2013. The effect of preprocessing by sequence-independent, single-primer amplification (SISPA) on metagenomic detection of viruses. Biosecur. Bioterror. 11 (Suppl. 1), S227–S234. <https://doi.org/10.1089/bsp.2013.0008>.
- Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with bowtie 2. Nat. Methods 9, 357–359. <https://doi.org/10.1038/nmeth.1923>.
- Li, D., Li, Z., Zhou, Z., Li, Z., Qu, X., Xu, P., Zhou, P., Bo, X., Ni, M., 2016. Direct next-generation sequencing of virus-human mixed samples without pretreatment is favorable to recover virus genome. Biol. Direct 11, 3. <https://doi.org/10.1186/s13062-016-0105-x>.
- Li, Z., Yu, M., Zhang, H., Wang, H.Y., Wang, L.F., 2005. Improved rapid amplification of cDNA ends (RACE) for mapping both the 5' and 3' terminal sequences of paramyxovirus genomes. J. Virol. Methods 130, 154–156. <https://doi.org/10.1016/j.jviromet.2005.06.022>.
- Liu, X., Gorovsky, M.A., 1994. Mapping the 5' and the 3' ends of *Tetrahymena thermophila* mRNAs using RNA ligase mediated amplification of cDNA ends (RLM-RACE). Nucleic Acids Res. 22, 551. <https://doi.org/10.1093/nar/22.3.551>.
- Ma, J., Prince, A., Aagaard, K.M., 2014. Use of whole genome shotgun metagenomics: a practical guide for the microbiome-minded physician scientist. Semin. Reprod. Med. 32, 5–13. <https://doi.org/10.1055/s-0033-1361817>.
- Menéndez-Arias, L., 2002. Molecular basis of fidelity of DNA synthesis and nucleotide specificity of retroviral reverse transcriptases. Prog. Nucleic Acid Res. Mol. Biol. 71, 91–147. [https://doi.org/10.1016/S0079-6603\(02\)71042-8](https://doi.org/10.1016/S0079-6603(02)71042-8).
- Menéndez-Arias, L., Sebastián-Martín, A., Álvarez, M., 2017. Viral reverse transcriptases. Virus Res. <https://doi.org/10.1016/j.virusres.2016.12.019>.
- Meng, C., Qiu, X., Jin, S., Yu, S., Chen, H., Ding, C., 2012. Whole genome sequencing and biological characterization of Duck/JS/10, a new lentogenic class I Newcastle disease virus. Arch. Virol. 157, 869–880. <https://doi.org/10.1007/s00705-012-1248-4>.
- Miller, P.J., Kim, L.M., Ip, H.S., Afonso, C.L., 2009. Evolutionary dynamics of Newcastle disease virus. Virology 391, 64–72. <https://doi.org/10.1016/j.viro.2009.05.033>.
- Munir, M., Linde, A.M., Zohari, S., Ståhl, K., Baule, C., Engström, B., M Renström, L.H., Berg, M., 2011. Whole genome sequencing and characterization of a virulent Newcastle disease virus isolated from an outbreak in Sweden. Virus Genes 43, 261–271. <https://doi.org/10.1007/s11262-011-0636-2>.
- Pinard, R., de Winter, A., Sarkis, G.J., Gerstein, M.B., Tartaro, K.R., Plant, R.N., Egholm, M., Rothberg, J.M., Leamon, J.H., 2006. Assessment of whole genome amplification-induced bias through high-throughput, massively parallel whole genome sequencing. BMC Genomics 7, 216. <https://doi.org/10.1186/1471-2164-7-216>.
- Qiu, X., Yu, Y., Yu, S., Zhan, Y., Wei, N., Song, C., Sun, Y., Tan, L., Ding, C., 2014. Development of strand-specific real-time RT-PCR to distinguish viral RNAs during Newcastle disease virus infection. Sci. World J. 2014. <https://doi.org/10.1155/2014/934851>.
- Satharasinghe, D.A., Murulitharan, K., Tan, S.W., Yeap, S.K., Munir, M., Ideris, A., Omar, A.R., 2016. Detection of inter-lineage natural recombination in avian paramyxovirus serotype 1 using simplified deep sequencing platform. Front. Microbiol. 7, 1–14. <https://doi.org/10.3389/fmicb.2016.01907>.
- Schaefer, B.C., 1995. Revolutions in rapid amplification of cDNA ends: new strategies for polymerase chain reaction cloning of full-length cDNA ends. Anal. Biochem. <https://doi.org/10.1006/abio.1995.1279>.
- Shi, S.H., Huang, Y., Cui, S.J., Cheng, L.F., Fu, G.H., Li, X., Chen, Z., Peng, C.X., Lin, F., Lin, J.S., Su, J.L., 2011. Genomic sequence of an avian paramyxovirus type 1 strain isolated from Muscovy duck (*Cairina moschata*) in China. Arch. Virol. 156, 405–412. <https://doi.org/10.1007/s00705-010-0866-y>.
- Shi, X., Jarvis, D.L., 2006. A new RACE method for extremely GC-rich genes. Anal. Biochem. 356, 222–228. <https://doi.org/10.1016/j.ab.2006.06.028>.
- Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D., Higgins, D.G., 2011. Fast, scalable generation of high quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. 7.
- Temmam, S., Monteil-Bouchard, S., Robert, C., Pascalis, H., Michelle, C., Jardot, P., Charrel, R., Raoult, D., Desnues, C., 2015. Host-associated metagenomics: a guide to generating infectious RNA Viromes. PLoS One 10. <https://doi.org/10.1371/journal.pone.0139810>.
- Thomson, E., Ip, C.L.C., Badhan, A., Christiansen, M.T., Adamson, W., Ansari, M.A., Bibby, D., Breuer, J., Brown, A., Bowden, R., Bryant, J., Bonsall, D., Silva, D., Hinds, C., Hudson, E., Klenerman, P., Lythgow, K., Mbisa, J.L., Mclauchlan, J., Myers, R., Piazza, P., Roy, S., Trebes, A., 2016. Comparison of Next-Generation Sequencing Technologies for Comprehensive Assessment of Full-Length Hepatitis C Viral Genomes. pp. 54. <https://doi.org/10.1128/JCM.00330-16.Editor>.
- Tyler, A.D., Christianson, S., Knox, N.C., Mabon, P., Wolfe, J., Van Domselaar, G., Graham, M.R., Sharma, M.K., 2016. Comparison of sample preparation methods used for the next-generation sequencing of mycobacterium tuberculosis. PLoS One 11. <https://doi.org/10.1371/journal.pone.0148676>.
- Van Borm, S., Rosseel, T., Steensels, M., Van den Berg, T., Lambrecht, B., 2013. What's in a strain? Viral metagenomics identifies genetic variation and contaminating circoviruses in laboratory isolates of pigeon paramyxovirus type 1. Virus Res. 171, 186–193. <https://doi.org/10.1016/j.virusres.2012.11.017>.
- Van Borm, S., Wang, J., Granberg, F., Colling, A., 2016. Next-generation sequencing workflows in veterinary infection biology: towards validation and quality assurance. Rev. Sci. Tech. 35, 67–81. <https://doi.org/10.20506/rst.35.1.2418>.
- Wei, D., Yang, B., Li, Y., Lin, Xue, C., Fang, Chen, Z., nan, Bian, H., 2008. Characterization of the genome sequence of an oncolytic Newcastle disease virus strain Italian. Virus Res. 135, 312–319. <https://doi.org/10.1016/j.virusres.2008.03.003>.