Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Complete-Genome Analysis of Beak and Feather Disease Virus in Iran: Evidence of Recombination and Selection

Sanaz Majidi
SANA Institute for Avian Health and Diseases Research
Mohammad Hossein Mokhtarian
SANA Institute for Avian Health and Diseases Research
Seyed Hadi Pourbakhsh
SANA Institute for Avian Health and Diseases Research
Peyman Younesi
SANA Institute for Avian Health and Diseases Research
Ali Salavati
Faculty of Veterinary Medicine, University of Tehran
Marva Mehdizadeh Jafari
SANA Avian Hospital
Jamshid Razmyar
jrazmyar@ut.ac.ir

Faculty of Veterinary Medicine, University of Tehran,

Research Article

Keywords: Complete Genome, Beak and feather disease virus (BFDV), Recombination, phylogeny, Iran

Posted Date: November 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11772812/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations:
No competing interests reported.

Figures are available in the Supplementary Files section.

Complete-Genome Analysis of Beak and Feather Disease Virus in Iran: Evidence of Recombination and Selection

Sanaz Majidi(https://orcid.org/0000-0003-0365-7713)[1]

Mohammad Hossein Mokhtarian(https://orcid.org/0000-0003-4511-6376)[1]

Seyed Hadi Pourbakhsh(https://orcid.org/0009-0001-1098-5232)[2]

Peyman Younesi(https://orcid.org/0009-0004-5902-7571)[1]

Ali Salavati[3]

Marva Mehdizadeh Jafari[4]

Jamshid Razmyar(https://orcid.org/0000-0002-1247-4591)[5*]

[1]SANA Institute for Avian Health and Diseases Research, Department of Viral Diseases, Tehran, Iran

[2]SANA Institute for Avian Health and Diseases Research, Department of Bacterial Diseases, Tehran, Iran

[3]Student Association of Avian Health and Diseases, Faculty of Veterinary Medicine, University of Tehran, Tehran, Iran.

[4]Molecular Biology Department, SANA Avian Hospital, Tehran, Iran.

[5]Department of Avian Health and Diseases, Faculty of Veterinary Medicine, University of Tehran, Tehran, Iran.

[*]Corresponding author: jrazmyar@ut.ac.ir

Address: Department of Avian Health and Diseases, Faculty of Veterinary Medicine, University of Tehran, Tehran, Iran. Tel:+982144121314-330

Declaration of generative AI

During the preparation of this work, the authors used OpenAI's ChatGPT in order to check grammar and improve language clarity. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Abstract

Beak and feather disease virus (BFDV) is regarded as a significant threat to avian biodiversity and the exotic bird trade. Despite its impact, complete-genome data from the Middle East remain limited. This study reports four complete BFDV genomes from Iran, deposited in GenBank. BFDV DNA was detected in 222 of 400 (55.5%) feather samples submitted to a specialist avian referral hospital; a

detection rate among submitted samples and not a national prevalence estimate; the highest rate was observed in Cockatiels *(Nymphicus hollandicus)* at 76%.

Recombination analysis suggested genomic mosaicism in SBFDV-4 sequence, a putative recombinant that showed closest sequence similarity to Polish and Thai lineages in the putative major and minor parental regions, respectively, consistent with putative parental-like contributions. Furthermore, the replication-associated protein (Rep) was highly conserved (86% to 99% amino acid identity to the reference) with no substitutions observed in the annotated motifs examined, whereas the capsid protein (Cap) showed amino acid variability at its N-terminus, a region annotated as containing the nuclear localization signal and predicted to contain epitopes; possible effects on antigenicity warrant experimental validation. In a phylogeny of the 548 available complete BFDV genomes, the four Iranian genomes were placed within Clade I.a but did not form a monophyletic group. These findings provide baseline genomic data that may inform future surveillance and highlight the need for further sampling; the contribution of international trade and the extent of recombination were not directly assessed.

Keywords: Complete Genome, Beak and feather disease virus (BFDV), Recombination, phylogeny, Iran

1. INTRODUCTION

Beak and feather disease virus (BFDV) is a member of the *Circovirus* genus in the *Circoviridae* family. It is the cause of psittacine Beak and feather disease (PBFD), a severe and often fatal condition in psittacine birds [1, 2]. The virus has a compact, circular, single-stranded DNA (ssDNA) genome of about two kilobases. This genome encodes two major proteins: the replication-associated protein (Rep) on the virion-sense strand and the capsid protein (Cap) on the complementary strand[3, 4]. BFDV was first described in the 1970s. The disease is now recognized worldwide, and international trade in exotic birds, both legal and illegal, has been implicated in its spread[5, 6].

The clinical presentation of BFDV infection is highly variable, ranging from subclinical carriage to severe, chronic disease characterized by symmetrical feather loss, beak and claw dystrophy, and deformities[7, 8]. Crucially, BFDV-induced immunosuppression renders infected birds susceptible to

secondary infections, which is a primary cause of mortality[9]. Asymptomatic carriers play a significant role in the silent spread of the virus, complicating control efforts and posing a substantial threat to aviculture and the conservation of endangered parrot populations [10, 11].

A key factor underpinning the successful global spread of BFDV is its significant genetic plasticity. The virus exhibits a high mutation rate and is prone to recombination, events that have been reported in the cap gene and proposed to contribute to host-switching and immune evasion[12, 13]. Studies from Australia, New Zealand, Poland, and China have revealed a complex landscape of co-circulating genotypes and the emergence of novel strains, often linked to captive breeding facilities and the introduction of new birds[3, 14-16]. While phylogenetic analyses based on individual genes, such as *rep* or *cap*, have provided insights, complete-genome sequencing offers a higher-resolution view of viral evolution, improves the detection of recombination events, and can inform phylogeographic inference.

Despite the global recognition of BFDV, epidemiological data remain uneven, with significant surveillance gaps in many regions, including parts of the Middle East. In Iran, the presence of BFDV has been confirmed by molecular and histopathologic methods[17, 18]. However, comprehensive molecular characterization based on complete-genome sequences has been limited. Given the critical role of geographic isolation and host species in shaping BFDV diversity, it is important to characterize the BFDV strains detected in Iran's psittacine populations. Such knowledge may help in assessing the risk of viral spillover, understanding local evolutionary dynamics, and informing national biosecurity and conservation strategies. Significant progress has been made in documenting the molecular status of BFDV within Iranian avian populations. A previous study by Dolatyabi et al. (2022) surveyed over a thousand birds, providing foundational insights into viral prevalence and mutation patterns by analyzing specific *rep* and *cap* gene segments[19]. While their work elucidated critical amino acid substitutions and codon usage biases, the reliance on partial sequences inherently limits a comprehensive understanding of the virus's broader genomic architecture.

To address this knowledge gap, the present study employs complete genome sequencing to characterize BFDV sequences from Iran. By moving beyond sub-genomic analysis, we aimed to evaluate the genetic

diversity, putative recombination events, and phylogenetic positioning of the viruses detected within a global context. These findings provide genomic data on BFDV in Iran and may contribute to international surveillance and management of this significant avian pathogen.

2. MATERIALS AND METHODS

2.1. Sample Collection

Between 2023 and 2025, we conducted a cross-sectional study at the SANA Avian Hospital, collecting a total of 400 feather samples from 11 psittacine species presented to the hospital.

2.2. Molecular Characterization and Complete-Genome Sequencing

Total genomic DNA was isolated from the feather follicles of the BFDV-positive samples selected for sequencing using the MBST DNA extraction kit (MBST, Iran). We employed a multiprimer PCR strategy using four previously published primer sets (Table 1) to generate overlapping amplicons spanning the entire BFDV genome, including both the *rep* and *cap* genes.

The PCR amplifications were carried out in 25 µL reactions containing 100 ng of template DNA, 12.5 µL of 2× Red Master Mix (Ampliqon, Denmark), and 10 pmol of each primer. The PCR cycling included an initial denaturation at 95°C for 5 minutes, followed by 35 cycles of 95°C for 30 seconds, 58°C for 30 seconds, and 72°C for 1 minute, concluding with a final extension at 72°C for 10 minutes. The resulting amplicons were visualized on a 1.5% agarose gel to confirm specificity and size before being purified. We selected a subset of positive samples for detailed genetic characterization. To include genomes from different host species, one strongly positive sample from each of four key species was chosen: the African Grey Parrot (Psittacus erithacus), Lovebird (Agapornis roseicollis), Cockatiel (Nymphicus hollandicus), and Alexandrine parakeet (Psittacula eupatria). These four species were selected to include genomes from different hosts among the BFDV-positive samples. Because only one genome was analyzed per species, these genomes are not intended to represent within species or national BFDV diversity.The four distinct PCR products from each selected host species were submitted to Macrogen Inc. (Seoul, South Korea) for bidirectional Sanger sequencing. Sequencing reads were mapped and assembled against the BFDV reference genome (Accession Number NC_001944.1) using

UGENE (version 52.1). These complete-genome sequences form the basis for our subsequent phylogenetic and evolutionary analyses.

Table 1. Comprehensive list of primers used for PCR amplification and bidirectional Sanger sequencing of the Rep and Cap genes. Product lengths and target regions are specified to ensure complete coverage of the viral genome.

Primer name	Gene	Sequence	Product length (bp)	Target	Ref.
Rep-F Rep-R	(*rep*) Replication Protein	AACCCTACAGACGGCGAG GTCACAGTCCTCCTTGTACC	717	initial screening & sequencing	[20]
Cap-F Cap-R	(*cap*) Capsid Protein	GAGGGCTGTAGTCTGTGTCG TACCGCAGACGTCACATCAG	799	Sequencing	
Cap-Rep F Cap-Rep R	Cap-Rep junction	ACCGTCGAAAGGTGCCAGCG ACGGCAGCAACAGCTCCGTC	956	Sequencing	
Rep-Cap F Rep-Cap R	Rep-Cap junction	CGGAGCTGTTGCTGCCGTGA ACCCGCTGGCACCTTTCGAC	1079	Sequencing	

2.3. Geographic and Heatmap Visualization of Sequence Identity

To illustrate the geographic distribution and comparative sequence similarity of four newly sequenced BFDVs, we generated both choropleth world maps and a heatmap. BLASTn (https://blast.ncbi.nlm.nih.gov/Blast.cgi) was used to compute nucleotide percent identity between each sequence and publicly available GenBank sequences, and the results were grouped by country of origin. Data preprocessing and aggregation were performed in Python (v3.11) using the pandas library. Choropleth maps were created for each sequence using plotly.express with the Natural Earth projection, applying the perceptually uniform “magma” color scale and consistent color normalization across maps for direct comparison. A heatmap was generated using seaborn, showing country–sequence mean identity values in matrix form, with the same color scale applied. Country ordering was based on descending mean identity to emphasize geographic patterns.

2.4. Complete-Genome Phylogenetic Inference

To reconstruct the evolutionary relationships of Beak and feather disease virus (BFDV; TaxID: 77856), 548 complete nucleotide sequences (including the four Iranian genomes) were retrieved from the NCBI Virus database. Sequence identifiers were standardized using a unified nomenclature that incorporated accession numbers, host species, geographic locations, and collection dates. The *Raven circovirus* (Accession: DQ146997.1) was utilized as an outgroup to reroot the phylogeny [21].

Multiple sequence alignment was performed using Clustal Omega [22]. To reduce researcher-induced bias and enhance reproducibility, automated alignment trimming was executed via ClipKIT (https://clipkit.genomelybio.com) using smart-gap (default) mode [23]. Phylogenetic relationships were inferred using the maximum-likelihood (ML) method implemented in IQ-TREE version 3. The best-fitting nucleotide substitution model was selected automatically using ModelFinder (-m MFP) based on the Bayesian information criterion (BIC), allowing objective determination of the most appropriate substitution model and associated parameters, including base frequencies and among-site rate heterogeneity [24, 25]. Because the trimmed alignment retained invariant sites, ascertainment bias correction was not applied. Statistical support for inferred clades was assessed using 1,000 ultrafast bootstrap replicates (-B 1000). To reduce potential overestimation of bootstrap support and improve reliability in this large dataset, a bootstrap-based nearest-neighbor interchange (BNNI) correction was applied during bootstrap analysis (--bnni), an approach shown to provide support values comparable to those of standard non-parametric bootstrap while being computationally efficient. The maximum-likelihood tree produced by IQ-TREE was later rerooted in FigTree using *Raven circovirus* (DQ146997.1), without altering topology or branch support.

2.5. Mutation Profile Analysis of Iranian BFDV Sequences (2025)

Sequence Datasets and Comparative Reference Strains. The molecular analysis was conducted using eight proteins of four Iranian BFDV sequences evaluated in 2025: four encoding the putative Replication-associated protein (rep/ORF1) (Accession Nos. YBO34886, YBO34887, YBO34889, YBO34891) and four encoding the Capsid protein (cap/ORF2) (Accession Nos. YBO34885, YBO34888, YBO34890, YBO34892). These sequences were compared against the prototypic Australian reference strains for Rep (AAC69861.1) and Cap (AAC69862.1).

To evaluate the Iranian sequences within the framework of global BFDV diversity, a comprehensive dataset of amino acid sequences from diverse geographical regions, including Australia, Brazil, China, and Southern Africa, was integrated into the substitution profiling. Furthermore, the temporal gap between the 1998 reference sequences and the 2025 Iranian sequences was analyzed to contextualize the accumulation of genetic substitutions and the potential role of genetic drift over 27 years.

Amino Acid Sequence Analysis. In silico Functional Domain Mapping. To locate amino acid variations relative to annotated functional regions, substitutions were compared to the Australian reference strain and mapped to annotated domains. For ORF1 (*rep*), this included motifs associated with Rolling-Circle Replication and enzymatic activities (ATPase/GTPase), such as the GXXGXGKS Walker A motif. For ORF2 (*cap*), the analysis focused on the N-terminal bipartite Nuclear Localization Signals (NLS), given its reported roles in nuclear localization and antigenicity.

2.6. Recombination and pairwise identity

To investigate potential recombination as a source of the observed genetic diversity, the complete genome sequences were screened for recombination events using RDP5 (Recombination Detection Program). Seven algorithms (RDP, GENECONV, BootScan, MaxChi, Chimaera, SiScan, and 3Seq) were employed. A recombination event was considered statistically significant only if it was detected by at least five of these algorithms.

To further examine the putative recombinant SBFDV-4, Simplot (version 3.5.1) was utilized. A sliding window analysis (200 bp window, 20 bp step) was performed to map nucleotide identity between the recombinant virus and its putative major and minor parental lineages. For the remaining Iranian sequences (SBFDV-1, -2, -3), the closest relatives identified in the preliminary phylogenetic analysis, along with representative sequences from major global clades, were included in the Simplot analysis for comparative context. This integrated approach allowed approximate localization of putative breakpoints and visualization of similarity to the putative parental-like lineages across the genome.

3. RESULTS

3.1. Molecular Characterization ,Complete-Genome Sequencing

proportion of positive specimens differed markedly among host species. Cockatiels (Nymphicus hollandicus) accounted for the largest share of detections, with 146 of 192 samples testing positive. Lower but still substantial detection rates were recorded in Rosy-faced Lovebirds (Agapornis roseicollis; 27/43) and Alexandrine Parakeets (Psittacula eupatria; 21/48). Species-specific results are summarized in Table 2. Complete genomes were obtained from one BFDV-positive sample of each of four host species, including the African Grey Parrot (Psittacus erithacus), Rosy-faced Lovebird (Agapornis roseicollis), Cockatiel (Nymphicus hollandicus), and Alexandrine Parakeet (Psittacula eupatria), providing a dataset that included different host species for phylogenetic and evolutionary analyses.

Table 2. Comparative screening of BFDV infection across eleven psittacine species. Out of 400 submitted samples.

Species	Number	Positive BFDV
Cockatiel (*Nymphicus hollandicus*)	192	146
Alexandrine parakeet (*Psittacula eupatria*)	48	21
Lovebird (*Agapornis roseicollis*)	43	27
Green-cheeked parakeet (*Pyrrhura molinae*)	36	7
Grey parrot (*Psittacus erithacus*)	29	16
Budgerigar (*Melopsittacus undulatus*)	15	3
Monk parakeet (*Myiopsitta monachus*)	15	2
Lorikeet (*Trichoglossus moluccanus*)	7	0
Macaw (*Ara macao*)	7	0
Cockatoo (*Cacatua galerita*)	5	0
Macaw (*Ara ararauna*)	3	0

3.2. Identification and Genomic Submission

Full-genome characterization was performed for four Iranian BFDV sequences (Table 3). The study represented diverse avian hosts, including the African Grey Parrot (SBFDV-4; PX470187), Lovebird (SBFDV-1; PX470188), Cockatiel (SBFDV-2; PX470189), and Alexandrine Parakeet (SBFDV-3; PX470190). The successful sequencing and GenBank submission of these sequences provide genomic

data from four psittacine hosts in Iran; with one genome per species, host-associated genetic variation cannot be assessed from these data.

Table 3. Detailed information on the four Iranian Beak and feather disease virus (BFDV) nucleotide and protein sequences characterized by complete-genome sequencing.

Sequence ID	Identifier	Host	Nucleotide accession no.	Protein accession no.
Seq 1	Lovebird/IR/SBFDV-1/2025	*Agapornis roseicollis*	PX470188	Cap: YBO34888 Rep: YBO34887
Seq 2	Cockatiel/IR/SBFDV-2/2025	*Nymphicus hollandicus*	PX470189	Cap: YBO34890 Rep: YBO34889
Seq 3	Alexandrine-parakeet/IR/SBFDV-3/2025	*Psittacula eupatria*	PX470190	Cap: YBO34892 Rep: YBO34891
Seq 4	Grey parrot/IR/SBFDV-4/2025	*Psittacus erithacus*	PX470187	Cap: YBO34885 Rep: YBO34886

3.3. Geographic Distribution and Comparative Sequence Identity Analysis

The genetic relationship between the four newly sequenced Iranian BFDV sequences and global strains was evaluated using nucleotide percent identity calculations, visualized in a heatmap (Figure 1) and choropleth maps (Supplementary Figure 1).

The Iranian sequences (SBFDV-1 to SBFDV-4) showed a range of nucleotide identity with sequences retrieved from GenBank, with mean identity values spanning from approximately 85% to 92.55%. The heatmap analysis indicated that the highest mean sequence identity for the Iranian sequences SBFDV-1, SBFDV-2, SBFDV-3, and SBFDV-4 was with strains from the USA (92.55%), Italy (92.53%), South Africa (91.77%), and the USA (91.44%), respectively.

All four Iranian sequences showed 88.56% to 90.68% mean nucleotide identity with Saudi Arabian strains; this similarity alone does not establish a shared evolutionary trajectory or recent dispersal within the Middle East. In contrast, the lowest identity values (approximately 85%) were consistently observed in comparison with strains from China, indicating lower mean nucleotide identity with the Chinese sequences available in GenBank. SBFDV-4, detected from an African Grey Parrot, showed lower mean identity than the other three Iranian sequences across most countries in the heatmap.

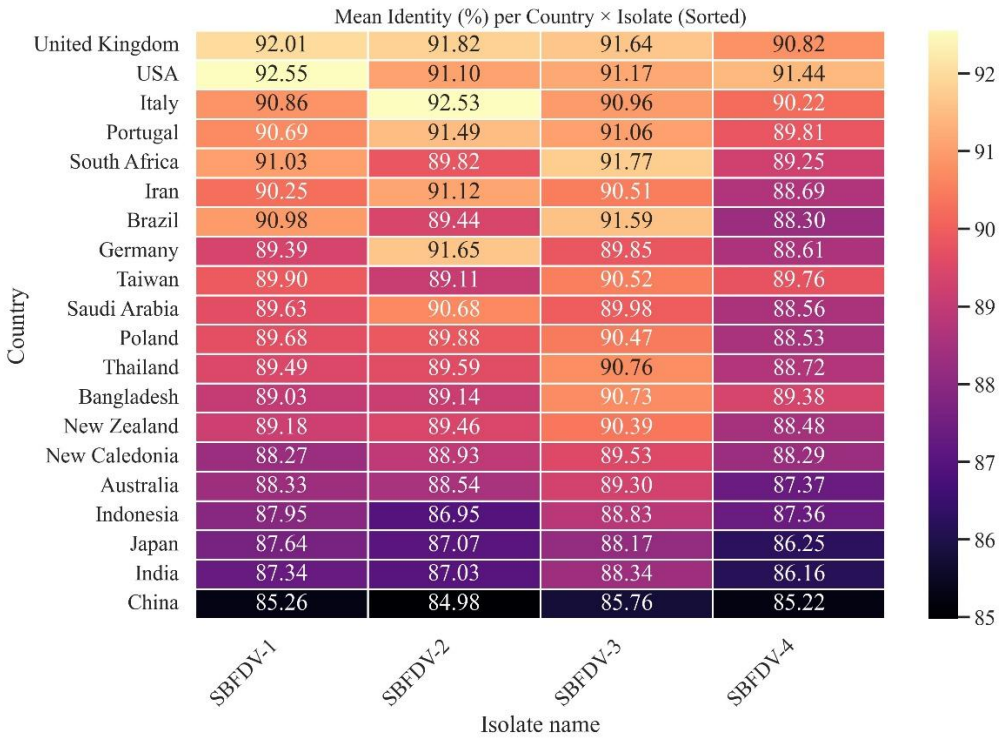


Figure 1 **Heatmap of mean sequence identity between BFDV isolates and GenBank sequences by country.** *heatmap showing the mean nucleotide percent identity between each BFDV isolate and GenBank reference sequences, aggregated by country of origin. Rows represent countries, and columns represent the four isolates. Each cell is color-coded using the "magma" colormap to indicate identity values, with darker shades representing lower similarity. Numeric values are annotated within cells, and countries are ordered by descending mean identity to highlight geographic trends. The heatmap was generated using Seaborn in Python and exported as a high-resolution image.*

3.4. Complete-Genome Phylogenetic Inference

To determine the evolutionary placement of the Iranian sequences, a Maximum Likelihood (ML) phylogenetic tree was constructed using 548 available complete BFDV nucleotide sequences, including the four Iranian genomes (Figure 2). The phylogenetic tree was rooted with *Raven circovirus* as the outgroup, and the TN+F+R5 nucleotide substitution model was selected as the best-fit model by ModelFinder based on Bayesian Information Criterion (BIC).

In this dataset, the Iranian SBFDV sequences were positioned within Clade I.a but did not form a unified monophyletic group. Their distribution across different subgroups suggests genetic variation among the Iranian viruses. SBFDV-3, obtained from an Alexandrine parakeet, and SBFDV-4, obtained from an African grey parrot, grouped together in a subgroup and shared their closest phylogenetic relationships with sequences from Taiwan and the USA. Conversely, SBFDV-1 and SBFDV-2, from a lovebird and cockatiel, respectively, occupied separate branches within Clade I.a. In this tree, SBFDV-2 was most closely related to Saudi Arabian sequences, whereas SBFDV-1 was most closely related to Australian sequences; these are closest relatives among the sampled genomes and do not indicate geographic origin**.**

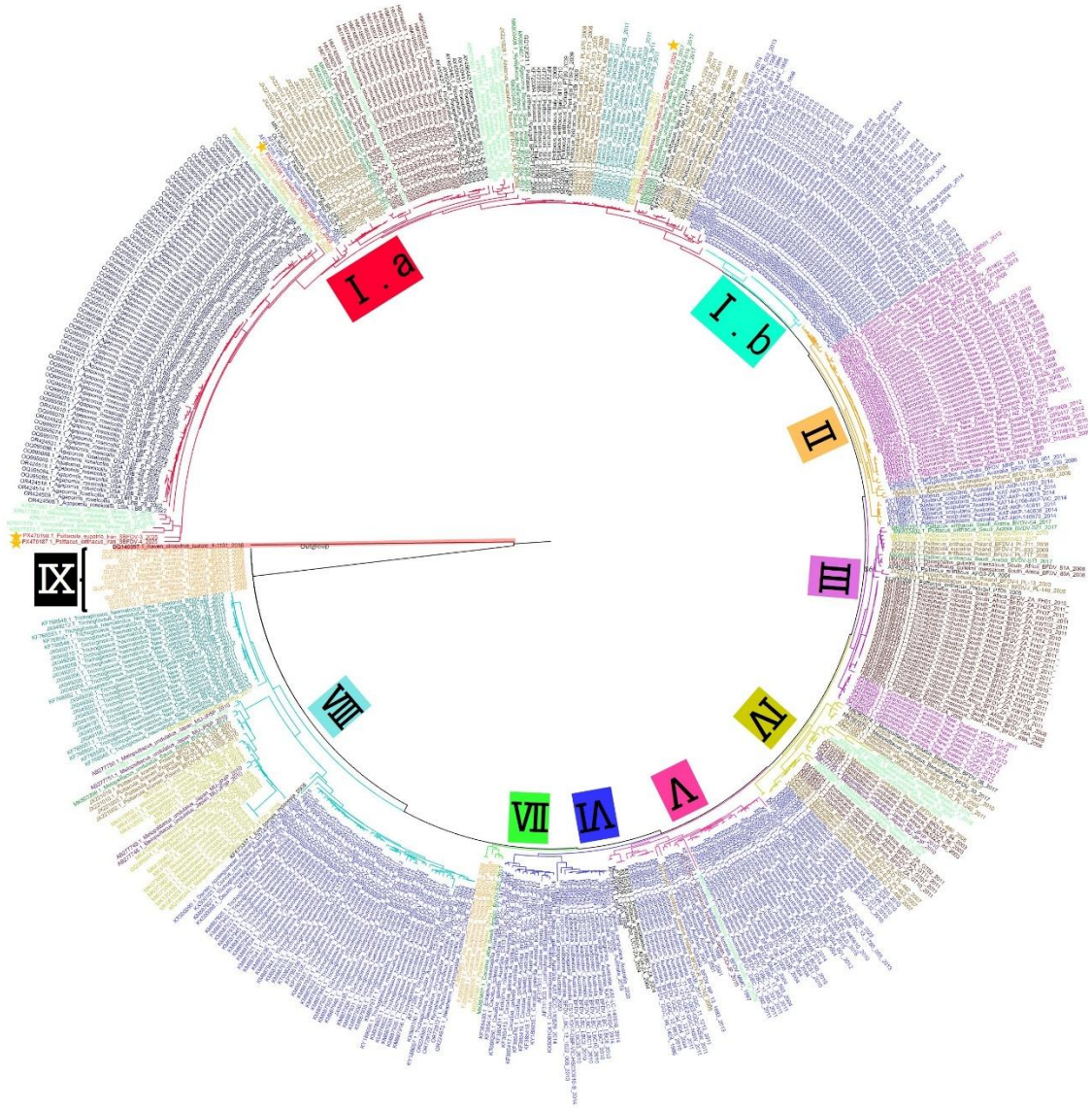


Figure 2 ***Global Phylogenetic Landscape of BFDV featuring Iranian Isolates.*** *The evolutionary relationships were reconstructed using 548 complete nucleotide sequences of Beak and Feather Disease Virus (BFDV) retrieved from the NCBI Virus database. Multiple sequence alignment was performed using Clustal Omega followed by automated trimming via ClipKIT to ensure high-fidelity homology assessment. The tree was constructed using IQ-TREE version 3, identified as the optimal parametric fit for the Circoviridae family. Raven circovirus (Accession: DQ146997.1) was utilized as the outgroup to root the phylogeny. The Iranian isolates characterized in the current study (SBFDV-1 to SBFDV-4) are highlighted in red text and marked with a star symbol, illustrating their distribution across distinct global clades (I–IX). Branch visualization and annotation were executed using the FigTree v1.4.4 (available at http://tree.bio.ed.ac.uk/software/figtree/)*

3.5. Mutation Profile of Iranian Sequences Compared to the Reference Strain

Rep Protein (ORF1) Mutation Profile. The Rep protein is historically recognized as the most conserved region of the Beak and feather disease virus (BFDV) genome, a characteristic attributed to the intense purifying selective pressure required to maintain its replicative function. In the present study, the four Iranian sequences (YBO34886, YBO34887, YBO34889, and YBO34891) showed high amino

acid identity (86% to 99%) relative to the reference sequence (AAC69861.1), consistent with the high conservation levels reported in global BFDV studies. Annotated functional motifs were examined using the UniProt database (Accession: Q9YUD3). Specifically, the P-loop (Walker A motif) sequence (GXXGXGKS), essential for ATP binding, showed no substitutions in any of the four Iranian sequences. This includes the Lysine (K) residue, which is reported to be important for the enzyme's catalytic activity; previous functional studies reported that mutations at this site significantly reduce both ATPase and GTPase activities [26].

Furthermore, other motifs annotated in UniProt (including the Walker B and the novel GYDG motifs) showed no substitutions. These regions have been implicated in the coordination of the magnesium ion (Mg^{2+}) and overall enzymatic efficiency during the initiation of rolling-circle replication (RCR) [26, 27]. While minor amino acid substitutions were observed, they were located outside the annotated motifs examined here and may represent ordinary sequence drift as described for other geographical lineages; their functional relevance was not assessed. Overall, the high amino acid identity between the 1998 reference strain and these 2025 Iranian sequences suggests conservation of the Rep protein within this limited dataset (four Iranian sequences compared with one reference), consistent with the evolutionary constraints reported for BFDV; functional integrity was not experimentally assessed.

Cap Protein (ORF2) Mutation Profile. The Cap protein, encoded by ORF2, exhibits a notably lower degree of conservation than the Rep protein, a divergence that may reflect selection pressure exerted by the host's immune system. In the present study, the four Iranian Cap sequences (YBO34885, YBO34888, YBO34890, and YBO34892) showed multiple amino acid substitutions, particularly within the N-terminus. This domain, as characterized in the UniProt database (Accession: Q9YUC8), contains the bipartite NLSs consisting of Arginine (Arg)-rich clusters. These findings align with previous research describing the N-terminus as the primary antigenic determinant and the most variable region of the cap gene [28, 29].

The substitutions observed in the Iranian sequences occur within a region that includes the annotated NLS and that has been predicted to contain major epitopes. Such substitutions could potentially alter the virus's antigenic profile and might also influence NLS-mediated nuclear trafficking and the

hypothesized co-translocation of the Rep protein (1); these possibilities were not tested in this study and warrant experimental validation. Compared with the 1998 Australian reference (AAC69862.1) and other global strains, the 2025 Iranian sequences show a degree of Cap divergence in line with that typically described for BFDV. This level of heterogeneity is consistent with the evolutionary patterns observed in Australian and South African lineages, where Cap variability has been associated with antigenically distinct subgroups [13]. Ultimately, the substitutions observed within the UniProt-annotated N-terminal region indicate heterogeneity among the four Iranian Cap sequences examined; whether this reflects adaptation, for example to local host-pathogen interactions, is a hypothesis that cannot be addressed with the present data.

3.6. Genomic Mosaicism and Sequence Conservation

The recombination and Simplot analyses are consistent with a contribution of recombination to the evolutionary history of the Iranian genomes examined here. A prominent finding is that SBFDV-4 (Grey Parrot/2025) was flagged as a putative recombinant, exhibiting the closest sequence similarity to Polish and Thai lineages in the recombination analysis, consistent with putative major (Polish-like; JX221003.1) and minor (Thai-like; FJ685989.1) parental contributions (Supplementary Table 1, event 42). These assignments reflect similarity to the closest available sequences and do not establish the geographic origin or transmission route of the parental viruses. Additionally, the analysis flagged SBFDV-2 as the sequence most closely related to the putative minor parental contribution for a Polish strain (Supplementary Table 1, event 48; the Polish sequence is dated 2006 in its label, earlier than SBFDV-2, so SBFDV-2 can represent only a related lineage); this does not establish genetic exchange between Iranian and European virus populations or its direction.

Simplot similarity curves indicate that the putative recombination signal was located mainly in the cap region (Figure 3). Similarity shifts in this region could reflect recombination or higher sequence variability; adaptation to host immune systems cannot be inferred from these data. Conversely, the *rep* gene and the Origin of Replication (Ori) maintain high conservation (Rep amino acid identity, 86% to 99%) and stability in Simplot profiles, which is consistent with functional constraint on replication-associated regions; no formal selection analysis was performed. These observations suggest greater

variability in the cap region than in rep among the four genomes examined; the frequency of recombination and the dynamics of Iranian BFDV populations cannot be inferred from four genomes.

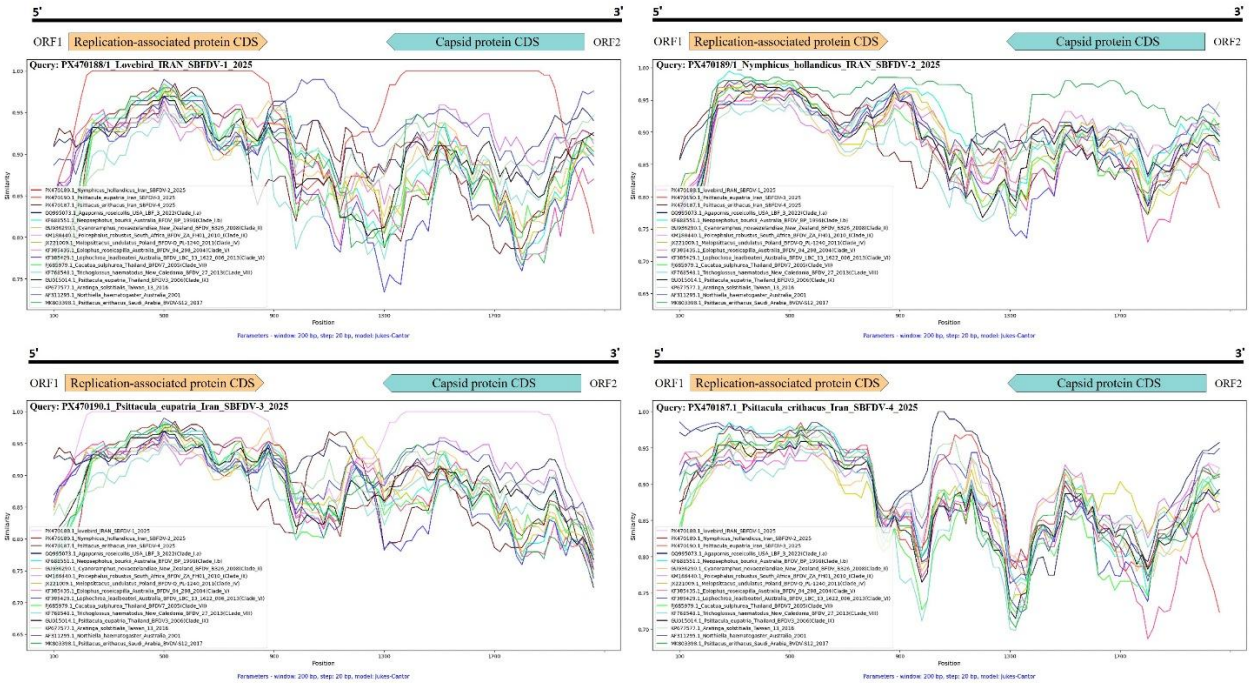


Figure 3 ***Genome-wide recombination profiling of Iranian BFDV isolates using Simplot analysis.*** *Each panel (A–D) illustrates a sliding window analysis of nucleotide identity between one Iranian isolate (designated as the Query) and a representative set of global reference sequences. (A) SBFDV-1, (B) SBFDV-2, (C) SBFDV-3, and (D) SBFDV-4. Pairwise similarity, plotted on the Y-axis, is represented as a normalized score ranging from 0 to 1.Panel (D) highlights a significant genomic mosaicism in SBFDV-4, showing distinct crossovers in similarity between Polish (JX221003.1) and Thai (FJ685989.1) lineages. These crossovers correspond to the recombination breakpoints identified by the RDP4 suite, primarily localized within the ORF2 (Cap) gene region. The high conservation of the ORF1 (Rep) gene across all panels reflects strong purifying selection, while the fluctuations in the Cap region indicate hotspots for genetic diversification and host-adaptive evolution*.

4. DISCUSSION

Supported by complete-genome sequences from Iran deposited in NCBI, this study describes the genetic diversity and evolutionary relationships of the four BFDV genomes from Iran examined here. The overall detection frequency of 55.5% observed in our sample set is similar to the 58% rate reported by Dolatyabi et al. [19] in Iran, suggesting that BFDV DNA is frequently detected in samples submitted for testing in Iran, although sampling frames may differ between studies. Notably, the high detection frequency in Cockatiels (*Nymphicus hollandicus*) is consistent with reports describing this species as a frequent carrier and a possible contributor to viral dissemination within the pet trade [19]. Lower detection rates were reported by Mohamed et al. in Saudi Arabia; differences in sampling frame, host

species composition and diagnostic approach may contribute, and the present data do not allow conclusions about breeding practices, transmission or viral load. Species-specific differences in detection warrant further investigation [30].

The geographic differences in mean nucleotide identity observed in our heatmap analysis can be considered alongside the global genomic patterns described by earlier research [30]. While BFDV is highly mobile due to global commerce, specific regions often harbor unique lineages due to founder effects. The lower mean nucleotide identity (approximately 85%) observed between the Iranian sequences and Chinese sequences is compatible with this idea; however, mean BLASTn identity by country cannot by itself indicate the origin or route of viral introductions into Iran, and no trade data were analyzed. These results are consistent with the findings of Mohamed et al. (2021), who noted that Middle Eastern BFDV strains often share greater identity with European and Australian sequences than with those from neighboring Asian territories [30]. Compared with the regional study by Dolatyabi et al. (2022) [19], the use of complete-genome data in this work provides whole-genome information on four Iranian viruses that complements partial-gene data and may inform future genomic surveillance.

To support a reproducible phylogenetic reconstruction, we employed automated trimming via ClipKIT and selected the optimal substitution models through ModelFinder. This approach reduces the subjectivity of manual editing. In the 548-genome phylogeny analyzed here, the four Iranian genomes were not monophyletic but were distributed within the global phylogenetic structure; this topology alone cannot indicate patterns of international trade or connectivity. While all four sequences clustered within the major Clade I.a, they separated into distinct subgroups whose closest relatives in this dataset came from different countries. SBFDV-1 and SBFDV-2 formed separate subgroups most similar to sequences from Australia and Saudi Arabia, respectively, whereas SBFDV-3 and SBFDV-4 formed a distinct subgroup most closely related to sequences from Taiwan and the USA. This mixed geographic pattern is comparable to the global lineage distribution discussed by Varsani et al. (2011) [13]. This pattern does not by itself indicate dispersal routes. Our work provides genomic data that may support future monitoring of BFDV evolution in Iran.

Analysis of the Iranian sequences shows contrasting levels of sequence conservation between Rep and Cap. The Rep protein (ORF1) showed high conservation (86% to 99% identity), with no substitutions observed in the annotated motifs examined, such as the Walker A motif. This conservation is consistent with the functional constraints of the *Circoviridae* replication machinery described by Shah et al. (2023) (23). Conversely, the Cap protein (ORF2) showed greater variability, especially at the N-terminus. This variability may reflect antigenic drift. One hypothesis is that such variability could contribute to immune escape, but no immunological or functional data were generated here, so this remains to be tested experimentally.

Finally, our use of complete-genome sequencing provides a complementary extension to earlier studies. While previous Iranian research focused on independent evaluations of the *rep* and *cap* genes, our integrated approach allowed screening for recombination signals across the whole genome, which single-gene analyses may not capture. Using RDP5 and Simplot, SBFDV-4 was identified as a putative recombinant exhibiting closest sequence similarity to Polish and Thai lineages, consistent with putative parental-like contributions; the geographic origin and direction of any exchange cannot be inferred. The observed recombination signals indicate the possible involvement of multiple BFDV lineages in the evolution of the Iranian BFDV genomes examined in this study. Simplot similarity curves were consistent with a recombination signal in the cap region, although a single putative recombinant is insufficient to define a recombination hotspot. Our data are consistent with a possible contribution of recombination to the diversity of the Iranian genomes examined; whether recombination influences host range or the emergence of novel variants cannot be determined from these data.

While this study provides genomic data on BFDV in Iran, some limitations should be acknowledged. First, although our screening involved 400 samples across 11 psittacine species, the sampling was primarily focused on birds referred to SANA Avian Hospital. Because the samples were obtained from birds referred to a specialized avian hospital (SANA), the sample set may not reflect the general bird population. Consequently, the observed detection frequency of 55.5% is likely influenced by referral bias inherent to clinic- and laboratory-based submissions and should not be interpreted as a prevalence estimate for the broader population of healthy captive or wild birds in Iran. Future studies incorporating

large-scale, structured surveillance of asymptomatic captive birds and systematic sampling of wild populations would provide a more holistic view of BFDV dynamics and potential spillover events in the Middle East. In addition, only four complete genomes (one per host species) were analyzed, so the recombination, phylogenetic and Rep/Cap variability findings are sequence-based and descriptive; possible effects on antigenicity, nuclear trafficking or host adaptation were not tested experimentally.

4.1. Conclusion

In conclusion, this study contributes four complete-genome sequences of Beak and feather disease virus (BFDV) from Iranian psittacines to the global BFDV genomic database. By integrating complete-genome sequencing with phylogenetic and recombination analyses, we identified genetic relationships among the Iranian BFDV genomes and reference sequences from different geographic regions. The use of ClipKIT for consistent sequence trimming and ModelFinder for selection of appropriate substitution models supported the reproducibility of the phylogenetic analyses.

The detection of genomic mosaicism in the genome designated SBFDV-4 suggests that recombination may have contributed to the evolutionary history of the BFDV genomes examined in this study. However, the limited number of available genomes and the absence of detailed epidemiological and trade-related information preclude definitive conclusions regarding the geographic origins, transmission routes, or directionality of viral movement. Similarly, the observed variability at the N-terminus of the Cap protein indicates that this region may be genetically variable and warrants further investigation to determine its potential effects on antigenicity, host interactions, and immune recognition.

The phylogenetic relationships identified among the Iranian and reference genomes provide baseline information that may inform future BFDV classification, nomenclature, and genomic surveillance. These data may also support the development of regionally informed biosecurity and monitoring strategies, although additional sampling from domestic, captive, and wild psittacine populations is required to assess the broader distribution and diversity of BFDV in Iran. Overall, this study provides baseline genomic data for BFDV research in Iran and highlights the value of continued, context-specific surveillance for monitoring viral evolution and informing disease-control efforts.

12. REFERENCES

1. Ritchie BW, Niagro FD, Lukert PD, Steffens III WL, Latimer KS. 1989 Characterization of a new virus from cockatoos with psittacine beak and feather disease. Virology.171(1):83–8. 0042-6822.
2. Todd D. 2000 Circoviruses: immunosuppressive threats to avian species: a review. Avian pathology.29(5):373–94. 0307-9457.
3. Bassami M, Berryman D, Wilcox G, Raidal S. 1998 Psittacine beak and feather disease virus nucleotide sequence analysis and its relationship to porcine circovirus, plant circoviruses, and chicken anaemia virus. Virology.249(2):453–9. 0042-6822.
4. Niagro F, Forsthoefel A, Lawther R, Kamalanathan L, Ritchie B, Latimer K, et al. 1998 Beak and feather disease virus and porcine circovirus genomes: intermediates between the geminiviruses and plant circoviruses. Archives of virology.143(9):1723–44. 0304-8608.
5. Fogell DJ, Martin RO, Bunbury N, Lawson B, Sells J, McKeand AM, et al. 2018 Trade and conservation implications of new beak and feather disease virus detection in native and introduced parrots. Conservation Biology.32(6):1325–35. 0888-8892.
6. Fogell DJ, Martin RO, Groombridge JJ. 2016 Beak and feather disease virus in wild and captive parrots: an analysis of geographic and taxonomic distribution and methodological trends. Archives of virology.161(8):2059–74. 0304-8608.
7. Latimer KS, Rakich PM, Kircher IM, Ritchie BW, Niagro FD, Steffens Iii W, et al. 1990 Extracutaneous viral inclusions in psittacine beak and feather disease. Journal of Veterinary Diagnostic Investigation.2(3):204–7. 1040-6387.
8. Pass D, Perry R. 1984 The pathology of psittacine beak and feather disease. Australian Veterinary Journal.61(3):69–74. 0005-0423.
9. Rahaus M, Wolff MH. 2003 Psittacine beak and feather disease: a first survey of the distribution of beak and feather disease virus inside the population of captive psittacine birds in Germany. Journal of Veterinary Medicine, Series B.50(8):368–71. 0931-1793.
10. Amery-Gale J, Marenda MS, Owens J, Eden PA, Browning GF, Devlin JM. 2017 A high prevalence of beak and feather disease virus in non-psittacine Australian birds. Journal of Medical Microbiology.66(7):1005–13. 0022-2615.
11. Kundu S, Faulkes CG, Greenwood AG, Jones CG, Kaiser P, Lyne OD, et al. 2012 Tracking viral evolution during a disease outbreak: the rapid and complete selective sweep of a circovirus in the endangered Echo parakeet. Journal of virology.86(9):5221–9. 0022-538X.
12. Raidal S, Sarker S, Peters A. 2015 Review of psittacine beak and feather disease and its effect on Australian endangered species. Australian Veterinary Journal.93(12):466–70. 0005-0423.
13. Varsani A, Regnard GL, Bragg R, Hitzeroth II, Rybicki EP. 2011 Global genetic diversity and geographical and host-species distribution of beak and feather disease virus isolates. Journal of General Virology.92(4):752–67. 0022-1317.
14. Julian L, Piasecki T, Chrząstek K, Walters M, Muhire B, Harkins GW, et al. 2013 Extensive recombination detected among beak and feather disease virus isolates from breeding facilities in Poland. Journal of General Virology.94(5):1086–95. 0022-1317.
15. Ma J, Tian Y, Zhang M, Wang W, Li Y, Tian F, et al. 2019 Identification and characterization of novel genotypes of psittacine beak and feather disease virus from budgerigar in China. Transboundary and Emerging Diseases.66(5):1827–33. 1865-1674.
16. Massaro M, Ortiz-Catedral L, Julian L, Galbraith JA, Kurenbach B, Kearvell J, et al. 2012 Molecular characterisation of beak and feather disease virus (BFDV) in New Zealand and its implications for managing an infectious disease. Archives of virology.157(9):1651–63. 0304-8608.

17. Ghorani M, Rezapanah M. 2020 Molecular Characterization of Psittacine Beak and Feather Disease Virus from Cockatiels (Nymphicus Hollandicus) in Iran. Iranian Journal of Virology.14(1):47–52.
18. Razmyar J, Dezfoulian O, Basami M, ZAMANI A, PEYGHAMBARI S. 2008 Psittacine beak and feather disease in Iran, molecular and histopathologic detection.
19. Dolatyabi S, Peighambari SM, Razmyar J. 2022 Molecular detection and analysis of beak and feather disease viruses in Iran. Frontiers in Veterinary Science.9:1053886. 2297-1769.
20. Peters A, Patterson EI, Baker BG, Holdsworth M, Sarker S, Ghorashi SA, et al. 2014 Evidence of psittacine beak and feather disease virus spillover into wild critically endangered orange-bellied parrots (Neophema chrysogaster). The Journal of Wildlife Diseases.50(2):288–96. 1943-3700.
21. Varsani A, Harrach B, Roumagnac P, Benkő M, Breitbart M, Delwart E, et al. 2024 2024 taxonomy update for the family Circoviridae: A. Varsani et al. Archives of virology.169(9):176. 0304-8608.
22. Sievers F, Higgins DG. 2018 Clustal Omega for making accurate alignments of many protein sequences. Protein Science.27(1):135–45. 0961-8368.
23. Steenwyk JL, Buida III TJ, Li Y, Shen X-X, Rokas A. 2020 ClipKIT: A multiple sequence alignment trimming software for accurate phylogenomic inference. PLoS biology.18(12):e3001007. 1544-9173.
24. Shah PT, Wang J, Liu Y, Hussain B, Ma Z-H, Wu C, et al. 2023 The phylogenetic and phylogeographic landscape of the beak and feather disease virus, 1996–2022. Infection, Genetics and Evolution.112:105442. 1567-1348.
25. Wong TK, Ly-Trong N, Ren H, Demotte P, Baños H, Roger AJ, et al. 2026 IQ-TREE 3: phylogenomic inference software using complex evolutionary models. Molecular Biology and Evolution.43(5):msag117. 0737-4038.
26. Huang S-W, Liu H-P, Chen J-K, Shien Y-W, Wong M-L, Wang C-Y. 2016 Dual ATPase and GTPase activity of the replication-associated protein (Rep) of beak and feather disease virus. Virus research.213:149–61. 0168-1702.
27. Chen J-K, Hsiao C, Wu J-S, Lin S-Y, Wang C-Y. 2019 Characterization of the endonuclease activity of the replication-associated protein of beak and feather disease virus: J. Chen et al. Archives of virology.164(8):2091–106. 0304-8608.
28. Heath L, Martin DP, Warburton L, Perrin M, Horsfield W, Kingsley C, et al. 2004 Evidence of unique genotypes of beak and feather disease virus in southern Africa. Journal of virology.78(17):9277–84. 0022-538X.
29. Sarker S, Ghorashi SA, Forwood JK, Bent SJ, Peters A, Raidal SR. 2014 Phylogeny of beak and feather disease virus in cockatoos demonstrates host generalism and multiple-variant infections within Psittaciformes. Virology.460:72–82. 0042-6822.
30. Mohamed M, El-Sabagh I, Vashi Y, Jagrit V, Salem MAE, Al-Ramadan M, et al. 2021 Analysis of the beak and feather disease viral genome indicates evidence of multiple introduction events into Saudi Arabia. Virus Research.295:198279. 0168-1702.

Statements & Declarations

Funding: This work was supported by the Student Association of Avian Health and Diseases [project number ECV-111-140].

Competing Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Author Contributions

S.M.: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization. M.H.M.: Software, Formal analysis, Methodology, Investigation, Data curation, Writing – review & editing, Visualization. S.H.P.: Validation, Resources, Writing – review & editing. P.Y. and A.S.: Writing – original draft. M.M.J.: Investigation, Resources, Data curation. J.R.: Conceptualization, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition. All authors reviewed the manuscript.

Data availability

The complete genome sequences generated in this study have been deposited in GenBank under accession numbers PX470187, PX470188, PX470189 and PX470190. The protein sequences are available under accession numbers YBO34885–YBO34892. All other data supporting the findings of this study are included in the article and its supplementary files.

Ethical approval

All procedures involving animals in this study were conducted in accordance with the relevant guidelines and regulations. The study protocol was reviewed and approved by the Ethics Committee of the SANA Institute for Avian Health and Diseases Research (Approval Code: **SANA-REC-2025-1061**). Specialized veterinarians collected samples during routine clinical examinations at SANA Avian Hospital, and informed consent was obtained from the bird owners for the use of diagnostic samples in this research.

Acknowledgements

The authors would like to thank the SANA Institute for Avian Health and Diseases Research for providing laboratory facilities and technical support.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTable.docx
- Supplementaryfigure.jpg
- Figures.pdf