

Characterization of Porcine Parvovirus 6 detected in domestic pig in Nigeria using Metagenomic Sequencing, Immuno-informatics and Capsid Structure Modeling

Olusoji Abiola

University of Ibadan

Anyebe Onoja

University of Ibadan

Uwem George

Lead City University

Temitope Faleye

Arizona State University

Ijeoma Ifeora

University of Nigeria

Oladipo Omotosho

University of Ibadan

Oluseyi Olayinka

University of Ibadan

Lander Coninck

KU Leuven

Toluwani Ajileye

Federal University Lokoja

Toluwanimi Akinleye

University of Ibadan

Bolutife Popoola

University of Ibadan

Arthur Oragwa

University of Jos

Olusegun Fagbohun

University of Ibadan

Temidayo Omobowale

University of Ibadan

Jelle Matthijnssens

KU Leuven

Olubusuyi Adewumi

adewumi1@hotmail.com

Research Article

Keywords: Porcine parvovirus 6, Metagenomic Sequencing, pigs, Nigeria, Africa

Posted Date: August 3rd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10347038/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

We report the first detection and genomic characterization of Porcine Parvovirus 6 (PPV6) (*Copiparvovirus ungulate4*) in Africa, identified in a pig from Nigeria in 2021. Metagenomic sequencing of spleen tissue revealed a novel PPV6 variant (designated NGRA4) belonging to the genus *Copiparvovirus*. Phylogenetic analysis of coding-complete genomes demonstrated that NGRA4 forms a distinct lineage that is ancestral to previously reported PPV6 variants globally. The NS1 protein exhibited 96.8% amino acid similarity to its closest relative in GenBank and contained 22 amino acid substitutions, half of which were in the helicase and origin-binding domains. Structural modeling of the capsid protein identified three domains, separated by polyglycine tracts and designated here as domains A, B and C (D-A, D-B, and D-C). D-A appeared to be the most rapidly evolving region, whereas D-B and D-C showed structural similarity to phospholipase A2 and parvovirus capsid proteins, respectively. Three surface-exposed amino acid substitutions in the D-C region (cluster X: A790P, T798S, and T931N) mapped to predicted B-cell epitopes and spatially overlapped with neutralizing antibody binding sites identified in the capsid of parvovirus B19 (pdb id: 6nn3), suggesting immune-driven evolution. Capsid structure prediction suggests a unique feature that bridges adjacent three-fold axes through the two-fold axis. Collectively, our findings suggest NGRA4 is a highly divergent PPV6 lineage that is under immunological pressure and detected for the first time in Nigeria and Africa. The distinctive genetic, immunologic and structural features of NGRA4 highlight the need for enhanced surveillance of PPV6 in African swine populations to better understand its epidemiology, potential role in reproductive failure and organ pathology, and implications for future control strategies.

Introduction

Parvoviruses belong to the family *Parvoviridae*, whose members infect a wide range of hosts, including humans and animals. They are globally distributed, and notable members of the family include parvovirus B19 and canine parvovirus, both of which have been associated with large-scale human epidemics and animal panzootics, respectively (Mietzsch et al., 2019; López-Astacio et al., 2023). The family is currently divided into three subfamilies: *Parvovirinae*, *Densovirinae*, and *Hamaparvovirinae*. The numbered porcine parvoviruses (PPV1–PPV8) are classified within the subfamilies *Parvovirinae* and *Hamavirinae*, specifically in the genera *Protoparvovirus*, *Tetraparvovirus*, *Copiparvovirus*, and *Chaphamavirus* (Table S1) (Vargas-Bermudez et al., 2023).

Parvovirus virions are non-enveloped, possess icosahedral symmetry, and have a diameter of approximately 25 nm. Their genome consists of a ~ 4–6 kb single-stranded linear DNA molecule flanked by inverted terminal repeats (ITRs) and encode two major open reading frames (ORFs) (Kim et al., 2022). ORF1 encodes the multifunctional nonstructural protein 1 (NS1), which oligomerizes and contains four functional domains: the origin-binding domain (OBD), which recognizes and binds the viral origin of replication; the oligomerization domain (OD), which mediates NS1 multimerization and assembly of the replication complex; the helicase domain (HD), which unwinds DNA and provides ATPase activity required for genome replication; and the transactivation domain (TD). Together, these domains perform functions essential for viral replication, host-cell modulation, and potentially virus–host adaptation (López-Astacio et al., 2023; Xie et al., 2023). Current taxonomic criteria stipulate that parvoviruses belonging to the same genus and species should share greater than 35% and 85%, respectively in amino acid identity in the NS1 protein (Pénzes et al., 2020). ORF2 encodes the structural proteins VP1 to VP2 (or VP4, depending on the genus). These proteins differ in length, with the larger capsid protein (VP1) typically containing an additional N-terminal unique region. In many parvoviruses, this

region possesses phospholipase A2 (PLA2) activity, which is essential for endosomal escape during viral entry into host cells (Mietzsch et al., 2019; López-Astacio et al., 2023). The structural proteins share a common C-terminus, and 60 copies assemble to form the viral icosahedral capsid. The capsid contains surface-exposed variable loops that frequently accumulate amino acid substitutions, some of which can alter antigenicity and contribute to immune escape (Mietzsch et al., 2019; López-Astacio et al., 2023).

Among the numbered porcine parvoviruses (PPV1–PPV8), PPV1 is the best characterized and the only virus unequivocally associated with clinical disease in pigs. In particular, PPV1 infection during early gestation in immunologically naïve sows and gilts can result in stillbirth, mummification, embryonic death, and infertility (SMEDI) syndrome (Truyen and Streck, 2012). Like other parvoviruses, PPV1 preferentially infects actively dividing cells, leading to more severe disease in fetuses and young pigs, while infections in adult animals are often mild or subclinical (Kailasan et al., 2015; Kim et al., 2022). The virus is thermostable in the environment, a characteristic that contributes to its persistence, infectivity, and transmission (Eterpi et al., 2009). In addition, PPV1 is considerably resistant to many disinfectants, including ethanol at concentrations of up to 70%. Infected pigs can shed the virus through feces, nasal secretions, and oral secretions, facilitating transmission within and between herds. Fomites also play an important role in viral spread both within and between farms (Truyen and Streck, 2012). Transmission may be further enhanced by the introduction of asymptomatically infected gilts/sow, the use of infected semen for artificial insemination, or breeding with infected boars. Furthermore, seronegative boars may become infected during mating (Szelei et al., 2006; Foerster et al., 2016). Commercially available vaccines provide effective protection against PPV1-associated disease (Vargas-Bermudez et al., 2023).

Although PPV1 remains the most characterized porcine parvovirus, increasing attention has been directed towards other members of the group, including PPV6, because of their detection in a variety of clinical settings. PPV6 was first identified in aborted pig fetuses in China in 2014 and has subsequently been reported in North America, Europe, and Asia (Ni et al., 2014; Schirtzinger et al., 2015; Cui et al., 2017). However, it has yet to be reported in Africa. Here, we describe a PPV6 variant detected in the spleen of a deceased pig from Ogun State in southwestern Nigeria.

Materials and methods

Sample collection, transport and processing

The sample from which the PPV6 genome described in this study was detected was collected from the Oke-Aro Farm Settlement in Ogun State, Nigeria. The piggery contains approximately 5,000 pens and occupies about 30 hectares of land. The viral genome was detected in a spleen sample collected during the necropsy of a deceased domestic pig from the farm in August 2021. Prior to death, the pig exhibited anorexia and a temperature of 39°C, suggesting a possible systemic infection. Because the spleen functions in filtering circulating blood for pathogens, spleen tissue was collected for analysis.

After collection, the sample was transported at +4°C to the NeuroAIDS Molecular Virology Laboratory, Department of Virology, College of Medicine, University of Ibadan, Nigeria where it was stored at -20°C until shipment on dry ice to the Laboratory of Viral Metagenomics at the Rega Institute, KU Leuven, Belgium. DNA extraction and metagenomic sequencing were performed at the Rega Institute.

Nucleic Acid extraction and Next generation Sequencing

The splenic sample was homogenized, and viral genome sequencing was performed using the NetoVIR protocol as previously described (Conceição-Neto et al., 2015). Briefly, the homogenate was centrifuged at $17,000 \times g$ for 3 minutes, after which the supernatant was passed through a $0.8\text{-}\mu\text{m}$ filter. The resulting filtrate was subjected to nuclease treatment, followed by nucleic acid extraction using the Qiagen Total Nucleic Acid Extraction Kit. Extracted nucleic acids were then randomly amplified using the Whole Transcriptome Amplification Kit 2 (Sigma-Aldrich). Sequencing libraries were prepared using the Nextera XT DNA Library Preparation Kit and subsequently sequenced on an Illumina NextSeq 500 platform using paired-end chemistry (2×150 bp).

Sequence assembly

Raw sequencing reads were quality-filtered and trimmed using Trimmomatic v0.36 and subsequently assembled *de novo* with metaSPAdes v3.15.3 on the KBase platform (Arkin et al., 2018). The resulting contigs were then identified through a BLASTn search against the GenBank database (Sayers et al., 2021). Sequencing depth and genome coverage were estimated by mapping the reads to the assembled genome using Bowtie2 v2.3.2 (Langmead and Salzberg, 2012). All software was run using default parameters.

Virus identification and typing

The recovered contig was first used as a query in a BLASTn search of the GenBank database. Open reading frames (ORFs) within the contig were then predicted using the Orffinder software (<https://www.ncbi.nlm.nih.gov/orffinder/>). The nucleotide sequences of the predicted ORFs were subsequently queried against the GenBank database using BLASTn, while their predicted protein products were analyzed using BLASTp searches against the same database.

To investigate the evolutionary relationship of the recovered virus, the contig was also used as a query in a BLASTn search of the NCBI Virus database. All matching sequences with coding-complete genomes and associated metadata, including country and year of collection, were downloaded and compiled into a reference dataset. The contig identified in this study was then added to this dataset. Multiple sequence alignment of this dataset was performed using MAFFT (Katoh et al., 2019). The resulting alignment was used to infer a maximum-likelihood phylogenetic tree in IQ-TREE (Nguyen et al., 2015). Prior to tree reconstruction, the best-fitting nucleotide substitution model was selected using ModelFinder (Kalyaanamoorthy et al., 2017). The tree was annotated using iTOLv7.6 (Letunic and Bork, 2024). To assess the suitability of the dataset for molecular clock analysis, its temporal signal was evaluated by performing a root-to-tip regression of genetic divergence against sampling date using TreeTime (Sagulenko et al., 2018).

Structural and Immuno-informatic analysis

Protein structures were predicted using AlphaFold3 (Abramson et al., 2024). The parvovirus capsid structure was modeled using the oligomer generator available in ViperDB (Montiel-Garcia et al., 2021). Structural homology searches were performed using Foldseek (van Kempen et al., 2024), and linear B-cell epitopes were predicted

using BepiPred 2.0 (Jespersen et al., 2017). Structural visualization, annotation, and alignment were conducted using ChimeraX v1.8 (Meng et al., 2023).

Results

Genome assembly, annotation, and taxonomic classification

To determine the identity, genomic organization, and taxonomic placement of the virus detected in the spleen of the deceased pig by metagenomic sequencing, we performed genome assembly, sequence similarity searches, ORF prediction, and pairwise identity analyses. A contig (subsequently referred to as NGRA4) with length of 6277nt (GC% = 47%) was assembled from 0.16% (16,384/10,273,502) of the trimmed Illumina reads (mean coverage = 368x). A BLASTn search of the GenBank database using NGRA4 as query showed that a porcine parvovirus 6 (PPV6) variant (JSTZ20181121-431; MZ577040.1) detected in China in 2018, was most closely related (Table 1). Two major ORFs (1989nt; 662aa and 3543nt;1180aa) were predicted by ORFfinder (Fig. 1A). BLASTp search showed that the protein product of ORF1 (662aa) encoded the nonstructural protein (NS1) while that of ORF2 (1180aa) encoded a capsid protein (Table 1). Pairwise identity of NS1 protein showed 96.8% similarity with QQD86394.1 (a PPV6 and member of the genus *Copiparvovirus*) (Table 1). In accordance with genus and species demarcation criteria, our data showed that the contig described in this study was PPV6, a member of the genus *Copiparvovirus* and Species *Copiparvovirus unguilate*4.

Analysis of ORF1 (NS1) amino acid substitutions and structural localization

To investigate the presence and potential structural and functional significance of amino acid variation within the NS1 protein, we analyzed sequence conservation and mapped amino acid substitutions onto the predicted protein structure. Specifically, we aligned NGRA4 with the five closest sequences available in GenBank and identified 22 amino acid substitutions in the NS1 protein (Figure S1). Notably, 50% (11/22) of these substitutions were located within the transactivation domain (residues 510–662) of NS1, while the remaining 50% were distributed across the helical domain (HD) and origin-binding domain (OBD) (residues 1–509; Figs. 2 and S1). Interestingly, of the eleven substitutions located within the HD and OBD, ten were surface-exposed in the predicted heptameric structure, with six in the HD and four in the OBD (Fig. 2). The remaining substitution, I448L, is located within the HD but is not surface-exposed.

Analysis of ORF2 (VP1) sequence divergence, structural organization, and antigenic features

To understand the evolutionary and structural significance of amino acid variation in the VP1 capsid protein, we performed comparative sequence, structural, and epitope analyses. Alignment of the five GenBank sequences most closely related to NGRA4 revealed a greater accumulation of amino acid substitutions upstream of the first polyglycine tract in VP1 than in the remainder of the protein (Figure S2). This observation was further supported by similarity analysis of the coding-complete sequences (Fig. 1B). Importantly, inspection of the mapped sequencing reads (Figure S3) indicated that the observed divergence was unlikely to be an artifact of poor sequencing coverage or assembly errors in this region. Consistent with this interpretation, examination of the

aligned sequences revealed no evidence of frameshifts or premature stop codons. Rather, the observed divergence included two in-frame deletions encompassing residues 127–128 and 357–363 relative to the corresponding regions of the closely related GenBank sequences (Figure S2).

Subsequently, we predicted the structure of the VP1 protein and identified three distinct domains, designated A, B, and C (hereafter referred to as D-A, D-B, and D-C, respectively) (Fig. 3A). Domain boundaries were delineated based on the overlap between polyglycine tracts (Figure S2) and AlphaFold-predicted unstructured regions located between well-defined structural domains (Fig. 3A). Subsequent domain-specific Foldseek searches revealed that D-A had no detectable structural homolog in the Protein Data Bank (PDB). In contrast, D-B and D-C showed the highest structural similarity to phospholipase A2 (PDB ID: 4AUP) and the capsid protein of parvovirus B19 (PDB ID: 6NN3), respectively (Fig. 3A). Interestingly, the 337-amino-acid region corresponding to D-A in the predicted structure also coincided with the region exhibiting the highest density of amino acid substitutions in the sequence alignment (Figure S2) and the greatest sequence divergence in the similarity analysis (Fig. 1B).

Among the numbered porcine parvoviruses (PPV1–PPV8), the only experimentally resolved capsid structure is that of PPV1 (PDB ID: 1K3V; Figure S4A), which belongs to the genus *Protoparvovirus* (Table S1). Since no experimentally resolved structure is currently available for any member of the genus *Copiparvovirus*, to which PPV6 belongs, we used the D-C region of NGRA4 VP1 (Fig. 4A) to predict the capsid structure (Fig. 4F). Mapping predicted B-cell epitopes and amino acid substitutions onto the three-fold axis of symmetry (Fig. 4B) of the predicted capsid identified five amino acid substitutions (A790P, T798S, T868Q, V869A, and T931N) that were both surface-exposed and located within predicted B-cell epitopes (Figs. 4C, 4D and 4E). Of these, the substitutions in cluster X (residues 790, 798, and 931) appeared to be unique to NGRA4, whereas T868Q and V869A were not (Figure S2). Structural comparison of the D-C region of NGRA4 VP1 with its closest structural match in the PDB (parvovirus B19 capsid protein [PDB ID: 6NN3]) revealed that residues 790 and 798 of NGRA4 VP1 spatially overlapped with a previously experimentally characterized neutralizing antibody footprint in parvovirus B19. This footprint comprises the set of capsid residues directly contacted by the antibody, specifically the heavy-chain binding region of the Fab fragment (Fig. 5).

Phylogenetic analysis and lineage assessment

To evaluate the evolutionary relationship of NGRA4 to previously described PPV6 variants, we performed phylogenetic analyses using all publicly available PPV6 genomic sequences as of June 1, 2026. The analyses suggested that NGRA4 may represent a previously undescribed PPV6 lineage, as it emerged as an outgroup in the phylogenetic reconstructions (Fig. 6). Importantly, as of June 1, 2026, no PPV6 genome from Nigeria or elsewhere in Africa was publicly available in the NCBI Virus database. Therefore, NGRA4 appears to represent the first PPV6 genome reported from both Nigeria and Africa. Root-to-tip regression analysis of genetic divergence against sampling date revealed an extremely weak temporal signal ($R^2 = 0.001$), indicating that sampling dates accounted for virtually none of the observed genetic variation. Given this lack of temporal structure, the dataset was considered unsuitable for reliable molecular clock inference, and time-scaled phylogenetic analyses were therefore not pursued further (Figure S5).

Table 1
Results of BLASTn and BLASTp searches of GenBank using the NGRA4 contig identified in this study as the query sequence.

ORF			Closest Hit						
			Genomic region	Size	GC%	Accession number	% similarity	Length (nt)	Country/year
	Contig		6,277nt	47	MZ577040.1	90.4	6,400	China/2018	PPV6
	cCDS		5,539nt	47	OR355617.1	91.0	5,607	Colombia/2021	PPV6
1	NS1	nt	1,989nt	47	MK378372.1	94.9	5,897	China/2017	PPV6
		Aa	662	N/A	QQD86394.1	96.8	662	China/1998	PPV6
2	VP1	nt	3,543nt	48	OR355617.1	89.3	5,607	Colombia/2021	PPV6
		Aa	1,180	N/A	AOE23860.1	86.8	1,189	Poland/NA	PPV6

Discussion

Previous reports across Africa have identified circulation of PPV1 to PPV4 (Blomstrom *et al.*, 2012, Molini *et al.*, 2022) with no reports hitherto on the circulation of PPV5 to PPV8. In this study, we report the detection of PPV6 in a domestic pig from Nigeria (Table 1). Based on currently publicly available genomic sequence data, this appears to be the first documented detection of PPV6 in both Nigeria and Africa. We also show that the detected variant might be the first documented member of a lineage that is ancestral to all previously described PPV6 lineages (Fig. 6). The marked divergence observed in the D-A region of the capsid protein (Fig. 1B), together with the presence of surface-exposed amino acid substitutions located within predicted B-cell epitopes (Fig. 4), suggests that this virus may be evolving under selective pressure, potentially driven by host immune responses. Furthermore, the closest coding-complete genome in GenBank differs from NGRA4 by approximately 10% at the nucleotide level (Table 1). Taken together with the phylogenetic and structural evidence presented here, these findings suggest that NGRA4 may belong to a divergent PPV6 lineage that has thus far escaped detection. Although NGRA4 was identified in a pig from Nigeria, additional closely related genomes will need to be sequenced and made publicly available before it can be determined with confidence whether this lineage is established and circulating within Nigeria, elsewhere in Africa, or in other regions of the world.

The NS1 protein is a multifunctional protein that serves as a key taxonomic marker for parvoviruses. It oligomerizes and performs functions essential for viral replication, modulation of host-cell processes, and potentially virus–host adaptation (López-Astacio *et al.*, 2023; Xie *et al.*, 2023). The NS1 protein of NGRA4 shares 96.8% amino acid identity with its closest match in GenBank (Table 1), confirming that NGRA4 belongs to the genus *Copiparvovirus*, species *Copiparvovirus ungulate4* and is a PPV6. We observed that 50% (11/22) of the amino acid substitutions identified in the NGRA4 NS1 protein were located within the HD and OBD domains. In the predicted heptameric structure, all but one of these substitutions were surface-exposed (Fig. 2). Given previous evidence that NS1 interacts with multiple host cellular proteins (Chen *et al.*, 2024), these surface-exposed substitutions may define or modify interaction surfaces involved in protein–protein interactions. Although the functional consequences of these substitutions remain unknown, the distinct phylogenetic position

of NGRA4 (Fig. 6) suggests that they may represent adaptive changes that contributed to the evolutionary success and fitness of this lineage.

Classically, parvovirus VP1 proteins are composed of two major domains: an N-terminal unique region containing a phospholipase A2 (PLA2) motif, followed by the capsid domain (VP2), which forms the icosahedral shell of the virion (Fig. 3B) (Mietzsch et al., 2019; López-Astacio et al., 2023). Structural prediction of the NGRA4 VP1 protein revealed the presence of three domains separated by polyglycine tracts and designated here as D-A, D-B, and D-C (Figs. 3A). Our analysis suggests that D-B and D-C correspond to the canonical VP1u and VP2 domains, respectively. In PPV1 (Fig. 3B) and other well-characterized parvoviruses, VP1u and VP2 are separated by a polyglycine tract that confers structural flexibility and is thought to facilitate the externalization of VP1u from the capsid during viral entry, thereby enabling PLA2-mediated endosomal escape (López-Astacio et al., 2023; Mietzsch et al., 2019). In contrast, the NGRA4 VP1 protein contains two polyglycine tracts, which appear to demarcate, or lie near, the boundaries between D-A and D-B and between D-B and D-C, respectively (Fig. 3A). The presence of these two polyglycine tracts suggests that the structural flexibility required for domain externalization may not be restricted to D-B alone. Instead, it raises the possibility that D-A, like D-B, may undergo conformational rearrangements or externalization during the viral life cycle. Although experimental studies will be required to test this hypothesis, the positioning of these polyglycine tracts (Fig. 3A) suggest that they play important roles in the structural dynamics and function of the PPV6 capsid.

Since a search of the PDB yielded no experimentally resolved structure for any member of the genus *Copiparvovirus*, to which PPV6 belongs, we predicted the capsid structure of NGRA4 (Fig. 4F). As with other parvoviruses, the predicted NGRA4 capsid exhibits clearly defined 5-, 3-, and 2-fold axes of symmetry. The structure also shows a depression surrounding the 3-fold axis, similar to what is observed in PPV1 and B19 (Figures S4A and S4B). However, unlike PPV1 and B19, NGRA4 appears to contain a bridge-like feature connecting regions around the 3-fold axis that extends toward and appears to traverse the 2-fold axis (Fig. 4F). We note that this predicted structure may not fully reflect the native virion architecture because experimentally resolved parvovirus capsids typically consist of approximately 10% VP1 and 90% VP2 among the 60 capsid subunits. This stoichiometry was not incorporated into our model. Instead, for several reasons, the NGRA4 capsid was predicted using only the D–C conformation (equivalent to VP2). First, it is unclear whether NGRA4 adopts the same ~ 1:9 VP1:VP2 stoichiometry observed in other parvoviruses. Second, even if it does, it is not known whether the VP1-derived subunits would correspond to D–A, D–B, and D–C conformations, only D–B and D–C, or another combination of conformational states.

Our findings identified three amino acid substitutions (A790P, T798S, and T931N [cluster X]) that are located within predicted B-cell epitopes, are surface-exposed, and are unique to NGRA4 (Fig. 4). Furthermore, comparative structural analysis shows that A790P and T798S are juxtaposed with the footprint of neutralizing antibodies targeting the VP2 protein of parvovirus B19 (Fig. 5). Collectively, these observations suggest that residues in cluster X may be under immune selection pressure and could form part of a neutralizing antibody epitope. Notably, the residues in cluster X are contributed by two different capsid monomers (Fig. 4). A similar architecture is observed in the footprint of the heavy chain of a neutralizing antibody bound to VP2 of parvovirus B19 (PDB: 6NN3), which also spans two adjacent monomers (Figures S4C–S4G). This raises the possibility that, if cluster X lies within a neutralizing antibody footprint, antibody-mediated neutralization may occur either by blocking attachment to a cellular receptor or co-receptor, or by crosslinking adjacent capsid monomers and thereby restricting the conformational changes required for capsid disassembly. In this context, our findings also

suggest that the receptor-binding domain of the putative PPV6 cell surface receptor may overlap with, or be proximal to, cluster X, as has been demonstrated for canine parvoviruses (Lee et al., 2019; López-Astacio et al., 2023). Consequently, cluster X may therefore be important not only for guiding fundamental studies of PPV6 biology, but also for informing the development of potential countermeasures.

In conclusion, our findings highlight the need for strengthened viral genomic surveillance, expanded genomics capacity, enhanced countermeasure development, and broader biosecurity initiatives in Nigeria and across the African continent. However, since only a single specimen was analysed, no inference can be made regarding prevalence, geographical distribution, or disease association. Furthermore, although PPV6 was detected in the spleen, causality cannot be inferred because metagenomic sequencing detects viral nucleic acids rather than directly indicating disease causation. Given the paucity of data on the circulation of porcine parvoviruses (PPVs), the limited information available for PPV1–4, and the absence of data for PPV5–8, there is a pressing need for systematic and comprehensive investigations. Such studies should aim to characterize the prevalence, co-infection patterns, and evolutionary diversity of PPV1–8 in swine herds in Nigeria and across Africa, thereby providing a stronger foundation for understanding their epidemiology, evolution, and potential impact on animal health.

Declarations

Ethics Approval and Consent to Participate

The collection of the specimen used in this study complied with institutional, national, and international guidelines for the care and use of animals in research. The spleen sample analyzed was obtained during routine necropsy of a deceased domestic pig at the Oke-Aro Farm Settlement, Ogun State, Nigeria. No animal was euthanized for research purposes. Permission for specimen collection and use was granted by the farm management in accordance with local livestock regulations. Ethical approval for this study was obtained from the Animal Care and Use Research Ethics Committee, University of Ibadan (Approval No.UI-ACUREC/099–0624/27). Informed consent was obtained from the owners of the farm prior to sample collection, transport, and subsequent laboratory analyses. All procedures, sample handling, and reporting adhered to the ARRIVE 2.0 (Animal Research: Reporting of In Vivo Experiments) guidelines to ensure transparency, reproducibility, and compliance with established standards of animal welfare.

Competing interests

The authors declare that they have no competing interests.

Consent for publication

Not applicable

Funding

Author Contribution

All authors contributed to the preparation and processing of samples for the study. ABO, UEG, TOCF, OOO, OAO, and OMA developed the study design. LDC and JM performed laboratory analysis, while TOCF, UEG, IMI, OAO, ABO, TGA, TEA, and AOO analyzed the data. TOCF and OJA wrote the initial draft of the manuscript. OOO, OAF, TOO, JM, and OMA supervised the study, while all authors reviewed and edited the manuscript. All the authors have read and approved the final version of the manuscript.

Acknowledgments

During the preparation of this manuscript, ChatGPT (OpenAI) was used to assist with language editing and improving clarity. The authors reviewed and edited all content and take full responsibility.

Data Availability

The datasets generated and/or analyzed during the current study are available in GenBank and SRA under accession numbers PV935222 and PRJNA1288516, respectively.

References

1. Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, Ronneberger O, Willmore L, Ballard AJ, Bambrick J, Bodenstein SW, Evans DA, Hung CC, O'Neill M, Reiman D, Tunyasuvunakool K, Wu Z, Žemgulytė A, Arvaniti E, Beattie C, Jumper JM. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*. 2024;630(8016):493–500. <https://doi.org/10.1038/s41586-024-07487-w>.
2. Arkin AP, Cottingham RW, Henry CS, Harris NL, Stevens RL, Maslov S, Dehal P, Ware D, Perez F, Canon S, Sneddon MW, Henderson ML, Riehl WJ, Murphy-Olson D, Chan SY, Kamimura RT, Kumari S, Drake MM, Brettin TS, Glass EM, Yu D. KBase: The United States Department of Energy Systems Biology Knowledgebase. *Nat Biotechnol*. 2018;36(7):566–9. <https://doi.org/10.1038/nbt.4163>.
3. Blomström AL, Ståhl K, Masembe C, Okoth E, Okurut AR, Atmnedi P, Kemp S, Bishop R, Belák S, Berg M. Viral metagenomic analysis of bushpigs (*Potamochoerus larvatus*) in Uganda identifies novel variants of Porcine parvovirus 4 and Torque teno sus virus 1 and 2. *Virology*. 2012;9:192. <https://doi.org/10.1186/1743-422X-9-192>.
4. Chen S, Liu F, Yang A, Shang K. For better or worse: crosstalk of parvovirus and host DNA damage response. *Front Immunol*. 2024;15:1324531. <https://doi.org/10.3389/fimmu.2024.1324531>.
5. Conceição-Neto N, Zeller M, Lefrère H, De Bruyn P, Beller L, Deboutte W, Yinda CK, Lavigne R, Maes P, Van Ranst M, Heylen E, Matthijnsens J. Modular approach to customise sample preparation procedures for viral metagenomics: a reproducible protocol for virome analysis. *Sci Rep*. 2015;5:16532. <https://doi.org/10.1038/srep16532>.
6. Cui J, Biernacka K, Fan J, Gerber PF, Stadejek T, Opriessnig T. Circulation of Porcine Parvovirus Types 1 through 6 in Serum Samples Obtained from Six Commercial Polish Pig Farms. *Transbound Emerg Dis*.

- 2017;64(6):1945–52. <https://doi.org/10.1111/tbed.12593>.
7. Eterpi M, McDonnell G, Thomas V. Disinfection efficacy against parvoviruses compared with reference viruses. *J Hosp Infect.* 2009;73(1):64–70. <https://doi.org/10.1016/j.jhin.2009.05.016>.
 8. Foerster T, Streck AF, Speck S, Selbitz HJ, Lindner T, Truyen U. An inactivated whole-virus porcine parvovirus vaccine protects pigs against disease but does not prevent virus shedding even after homologous virus challenge. *J Gen Virol.* 2016;97(6):1408–13. <https://doi.org/10.1099/jgv.0.000446>.
 9. Jespersen MC, Peters B, Nielsen M, Marcatili P. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Res.* 2017;45(W1):W24–9. <https://doi.org/10.1093/nar/gkx346>.
 10. Kailasan S, Agbandje-McKenna M, Parrish CR. Parvovirus Family Conundrum: What Makes a Killer? *Annual Rev Virol.* 2015;2(1):425–50. <https://doi.org/10.1146/annurev-virology-100114-055150>.
 11. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermini LS. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods.* 2017;14(6):587–9. <https://doi.org/10.1038/nmeth.4285>.
 12. Katoh K, Rozewicki J, Yamada KD. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform.* 2019;20(4):1160–6. <https://doi.org/10.1093/bib/bbx108>.
 13. Kim SC, Kim JH, Kim JY, Park GS, Jeong CG, Kim WI. Prevalence of porcine parvovirus 1 through 7 (PPV1-PPV7) and co-factor association with PCV2 and PRRSV in Korea. *BMC Vet Res.* 2022;18(1):133. <https://doi.org/10.1186/s12917-022-03236-1>.
 14. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012;9(4):357–9. <https://doi.org/10.1038/nmeth.1923>.
 15. Lee H, Callaway HM, Cifuentes JO, Bator CM, Parrish CR, Hafenstein SL. Transferrin receptor binds virus capsid with dynamic motion. *Proc Natl Acad Sci USA.* 2019;116(41):20462–71. <https://doi.org/10.1073/pnas.1904918116>.
 16. Letunic I, Bork P. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Res.* 2024;52:W78–82. <https://doi.org/10.1093/nar/gkae268>.
 17. López-Astacio RA, Adu OF, Lee H, Hafenstein SL, Parrish CR. The Structures and Functions of Parvovirus Capsids and Missing Pieces: the Viral DNA and Its Packaging, Asymmetrical Features, Nonprotein Components, and Receptor or Antibody Binding and Interactions. *J Virol.* 2023;97(7):e0016123. <https://doi.org/10.1128/jvi.00161-23>.
 18. Meng EC, Goddard TD, Pettersen EF, Couch GS, Pearson ZJ, Morris JH, Ferrin TE. UCSF ChimeraX: Tools for structure building and analysis. *Protein science: publication Protein Soc.* 2023;32(11):e4792. <https://doi.org/10.1002/pro.4792>.
 19. Mietzsch M, Péntzes JJ, Agbandje-McKenna M. Twenty-Five Years of Structural Parvovirology. *Viruses.* 2019;11(4):362. <https://doi.org/10.3390/v11040362>.
 20. Molini U, Franzo G, Settypalli TBK, Hemberger MY, Khaiseb S, Cattoli G, Dundon WG, Lamien CE. Viral Co-Infections of Warthogs in Namibia with African Swine Fever Virus and Porcine Parvovirus 1. *Animals: open access J MDPI.* 2022;12(13):1697. <https://doi.org/10.3390/ani12131697>.
 21. Montiel-Garcia D, Santoyo-Rivera N, Ho P, Carrillo-Tripp M, Ili CLB, Johnson JE, Reddy VS. *Nucleic Acids Res.* 2021;49(D1):D809–16. <https://doi.org/10.1093/nar/gkaa1096>. VIPERdb v3.0: a structure-based data analytics platform for viral capsids.

22. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. Mol Biol Evol. 2015;32(1):268–74. <https://doi.org/10.1093/molbev/msu300>. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies.
23. Ni J, Qiao C, Han X, Han T, Kang W, Zi Z, Cao Z, Zhai X, Cai X. Identification and genomic characterization of a novel porcine parvovirus (PPV6) in China. Virol J. 2014;11:203. <https://doi.org/10.1186/s12985-014-0203-2>.
24. Péntzes JJ, Söderlund-Venermo M, Canuti M, Eis-Hübinger AM, Hughes J, Cotmore SF, Harrach B. Reorganizing the family Parvoviridae: a revised taxonomy independent of the canonical approach based on host association. Arch Virol. 2020;165(9):2133–46. <https://doi.org/10.1007/s00705-020-04632-4>.
25. Sagulenko P, Puller V, Neher RA. TreeTime: Maximum-likelihood phylodynamic analysis. Virus Evol. 2018;4(1):vex042. <https://doi.org/10.1093/ve/vex042>.
26. Sayers EW, Cavanaugh M, Clark K, Pruitt KD, Schoch CL, Sherry ST, Karsch-Mizrachi I. GenBank Nucleic acids Res. 2021;49(D1):D92–6. <https://doi.org/10.1093/nar/gkaa1023>.
27. Schirtzinger EE, Suddith AW, Hause BM, Hesse RA. First identification of porcine parvovirus 6 in North America by viral metagenomic sequencing of serum from pigs infected with porcine reproductive and respiratory syndrome virus. Virol J. 2015;12:170. <https://doi.org/10.1186/s12985-015-0401-6>.
28. Sun Y, Klose T, Liu Y, Modrow S, Rossmann MG. Structure of Parvovirus B19 Decorated by Fabs from a Human Antibody. J Virol. 2019;93(9):e01732–18. <https://doi.org/10.1128/JVI.01732-18>.
29. Szelei J, Zadori Z, Tijssen P. Porcine parvovirus. In: Kerr JR, Cotmore SF, Bloom ME, Linden RM, Parrish CR, editors. Parvoviruses. London, UK: Hodder Arnold; 2006. pp. 435–45.
30. Truyen U, Streck AF. Porcine parvovirus. In: Zimmerman JJ, Karriker LA, Ramirez A, Schwartz KJ, Stevenson GW, editors. Diseases of Swine, tenth ed. Chichester, West Sussex: Wiley-Blackwell; 2012.
31. van Kempen M, Kim SS, Tumescheit C, Mirdita M, Lee J, Gilchrist CLM, Söding J, Steinegger M. Fast and accurate protein structure search with Foldseek. Nat Biotechnol. 2024;42(2):243–6. <https://doi.org/10.1038/s41587-023-01773-0>.
32. Vargas-Bermudez DS, Mogollon JD, Franco-Rodriguez C, Jaime J. The Novel Porcine Parvoviruses: Current State of Knowledge and Their Possible Implications in Clinical Syndromes in Pigs. Viruses. 2023;15(12):2398. <https://doi.org/10.3390/v15122398>.
33. Xie Q, Wang J, Gu C, Wu J, Liu W. Structure and function of the parvoviral NS1 protein: a review. Virus Genes. 2023;59(2):195–203. <https://doi.org/10.1007/s11262-022-01944-2>.

Figures

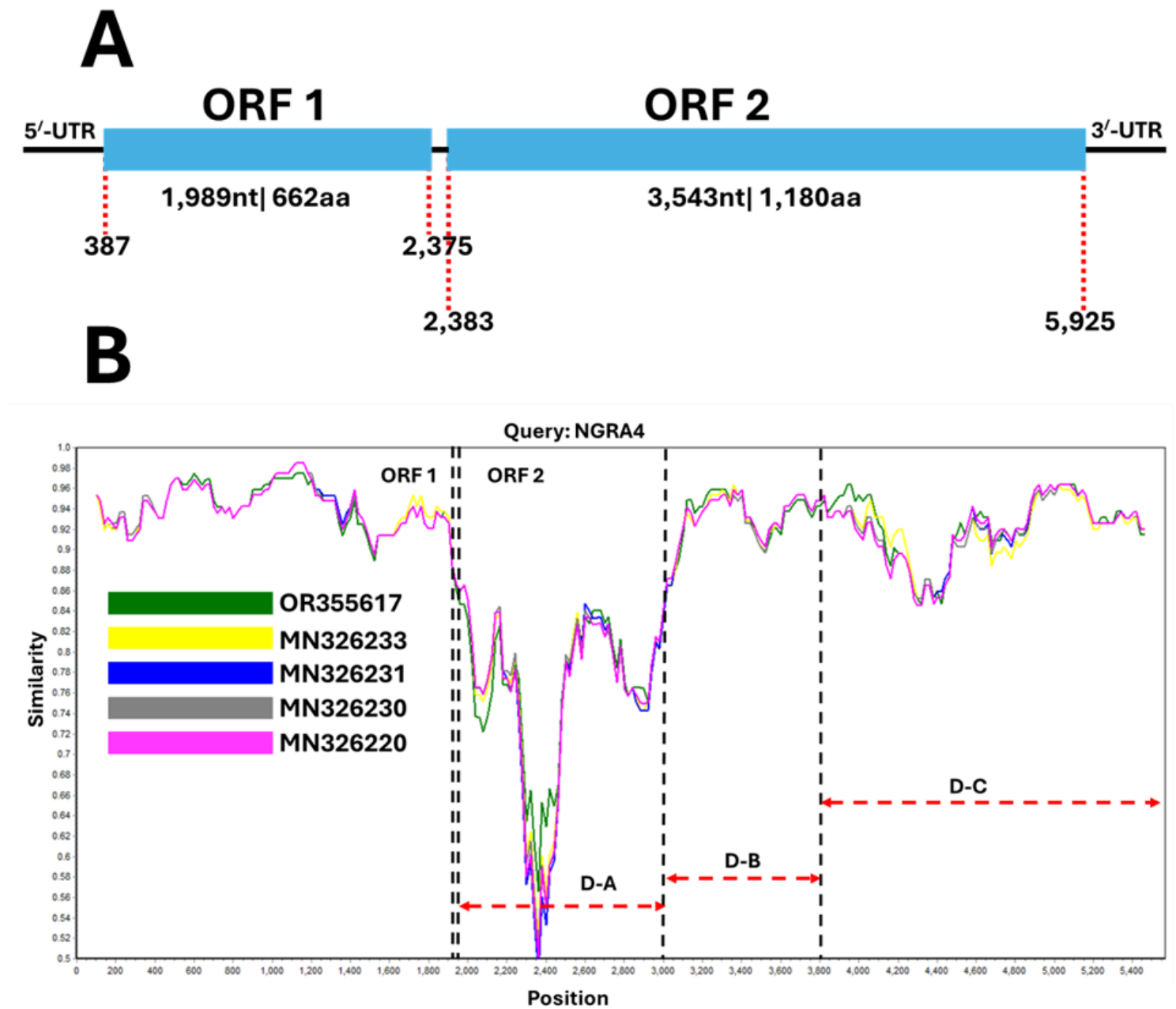


Figure 1

(Panel A) Schematic representation of the predicted coding potential of NGRA4 generated using ORFfinder. The two predicted open reading frames (ORFs) are shown. (Panel B) Similarity plot comparing the complete coding sequence (cCDS) of NGRA4 with the five most similar PPV6 sequences identified in GenBank. The double black dashed lines indicate the boundary between ORF1 and ORF2. The single black dashed lines within ORF2 indicate the boundaries of domains D-A, D-B, and D-C.

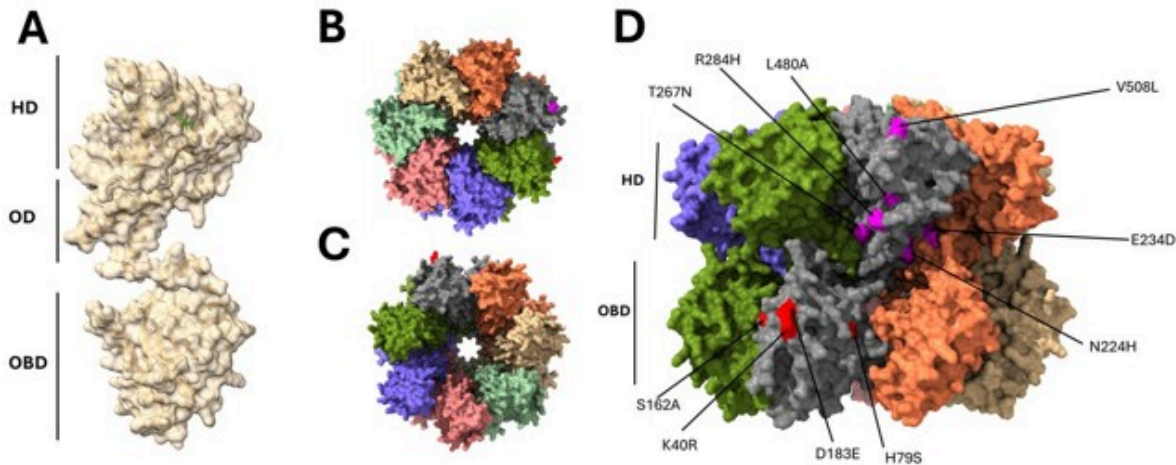
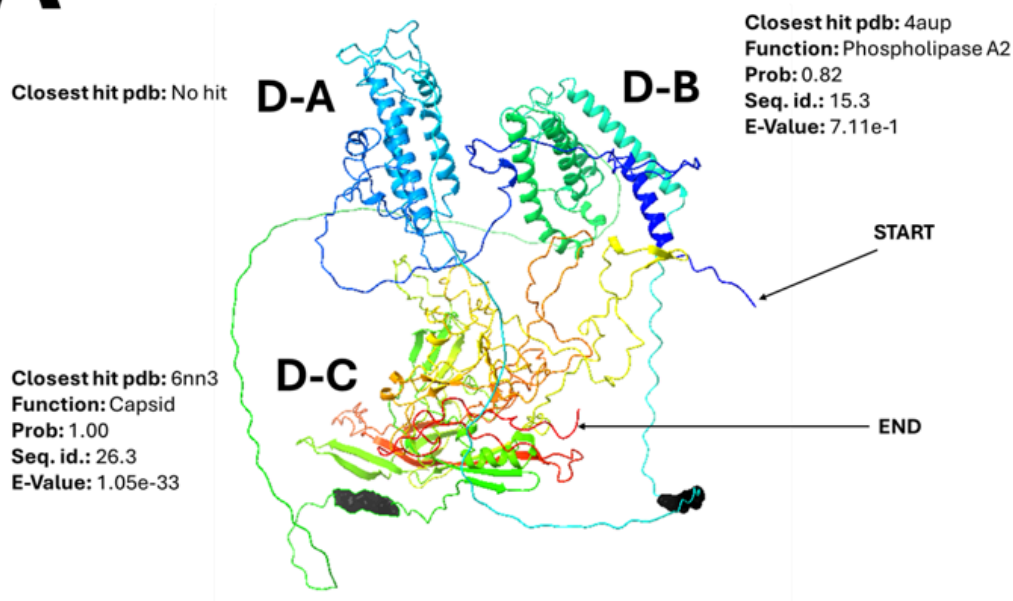
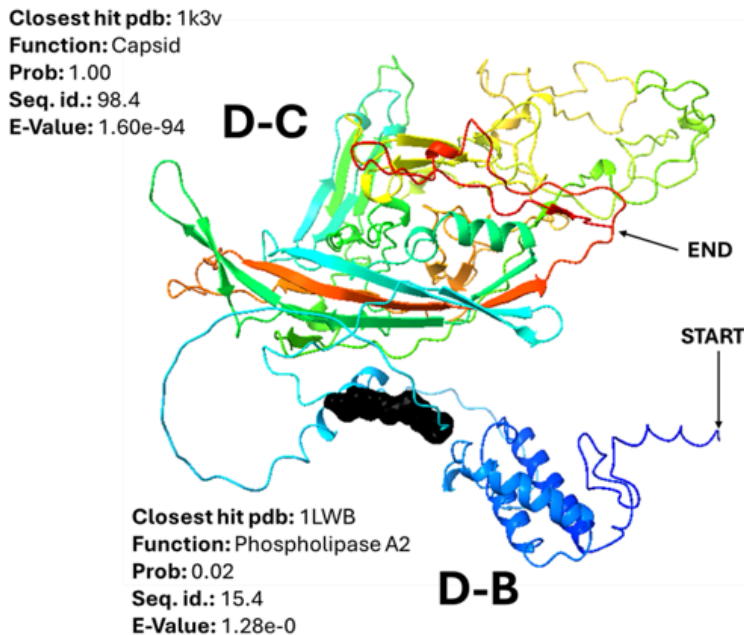


Figure 2

(Panel A) Predicted structure of the NS1 (ORF1) monomer. Three of the four domains described by Xie *et al.* (2023) are shown. The fourth domain, the transactivation domain (TD), is disordered in the predicted structure and is therefore not shown. The displayed domains include the origin-binding domain (OBD), oligomerization domain (OD), and helicase domain (HD). (Panels B–D) Predicted structure of the NS1 heptamer, which is typically assembled around DNA during unwinding (Xie *et al.*, 2023). Panel B shows the top view from the HD side, panel C shows the bottom view from the OBD side, and panel D shows a side view of the heptamer with amino acid substitutions identified in NGRA4 (Figure S1) mapped onto one monomer. All but one of the substitutions are located on the exterior surface of the heptamer and occur in both the HD (purple) and OBD (red).

A**B****Figure 3**

(Panel A) Predicted structure of the ORF2 protein encoded by NGRA4. The model is colored from the N-terminus (blue) to the C-terminus (red). The three domains, designated D-A, D-B, and D-C from the N- to C-terminus, are labeled accordingly. D-A does not exhibit similarity to any currently available structure in the Protein Data Bank (PDB). D-B is structurally similar to phospholipase A2, whereas D-C is similar to the capsid protein of parvovirus B19. (Panel B) Predicted structure of the VP1 protein (XSS91364.1) encoded by a PPV1 (genus *Protoparvovirus*)

genome (PP657027.1) available in GenBank. Only two domains, corresponding to D-B and D-C in panel A, were identified. D-B is structurally similar to phospholipase A2, whereas D-C is similar to the capsid protein of porcine parvovirus 1. In both panels, polyglycine stretches (≥ 5 residues) are displayed in surface representation and colored black.

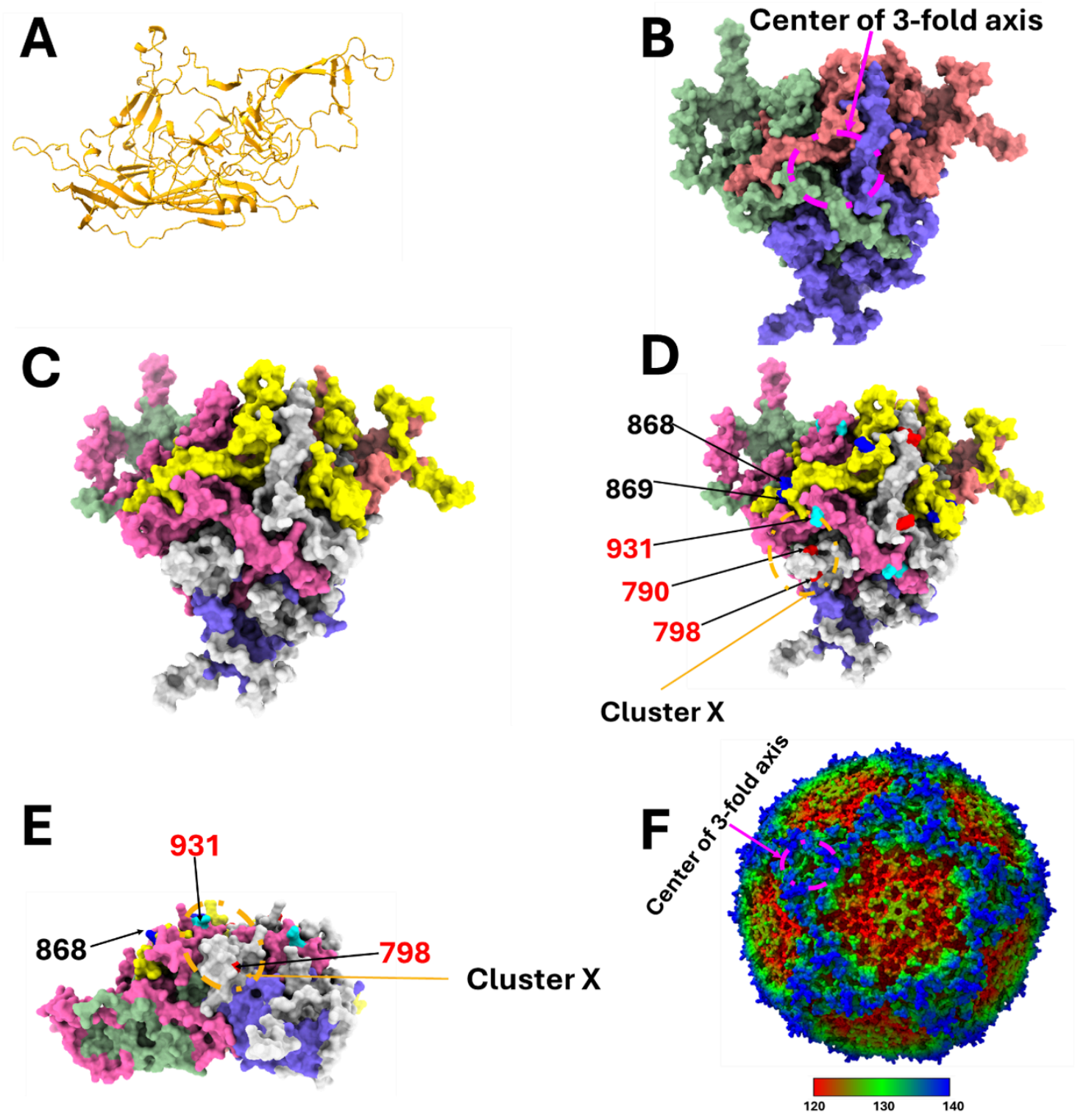


Figure 4

Predicted structure of the NGRA4 capsid. (Panel A) Ribbon representation of D-C of VP1. (Panel B) Trimeric assembly of domain C forming the threefold axis of symmetry of the capsid. Individual monomers are shown in different colors. The center of the threefold axis is highlighted by a purple oval. (Panel C) Same structure as in panel B with predicted B-cell epitopes overlaid. (Panel D) Surface-exposed amino acid substitutions located within predicted B-cell epitopes are highlighted. The orange dashed oval indicates a region designated Cluster X, which contains surface-exposed amino acid substitutions unique to NGRA4 (shown in red). Panels B–D show the trimer from the top view, whereas panel E shows the side view. (Panel F) Predicted NGRA4 capsid structure colored according to surface topology. The center of the threefold axis is highlighted by a purple oval, as in panel B.

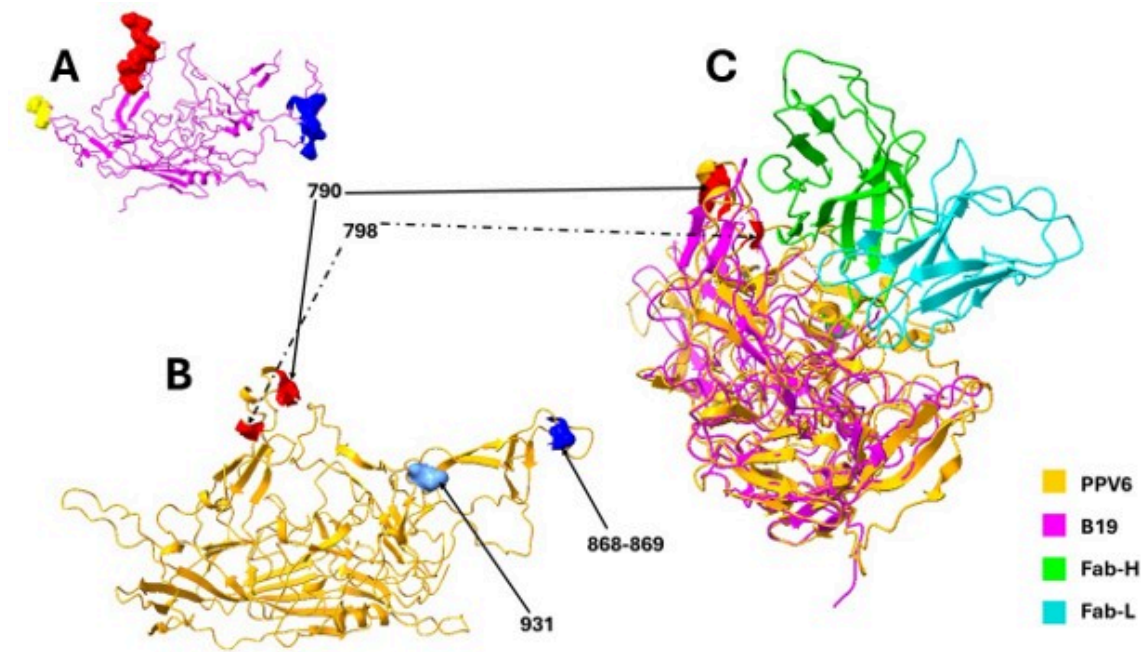


Figure 5

(Panel A) Structure of parvovirus B19 VP2 (PDB: 6NN3) with experimentally determined neutralizing antibody contact regions (Sun *et al.*, 2019) highlighted in yellow, red, and blue. (Panel B) Domain C of the NGRA4 VP1 protein. Surface-exposed amino acid substitutions located within predicted B-cell epitopes are highlighted in red, blue, and cornflower blue. (Panel C) Structural alignment of the capsid proteins shown in panels A and B together with the heavy and light chains of the antigen-binding fragment (Fab) of the neutralizing antibody whose footprint is shown in panel A (PDB: 6NN3). Note the overlap between the region highlighted in red in panel B and the heavy-chain binding footprint observed in panel C.

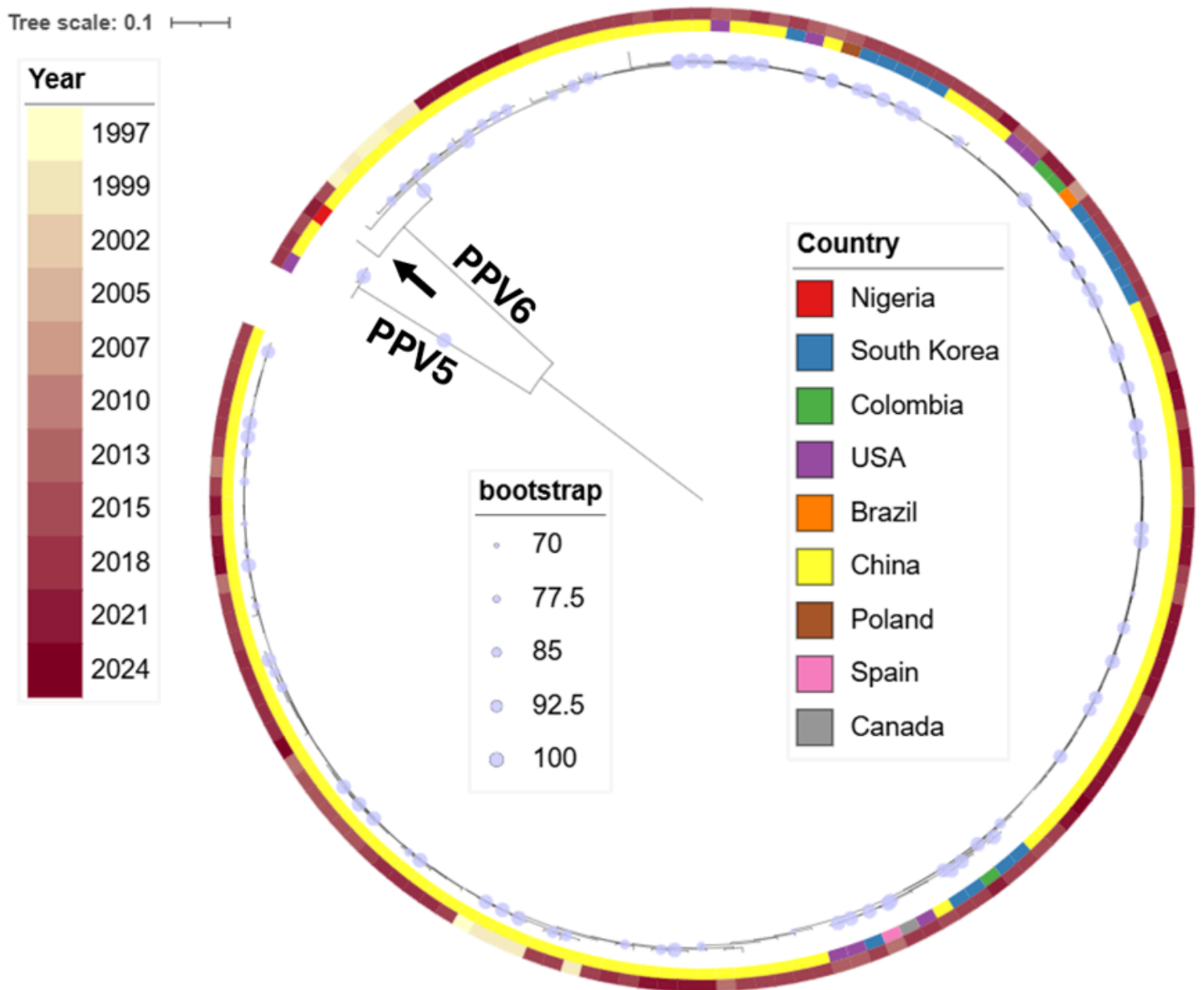


Figure 6

Maximum-likelihood phylogenetic tree inferred from NGRA4 and 150 PPV6 genomes identified through a BLASTn search of the NCBI Virus database using NGRA4 as the query sequence. Three PPV5 genomes were included as outgroups to root the tree. NGRA4 (black arrow) forms a distinct basal lineage relative to all previously available PPV6 genomes. Concentric annotation rings indicate country (inner), and year of sample collection (outer). Bootstrap support values $\geq 70\%$ are shown at nodes. GenBank accession numbers for all sequences included in the analysis are provided in Table S2.

Supplementary Files

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

- [Supplementary.docx](#)