AsPac J. Mol. Biol. Biotechnol. 2024
Vol. 32 (4) (Special Issue) : 94-113

Complete genome sequence of local Malaysian Fowl adenovirus serotype 8b UPM T221 strain

Bahiyah Azli[a], Nur Farhana Salim[a], Mohd Hair-Bejo[a,c], Norfitriah Mohamed Sohaimi[a,d], Nor Asilah Wati Abdul Hamid[e,f], Nurulfiza Mat Isa[a,b*]

[a]*Laboratory of Vaccine and Biomolecules, Institute of Bioscience, Universiti Putra Malaysia Malaysia*
[b]*Department of Cell and Molecular Biology, Faculty of Biotechnology and Biomolecular Science, Universiti Putra Malaysia, 43400 Serdang Malaysia*
[c]*Department of Veterinary Pathology and Microbiology, Faculty of Veterinary Medicine, Universiti Putra Malaysia, 43400 Serdang Malaysia*
[d]*Department of Veterinary Laboratory Diagnostic, Faculty of Veterinary Medicine, Universiti Putra Malaysia, 43400 Serdang Malaysia*
[e]*Department of Communication Technology and Network, Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, 43400 Serdang Malaysia*
[f]*Laboratory of Computational Science and Mathematical Physics, Institute of Mathematical Research, Universiti Putra Malaysia, 43400 Serdang Malaysia*

Received 11th October 2024 / Accepted 24th December 2024 / Published 31st December 2024

Abstract. Recently, Fowl adenovirus serotype 8b (FAdV-8b) infection has emerged as a serious threat to Malaysia's poultry industry, acting as a causal agent of Inclusion Body Hepatitis (IBH) with a mortality rate of 10-30% among chickens. In this study, we isolated an FAdV strain from the liver of an IBH-positive dead commercial broiler chicken in Tawau, Sabah, and subjected it to pathogenicity analysis and Whole genome sequencing (WGS). Upon inoculation of UPM T221 isolate into Specific pathogen-free chicken embryonated eggs, slow mortality pattern of 6- to 12-days post-infection (dpi) was recorded, with nil gross lesions of both chorioallantoic membrane (CAM) and liver observed during harvesting. Upon subjection to WGS, the genome of UPM T221 was found to be 44722 bp in length with 58.1% GC content, 37 coding sequence (CDS), identifying the isolate as a strain from FAdV-8b of the FAdV-E. Interestingly, the left ORF regions and central genes showed a higher potential for significant genetic divergence with Single nucleotide polymorphism (SNPs) and InDels gaps. These variants are displayed in the within the structural capsid of UPM T221, such as hexon and penton, as well as in virion replication processes such as DNA polymerase and pTP. These findings of less virulence, non-pathogenic UPM T221 profile proposed the strain as a suitable live attenuated candidate vaccine, also contributing to the current understanding of the genetic diversity of FAdV for developing autogenous vaccines or diagnostic materials especially against local strain infections.

Keywords: avian; fowl adenovirus; FAdV-8b; genomic; WGS

INTRODUCTION

Fowl adenovirus (FAdV) is an avian-susceptible pathogen classified under the genus *Aviadenovirus* within the *Adenoviridae* family. This pathogen is subgrouped into five groups of FAdV-A to FAdV-E, and further subtypes into 12 serotypes, determined by its restriction enzyme digestion patterns and cross-neutralisation test results of the isolated strain (Adair *et al*., 1979; Meulemans *et al*., 2001; Pallister *et al*., 1996). FAdV is characterised with a double stranded DNA (dsDNA), non-enveloped genome, with size ranging from 43-45 kilo basepair (kbp). The genomes encode non-structural proteins involved in enzymatic processes, and three structural proteins,

*Author for correspondence: Nurulfiza Mat Isa, Laboratory of Vaccine and Biomolecules, Institute of Bioscience, Universiti Putra Malaysia 43400 Serdang Malaysia
Email – nurulfiza@upm.edu.my

specifically the fibre, hexon and penton proteins. Fibre protein is annotated to be involved in the primary viral attachment, hexon acting as the major surface-exposed capsid structure that harbours the group- and type-specific antigenic determinants (Sohaimi & Hair-Bejo, 2021; Steer *et al*., 2009), while penton forms base that provides stability during penetration and viral entry into the host cells (Sheppard & Trist, 1992).

Outbreaks of FAdV causes had shown severe economic impact of losses and food security issue for a nation, especially in the poultry industry (Kiss *et al*., 2021; Schachner *et al*., 2018). This pathogen has been profiled as the primary aetiological agent of several avian diseases such as Hyper hydropericadium syndrome (HHS), Inclusion body hepatitis (IBH), Avian gizzard erosion (AGE), and more. Interestingly, different diseases manifestation is observed depending on the serotypes infecting the flocks, as well as different susceptibility profile across chicken species and ages. The FAdV-C is reported to be responsible for HHS outbreaks (Mo *et al*., 2019; Niu *et al*., 2022), FAdV-D and -E isolates lead to IBH (Alvarado *et al*., 2007; Hair-Bejo, 2005; Itakura *et al*., 1974; Mase *et al*., 2012; Norina *et al*., 2016; Schachner *et al*., 2019), AGE only in FAdV-A and FAdV-B-infected chickens (Grafl *et al*., 2018). The first FAdV outbreak case was reported back in 2005 (Hair-Bejo, 2005), with FAdV-8b being the most pre-dominant serotypes in Malaysia for the past two decades. To date, the number of outbreaks cases reported worldwide is increasing, suggesting necessary biological interventions is in dire needs. The most common method to act as control and preventive measures against viral infections are application of diagnostic protocols and development of vaccines for vaccination, poses significant importance in combating future infections.

Whole genome sequencing (WGS) is one of the bioinformatics protocols often employed in investigating the genetic basis of our organism of interest, such as virulence, resistance, and more. The advancement of WGS and its thorough pipeline constructions has revolutionised the understanding of pathogens evolution, epidemiology, as well as better identification as compared to the conventional PCR method. This method which involved Next-generation sequencing (NGS) to produce libraries of genomic data offers holistic and comprehensive approach towards full genome biological functions and its profile (Kapgate *et al*., 2015; Slatko *et al*., 2018; Voelkerding *et al*., 2009). The employment of WGS to elucidate the virulence profile of avian viruses such as Avian reoviruss (Narvaez *et al*., 2023), Newcastle disease virus (Eid *et al*., 2022), Avian influenza (Ip *et al*., 2023), as well as FAdV (Chen *et al*., 2022; Huang *et al*., 2019; Ojkic & Nagy, 2000; Rashid *et al*., 2024), had been widely adopted across the regions, enables the identification and comparison between and within FAdV serotypes. Also, correct identification of serotypes involved in IBH outbreaks and knowledge of the complete sequence of the FAdV-8b UPM T221 which was retrieved from the swollen, discoloured liver of 27-days old dead commercial broiler chicken from an outbreak in Sabah (2017) is crucial, as it will provide valuable insights on the molecular evolution and vaccination strategies as also seen adopted in bacteriology (Serruto *et al*., 2009; Shamanna *et al*., 2024)Therefore, the aim of this study is to perform the whole genome sequencing onto the FAdV-8b strain UPM T221 isolate, and further establish its genomic organisation profile against the 1st archived complete local Malaysian FAdV-8b strain, UPM04217 (Alemnesh *et al*., 2012; Juliana *et al*., 2014; Mat Isa *et al*., 2019). This study represents the first comprehensive genomic comparison between non-pathogenic UPM T221 and pathogenic UPM04217, providing novel insights into the genetic basis of reduced virulence in potential vaccine candidates and the molecular drivers of pathogenicity in FAdV-8b. The WGS of FAdV-8b UPM T221, along with its significant profile of genetic changes, phylogenetic information, alteration of gene features via comparison against a pathogenic local FAdV-8b reference were investigated and reported. The findings would provide insights on the reduced virulence profile exhibited by FAdV-8b UPM T221 depicted by *in vivo* studies in SPF chicken embryonated eggs, also potentially highlighting key gene players involving in the alteration of FAdV strain's virulence prowess.

MATERIALS AND METHODS

Field isolation

The field virus isolate used in this studywas retrieved from Faculty of Veterinary Medicine, Universiti Putra Malaysia. The isolate was retrieved from the liver of a dead 27-days old commercial broiler chicken of a Sabah (2017) outbreak, with recorded mortality of 2%. The dead commercial broiler chicken was reported with physical appearance of ruffled feathers, poor performance, severe depressions, low body weight. Meanwhile, the harvested liver was observed with major friable, discoloured gross lesions, upon necropsy analysis. The harvested liver sample was then homogenised thoroughly in 1X PBS (Life Technologies™, USA) to produce 50% w/v suspension. Next, the homogenate suspension was centrifuged with 32R centrifuge (Hettich Universal, Germany) at 350 xg for 5 mins at 4°C, post-three sets of freeze-thaw cycles. The retrieved supernatants were subsequently filtered using 0.22 μM polyethersulfone (PES) membrane syringe filters (Microlab Scientific Inc., China).

Pathogenicity analysis in specific pathogen-free chicken embryonated eggs

The SPF chicken embryonated eggs (CEE) used in this study were purchased from Malaysian Vaccine Pharmaceutical® Sdn. Bhd. (Malaysia). Then, 0.1 mL $10^{8.79}$ EID50/mL UPM T221 filtered virus inoculum from the field were inoculated into the SPF CEE were inoculated with using a 1cc/mL syringe (Terumo, China) with 25G sterilised needle (Teruma, China) via the chorioallantoic membrane (CAM) route. The inoculated eggs were then incubated in 37°C, 5% CO_2 sterile incubator. Following inoculation, the mortality rate of SPF CEE was monitored with 24-hours interval and recorded in a tabulated format across the designated observation period. Mortality and gross lesions, particularly in the liver and CAM were assessed to evaluate the pathogenicity of UPM T221. Positive indicators of IBH infection included liver discolouration, friability, and thickening or cloudiness of the CAM, which deemed infected with FAdV-8b.

DNA extraction

The total viral DNA extraction was performed by subjecting 200 μL freshly prepared inoculum to InnuPREP® viral DNA/RNA extraction kit (AnalytikJena, Germany) following the provided guidelines by the manufacturer accordingly, with final extracted viral DNA volume of 50 mL. Then, the viral DNA was proceed into DNA concentration and purity determination using the Nanodrop BioSpectrophotomer® (Eppendorf, Germany). Only extracted viral DNA with high yield concentration in ng/μL and good 260mm/280mm purity of 1.8~2.2, was further subjected into further downline sequencing and bioinformatics analysis.

Whole genome sequencing

45 μL of purified viral genomic DNA was subjected to WGS using DNBSEQ-G400 sequencer (MGI, China). Library preparation was carried out using the MGIEasy FS DNA Library Prep Set Kit (MGI, China) and generated 150 bp paired-end reads by following the manufacturer's guidelines. Related data were deposited into the NCBI Sequence Read Archive (SRA; Accession No.: SUB14782021). The summary of Next generation-sequencing (NGS) and WGS bioinformatics pipeline in this study is shown in Figure 1.

Reads pre-processing

The quality control checking of reads was performed via FastQC v0.12.1 (Andrew, 2010). The raw reads produced were subjected to pre-processing to filter low-quality sequences, by performing adaptor removal of Forward and Reverse, using SeqPrep v1.2 (jstjohn, 2016), and filter the sequences with parameters of Phred33, base quality <20, and 0.1% N bases as high-quality sequence via Sickle v1.33 (Joshi & Fass, 2011).

Reads-based mapping analysis

The SPAdes v3.15.5 (Bankevich *et al.*, 2012)and short read assembly Velvet v1.2.10 (Zerbino & Birney, 2008) were used to assemble the clean reads into contigs and scaffolds, via reference-based assembly of De Bruijn graph algorithm.

FAdV-8b UPM04217 strain sequence (Accession No.: KU517714.1) was retrieved from the GenBank of National Centre of Biotechnology Information (NCBI) and selected as the reference for the genome assembly. The scaffold and re-assembly processes were done to create full consensus sequences of UPM T221. The assembly output includes reads mapping ratio, depth, and coverage against the reference UPM04217. Next, the circular genome data visualisation was performed via Artemis DNAplotter (Carver *et al.*, 2009). Also, the repeat regions within the FAdV genomes were identified using the online Tandem Repeats Finder v4.09.1 (Benson, 1999) Additionally, the final consensus sequence of UPM T221 was subjected to coding sequence (CDS) prediction via Prokka v1.14.5 (Seemann, 2014) and ORFFinder v1.8 (Chokyotager, 2021).

Genomic features analysis

The final consensus sequences of UPM T221 isolate was subjected to annotation against local BLASTN program (McGinnis & Madden, 2004), aligning against the non-redundant (nr) database. Then, the syntenic regions schematic Circos diagram between UPM T221 and reference genome was visualised using GGisy v1.0 tool (Sannrone, 2017) which utilised the Circlize CRAN R package (Zhang *et al.*, 2013). Also, the collinearity alignment value between UPM T221 and the genome of reference was plotted into dot plots via ggplot2 CRAN R (Wickham, 2016). Following annotation, the complete sequence of UPM T221 was subjected to synteny analysis against UPM04217 as reference, in order to identify the conservation blocks of genes and other genomic features shared between the isolates via Mauve v2.4.0 tool, utilising the progressiveAlignment setting (Darling *et al.*, 2004).

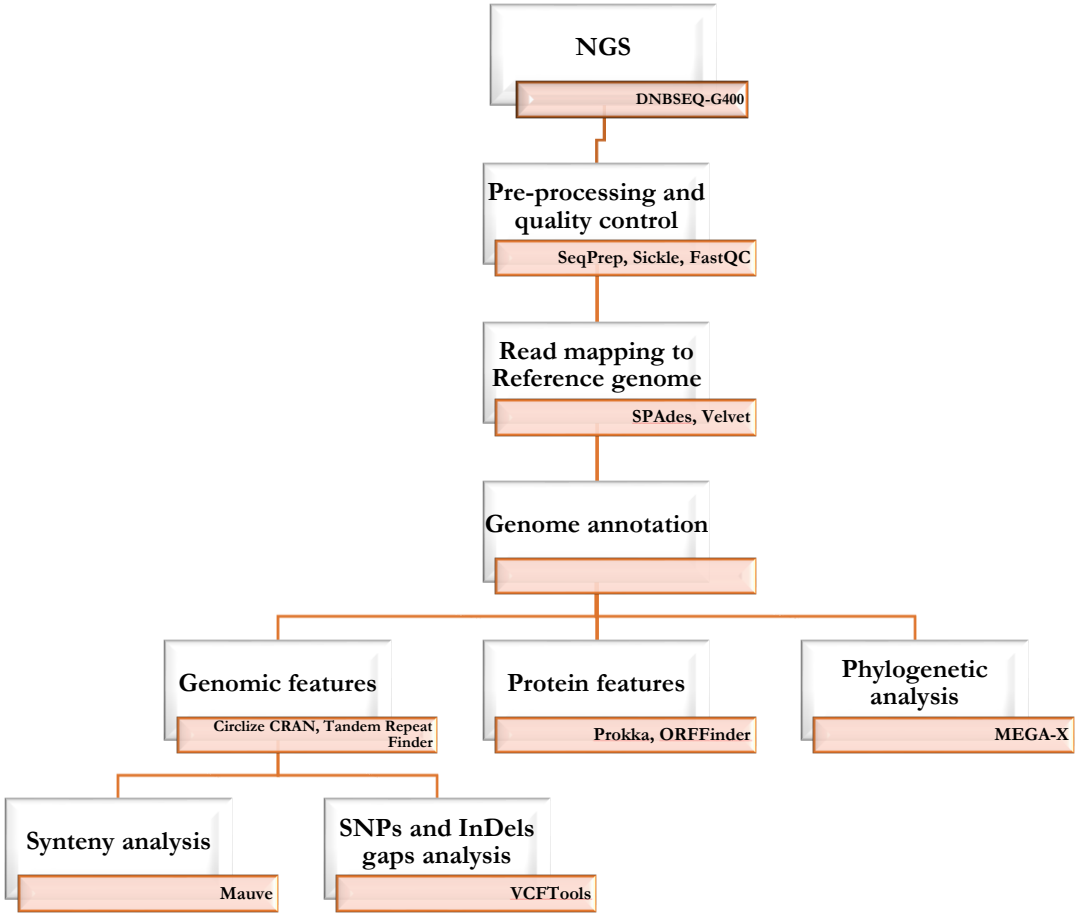

Figure 1. Summary of Next-generation sequencing and Whole genome sequencing bioinformatics pipeline in this study.

To identify the Single nucleotide polymorphisms (SNPs), insertions, deletions (InDels) gaps, between the genomes of UPM T221 and reference genome, high-quality paired-end reads were aligned using BWA v0.7.17 (Li & Durbin, 2010). Then, the alignment files were sorted, indexed, and 'variant called' using samtools v1.12 (Danecek *et al*., 2021). Then, the bcftools v1.12 converts the data into a Variant Call Format file (.vcf) output, which lists the identified variants. The identified SNPs and InDels gaps were subjected to filtration to remove low-quality variants with Phred score < 30 (Q30), using VCFtools v0.1.16 (Danecek *et al*., 2011). For visualisation, lollipop plots of SNPs and InDels regions were created using the ggplot2 package from CRAN in R.

Phylogenetic analysis

Multiple sequence alignments (MSA) were applied onto 33 complete genome sequences in FASTA format of all 12 FAdV serotypes and one Duck adenovirus (DAdV) were retrieved from the GenBank of NCBI (Table 1) using MEGA-X v10.2.6 (Kumar *et al*., 2018), employing the ClustalW algorithm (Thompson *et al*., 1994). The phylogenetic tree of the complete whole genome sequence was generated using Maximum-likelihood (ML) analysis with 1000 bootstrap replicates setting, DAdV as the outrgroup of the analysis.

Table 1. Fowl adenovirus strains from species A-E used in the phylogenetic tree construction of complete whole genome sequence analysis.

No	Isolate	Accession No.	Serotype	Group	Origin	Reference
1	SD1356	MG712775.1	8b	E	China	Huang *et al*., 2019
2	764	KT862811.1	8b		UK	Marek *et al*., 2016
3	UPM04217	KU517714.1	8b		Malaysia	Unpublished
4	UPM T221**	N/A	8b		Malaysia	This study
5	QD2016	MF577036.1	8b		China	Unpublished
6	TR59	KT862810.1	8a		Japan	Marek *et al*., 2016
7	YR36	KT862809.1	7		Japan	Marek *et al*., 2016
8	CR119	KT862808.1	6		Japan	Marek *et al*., 2016
9	380	KT862812.1	11	D	UK	Marek *et al*., 2016
10	SR49	KT862807.1	3		Japan	Marek *et al*., 2016
11	SR48	KT862806.1	2		Japan	Marek *et al*., 2016
12	685	KT862805.1	2		Japan	Marek *et al*., 2016
13	A-2A	AF083975.2	9		Canada	Cao *et al*., 1998
14	GX2017-01	MN577977.1	4	C	China	Rashid *et al*., 2020
15	AH170721	MW699358.1	4		China	Unpublished
16	GD616	MW509553.1	4		China	Unpublished
17	GX2019-014	MW448476.1	4		China	Unpublished
18	FaDV4/CN/SD-1	MW167778.1	4		China	Unpublished
19	HN19/FAdV-4	MT635844.1	4		China	Unpublished
20	GX2019-09	MN577985.1	4		China	Rashid *et al*., 2020
21	GX2018-08	MN577984.1	4		China	Rashid *et al*., 2020
22	GX2017-02	MN577978.1	4		China	Rashid *et al*., 2020
23	AHFY19	MN542422.1	10		China	Unpublished
24	GDMZ	MG856954.1	4		China	Unpublished
25	NIVD2	MG547384.1	4		China	Zou *et al*., 2018
26	HLJFAd15	KU991797.1	4		China	Unpublished
27	B1-7	KU342001.1	4		India	Unpublished
28	WHRS	OM836676.1	5	B	China	Marek *et al*., 2013
29	340	NC_21221.1	5		Ireland	Marek *et al*., 2013
30	D2453/1/10-12/13/UA	MT500572.1	5		Hungary	Homonnay *et al*., 2021
31	LYG	MK757473.1	5		China	Unpublished
32	CELO	U46933.1	1	A	Austria	Chiocca *et al*., 1996
33	DAdV-JSXZ	OP784579.1	DAdV		China	Unpublished

**Isolates used in this study

RESULTS

Pathogenicity analysis

The inoculation of FAdV UPM T221 inoculum into viable eggs was observed within 12-dpi window timeframe. In every passage of both FAdV-8b isolates, the mocked PBS-inoculated eggs and non-inoculated eggs are observed with viable and normal eggs development throughout experiment. Meanwhile, the eggs inoculated with UPM T221 were exhibited with a delayed mortality profile, despite achieving 100% cumulative mortality at 13-dpi, except for Passage 1. In the first passage batch, only 40% mortality was achieved with death range from 3- to 4-dpi, which suggested a low virulence profile of the UPM T221 isolate. From P2-P5, the mortality of UPM T221-inoculated eggs started at the average of 6-dpi, and peaks at 8- to 9-dpi (Table 2). Throughout trial, all performed passages of P2-P5 tabulated with relatively similar mortality profile, regardless of the number of SPF eggs inoculated per batch. Upon harvesting, all embryos in the control group were observed to be normal without any lesions throughout the passages, exhibiting a normal liver (Figure 2A), and thin and transparent CAM (Figure 2B). The embryo of eggs inoculated with UPM T221 were reported with mild gross lesions, observed to have glistening, firm and brick slight pale red coloured liver tissues (Figure 2C) and thin, clear CAM (Figure 2D) similarly with the control group. The nil lesions observation was exhibited consistently throughout P2 until P5 of UPM T221 isolate propagation.

Table 2. Cumulative mortality frequency of Specific pathogen-free chicken embryonated eggs infected with FAdV strain UPM T221 isolate.

Passage	Cumulative mortality, days post-inoculation (dpi)													Total death	Mortality, %
	0	1	2	3	4	5	6	7	8	9	10	11	12		
P1	0[a]/10[b]	0/10	0/10	2/10	4/10	4/10	4/10	4/10	4/10	4/10	4/10	4/10	4/10	4/10	40
P2	0[a]/3[b]	0/3	0/3	0/3	0/3	0/3	1/3	2/3	3/3	-	-	-	-	3/3	100
P3	0/18	0/18	0/18	0/18	0/18	3/18	3/18	3/18	5/18	10/18	12/18	14/18	18/18	18/18	100
P4	0/4	0/4	0/4	0/4	0/4	0/4	0/4	0/4	2/4	2/4	3/4	4/4	-	4/4	100
P5	0/92	0/92	0/92	0/92	0/92	14/92	17/92	28/92	37/92	63/92	76/92	85/92	92/92	92/10	100

[a] the number of infected embryos that died in the passage; [b] the total number of eggs inoculated in the passage

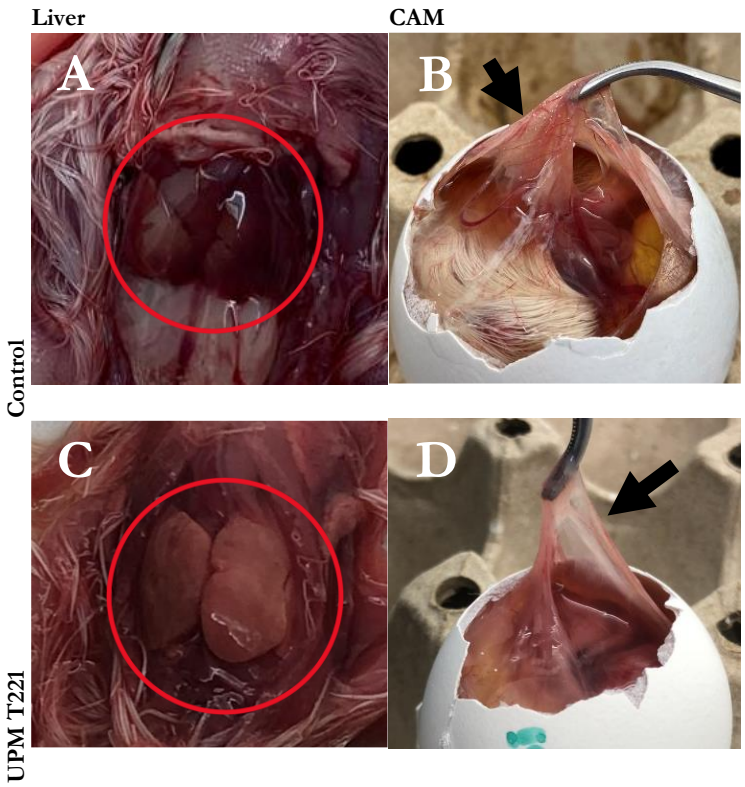

Figure 2. (A, C) Liver and (B, D) chorioallantoic membrane (CAM) lesions in control andFAdV strain UPM T221-inoculated Specific pathogen-free chicken embryonated eggs.

Raw and clean reads statistics analysis

The sequencing results for the FAdV UPM T221 of Passage 5 sample indicate strong read quality of both raw and clean reads metrics (Table 3). In the raw read analysis, UPM T221 generated a total of 3,188,045 reads for both R1 and R2, with high-quality scores across both read lengths. For R1, the Q20 and Q30 values were 95.86% and 86.95%, respectively, while R2 had slightly lower quality scores of 90.87% (Q20) and 78.70% (Q30). Next, the pre-processing into clean reads resulted in 2,992,780 total clean reads for both R1 and R2. The clean read quality remained high, with R1 achieving Q20 and Q30 scores of 96.04% and 87.31%, respectively. R2 also exhibited increased in quality, with Q20 and Q30 values of 91.93% and 80.00%. These quality metrics reflect robust sequencing integrity, demonstrating minimal data loss between raw and clean read post pre-processing. The overall high-quality reads will contribute to the accurate assembly and downstream genomic annotation of the FAdV UPM T221 strain and subsequent comparative analysis.

WGS and phylogenetic analysis

The read-based mapping results of FAdV UPM T221 against FAdV-8b UPM04217, the first local Malaysian FAdV-8b strain, showed substantial variation in depth. Upon read-based mapping, UPM T221 successfully exhibited a significantly high average depth of 3908.17 and a median depth of 3909.0, with 99.68% genome coverage, indicating UPM04217 as the best reference sequence selection. Upon analysis, the genome of UPM T221 was sequenced for 44722 bp, reported with GC content of 58.1%, and contained 37 predicted consensus CDS across the genome. The circular representation of UPM T221 genome visualised from DNAplotter is shown in Figure 3. Additionally, the Inverted terminal repeats (ITR) and Tandem repeat 1 (TR-1) of UPM T221 were approximately located at 37556 – 37662 bp with 54 bp, and 38227 – 38318 bp with 33 bp in length.

Table 3. Raw and clean reads statistics of sequence non-pathogenic FAdV strain UPM T221 isolate.

Sample	Raw reads statistics				Clean reads statistics			
	Read length	Total no of raw reads	Q20	Q30	Read length	Total no of clean reads	Q20	Q30
UPM T221 R1	150	3188045	95.86%	86.95%	150	2992780	96.04%	87.31%
UPM T221 R2	150	3188045	90.87%	78.70%	150	2992780	91.93%	80.00%

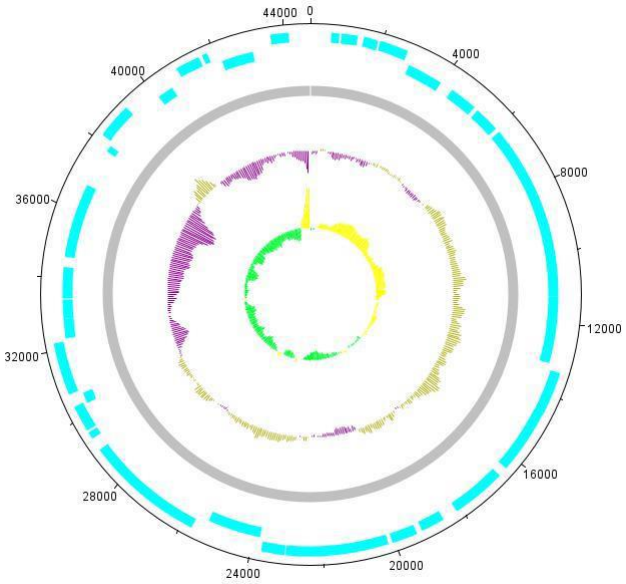

Figure 3. Circular representation of FAdV-8b strain UPM T221 isolate complete genome. Tracks from the Circos figure, going inwards, represented the (1) forward CDS, (2) reverse CDS, (3) genome coverage, (4) GC content, and (5) GC skew. **GC content and GC skew displayed in 1000 window size and 100 step size

The local alignment via the BLASTN tool of NCBI reported UPM T221 isolate with FAdV-8b SD1356 (98.13%, E-value= 0, coverage= 99%), FAdV-8b SD2009 (98.13%, E-value= 0, coverage= 99%), and FAdV-E HN1472 (97.11%, E-value= 0, coverage= 99%). Meanwhile, BLASTN only reported a percentage similarity of 98.98% of UPM T221 alignment against UPM04217 (E-value= 0, coverage= 98%). In short, the alignment against UPM04217 was reported with a similarity percentage of ~99%, yet with slightly low query coverage, possibly due to the 650 bp difference between UPM T221 and UPM04217. Next, the phylogenetic tree was constructed using complete genome sequences of UPM T221 and previous characterised FAdV isolates via ClustalW algorithm with 1000 bootstraps setting in MEGA-X. Figure 4 displays UPM T221 to have close relationship with other FAdV-8b isolates, being in the same dendogram branch with UPM04217, and both China's SD1356, and QD2016 strains.

Genomic variant analysis

The similarity comparison between UPM T221 and UPM04217 was performed by subjecting them to multiple genome alignment via the progressive Alignment tool of Mauve. In this stage, UPM T221 and UPM04217 are selected as the query and reference sequence, respectively, to compare the genomic profile between the non-pathogenic UPM T221 and pathogenic UPM04217.

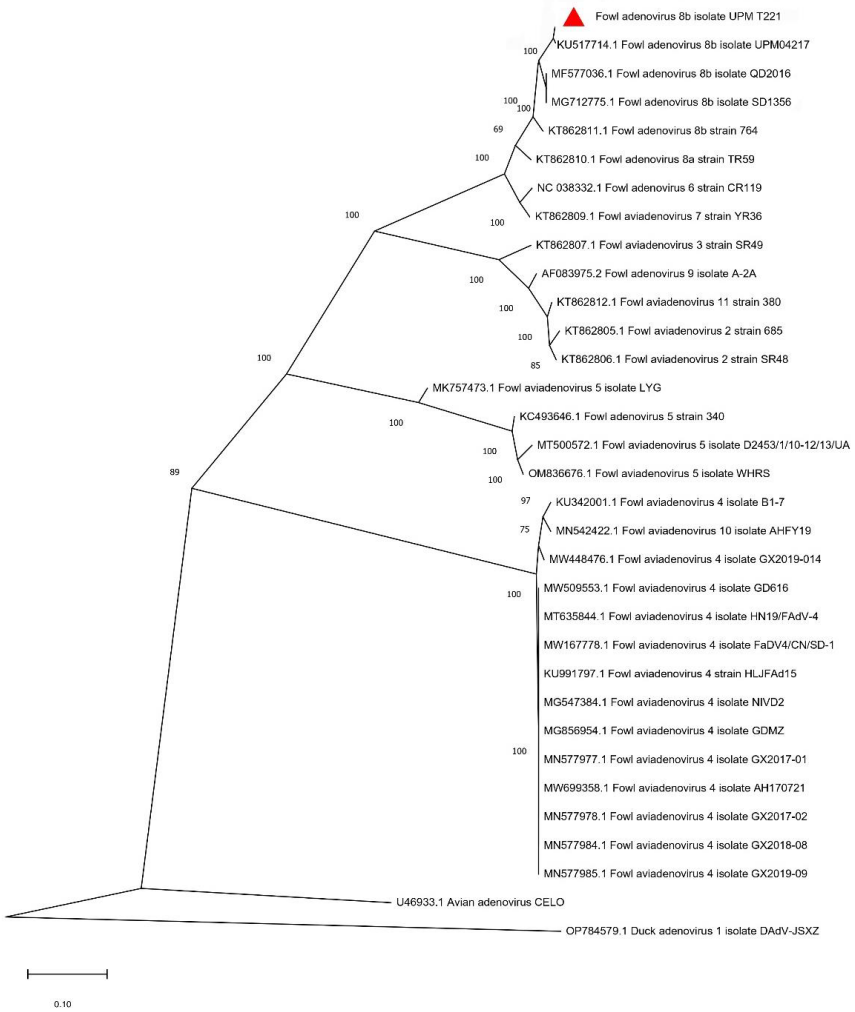

Figure 4. Phylogenetic analysis and sequence identity of complete genome sequence of FAdV and DAdV, performed via MEGA-X software with Maximum-likelihood (ML) algorithm applied with 1,000 bootstraps, with DAdV-JSXZ is set as the outgroup. The red triangle indicates FAdV-8b UPM T221 complete genome sequence.

The similarity percentage between UPM T221 and UPM04217 isolates was 43361/45190, approximately 95.95%. Figure 5 displays the Chord diagram of UPM T221 and UPM04217 with tracks for analyses. Similarly, the alignment statistics retrieved from Mauve (Table 4) reported both UPM T221 and UPM04217 with only one LCB. Besides, the generated dot plot of collinearity in Figure 5 also further confirms the serotypes of both isolates which comprise of similar genome organisation and arrangements. Also, the alignment between both isolates was tabulated with two Blocks, highlighting a major InDels in the alignment. As compared to UPM T221 alignment, this analysis was hypothetically predicted to produce 26 'broken' proteins and only 12 complete proteins. According to Mauve, the number of SNPs identified between UPM T221-UPM04217 is approximately 213. Also 28/51 and 23/51 gaps are identified in UPM T221 and UPM04217, respectively. Lastly, the alignment reported a LCB weight of 44205 bp, indicating high confidence in the LCB analysis performed.

Predicted ORF analysis

The ORFs of UPM T221 were predicted using ORFFinder and Prokka. Both tools reported similar numbers of ORFs for the UPM T221 genome, precisely 37. All predicted ORFs, their annotations, and aa lengths are recorded in Table 5.

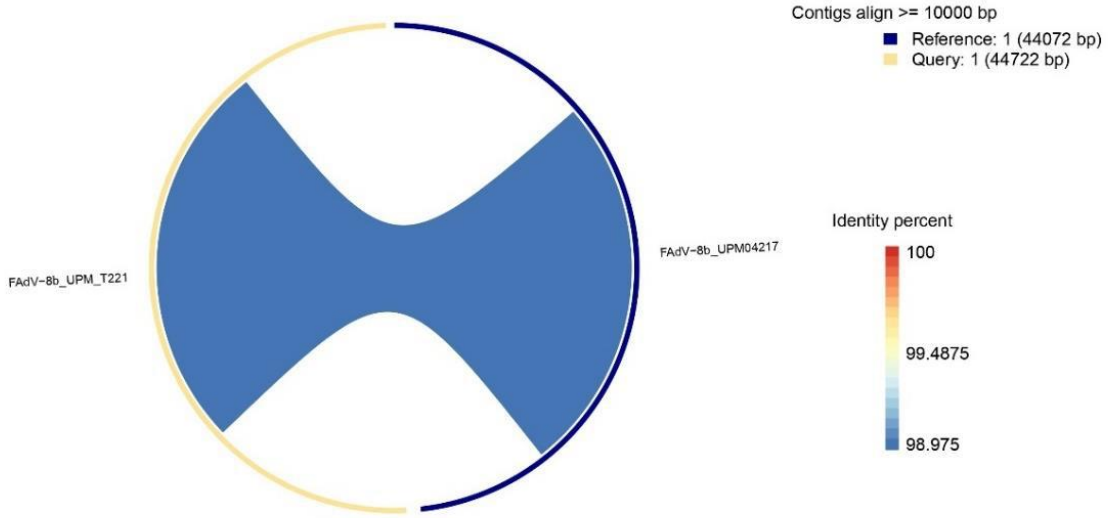

Figure 5. Chord diagram comparing FAdV-8b UPM T221 and UPM04217.

Table 4. Multiple genome alignment statistics via Local collinear blocks rearrangement between FAdV-8b UPM T221 and UPM04217.

Statistics	**FAdV-8b UPM T221 (assembly) and UPM04217 (reference)**
Number of assembly bases, bp	44722
Number of reference bases, bp	44072
Number of LCB	1
Number of Blocks	2
Number of CDS	
Complete	26
Broken	12
Variants	
Gaps in assembly	28
Gaps in reference	23

Left end genome proteins

The left genome sequences involved the arrangement of 11 proteins, from ORF0 to ORF12. In this region, the ORF0, ORF1, ORF1C, ORF2, ORF14, ORF13, and ORF12 of UPM T221 were predicted with relatively similar aa sizes. However, ORF1B and ORF14A were observed only in UPM T221, whereas ORF1A/B and ORF24 were present exclusively in UPM04217, suggesting potential functional divergence that might influence pathogenicity. The presence of either ORF1B or ORF1A/B, as well as ORF14A or ORF24, in one of the FAdV-8b strains, may indicate functional homologous or complementation between these proteins and their counterparts.

Table 5. Predicted open reading frame amino acid features of the non-pathogenic FAdV-8b UPM T221 and pathogenic UPM04217.

Gene	Category	UPM T221	UPM04217
ORF0	Left end	94	77
ORF1	Left end	153	153
ORF1B	Left end	81	81
ORF1C	Left end	88	73
ORF2	Left end	268	268
ORF14A	Left end	232	232
ORF14	Left end	115	115
ORF13	Left end	268	268
ORF12	Left end	251	251
pIVa2	Core	402	402
DNA polymerase	Essential enzyme	1302	1304
pTP	Core	642	264
			186
52/55 kDa	N/A	410	410
pIIIa	Minor cement	588	588
Penton	Major structural	509	509
pVII	Core	78	-
pX	N/A	215	220
pVI	Minor cement	228	228
Hexon	Major structural	947	947
Protease	Core	208	208
DNA-binding protein	Essential enzyme	517	468
Hexon assembly 100K	N/A	1099	346
			751
33K	N/A	73	73
pVIII	Minor cement	242	242
U exon	N/A	99	99
Fibre	Major structural	487	487
ORF22	Right end	192	192
ORF20A	Right end	186	186
ORF20	Right end	303	303
ORF19	Right end	714	773
ORF28	Right end	62	62
ORF29	Right end	55	66
GAM-1	Right end	277	131
			144
ORF17	Right end	156	156
ORF11	Right end	229	229
ORF11A	Right end	46	46
ORF23	Right end	312	312
ORF25	Right end	171	171
Total		37	40

Central adenoviral genome proteins

The central adenoviral genome sequences stretched from pIVa2 until fibre, summing up into 17 proteins with vital functions in adenovirus virulence and pathogenic profile. This region consists of three minor cement proteins, three major structural proteins, four core proteins, and essential enzymes. Among the three major capsid proteins, only Hexon was predicted with relatively similar aa size with hexon of UPM04217. Meanwhile, both penton and fibre of UPM T221 are shorter in size, approximately 43 aa and 35 aa. Meanwhile, the minor cement proteins (pVI and pVIII) located between the structural capsid and core proteins in FAdV-8b had equal aa sizes in both genomes, except for pIIIa with 48 aa longer than UPM04217. The slight difference in pIIIa minor cement protein may affected the genome stability via interaction of the internal proteins and structural capsid proteins and potentially to carry out different vital function internally between both strains. Interestingly, the Hexon assembly 100K protein was only predicted in UPM T221 but not in UPM04217. The lack of this protein involved during hexon folding may occur due to disruption of amino acid translation, or technical assembly of genome during pre-processing in previous UPM04217 annotations. Also, the 33K protein in UPM T221 was 70 aa shorter than in UPM04217, suggesting functional divergence, approximately the size of 33K in UPM0417 is double in aa residues length as compared to UPM T221. Next, among the four essential core proteins- pIVa2, pTP, pVII, and protease, only pTP was seen 12 aa longer in UPM T221 compared to UPM04217, which might lead to an alteration of the viral replication process of the non-pathogenic UPM T221. The remaining UPM T221 core proteins aa residues are relatively similar in size.

Right end genome proteins

The right genome sequences consist of 13 genes, rearranged from ORF22 until ORF25. Only five right genome ORFs between both genomes had similar aa size (ORF22, ORF17, ORF11, ORF23, and ORF25), while one right genome gene, ORF11A, was present in UPM T221 but not in UPM04217. Upon inspection, UPM T221 has longer ORF20A (186 aa) and ORF20 (303 aa) than UPM04217 (174 aa and 224 aa). Conversely, UPM04217 has longer ORF19 (773 vs 714 aa), ORF28 (106 vs 62 aa), and ORF29 (67 vs 55 aa) when compared against UPM T221. Interestingly, the GAM-1 protein of UPM04217 was found as three different ORF frames, roughly 132, 145, and 87 aa in length, which summed into 364 aa residues. In contrast, UPM T221 was only predicted with a single 277 aa frame ORF of GAM-1. Additionally, the presence of ORF11A in UPM T221 and ORF33 in UPM04217 alternatively suggesting that these proteins may perform similar functions as homologous gene, potentially compensating for each other's absence in the respective strains.

Predicted InDels and gaps

According to Mauve, 28 and 23 gap counts are present in UPM T221 and UPM04217 upon alignment. The lollipop plots in Figure 6A and 6B shows higher gap counts in the left and right end genome of UPM T221 than UPM04217, approximately <20000 bp and >30000 bp regions, which reside in essential enzymes of FAdV. This suggested an insertion event presence in UPM T221 within this area of genome. Meanwhile, lower gap counts were identified in the central genome region of UPM T221, indicating potential variations of the functionality of genes such as core, structural proteins, minor cements and more residing here. Upon further analysis, however, only 12/28 and 4/23 gaps in respective genomes were observed to reside within the CDS. This may suggest that the remaining gaps are localised at the non-coding sequence region of each genome that is known to be involved in gene regulation. There is the presence of a deletion in UPM T221 identified in UPM04217 from 12423-12698 bp (275 bp), and a gap in UPM04217 indicates a 681 bp-insertion in UPM T221, approximately at 39299-39980 bp. As seen in Table 6, among the eight CDS, the gaps in UPM T221 are such in ORF19, pX, and pTP genes, while UPM04217 were tabulated with unique gaps in only 33K, pIIIa, and ORF29 genes.

Predicted SNPs

The SNPs retrieved from bcftools in UPM T221 were plotted in the lollipop plot across the 44722 bp length. According to Mauve, 213 SNPs were identified upon alignment between UPM T221 as the assembly and UPM04217 as the reference

sequence. As seen in Figure 7, abundance SNPs are displayed at the central genome to the right end of the genome, whereas sparse SNPs are reported plotted at the left end of the genome. Meanwhile, only 149/213 SNPs were determined to localise within the CDS of UPM T221. The Top 5 genes with SNPs are ORF25 (53), ORF19 (32), pTP (12), ORF11 (11), and 52/55 kDa (5), as tabulated in Table 7. These proteins involved in both structural and enzymatic roles such as DNA replication. Among the major structural genes, only fibre is reported with the SNPs of 1, potentially indicating similar structural capsid protein residues between both genomes. Among the minor cement proteins, only pIIIa was tabulated with SNPs of 3. Meanwhile, the core genes displayed with SNPs was pTP with 12, while the remaining has nil SNPs. Interestingly, seven out of 13 ORFs residing at the right end genome of UPM T221 are reported with SNPs.

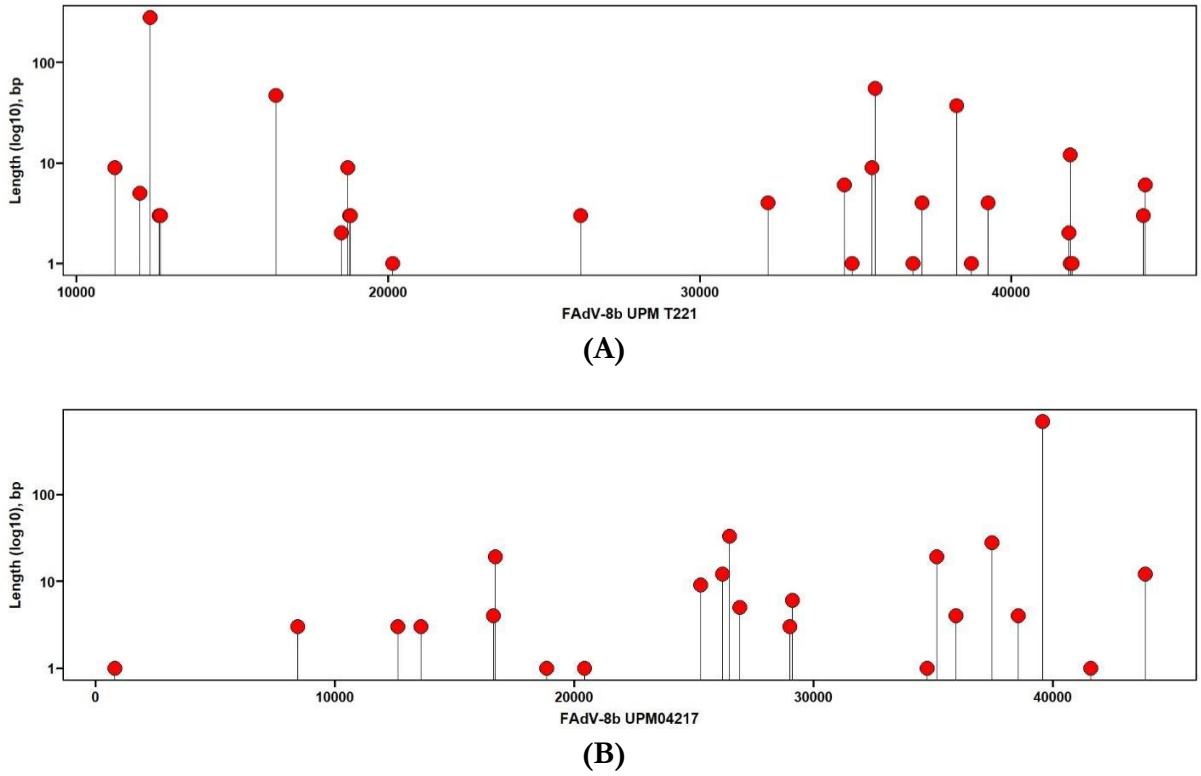

Figure 6. Distribution of gaps and its lengths in log10, across the (A) FAdV-8b UPM T221, and (B) UPM04217 genome, upon alignment with progressive Alignment of Mauve.

Table 6. Distribution of gaps across genes in FAdV-8b UPM T221 and UPM04217.

Gene	Category	No of gaps in UPM T221	No of gaps in UPM04217
DNA polymerase	Essential enzyme	1	-
pTP	Core	2	-
pIIIa	Minor cement	-	1
pX	N/A	3	-
33K	N/A	-	2
ORF19	Right end	4	-
ORF29	Right end	1	1
GAM-1	Right end	1	-
Total		12	4

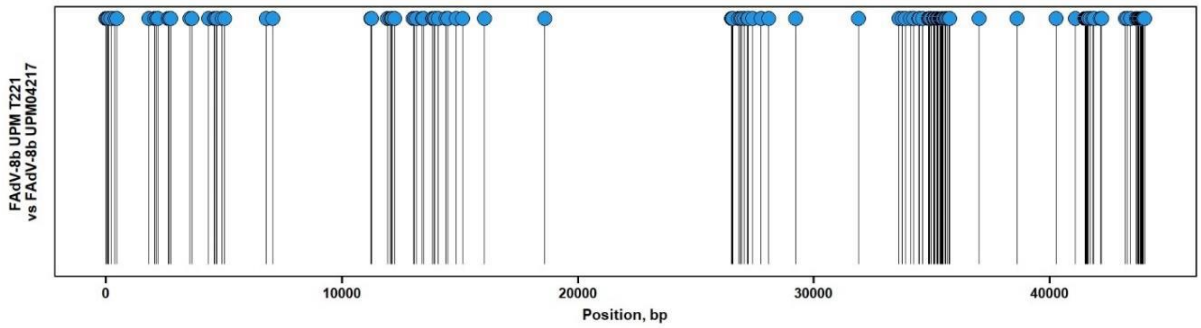

Figure 7. Lollipop plot of the distribution of SNPs between UPM T221 (assembly) and UPM04217 (reference).

Table 7. Distribution of SNPs across genes in UPM T221 (assembly) when compared against UPM04217 (reference).

Gene	Category	No of SNPs
ORF1C	Left end	1
ORF2	Left end	6
ORF14	Left end	1
ORF13	Left end	6
pIVa2	Core	2
DNA polymerase	Essential enzyme	4
pTP	Core	12
52/55 kDa	N/A	8
pIIIa	Minor cement	3
pX	N/A	1
Fibre	Major structural	1
ORF20	Right end	4
ORF19	Right end	32
GAM-1	Right end	1
ORF17	Right end	1
ORF11	Right end	11
ORF23	Right end	2
ORF25	Right end	53
Total		149

DISCUSSION

FAdV infection has been a major concern for the past several years, with increasing reports of IBH outbreak cases in poultry farmhouses (Sohaimi & Clifford, 2021). Major biological interventions such as diagnostic tools and vaccine developments are deemed vital. Identifying environmental isolates, especially from flocks with IBH clinical signs, is one of the most important steps of improvising and optimising the diagnostic protocols. Characterising novel isolates presence within the environment is mentioned as one of the appropriate methods to speed up the One Health Genomic objective for diagnostic and tracking the spread of outbreak for epidemiological data (Chakraborty & Barbuddhe, 2021). The genomic comparison of FAdV-8b UPM T221 and UPM04217 isolates offers critical insights into their evolutionary divergence, genomic features, and potential implications for vaccine development for FAdV-8b infection. Both strains, though genetically related, display distinct phenotypic characteristics, with UPM T221 being less virulent and non-pathogenic,

whereas UPM04217 was profiled with a pathogenic profile as stated in the published work of Alemnesh *et al.* (2012). Thus, in this study, we investigated the pathogenicity of the FAdV-8b UPM T221 in SPF CEE and its genome against UPM04217 in order to investigate the former's potential as a candidate. This analysis will also provide an in-depth examination of the genomic and protein features of the local FAdV-8b isolates, exploring the potential determinants of virulence.

Eight-day-old SPF chicken embryonated eggs inoculated with FAdV-8b UPM T221 showed a slower mortality curve, suggesting differences in virulence as compared to previously reported pathogenic FAdV-8b strains of UPM T27, UPM T1901, and UPM04217 (Ahmed *et al.*, 2021; Azli *et al.*, 2023; Mat Isa *et al.*, 2019), which has shorter infection time frame between 3- to 7-dpi. This elucidation is also supported in the nil gross lesions observed at the infected liver of inoculated SPF eggs, with brick red-coloured, firm, glistening and smooth texture, despite inoculated with high viral titre of $10^{8.79}$ EID50/mL. This indicates that FAdV-8b UPM T221 may elicit a milder pathogenic response, potentially due to mutations or deletions in virulence-associated genes (Grgić *et al.*, 2014; Walters *et al.*, 2002), yet further analysis is required. Interestingly, the survival of embryos beyond the typical acute phase further highlighted the low-virulence profile of UPM T221, suggesting its potential as a candidate for live attenuated vaccine development. This *in vivo* pathogenicity analysis is crucial, as not all field isolates are suitable for development into autogenous vaccine strains (Schachner *et al.*, 2018), as per practice in development of vaccine against Human Influenza A(Steinbrück *et al.*, 2014). Understanding the pathogenic potential of isolates helps identify field strains that are both safe and effective for vaccine development, particularly when targeting specific geographical outbreaks. While some field isolates, like FAdV-8b UPM T221, exhibit low pathogenicity, others may retain high virulence, posing risks when used as vaccine strains. Therefore, understanding the pathogenicity profile through in vivo studies allows for the careful selection of strains that are both safe and effective, ensuring that only attenuated isolates with minimal pathogenic effects are considered for vaccine development. This approach helps prevent the unintended introduction of virulent strains that could exacerbate disease in immunised flocks.

Next, the WGS results from the fifth passage of the UPM T221 sample revealed a high-quality read depth essential for genomic analysis. The raw read analysis indicated a total of 3,188,045 reads, with R1 achieving Q20 and Q30 values of 95.86% and 86.95%, respectively. Following pre-processing, a total of 2,992,780 clean reads were retained, demonstrating robust sequencing integrity and minimal data loss. The quality metrics suggest that the sequencing is suitable for accurate genomic assembly and subsequent analysis. The mapping of these clean reads against the local Malaysian isolate, FAdV-8b UPM04217, yielded more >98% coverage for UPM T221. This high coverage also affirms the completeness of the UPM T221 genome as a complete sequence, as well as affirms the reference selection of UPM04217, allowing for a detailed comparison between the two genomes.

The UPM T221 genome spans 44722 bp with a GC content of 58.1%, influencing various genomic characteristics, including replication and stability. The GC content is distributed unevenly across the genome, with lower averages observed at both ends and a central region exhibiting higher GC percentages. This distribution may correlate with the functional roles of specific genes within the genome, highlighting the potential for nucleotide properties to influence virulence. The local alignment using BLASTN revealed UPM T221 with 98.13% identity to FAdV-8b SD1356, whose close genetic relatedness was further confirmed through phylogenetic analysis using the complete genome sequences. To date, most FAdV isolates identified in Malaysia are predominantly reported as FAdV-8b strain, which supports the claim of predominant nature of different serotypes to reports as the primary aetiological agents of FAdV disease, according to country and continents (Niczyporuk & Kozdruń, 2022; Sohaimi *et al.*, 2018). Among the diverse serotypes circulating in Malaysia, FAdV-8b has consistently emerged as a predominant strain. UPM T1901 has been the latest addition to the list of identified FAdV-8b isolates since the first report of local cases in 2005 (Hair-Bejo, 2005). This notion facilitates the researchers in prioritising potential target genomic composition to act as a template in autogenous vaccines or

antiviral development as biosecurity to fight against local FAdV infections.

Next, the in-depth genomic alignment using the Mauve tool showed 28 gaps in UPM T221 and 23 gaps in UPM04217, approximately with very fewer gaps observed in the central genome region of UPM T221. This suggests possible InDels events in UPM T221, which may contribute to its non-pathogenic phenotype due to potential frameshift occurring (Kondo *et al*., 2016; Pagel *et al*., 2019; Sturdevant *et al*., 2010; Zhu *et al*., 2024), yet further analysis is needed. Interestingly, only 12/28 and 4/23 gaps in UPM T221 and UPM04217 were reported to reside within annotated CDS. This elucidates the potential of InDels events occurring in non-coding sequences, such as promoter or enhancer regions, to regulate the pathogenicity of an FAdV-8b strain (Ellingford *et al*., 2022; Ludwig, 2002). There are more insertion events in a non-coding sequence of the non-pathogenic UPM T221, which may lead to the less virulent state seen in the *in vivo* pathogenicity study, potentially actively regulating the expressions of vital FAdV genes. Meanwhile, UPM04217 has more insertion events at the CDS, which would affect the protein structure upon translation of the virus. Furthermore, the SNP analysis also underscores key genetic variations between the two isolates- UPM T221 exhibited 213 SNPs relative to UPM04217, with the majority clustered toward the left and central regions of the genome. Of these, only 149 SNPs were localised within the CDS of UPM T221, particularly in the structural and enzymatic genes essential for lipid metabolism, viral replication, and virion assembly. For instance, the ORF19 gene, a lipase enzyme responsible for lipid regulation in the liver of the host for viral survival (Pei *et al*., 2019), harboured a high number of SNPs (32) in UPM T221. Other significant SNPs were found in fibre (1), pTP (12), pIVa2 (2), and DNA polymerase (4), suggesting that these mutations might affect the replication efficiency and viral assembly in UPM T221, contributing to its less virulent phenotype. DNA polymerase and pTP are proteins involved in DNA replication, while pIVa2 is often elucidated to play a significant role in virion replication at the late stage of the FAdV life cycle (Kulanayake & Tikoo, 2021). Meanwhile, the absence of SNPs in other structural proteins of penton and hexon further suggests the conservation of these essential capsid proteins in both strains for serotype identification and characterisation.

The detailed comparison of protein features between UPM T221 and UPM04217 adds another dimension to understanding the differences in pathogenicity and potential vaccine efficacy of local FAdV-8b strains. UPM T221, identified as a non-pathogenic strain, and UPM04217, with pathogenic characteristics, exhibit distinct differences at the protein level that influence their behaviour within the host environment. The analysis of ORFs in UPM T221 revealed a total of 37 predicted CDS, with many showing significant differences in aa length and structure compared to the UPM04217. Notably, several ORFs at the left and right ends of the genome exhibited size variations that may influence the non-pathogenic profile of UPM T221. At the left end of the genome, the presence of ORF1B and ORF14A in UPM T221, and their absence in UPM04217, could indicate that these proteins play a role in modulating the host response or viral replication in a way that diminishes virulence as a FAdV. According to Corredor *et al*. (2006), ORF1A/B protein is homologous towards ORF1B and ORF24 to ORF14A, explaining the absence in UPM T221. These differences between the strains highlight the complexity of viral attenuation, where both the presence of additional genes and the loss of specific genes can lead to a less virulent phenotype- adopting the complementation of ORFs with the same functionality. Meanwhile, at the central region of the UPM T221 genome, which encodes major and minor capsid proteins and core proteins, shows high conservation of the hexon protein. This conservation aligns with hexon crucial role in viral structure, suggesting that structural integrity is preserved in UPM T221. However, the shorter penton and fibre proteins in UPM T221 may affect the virus's ability to assemble or bind to host cells efficiently. In particular, the fibre protein, which mediates viral attachment to host cell receptors, is known to be a key determinant of host specificity and tissue tropism in adenoviruses (Grgić *et al*., 2014; Guardado-Calvo *et al*., 2007; Günes *et al*., 2012; Sohaimi & Hair-Bejo, 2021; Walters *et al*., 2002). A shorter fibre protein may reduce the virus's ability to infect certain cell types, thereby contributing to its non-pathogenic behaviour.

Moreover, the shared properties suggest that the hexon protein in both strains will elicit a similar immune response, making UPM T221 an ideal candidate for a live attenuated vaccine, as it retains the major antigenic determinants necessary for immune protection. Also, core proteins, such as pTP, play a crucial role in viral DNA replication, and the longer pTP in UPM T221 suggests that replication may be slightly altered compared to UPM04217. While pTP is essential for initiating viral DNA replication, an extended version of this protein could affect the replication rate or efficiency, contributing to the slower viral spread observed in UPM T221 infections. This slower replication could give the host immune system more time to mount a defensive response, leading to a more controlled and less severe infection, as seen in the *in vivo* pathogenicity study in SPF CEE. Nevertheless, the conservation of minor cement proteins (pVI, pIIIa, and pVIII) across both genomes implies that these proteins, essential for capsid stability and genome packaging, remain functionally intact, ensuring the structural integrity of the virus in both pathogenic and non-pathogenic forms.

This study proposed the critical role of WGS in enhancing our understanding of FAdV-8b strains and their application in vaccine development. Through the use of WGS, the UPM T221 strain was identified as a promising live attenuated vaccine candidate, owing to its genomic stability, reduced mutation rate, and non-pathogenic characteristics. By analysing the complete genome of UPM T221, specific genetic markers and variations associated with its attenuated virulence were identified. These findings support its potential to induce protective immunity in poultry without causing disease, a key feature for vaccine safety. In contrast, comparative genomic studies of pathogenic strains, such as UPM04217, revealed a higher accumulation of SNPs and genomic InDels, such as DBPs, pTP, DNA polymerase, ORF19, that likely contribute to their heightened virulence. These differences in genetic makeup, particularly in regions coding for proteins involved in viral replication and host interaction, help to distinguish attenuated strains like UPM T221 from virulent ones, such as UPM04217. WGS not only enables the identification of these genetic variations but also facilitates a deeper exploration of the structural and functional elements within the genome. In the case of UPM T221, the identification of specific genomic deletions and SNPs, especially in regions associated with immune evasion and replication, provides valuable insights into the molecular mechanisms behind its reduced pathogenicity. These genomic features are crucial for the rational design of safer and more effective vaccines, as they highlight the specific genetic modifications that can attenuate a virus while preserving its ability to stimulate immunity. Consequently, WGS proves to be an indispensable tool in vaccine research, offering comprehensive data on the genetic factors that govern virulence and attenuation in FAdV-8b, thereby guiding future efforts in developing vaccines tailored to combat viral infections in poultry industries with similar predominant serotypes with Malaysia.

CONCLUSION

The application of WGS and comparative genomic analysis provides invaluable insights into the molecular mechanisms underlying viral pathogenicity and vaccine development, particularly for genetically complex viruses like FAdV-8b. WGS enables comprehensive sequencing of viral genomes, identifying genetic variations, mutations, and structural features critical to understanding phenotypic differences. In the comparative study between non-pathogenic UPM T221 and pathogenic UPM04217, key genomic differences were identified, shedding light on their divergent virulence profiles. UPM T221, with a slightly larger genome, lower mutation rate, and stable core proteins, demonstrated reduced virulence. The analysis identified 149 SNPs, mainly in critical genes like fiber, essential for viral structure and early host interaction, and structural variations in ORF19's lipase enzyme, indicating reduced pathogenic potential. Conversely, UPM04217 exhibited a higher number of SNPs and insertions in enzymatic proteins like pTP and GAM-1, alongside disruptive mutations, suggesting enhanced virulence through increased replication efficiency or immune evasion. These findings underline the potential of UPM T221 as a safe and

effective vaccine candidate due to its genomic stability, while UPM04217 provides insights into mechanisms driving FAdV-8b virulence, as well as establishing novel framework for leveraging comparative genomics in vaccine candidate identification, offering groundbreaking perspectives for controlling adenoviral diseases in poultry.

Future research will involve targeted *in silico* studies of protein structures and physicochemical properties to further validate the genomic findings, and *in vivo* candidate vaccine potency and efficacy studies. Additionally, functional assays to assess the impact of identified mutations on viral replication and host interactions will enhance our understanding of FAdV-8b pathogenicity. Expanding comparative genomic studies to include other FAdV-8b strains will also help refine criteria for vaccine candidate selection. Integrating advanced bioinformatics and artificial intelligence tools to predict host-pathogen interactions could provide deeper insights into immune evasion mechanisms and guide rational vaccine design. These efforts aim to optimize the development of safe, effective, and long-lasting vaccines, offering sustainable solutions for controlling IBH and other adenoviral diseases in poultry populations globally.

ACKNOWLEDGEMENTS

This study was funded by Ministry of Higher Education (MOHE) of Malaysia under the High Centre of Excellence Grant [Grant No.: 5220002].

CONFLICT OF INTEREST

The authors have declared that no conflict of interest exists.

REFERENCES

Adair, B. M., Curran, W. L., & McFerran, J. B. (1979). Ultrastructural studies of the replication of fowl adenoviruses in primary cell cultures. *Avian Pathology*, *8*(2), 133–144. https://doi.org/10.1080/03079457908418336

Ahmed, S., Mariatulqabtiah, A. R., Bejo, M. H., Omar, A. R., Ideris, A., & Mat Isa, N. (2021). Molecular markers and phylogenetic analysis of UPMT27, a field isolate of the Malaysian fowl adenovirus associated with inclusion body hepatitis. *Pertanika Journal of Science and Technology*, *29*(1). https://doi.org/10.47836/pjst.29.1.29

Alemnesh, W., Hair-Bejo, M., Aini, I., & Omar, A. R. (2012). Pathogenicity of Fowl Adenovirus in Specific Pathogen Free Chicken Embryos. *Journal of Comparative Pathology*, *146*(2–3), 223–229. https://doi.org/10.1016/j.jcpa.2011.05.001

Alvarado, I. R., Villegas, P., El-Attrache, J., Jensen, E., Rosales, G., Perozo, F., & Purvis, L. B. (2007). Genetic characterization, pathogenicity, and protection studies with an avian adenovirus isolate associated with inclusion body hepatitis. *Avian Diseases*, *51*(1), 27–32. https://doi.org/10.1637/0005-2086(2007)051[0027:GCPAPS]2.0.CO;2

Andrew, S. (2010). *FASTQC. A quality control tool for high throughput sequence data*. (Version 0.11.5). Available from http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Azli, B., Salim, N. F., Omar, A. R., Hair-Bejo, M., Sohaimi, N. M., & Isa, N. M. (2023). Molecular Characterisation of Partial Structural Genes of Fowl Adenovirus Serotype 8b UPMT1901 Field Strain Isolate Associated with the Inclusion Body Hepatitis in Malaysia's Commercial Broiler Chickens. *Pertanika Journal of Tropical Agricultural Science*, *46*(3), 1003–1026. https://doi.org/10.47836/pjtas.46.3.15

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., & Pevzner, P. A. (2012). SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology*, *19*(5), 455–477. https://doi.org/10.1089/cmb.2012.0021

Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Research*, *27*(2), 573–580. https://doi.org/10.1093/nar/27.2.573

Cao, J. X., Krell, P. J., & Nagy, E. (1998). Sequence and transcriptional analysis of terminal regions of the fowl adenovirus type 8 genome. *The Journal of General Virology*, *79 (Pt 10*, 2507–2516. https://doi.org/10.1099/0022-1317-79-10-2507

Carver, T., Thomson, N., Bleasby, A., Berriman, M., & Parkhill, J. (2009). DNAPlotter: circular and linear interactive genome visualization. *Bioinformatics*, *25*(1), 119–120. https://doi.org/10.1093/bioinformatics/btn578

Chakraborty, T., & Barbuddhe, S. (2021). Enabling One Health solutions through genomics. *Indian Journal of Medical Research*, *153*(3), 273. https://doi.org/10.4103/ijmr.IJMR_576_21

Chen, H., Li, M., Liu, S., Kong, J., Li, D., Feng, J., & Xie, Z. (2022). Whole-genome sequence and pathogenicity of a fowl adenovirus 5 isolated from ducks with egg drop syndrome in China. *Frontiers in Veterinary Science*, *9*, 961793. https://doi.org/10.3389/fvets.2022.961793

Chiocca, S., Kurzbauer, R., Schaffner, G., Baker, A., Mautner, V., & Cotten, M. (1996). The complete DNA sequence and genomic organization of the avian adenovirus CELO. *Journal of Virology*, *70*(5), 2939–2949. https://doi.org/10.1128/jvi.70.5.2939-2949.1996

Chokyotager. (2021). *ORFFinder* (v1.8). https://github.com/Chokyotager/ORFFinder

Corredor, J. C., Krell, P. J., & Nagy, É. (2006). Sequence Analysis of the Left End of Fowl Adenovirus Genomes. *Virus Genes*, *33*(1), 95–106. https://doi.org/10.1007/s11262-005-0031-y

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E.,

DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., & Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, *10*(2). https://doi.org/10.1093/gigascience/giab008

Darling, A. C. E., Mau, B., Blattner, F. R., & Perna, N. T. (2004). Mauve: Multiple Alignment of Conserved Genomic Sequence With Rearrangements. *Genome Research*, *14*(7), 1394–1403. https://doi.org/10.1101/gr.2289704

Eid, A. A. M., Hussein, A., Hassanin, O., Elbakrey, R. M., Daines, R., Sadeyen, J.-R., Abdien, H. M. F., Chrzastek, K., & Iqbal, M. (2022). Newcastle Disease Genotype VII Prevalence in Poultry and Wild Birds in Egypt. *Viruses*, *14*(10), 2244. https://doi.org/10.3390/v14102244

Ellingford, J. M., Ahn, J. W., Bagnall, R. D., Baralle, D., Barton, S., Campbell, C., Downes, K., Ellard, S., Duff-Farrier, C., FitzPatrick, D. R., Greally, J. M., Ingles, J., Krishnan, N., Lord, J., Martin, H. C., Newman, W. G., O'Donnell-Luria, A., Ramsden, S. C., Rehm, H. L., … Whiffin, N. (2022). Recommendations for clinical interpretation of variants found in non-coding regions of the genome. *Genome Medicine*, *14*(1), 73. https://doi.org/10.1186/s13073-022-01073-3

Grafl, B., Garcia-Rueda, C., Cargill, P., Wood, A., Schock, A., Liebhart, D., Schachner, A., & Hess, M. (2018). Fowl aviadenovirus serotype 1 confirmed as the aetiological agent of gizzard erosions in replacement pullets and layer flocks in Great Britain by laboratory and in vivo studies. *Avian Pathology*, *47*(1), 63–72. https://doi.org/10.1080/03079457.2017.1367364

Grgić, H., Krell, P. J., & Nagy, É. (2014). Comparison of fiber gene sequences of inclusion body hepatitis (IBH) and non-IBH strains of serotype 8 and 11 fowl adenoviruses. *Virus Genes*, *48*(1), 74–80. https://doi.org/10.1007/s11262-013-0995-y

Guardado-Calvo, P., Llamas-Saiz, A. L., Fox, G. C., Langlois, P., & van Raaij, M. J. (2007). Structure of the C-terminal head domain of the fowl adenovirus type 1 long fiber. *Journal of General Virology*, *88*(9), 2407–2416. https://doi.org/10.1099/vir.0.82845-0

Günes, A., Marek, A., Grafl, B., Berger, E., & Hess, M. (2012). Real-time PCR assay for universal detection and quantitation of all five species of fowl adenoviruses (FAdV-A to FAdV-E). *Journal of Virological Methods*, *183*(2), 147–153. https://doi.org/10.1016/j.jviromet.2012.04.005

Hair-Bejo, M. (2005). Inclusion body hepatitis in a flock of a commercial broilers chickens. *Journal of Veterinary Malaysia*, *31*, 23–26.

Homonnay, Z., Jakab, S., Bali, K., Kaszab, E., Mató, T., Kiss, I., Palya, V., & Bányai, K. (2021). Genome sequencing of a novel variant of fowl adenovirus B reveals mosaicism in the pattern of homologous recombination events. *Archives of Virology*, *166*(5), 1477–1480. https://doi.org/10.1007/s00705-021-04972-9

Huang, Q., Ma, X., Huang, X., Huang, Y., Yang, S., Zhang, L., Cui, N., & Xu, C. (2019). Pathogenicity and complete genome sequence of a fowl adenovirus serotype 8b isolate from China. *Poultry Science*, *98*(2), 573–580. https://doi.org/10.3382/ps/pey425

Ip, H. S., Uhm, S., Killian, M. L., & Torchetti, M. K. (2023). An Evaluation of Avian Influenza Virus Whole-Genome Sequencing Approaches Using Nanopore Technology. *Microorganisms*, *11*(2), 529. https://doi.org/10.3390/microorganisms11020529

Itakura, C., Yasuba, M., & Goto, M. (1974). Histopathological Studies on Inclusion Body Hepatitis in Broiler Chickens. *The Japanese Journal of Veterinary Science*, *36*(4), 329–340. https://doi.org/10.1292/jvms1939.36.329

Joshi, N., & Fass, J. (2011). *Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files [Software]* (v. 1.33). https://github.com/najoshi/sickle

jstjohn. (2016). *SeqPrep: Tool for stripping adaptors and/or merging paired reads with overlap into single reads*. https://github.com/jstjohn/SeqPrep

Juliana, M. A., Nurulfiza, M. I., Hair-Bejo, M., Abdul Rahman, O., & Aini, I. (2014). Molecular characterization of fowl adenoviruses isolated from inclusion body hepatitis outbreaks in commercial broiler chickens in Malaysia. *Pertanika Journal of Tropical Agriculture Science*, *37*(4), 483–497.

Kapgate, S. S., Barbuddhe, S. B., & Kumanan, K. (2015). Next generation sequencing technologies: tool to study avian virus diversity. *Acta Virologica*, *59*(1), 3–13. https://doi.org/10.4149/av_2015_01_3

Kiss, I., Homonnay, Z. G., Mató, T., Bányai, K., & Palya, V. (2021). An overview on distribution of fowl adenoviruses. *Poultry Science*, *100*(5), 101052. https://doi.org/10.1016/j.psj.2021.101052

Kondo, M., Hirai, H., Furukawa, T., Yoshida, Y., Suzuki, A., Kawaguchi, T., & Che, F.-S. (2016). Frameshift Mutation Confers Function as Virulence Factor to Leucine-Rich Repeat Protein from Acidovorax avenae. *Frontiers in Plant Science*, *7*, 1988. https://doi.org/10.3389/fpls.2016.01988

Kulanayake, S., & Tikoo, S. (2021). Adenovirus Core Proteins: Structure and Function. *Viruses*, *13*(3), 388. https://doi.org/10.3390/v13030388

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular evolutionary genetics analysis across computing platform. *Molecular Biology and Evolution*, *35*(6), 1547–1549. https://doi.org/10.1093/molbev/msy096

Li, H., & Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)*, *26*(5), 589–595. https://doi.org/10.1093/bioinformatics/btp698

Ludwig, M. (2002). Functional evolution of noncoding DNA. *Current Opinion in Genetics & Development*, *12*(6), 634–639. https://doi.org/10.1016/S0959-437X(02)00355-6

Marek, A., Kaján, G. L., Kosiol, C., Benkő, M., Schachner, A., & Hess, M. (2016). Genetic diversity of species Fowl aviadenovirus D and Fowl aviadenovirus E. *The Journal of General Virology*, *97*(9), 2323–2332. https://doi.org/10.1099/jgv.0.000519

Marek, A., Kosiol, C., Harrach, B., Kaján, G. L., Schlötterer, C., & Hess, M. (2013). The first whole genome sequence of a Fowl adenovirus B strain enables interspecies comparisons within the genus Aviadenovirus. *Veterinary Microbiology*, *166*(1–2), 250–256. https://doi.org/10.1016/j.vetmic.2013.05.017

Mase, M., Nakamura, K., & Minami, F. (2012). Fowl Adenoviruses Isolated from Chickens with Inclusion Body Hepatitis in Japan, 2009–2010. *Journal of Veterinary Medical Science*, *74*(8), 1087–1089. https://doi.org/10.1292/jvms.11-0443

Mat Isa, N., Mohd Ayob, J., Ravi, S., Mustapha, N. A., Ashari, K. S., Bejo, M. H., Omar, A. R., & Ideris, A. (2019). Complete genome sequence of fowl adenovirus-8b UPM04217 isolate associated with the inclusion body hepatitis disease in commercial broiler chickens in Malaysia reveals intermediate evolution. *VirusDisease*, *30*(3), 426–432. https://doi.org/10.1007/s13337-019-00530-9

Meulemans, G., Boschmans, M., van den Berg, T. P., & Decaesstecker, M. (2001). Polymerase chain reaction combined with restriction enzyme analysis for detection and differentiation of fowl adenoviruses. *Avian Pathology*,

30(6), 655–660. https://doi.org/10.1080/03079450120092143

Mo, K., Lyu, C., Cao, S., Li, X., Xing, G., Yan, Y., Zheng, X., Liao, M., & Zhou, J. (2019). Pathogenicity of an FAdV-4 isolate to chickens and its genomic analysis. *Journal of Zhejiang University-SCIENCE B*, *20*(9), 740–752. https://doi.org/10.1631/jzus.B1900070

Narvaez, S. A., Harrell, T. L., Oluwayinka, O., Sellers, H. S., Khalid, Z., Hauck, R., Chowdhury, E. U., & Conrad, S. J. (2023). Optimizing the Conditions for Whole-Genome Sequencing of Avian Reoviruses. *Viruses*, *15*(9), 1938. https://doi.org/10.3390/v15091938

Niczyporuk, J. S., & Kozdruń, W. (2022). Current epidemiological situation in the context of inclusion body hepatitis in poultry flocks in Poland. *Virus Research*, *318*, 198825. https://doi.org/10.1016/j.virusres.2022.198825

Niu, D., Feng, J., Duan, B., Shi, Q., Li, Y., Chen, Z., Ma, L., Liu, H., & Wang, Y. (2022). Epidemiological survey of avian adenovirus in China from 2015 to 2021 and the genetic variability of highly pathogenic Fadv-4 isolates. *Infection, Genetics and Evolution*, *101*, 105277. https://doi.org/10.1016/j.meegid.2022.105277

Norina, L., Norsharina, A., Nurnadiah, A. H., Redzuan, I., Ardy, A., & Nor-Ismaliza, I. (2016). Avian Adenovirus Isolated From Broiler Affected With Inclusion Body Hepatitis. *Avian Adenovirus Isolated From Broiler Affected With Inclusion Body Hepatitis*, *Volume 7*, 121–126. http://www.dvs.gov.my/dvs/resources/user_14/MJVR_V7N2/MJVR-V7N2-p121-126.pdf

Ojkic, D., & Nagy, É. (2000). The complete nucleotide sequence of fowl adenovirus type 8. *Microbiology*, *81*(7), 1833–1837. https://doi.org/10.1099/0022-1317-81-7-1833

Pagel, K. A., Antaki, D., Lian, A., Mort, M., Cooper, D. N., Sebat, J., Iakoucheva, L. M., Mooney, S. D., & Radivojac, P. (2019). Pathogenicity and functional impact of non-frameshifting insertion/deletion variation in the human genome. *PLOS Computational Biology*, *15*(6), e1007112. https://doi.org/10.1371/journal.pcbi.1007112

Pallister, J., Wright, P. J., & Sheppard, M. (1996). A single gene encoding the fiber is responsible for variations in virulence in the fowl adenoviruses. *Journal of Virology*, *70*(8), 5115–5122. https://doi.org/10.1128/JVI.70.8.5115-5122.1996

Rashid, F., Xie, Z., Wei, Y., Xie, Z., Xie, L., Li, M., & Luo, S. (2024). Biological features of fowl adenovirus serotype-4. *Frontiers in Cellular and Infection Microbiology*, *14*. https://doi.org/10.3389/fcimb.2024.1370414

Rashid, F., Xie, Z. Z., Zhang, L., Luan, Y., Luo, S., Deng, X., Xie, L., Xie, Z. Z., & Fan, Q. (2020). Genetic characterization of fowl aviadenovirus 4 isolates from Guangxi, China, during 2017-2019. *Poultry Science*, *99*(9), 4166–4173. https://doi.org/10.1016/j.psj.2020.06.003

Sanrone. (2017). *GGisy* (v1.0). https://github.com/Sanrrone/GGisy.git

Schachner, A., Gonzalez, G., Endler, L., Ito, K., & Hess, M. (2019). Fowl Adenovirus (FAdV) Recombination with Intertypic Crossovers in Genomes of FAdV-D and FAdV-E, Displaying Hybrid Serological Phenotypes. *Viruses*, *11*(12), 1094. https://doi.org/10.3390/v11121094

Schachner, A., Matos, M., Grafl, B., & Hess, M. (2018). Fowl adenovirus-induced diseases and strategies for their control – a review on the current global situation. *Avian Pathology*, *47*(2), 111–126. https://doi.org/10.1080/03079457.2017.1385724

Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, *30*(14), 2068–2069. https://doi.org/10.1093/bioinformatics/btu153

Serruto, D., Serino, L., Masignani, V., & Pizza, M. (2009). Genome-based approaches to develop vaccines against bacterial pathogens. *Vaccine*, *27*(25–26), 3245–3250. https://doi.org/10.1016/j.vaccine.2009.01.072

Shamanna, V., Srinivas, S., Couto, N., Nagaraj, G., Sajankila, S. P., Krishnappa, H. G., Kumar, K. A., Aanensen, D. M., & Lingegowda, R. K. (2024). Geographical distribution, disease association and diversity of Klebsiella pneumoniae K/L and O antigens in India: roadmap for vaccine development. *Microbial Genomics*, *10*(7). https://doi.org/10.1099/mgen.0.001271

Sheppard, M., & Trist, H. (1992). Characterization of the avian adenovirus penton base. *Virology*, *188*(2), 881–886. https://doi.org/10.1016/0042-6822(92)90546-2

Slatko, B. E., Gardner, A. F., & Ausubel, F. M. (2018). Overview of Next-Generation Sequencing Technologies. *Current Protocols in Molecular Biology*, *122*(1). https://doi.org/10.1002/cpmb.59

Sohaimi, N. M., & Clifford, U. C. (2021). Fowl adenovirus in chickens: Diseases, epidemiology, impact, and control strategies to the Malaysian poultry industry- a review. *Journal of World's Poultry Research*, *11*(3), 387–396. https://doi.org/10.36380/jwpr.2021.46

Sohaimi, N. M., & Hair-Bejo, M. (2021). A recent perspective on fiber and hexon genes proteins analyses of fowl adenovirus towards virus infectivity - A review. *Open Veterinary Journal*, *11*(4), 569. https://doi.org/10.5455/OVJ.2021.v11.i4.6

Sohaimi, N. M., Omar, A. R., & Ideris, A. (2018). Molecular detection and pathogenicity of fowl adenovirus. *International Journal of Agricultural Sciences and Veterinary Medicine*, *6*(1).

Steer, P. A., Kirkpatrick, N. C., O'Rourke, D., & Noormohammadi, A. H. (2009). Classification of fowl adenovirus serotypes by use of high-resolution melting-curve analysis of the hexon gene region. *Journal of Clinical Microbiology*, *47*(2), 311–321. https://doi.org/10.1128/JCM.01567-08

Steinbrück, L., Klingen, T. R., & McHardy, A. C. (2014). Computational Prediction of Vaccine Strains for Human Influenza A (H3N2) Viruses. *Journal of Virology*, *88*(20), 12123–12132. https://doi.org/10.1128/JVI.01861-14

Sturdevant, G. L., Kari, L., Gardner, D. J., Olivares-Zavaleta, N., Randall, L. B., Whitmire, W. M., Carlson, J. H., Goheen, M. M., Selleck, E. M., Martens, C., & Caldwell, H. D. (2010). Frameshift Mutations in a Single Novel Virulence Factor Alter the In Vivo Pathogenicity of Chlamydia trachomatis for the Female Murine Genital Tract. *Infection and Immunity*, *78*(9), 3660–3668. https://doi.org/10.1128/IAI.00386-10

Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, *22*(22), 4673–4680. https://doi.org/10.1093/nar/22.22.4673

Voelkerding, K. V, Dames, S. A., & Durtschi, J. D. (2009). Next-Generation Sequencing: From Basic Research to Diagnostics. *Clinical Chemistry*, *55*(4), 641–658. https://doi.org/10.1373/clinchem.2008.112789

Walters, R. W., Freimuth, P., Moninger, T. O., Ganske, I., Zabner, J., & Welsh, M. J. (2002). Adenovirus Fiber Disrupts CAR-Mediated Intercellular Adhesion Allowing Virus Escape. *Cell*, *110*(6), 789–799. https://doi.org/10.1016/S0092-8674(02)00912-1

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org

Zerbino, D. R., & Birney, E. (2008). Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research*, *18*(5), 821–829. https://doi.org/10.1101/gr.074492.107

Zhang, H., Meltzer, P., & Davis, S. (2013). RCircos: an R package for Circos 2D track plots. *BMC Bioinformatics*, *14*(1), 244.

https://doi.org/10.1186/1471-2105-14-244

Zhu, R., Wang, Y., Zhang, H., Yang, J., Fan, J., Zhang, Y., Wang, Y., Li, Q., Zhou, X., Yue, H., Qi, Y., Wang, S., Chen, T., Zhang, S., & Hu, R. (2024). Deletion of the B125R gene in the African swine fever virus SY18 strain leads to an A104R frameshift mutation slightly attenuating virulence in domestic pigs. *Virus Research*, *343*, 199343. https://doi.org/10.1016/j.virusres.2024.199343

Zou, X.-H., Bi, Z.-X., Guo, X.-J., Zhang, Z., Zhao, Y., Wang, M., Zhu, Y.-L., Jie, H.-Y., Yu, Y., Hung, T., & Lu, Z.-Z. (2018). ☆DNA assembly technique simplifies the construction of infectious clone of fowl adenovirus. *Journal of Virological Methods*, *257*, 85–92. https://doi.org/10.1016/j.jviromet.2018.04.001