Genetic and Glycosylation Variability of the GP5 Glycoprotein in NADC34 Variant Strains of *Betaarterivirus americense* from Lima, Peru

Molecular Diversity of the GP5 Glycoprotein in *Betaarterivirus americense*

- ✓ Felicita Karina Camargo Paredes [e] **Orcid:** 0000-0002-4771-1665
- ✓ Rony Cotaquispe Nalvarte [a] **Orcid:** 0000-0002-8273-4803
- ✓ Edgard De la Cruz Vásquez [b] **Orcid:** 0000-0002-9229-985X
- ✓ Julio Cesar Ecos Espino [c] **Orcid:** 0000-0003-2968-7208
- ✓ Miriam Legua Barrios [c] **Orcid:** 0000-0002-4993-0909
- ✓ Erick Llona García [c] **Orcid:** 0000-0002-7522-0193
- ✓ Ayda Liliana Reyes Ruiz [e] **Orcid: 0000-0002-0232-1890**
- ✓ Cesar Augusto Mendoza Yañez [e] **Orcid:** 0000-0003-4272-8633
- ✓ Merici Ingrid Medina Guerrero [b] **Orcid:** 0000-0003-0728-9997
- ✓ Jennifer Toskano Hurtado [d] jtoskano@corpmontana.com

[a] Research Professor, Universidad Privada Científica del Sur, Lima, Perú.
[b] Professor, Universidad Privada Autónoma de Ica, Ica, Perú.
[c] Professor, Universidad Privada San Juan Bautista, Ica, Perú.
[d] Animal Nutrition and Health Business Unit, Innovation and Development Area, Corporation Montana S.A., Lima, Peru.
[e] Professor, Universidad Nacional San Luis Gonzaga, Ica, Perú.

ABSTRACT

Lima is a major center of swine production in Peru; therefore, these data provide an initial insight into GP5 variability in a high-risk production setting. Accordingly, the aim of this study was to identify the genetic diversity of *Betaarterivirus americense* circulating in commercial pig farms in Lima. Phylogenetic inference was performed based on amino acid substitution analysis using the maximum likelihood method with BioNJ, N-glycosylation analysis was conducted using the NetNGlyc 1.0 server, and genetic diversity was assessed using DnaSP software. GP5 phylogeny supported clustering within sublineage 1.5 of the NADC34 variant. Significant amino acid substitutions were identified in immunologically relevant domains involved in the induction of neutralizing and non-neutralizing antibodies: 16/24 at N32 (S/G/R/E), 7/24 at L39 (F), 6/24 at Q40 (L/R), 4/24 at N43 (I), 15/24 at K57 (E/V/R), and 8/24 at N58 (K/E/V/R), located within GP5 epitopes A, B, and C. A total of nine potential N-glycosylation sites (N30, N32, N33, N34, N35, N44, N50, N51, and N57) were identified among the 24 strains. Of these, sites N34, N44, and N51 were partially conserved in most strains or fully conserved in strains 40 and 41, whereas complete loss was observed in strains 36, 37, 38, and 39. Genetic diversity analysis revealed 530 polymorphic sites, and Tajima's D test yielded a negative value ($D = -0.78746$), indicating high genetic variability within the Lima cohort. Overall, our findings from the 24 *Betaarterivirus americense* strains belonging to the NADC34 sublineage reflect the presence of diverse viral

variants in pig farms in Lima; broader and temporally resolved sampling is required to assess national-level patterns.

Keywords: *Betaarterivirus americense*; glycoprotein GP5; N-glycosylation; NADC34

INTRODUCTION

Porcine reproductive and respiratory syndrome (PRRS) is a viral disease of major sanitary importance and represents a serious threat to the global swine industry [13]. The etiological agent has been molecularly and structurally characterized, leading to the identification of two main species: *Betaarterivirus americense* and *Betaarterivirus europense* [3,5,6]. The viral genome consists of a positive-sense single-stranded RNA containing a 5′ cap structure and a 3′ polyadenylated tail. It includes a 5′ untranslated region (UTR), at least ten open reading frames (ORFs)—ORF1a, ORF1b, ORF2a, ORF2b, ORF3, ORF4, ORF5, ORF5a, ORF6, and ORF7—and a 3′ UTR [28,33]. ORF1a and ORF1b comprise approximately three-quarters of the genome and encode the replicase polyproteins pp1a and pp1a/b, which undergo extensive proteolytic processing by virus-encoded proteases to produce at least 16 nonstructural proteins (nsps) involved in viral replication and transcription [10,11]. The remaining ORFs, located in the 3′ terminal region, encode eight structural proteins, including four membrane-associated glycoproteins (GP2a, GP3, GP4, and GP5), three non-glycosylated membrane proteins (E, ORF5a, and M), and the nucleocapsid protein (N) [8,34].

Phylogenetic analysis of the ORF5 gene in *Betaarterivirus americense* has identified nine monophyletic lineages (L1–L9), along with multiple sublineages [6,21]. ORF5 encodes the major viral envelope glycoprotein (GP5), a transmembrane glycosylated protein exposed on the virion surface that plays a central role in viral attachment to host cell receptors and constitutes a key target of the humoral immune response [1,7,26]. GP5 comprises a signal peptide, two hypervariable regions (HVR1 and HVR2), three transmembrane domains (TM1, TM2, and TM3), a decoy epitope, the primary neutralizing epitope (PNE), an additional B-cell epitope, and two T-cell epitopes [37]. This protein is essential for viral infectivity and harbors critical immunological domains associated with viral neutralization, including antigenic determinants and N-glycosylation sites [18].

GP5 glycosylation represents an important immune evasion mechanism employed by several enveloped viruses to evade, block, or reduce the activity of neutralizing antibodies [2]. In particular, alterations in N-glycosylation sites associated with the primary neutralizing epitope significantly affect viral immunogenicity, rendering virions resistant to antibody-mediated neutralization in vivo [9]. In addition, high genetic variability is one of the main evolutionary mechanisms driving the emergence of new *Betaarterivirus americense* variants and plays a key role in the increasing number of circulating genotypes [15,16].

In Lima, following the molecular identification of lineage 1, sublineage 1.5, corresponding to the IA/2014/NADC34 variant in 2019 and 2025 [6,23], uncertainties remain regarding the genetic characteristics of viral strains and the diversity of lineages associated with reduced immunogenicity, which is linked to the high variability of *Betaarterivirus americense*. Therefore, the aim of this study was to characterize the diversity in the molecular architecture of glycoprotein 5 (GP5) in strains circulating in Lima, contributing to a broader understanding of genetic variation, molecular epidemiology, and providing empirical insight into the limitations faced in the development of effective vaccines.

MATERIALS AND METHODS

Biological material

This study analyzed 24 nucleotide sequences obtained from strains circulating in commercial pig farms in Lima, Peru [6].

RNA extraction, quantification, and conventional internal RT-PCR of the ORF5 gene

Procedures for viral RNA extraction, cDNA synthesis, quantification, and conventional internal RT-PCR amplification of the *Betaarterivirus americense* ORF5 gene were performed as previously described by Cotaquispe et al. [6].

Electrophoresis, purification, and sequencing

Amplified products were visualized using a multiSUB™ Choice horizontal electrophoresis system (Cleaver Scientific) with 1% molecular-grade agarose gels (SCL AG500, Cleaver Scientific), stained with Fluorescent Dye–DNA/Safe-Green™ (ABM®). An Opti-DNA 100 bp molecular weight marker (50 bp–1.5 kb; ABM®) was used, and gels were visualized using a UVP UV Solo Touch photodocumentation system (Analytik Jena). Gel fragments containing amplicons of the expected size were excised and purified using the innuPREP nucleic acid purification kit (Analytik Jena). Purified products were bidirectionally sequenced by Macrogen (South Korea) using the BigDye™ Terminator v3.1 Cycle Sequencing Kit (Life Technologies) on an ABI 3137 XL Genetic Analyzer (Life Technologies).

Phylogenetic analysis of glycoprotein 5 (GP5)

Lineage determination in this study was based exclusively on GP5 amino acid phylogeny; nucleotide-level phylogenetic analyses were previously reported by Cotaquispe et al. [6]. Sequencing chromatograms of GP5 generated in this study were analyzed using Chromas Lite® v2.6.6 and combined with 58 reference sequences retrieved from GenBank. Nine sequences

corresponded to Peruvian isolates (MH791386.1, MH791385.1, MH791383.1, MH791388.1, MH791379.1, MH791378.1, MH791376.1, MH791391.1, and MH791390.1), while the remaining sequences represented pathogenic strains isolated worldwide (MF326985.1, DQ009647.1, DQ306879.1, DQ306878.1, DQ306877.1, GQ857656.1, FJ349261.1, AY588319.1, DQ489311.1, GQ461593.1, EF484033.1, AY032626.1, EF635006.1, KF611905.1, KF287140.1, JQ715697.1, EU939312.1, GQ374442.1, JQ663556.1, HQ315837.1, FJ889130.1, JX512910.2, JQ663545.1, JN411262.1, AY743931.1, JF802085.1, GU047345.1, EU076704.1, AY366525.1, KC862567.1, AF325691.1, KC862577.1, DQ473474.1, AB288356.1, JN660150.1, EU807840.1, EU262603.1, AF184212.1, DQ176019.1, AF331831.1, AF176348.2, U87392.3, MF279253.1, EF536003.1, AF066183.4, DQ988080.1, AY424271.1, AF494042.1, and EF112445.1).

Phylogenetic trees were inferred based on amino acid substitution analysis using the maximum likelihood method with BioNJ and the Jones–Taylor–Thornton (JTT) substitution model [14]. Because sublineage placement may vary depending on the protein model and algorithm selected, node support values are reported for relevant internal nodes to assess tree robustness. Trees were drawn to scale, with branch lengths representing the number of substitutions per site. The final dataset comprised 196 amino acid positions. Bootstrap support values were calculated from 1,000 alignment replicates. Evolutionary analyses were conducted using DNAMAN v10 [32] and MEGA v6 [29].

Analysis of immunological domains of glycoprotein 5

Amino acid sequence analysis was performed using MEGA v6.0 to characterize GP5 structural features, including the signal peptide, decoy epitopes I and II, hypervariable regions I and II, the primary neutralizing epitope (PNE), transmembrane regions 1, 2, and 3, T-cell epitopes 1 and 2, and the B-cell epitope. This analysis enabled the identification of genetic variants, as mutations within viral protein structures may affect infectivity, pathogenicity, and viral persistence [18]. GP5 length was determined, amino acid substitutions in key immunological domains were identified, and potential N-glycosylation sites were predicted using the NetNGlyc 1.0 server [12]. Signal peptides were additionally predicted using SignalP v4.0 [23].

Genetic diversity analysis: Tajima's D test and DnaSP v6

Statistical analysis was based on the neutral theory of molecular evolution, which assumes that the evolution of slightly deleterious or advantageous, nearly neutral mutations is primarily influenced by mutation rates, selection coefficients, and population size. Tajima's D statistic [30] compares two estimates of genetic variation (θ) within a sequence set. The statistic has a mean of zero with a standard deviation of ± 1 ($-1 < D < +1$). The

observed D value was compared with the corresponding critical values reported in Table 2 [31], according to sample size and the confidence limit closest to the calculated D value. Negative deviations indicate departures from neutrality, whereas positive D values may suggest positive selection or population size reduction. DnaSP v6 was used to calculate Tajima's D values and to assess the genetic diversity of the analyzed isolates [25].

RESULTS

Serum samples were analyzed using an in-house kit with primers targeting the ORF5 gene, which encodes viral glycoprotein 5 (GP5). Primers were validated and standardized using vaccine strains from both viral species as positive controls, along with a commercial IDEXX positive control mixture (Figure 1). Amplification of the target gene yielded a 723 bp product, Figure 2. Sequencing of all partial amplicons confirmed the presence of the GP5-coding region, with lengths of 588 nucleotides corresponding to 196 functional amino acids.

Phylogenetic analysis of glycoprotein 5 (GP5)

Phylogenetic relationships among the study strains were inferred based on amino acid sequence alignments using representative GenBank reference strains from lineages 1, 5, and 8, as shown in Figure 3. In addition, nine isolates previously reported as Peruvian and belonging to the NADC34 sublineage 1.5 (1-7-4) variant were included. The generated phylogenetic tree showed 100% bootstrap support for the separation of *Betaarterivirus europense* and *Betaarterivirus americense*, as well as multiple branches corresponding to clusters formed by representative variants of species 2 within Lineage 1, including sublineage 1.1 USA (99-3584, 99-3298, and 98-3403), sublineage 1.5 (NADC34), sublineage 1.6 (NADC31), sublineage 1.8 (HENAN), and sublineage 1.9 (MN184A); Lineage 5, sublineage 5.1 (VR2332); and Lineage 8, sublineage 8.9 (Ingelvac) and sublineage 8.7 (JXA1). Representative variants of species 1 were also included, namely Lineage 1 sublineage 1.A from South Korea (SNU1 and KNU-07), sublineage 1.C (EuroPRRSV, CRESA-VP21, and Porcilis GP5), and Lineage 3 (BELGA-Lena). Genealogical analysis revealed that the study strains (turquoise branch) clustered with Peruvian isolates reported in 2015, 2016, and 2017 and showed a close evolutionary relationship with the NADC34 sublineage 1.5 variant originally described in North America, Figure 3.

Glycoprotein 5 (GP5) diversity

The translated amino acid sequences (196 aa) consisted of a signal peptide (aa 1–31), decoy epitope I (A) (aa 27–36), decoy epitope II (aa 177–193), which is involved in immune response modulation, the primary neutralizing epitope (B) (aa 37–44), the neutralizing epitope (C) (aa 51–60), transmembrane region 1 (aa 64–80), transmembrane region 2 (aa 94–100), transmembrane

region 3 (aa 108–125), T-cell recognition epitope 1 (aa 115–129), T-cell recognition epitope 2 (aa 148–161), a B-cell epitope (aa 177–196), and a stop codon at position 197, as shown in Figure 4. Multiple amino acid substitutions were identified within antigenic regions of the synthesized glycoprotein. Specifically, several neutral substitutions were observed within the 26 amino acids corresponding to the signal peptide. In contrast, nonsynonymous substitutions with relevant functional implications for the early antibody response mediated by decoy epitope I (A) were detected, particularly within hypervariable region 1. These substitutions included 16/24 at N32 (S/G/R/E), 3/24 at S34 (N/T), and 6/24 at S35 (N/I). Within decoy epitope II, substitutions were observed in 2/24 strains at T185 (P), a site involved in cross-neutralizing antibody responses generated against species 1 and 2. Regions responsible for the induction of neutralizing antibodies were also analyzed. The primary neutralizing epitope B (PNE) is characterized by a highly conserved region encompassing key antigenic determinants at positions 39, 40, and 43. Our results showed substitutions in 7/24 strains at L39 (F), 6/24 at Q40 (L/R), 4/24 at N43 (I), and 5/24 at I46 (V). The neutralizing epitope C, together with hypervariable region 2, includes antigenic determinants at amino acid positions 57 and 58; substitutions were detected in 15/24 strains at K57 (E/V/R) and in 8/24 at S58 (H/R/N), which may facilitate evasion of neutralizing antibodies.

GP5 amino acid analysis

The obtained sequences were analyzed to predict potential N-glycosylation sites within GP5, as these modifications may contribute to viral immune evasion by reducing the neutralizing capacity of epitopes. Glycosylation sites N33, N44, and N51 are located within major neutralizing epitopes. A total of nine potential N-glycosylation sites were identified (N30, N32, N33, N34, N35, N44, N50, N51, and N57), of which 20/24 isolates harbored sites N44 and N51. Furthermore, all isolates contained three or more N-glycosylation sites in their sequences. Notably, a previously unreported site (N57) was identified, and a higher number of glycosylated sites was observed in strains 40 and 41, Table 1. Additionally, signal peptide cleavage site prediction using SignalP v4.0 indicated enzymatic cleavage between residues 31 and 32 of the protein sequence, Table 1.

Statistical analysis: Tajima's D test and genetic diversity using DnaSP v6

Genetic diversity analysis using DnaSP revealed a high number of polymorphic nucleotide positions across the analyzed GP5 sequences, reflecting substantial sequence heterogeneity, with 530 polymorphic sites and 61 conserved sites (Table 2). Tajima's D test yielded a value of −0.78746, with no significant deviation from neutrality ($P > 0.10$). This negative value suggests high genetic variability or the presence of population subdivision with an excess of rare, potentially deleterious mutations.

DISCUSSION

The present study identified 24 strains belonging to the species *Betaarterivirus americense*. Phylogenetic inference was conducted using 58 reference sequences associated with major epidemic outbreaks worldwide, retrieved from public databases, revealing the presence of lineages 1, 5, and 8. GP5-based phylogenetic analysis supported the clustering of the study strains within Lineage 1, sublineage 1.5 of the NADC34 variant, consistent with GP5-based classification approaches and methodologies described by Nguyen et al. [20] and Shin et al. [27]. Our findings are in agreement with those reported by Ramírez et al. [24], who demonstrated that Peruvian isolates, based on RFLP typing, are related to the 1-7-4 variant of Lineage 1, specifically the NADC34 sublineage 1.5.

Previous studies of the GP5 structural protein have demonstrated that antigenic domains play a critical role in immune responses due to the occurrence of synonymous and nonsynonymous substitutions within key immunological regions involved in the induction of non-neutralizing antibodies—such as decoy epitopes I (A) and II—as well as neutralizing epitopes B and C, which collectively contribute to immune evasion and viral neutralization dynamics [19,22]. Overall, our results indicate varying degrees of amino acid substitution across major functional domains, including the signal peptide, transmembrane regions 1, 2, and 3, T-cell recognition epitopes (1 and 2), and the B-cell epitope, when compared with the reference strain MF326985.1_IA/2014/NADC34.

Specifically, substitutions observed in hypervariable region 1 of decoy epitope I (A)—16/24 at N32 (S/G/R/E), 3/24 at S34 (N/T), and 6/24 at S35 (N/I)—highlight substantial positional variability. Variability in this region has been associated with rapid and robust induction of non-neutralizing antibody responses [20]. Decoy epitopes located at aa 27–36 and aa 177–196 are known to function as immune evasion mechanisms: upon infection, these epitopes preferentially stimulate the early production of high titers of non-neutralizing antibodies, thereby delaying the development of effective neutralizing antibodies [37].

Furthermore, substitutions observed at positions 7/24 in L39 (F), 6/24 in Q40 (L/R), and 4/24 in N43 (I) indicate marked variability within the primary neutralizing epitope (B). Faaberg et al. [9] identified amino acids 38, 42, 43, and 44 as key neutralization determinants, and amino acids 39, 40, and 41 as antibody-binding sites; moreover, mutations at position 39 have been used to classify viral variants [37]. These observations collectively support the high genetic variability observed in our strains. Studies by Chen et al. [4] demonstrated that mutations in these regions can reduce *Betaarterivirus americense* sensitivity to neutralizing antibodies in vivo, potentially impairing immune clearance of viral particles and contributing to vaccine failure and increased morbidity in swine herds [15]. However, neutralization capacity was not evaluated in this study, and the

functional consequences of these substitutions for pathogenicity and vaccine escape require domain-specific selection analyses and experimental validation. Accordingly, further in-depth investigations are warranted.

Additionally, our data revealed 15/24 substitutions at K57 (E/V/R) and 8/24 at N58 (K/E/V/R) within neutralizing epitope C and its associated hypervariable region 2. These substitutions have been associated with evasion of neutralizing antibodies [20]. The substitution counts reported for epitopes A, B, and C are descriptive and intended to highlight positional variability within GP5. Formal enrichment or selection analyses—either domain- or site-specific—were not performed in this study and will be addressed in future work. Nonetheless, the observed sequence characteristics are consistent with virulence-associated patterns reported in reference databases and may facilitate the emergence of more pathogenic variants in the Lima region [20], although pathogenicity was not assessed here.

Another critical function of GP5 is its role in membrane protein glycosylation, a process involving the addition or removal of glycan residues at potential N-glycosylation sites. These sites were predicted using the NetNGlyc 1.0 server [12]. Although this tool may generate false positives, the reported incidence is low. NetNGlyc 1.0 identified nine plausible N-glycosylation sites (N30, N32, N33, N34, N35, N44, N50, N51, and N57) within the extravirion region of GP5 across the 24 analyzed strains. Data reported by Ansari et al. [2] suggest that hypoglycosylated wild-type viruses are less common than highly glycosylated strains, supporting the hypothesis of viral variant expansion in the present study. N-glycosylation is thought to mask one or more neutralizing epitopes, thereby preventing the induction of neutralizing antibodies when glycans are present [2].

The observed variability in the number of potential N-glycosylation sites among study strains is consistent with findings reported by Li et al. [17], Xie et al. [35], and Chen et al. [4]. In our dataset, strains 36, 37, 38, 39, 43, and 44 exhibited two to three predicted sites, whereas strains 40 and 41 displayed six sites. According to Ansari et al. [2], the three conserved N-glycosylation sites (N34, N44, and N51) contribute significantly to viral variability and failure of control by commercial vaccines, with glycan addition at N44 being particularly critical for enhancing antibody escape. In this study, strains 40 and 41 retained all three conserved sites, whereas strains 36, 37, 38, and 39 lacked them. Moreover, conserved sites N44 and N51 were present in 20 of the 24 isolates, and 18 isolates harbored more than three N-glycosylation sites.

Zhang et al. [36] reported that an increased number of N-glycosylation sites may reduce the immunogenicity of adjacent neutralizing epitopes, leading us to hypothesize that these isolates may evade immune recognition and neutralization by vaccine-induced antibodies. This observation is highly relevant for the design of effective vaccines against emerging variants, suggesting that next-generation vaccines targeting these regions may need to incorporate modifications in GP5

glycosylation patterns to enhance neutralizing antibody responses [2]. Nevertheless, additional studies are required to determine whether glycosylation site mutations directly affect epitope exposure, immune evasion, and viral pathogenicity [4].

Finally, genetic diversity analysis of the 24 strains using DnaSP revealed high variability among strains circulating in Lima, Peru, with 530 variable nucleotide sites. Tajima's D test indicated substantial population subdivision, yielding a negative D value of −0.78746 [24]. These findings underscore Lima's role as a major swine production center in Peru and provide an initial overview of GP5 diversity in a high-risk setting, as all sequences were obtained from commercial farms in the Lima region. Although metadata were limited, caution is warranted when extrapolating these findings at the national scale. Overall, our results demonstrate high GP5 diversity within the Lima cohort; broader sampling with temporal resolution is required to evaluate national patterns of viral evolution and spread. Such efforts may help explain reduced vaccine compatibility, although vaccine efficacy in Peru was not evaluated in this study.

CONCLUSIONS

- ✓ Phylogenetic analysis based on GP5 glycoprotein sequence alignment demonstrated that the 24 study strains are closely related to Peruvian reference strains and cluster within Lineage 1, sublineage 1.5 of the NADC34 variant of *Betaarterivirus americense*.
- ✓ Significant amino acid substitutions were identified within key immunological domains involved in the induction of neutralizing and non-neutralizing antibodies. These included substitutions at positions 16/24 at N32 (S/G/R/E), 7/24 at L39 (F), 6/24 at Q40 (L/R), 4/24 at N43 (I), 15/24 at K57 (E/V/R), and 8/24 at N58 (K/E/V/R), corresponding to epitopes A, B, and C of the GP5 glycoprotein.
- ✓ Nine potential N-glycosylation sites (N30, N32, N33, N34, N35, N44, N50, N51, and N57) were identified across the 24 strains. Among these, sites N34, N44, and N51 were partially conserved in most strains, fully conserved in strains 40 and 41, and completely absent in strains 36, 37, 38, and 39.
- ✓ Genetic diversity analysis of the ORF5 nucleotide sequences (588 nt) revealed 530 polymorphic nucleotide sites. Statistical analysis using Tajima's D test indicated high genetic variability within the Lima cohort, yielding a negative D value of −0.78746.

SUPPLEMENTARY MATERIAL

- ✓ Supplementary Material 1. Amino acid sequence alignment of glycoprotein 5 using MEGA 6.

FUNDING

✓ This work was carried out and funded at the facilities of the Animal Nutrition and Health Business Unit, Innovation and Development Area, Corporación Montana S.A.

ACKNOWLEDGMENTS

✓ The authors thank Montana S.A. for funding and support in conducting this research, and Bach. Segundo Del Águila Soto for logistical support and valuable suggestions.

REFERENCES

1. Allende R, Kutish GF, Laegreid W, Lu Z, Lewis TL, Rock DL, Friesen J, Galeota JA, Doster AR, Osorio FA. 2000. Mutations in the genome of porcine reproductive and respiratory syndrome virus responsible for the attenuation phenotype. *Archives of Virology* 145(6):1149–1161. https://doi.org/10.1007/s007050070115
2. Ansari IH, Kwon B, Osorio FA, Pattnaik AK. 2006. Influence of N-linked glycosylation of porcine reproductive and respiratory syndrome virus GP5 on virus infectivity, antigenicity, and ability to induce neutralizing antibodies. *Journal of Virology* 80(8):3994–4004. https://doi.org/10.1128/JVI.80.8.3994-4004.2006
3. Brinton MA, Gulyaeva AA, Balasuriya U, Dunowska M, Faaberg KS, Goldberg T, Leung F, Nauwynck HJ, Snijder EJ, Stadejek T, Gorbalenya AE. 2021. ICTV Virus Taxonomy Profile: Arteriviridae 2021. *Journal of General Virology* 102(8):001632. https://doi.org/10.1099/jgv.0.001632
4. Chen XW, Li L, Yin M, Wang Q, Luo WT, Ma Y, Pu ZH, Zhou JL. 2017. Cloning and molecular characterization of the ORF5 gene from a PRRSV-SN strain from Southwest China. *Microbial Pathogenesis* 112:295–302. https://doi.org/10.1016/j.micpath.2017.09.011
5. Cotaquispe NR, De la Cruz VE, Quispe CE, Sitaras N, Toskano HJ. 2022. Molecular characterization of porcine reproductive and respiratory syndrome virus based on the ORF7 nucleocapsid gene in swine farms in Lima, Peru. *Revista de Investigaciones Veterinarias del Perú* 33:e21638. https://doi.org/10.15381/rivep.v33i6.21638
6. Cotaquispe NR, Legua BM, Escajadillo LP, et al. 2025. Molecular characterization of the ORF5 (GP5) gene of the causative agent of porcine reproductive and respiratory syndrome detected in swine farms in Lima, Peru. *Revista Argentina de Microbiología*. https://doi.org/10.1016/j.ram.2025.07.005
7. Dea S, Gagnon CA, Mardassi H, Pirzadeh B, Rogan D. 2000. Current knowledge on the structural proteins of porcine reproductive and respiratory syndrome virus: comparison of North American and European isolates. *Archives of Virology* 145(4):659–688.

8. Dokland T. 2010. The structural biology of PRRSV. *Virus Research* 154(1–2):86–97. https://doi.org/10.1016/j.virusres.2010.07.029
9. Faaberg KS, Hocker JD, Erdman MM, Harris DL, Nelson EA, Torremorell M, Plagemann PG. 2006. Neutralizing antibody responses of pigs infected with natural GP5 N-glycan mutants of porcine reproductive and respiratory syndrome virus. *Viral Immunology* 19(2):294–304. https://doi.org/10.1089/vim.2006.19.294
10. Fang Y, Snijder EJ. 2010. The PRRSV replicase: exploring the multifunctionality of an intriguing set of nonstructural proteins. *Virus Research* 154:61–76.
11. Fang Y, Treffers EE, Li Y, Tas A, Sun Z, van der Meer Y, de Ru AH, van Veelen PA, Atkins JF, Snijder EJ, Firth AE. 2012. Efficient −2 frameshifting by mammalian ribosomes to synthesize an additional arterivirus protein. *Proceedings of the National Academy of Sciences USA* 109:E2920–E2928.
12. Gupta R, Brunak S. 2002. Prediction of glycosylation across the human proteome and the correlation to protein function. *Pacific Symposium on Biocomputing*:310–322.
13. Holtkamp DJ, Kliebenstein JB, Neumann EJ, Zimmerman JJ, Rotto HF, Yoder TK, Wang C, Yeske PE, Mowrer CL, Haley CA. 2013. Assessment of the economic impact of porcine reproductive and respiratory syndrome virus on United States pork producers. *Journal of Swine Health and Production* 21:72–84.
14. Jones DT, Taylor WR, Thornton JM. 1992. The rapid generation of mutation data matrices from protein sequences. *Computer Applications in the Biosciences* 8:275–282.
15. Li B, Fang L, Xu Z, Liu S, Gao J, Jiang Y, Chen H, Xiao S. 2009. Recombination in vaccine and circulating strains of porcine reproductive and respiratory syndrome virus. *Emerging Infectious Diseases* 15(12):2032–2035. https://doi.org/10.3201/eid1512.090390
16. Li B, Xiao S, Wang Y, Xu S, Jiang Y, Chen H, Fang L. 2009. Immunogenicity of the highly pathogenic PRRSV GP5 protein encoded by a synthetic ORF5 gene. *Vaccine* 27(13):1957–1963. https://doi.org/10.1016/j.vaccine.2009.01.098
17. Li B, Fang L, Guo X, Gao J, Song T, Bi J, He K, Chen H, Xiao S. 2011. Epidemiology and evolutionary characteristics of PRRSV in China between 2006 and 2010. *Journal of Clinical Microbiology* 49(9):3175–3183. https://doi.org/10.1128/JCM.00234-11
18. Martínez-Bautista NR, Sciutto-Conde E, Cervantes-Torres J, et al. 2018. Phylogenetic analysis of ORF5 and ORF7 of PRRSV and frequency of wild-type virus in Mexico. *Transboundary and Emerging Diseases* 65:993–1008. https://doi.org/10.1111/tbed.12831
19. Murtaugh MP, Stadejek T, Abrahante JE, Lam TT, Leung FC. 2010. The ever-expanding diversity of PRRSV. *Virus Research* 154(1–2):18–30. https://doi.org/10.1016/j.virusres.2010.08.015
20. Nguyen NH, Tran H, Nguyen TQ, et al. 2022. Phylogenetic analysis of PRRSV in Vietnam, 2021. *Virus Genes* 58(4):361–366. https://doi.org/10.1007/s11262-022-01912-w

21. Park J, Choi S, Jeon JH, Lee KW, Lee C. 2020. Novel lineage 1 recombinants of PRRSV isolated from vaccinated herds. *Archives of Virology* 165(10):2259–2277. https://doi.org/10.1007/s00705-020-04743-y
22. Plagemann PG. 2004. The primary GP5 neutralization epitope of North American PRRSV isolates. *Veterinary Immunology and Immunopathology* 102(3):263–275. https://doi.org/10.1016/j.vetimm.2004.09.011
23. Petersen TN, Brunak S, von Heijne G, Nielsen H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods* 8(10):785–786. https://doi.org/10.1038/nmeth.1701
24. Ramírez M, Bauermann FV, Navarro D, et al. 2019. Detection of PRRSV 1-7-4 type strains in Peru. *Transboundary and Emerging Diseases* 66:1107–1113. https://doi.org/10.1111/tbed.13134
25. Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, et al. 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution* 34(12):3299–3302. https://doi.org/10.1093/molbev/msx248
26. Shi M, Lam TT, Hon CC, et al. 2010. Phylogeny-based evolutionary and geographical dissection of North American PRRSV. *Journal of Virology* 84:8700–8711. https://doi.org/10.1128/JVI.02551-09
27. Snijder EJ, Kikkert M, Fang Y. 2013. Arterivirus molecular biology and pathogenesis. *Journal of General Virology* 94:2141–2163. https://doi.org/10.1099/vir.0.056341-0
28. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution* 30(12):2725–2729.
29. Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123(3):585–595.
30. Tajima F. 1993. Statistical analysis of DNA polymorphism. *Idengaku Zasshi* 68(6):567–595. https://doi.org/10.1266/jjg.68.567
31. Tajima F. 1993. Statistical analysis of DNA polymorphism. *Idengaku Zasshi* 68(6):567–595. https://doi.org/10.1266/jjg.68.567
32. Wang X, Yu G, Zhao J, Cui N, Yu Y, Fan H. 2019. Functional identification of *Corynespora cassiicola*-responsive miRNAs and their targets in cucumber. *Frontiers in Plant Science* 10:668. https://doi.org/10.3389/fpls.2019.00668
33. Wootton S, Yoo D, Rogan D. 2000. Full-length sequence of a Canadian porcine reproductive and respiratory syndrome virus isolate. *Archives of Virology* 145(11):2297–2323.
34. Wu WH, Fang Y, Rowland RR, Lawson SR, Christopher-Hennings J, Yoon KJ, Nelson EA. 2005. The 2b protein as a minor structural component of porcine reproductive and respiratory syndrome virus. *Virus Research* 114:177–181.

35. Xie J, Zhu W, Chen Y, Wei C, Zhou P, Zhang M, Huang Z, Sun L, Su S, Zhang G. 2013. Molecular epidemiology of PRRSV in South China from 2007 to 2011 based on ORF5 genetic analysis. *Microbial Pathogenesis* 63:30–36. https://doi.org/10.1016/j.micpath.2013.05.013
36. Zhang L, Feng Y, Martin DP, Chen J, Ma S, Xia P, Zhang G. 2017. Genetic diversity and phylogenetic analysis of the ORF5 gene of PRRSV from central China. *Research in Veterinary Science* 115:226–234.
37. Zhou L, Kang R, Ji G, Tian Y, Ge M, Xie B, Yang X, Wang H. 2018. Molecular characterization and recombination analysis of PRRSV emerged in southwestern China during 2012–2016. *Virus Genes* 54(1):98–110. https://doi.org/10.1007/s11262-017-1519-y

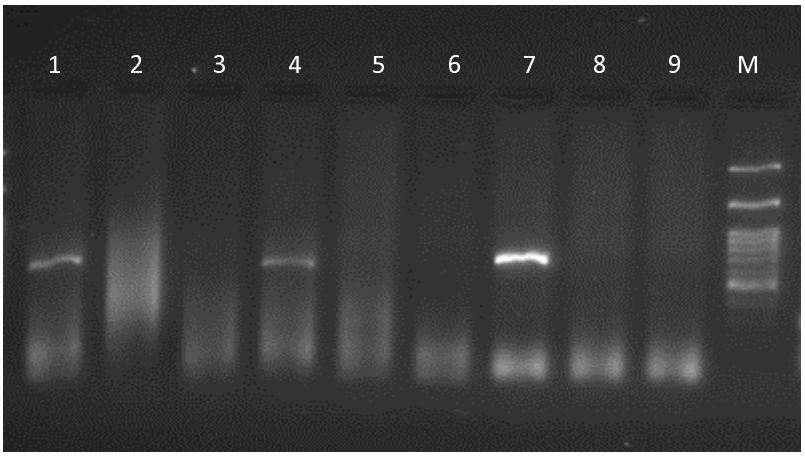


Figure 1. In-house conventional RT-PCR of the ORF5 gene designed in the laboratory. Lane 1 (ORF5 SKA): positive cDNA from the Nebraska Prime Pac® PRRS-2 strain; Lane 4 (ORF5-SKA): positive cDNA from the VP-046 SUIPRABAC® PRRS-1 strain; Lane 7 (ORF5-SKA): positive control cDNA mixture of PRRS-1 and PRRS-2 from the IDEXX commercial kit; Lanes 2, 5, and 8 (avian bronchitis): negative cDNA controls from IBV; Lanes 3, 6, and 9 (PCR blank): nuclease-free water; Lane M (ladder): 100 bp Plus Opti-DNA Marker ABM®. Image adapted from Cotaquispe et al., 2025.

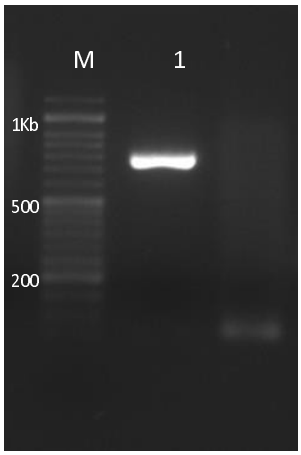


Figure 2. In-house conventional RT-PCR of the ORF5 gene designed in the laboratory. Lane 1 (ORF5-SKA): cDNA from a PRRS field-positive sample; Lane 2 (avian bronchitis): negative control cDNA from IBV; Lane M (ladder): 100 bp Plus Opti-DNA Marker ABM® (50 bp–1.5 kb).

Table 1. Results of the genotypic variability of 24 strains of *Betaarterivirus americense* from blood samples of pigs from the Lima region, Peru.

Cepa	GP5 (aa)	Genotyping	N-glycosylation						Signal peptide
			N° 1	2	3	4	5	6	
25	196	Lineage 1: Sublineage 1.5	30	33	44	51	57		1-31
26	196	Lineage 1: Sublineage 1.5	30	33	44	51	57		1-31
27	196	Lineage 1: Sublineage 1.5	30	33	44	51	57		1-31
28	196	Lineage 1: Sublineage 1.5	30	33	44	51	57		1-31
29	196	Lineage 1: Sublineage 1.5	30	33	44	51			1-31
30	196	Lineage 1: Sublineage 1.5	32	35	44	51	57		1-31
31	196	Lineage 1: Sublineage 1.5	32	35	44	51	57		1-31
32	196	Lineage 1: Sublineage 1.5	32	35	44	51	57		1-31
33	196	Lineage 1: Sublineage 1.5	32	35	44	51	57		1-31
34	196	Lineage 1: Sublineage 1.5	30	33	44	51	57		1-31
35	196	Lineage 1: Sublineage 1.5	30	33	44	51	57		1-31
36	196	Lineage 1: Sublineage 1.5	33	50					1-31
37	196	Lineage 1: Sublineage 1.5	33	50					1-31
38	196	Lineage 1: Sublineage 1.5	33	50					1-31
39	196	Lineage 1: Sublineage 1.5	33	50					1-31
40	196	Lineage 1: Sublineage 1.5	30	33	34	44	51	57	1-31
41	196	Lineage 1: Sublineage 1.5	30	33	34	44	51	57	1-31
42	196	Lineage 1: Sublineage 1.5	33	44	51	57			1-31
43	196	Lineage 1: Sublineage 1.5	44	51	57				1-31
44	196	Lineage 1: Sublineage 1.5	44	51	57				1-31
45	196	Lineage 1: Sublineage 1.5	32	33	44	51			1-31
46	196	Lineage 1: Sublineage 1.5	32	33	44	51			1-31
47	196	Lineage 1: Sublineage 1.5	32	33	44	51			1-31
48	196	Lineage 1: Sublineage 1.5	32	33	44	51	57		1-31

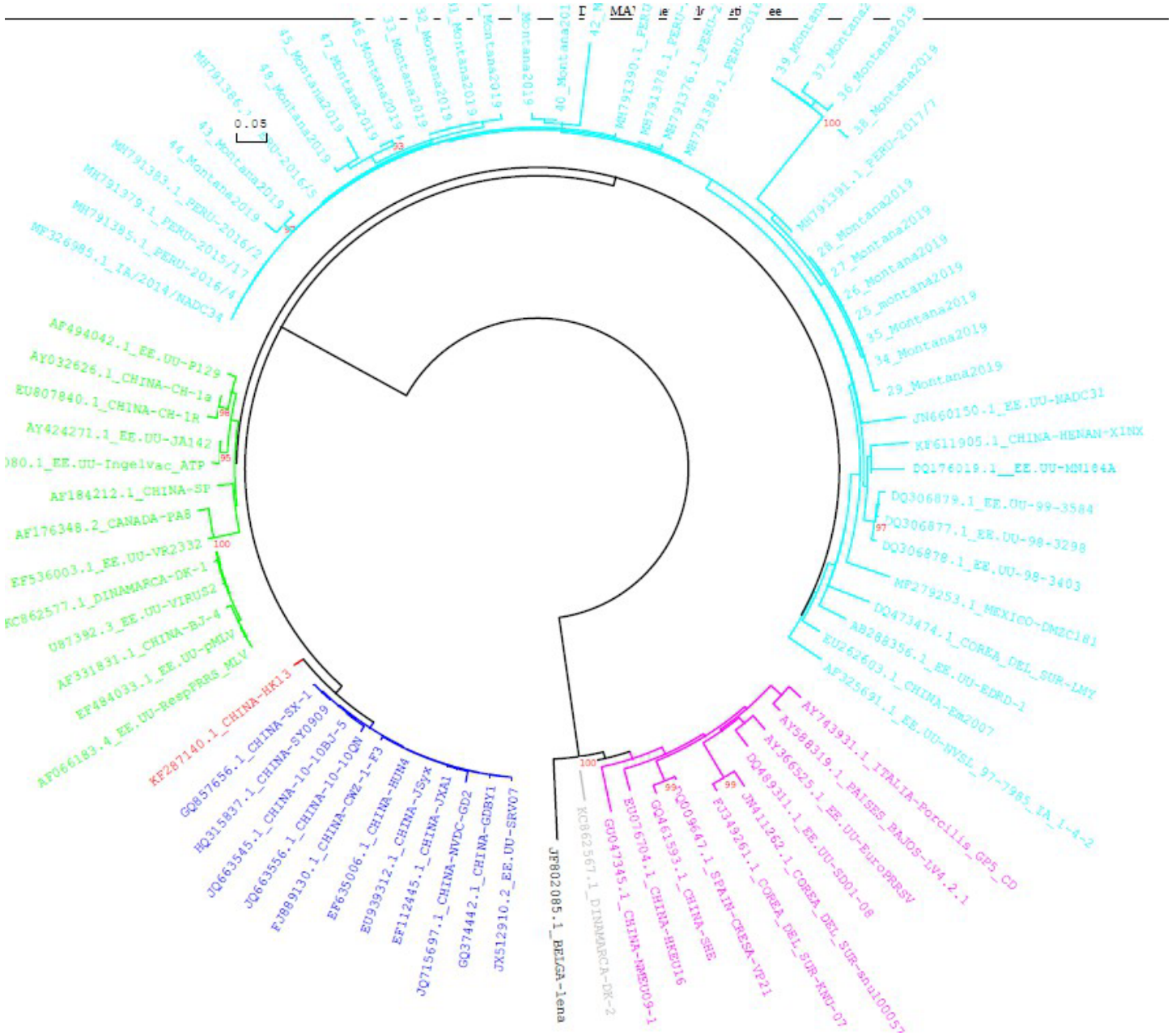

0.05
MH791386.1_PERU-2016/5
45_Montana2019
47_Montana2019
46_Montana2019
33_Montana2019
32_Montana2019
Montana2019
Montana2019
40_Montana2019
42_M
MH791390.1_PERU
MH791378.1_PERU-
MH791376.1_PERU-2
MH791388.1_PERU-2016
39_Montana
37_Montana
36_Montana2019
38_Montana2019
MH791391.1_PERU-2017/7
49_Montana2019
43_Montana2019
44_Montana2019
MH791383.1_PERU-2016/2
MH791379.1_PERU-2015/17
MH791385.1_PERU-2016/4
MF326985.1_IR/2014/NADC34
28_Montana2019
27_Montana2019
26_Montana2019
25_montana2019
35_Montana2019
34_Montana2019
29_Montana2019
JN660150.1_EE.UU-NADC31
KF611905.1_CHINA-HENAN-XINX
DQ176019.1__EE.UU-MN184A
DQ306879.1_EE.UU-99-3584
DQ306877.1_EE.UU-98-3298
DQ306878.1_EE.UU-98-3403
MF279253.1_MEXICO-DMZC181
DQ473474.1_COREA_DEL_SUR-LMY
AB288356.1_EE.UU-EDRD-1
EU262603.1_CHINA-Em2007
AF325691.1_EE.UU-NVSL_97-7985_IA_1-4-2
AY743931.1_ITALIA-Porcilis_GP5_CD
AY588319.1_PAISES_BAJOS-LV4.2.1
AY366525.1_EE.UU-EuroPRRSV
DQ489311.1_EE.UU-SD01-08
JN411262.1_COREA_DEL_SUR-KNU-07
EU349261.1_COREA_DEL_SUR-snu100057
FJ009647.1_SPAIN-CRESA-VP21
GQ461593.1_CHINA-SHE
EU076704.1_CHINA-HBEU16
GU047345.1_CHINA-HBEU09-1
KC862567.1_DINAMARCA-DK-2
JF802085.1_BELGA-lena
AF494042.1_EE.UU-P129
AY032626.1_CHINA-CH-1a
EU807840.1_CHINA-CH-1R
AY424271.1_EE.UU-JA142
080.1_EE.UU-Ingelvac_ATP
AF184212.1_CHINA-SP
AF176348.2_CANADA-PA8
AF261332.1_EE.UU-VR2332
EF536003.1_EE.UU-DK-1
KC862577.1_DINAMARCA-DK-1
U87392.3_EE.UU-VIRUS2
AF331831.1_CHINA-BJ-4
EF484033.1_EE.UU-pMLV
AF066183.4_EE.UU-RespPRRS_MLV
KF287140.1_CHINA-HK13
GQ857656.1_CHINA-SX-1
HQ315837.1_CHINA-SY0909
JQ663545.1_CHINA-10-10BJ-5
JQ663556.1_CHINA-10-10QN
FJ889130.1_CHINA-CWZ-1-F3
EF635006.1_CHINA-HUN4
EU939312.1_CHINA-JSyX
EF112445.1_CHINA-JXA1
JQ715697.1_CHINA-NVDC-GD2
GQ374442.1_CHINA-GDBY1
JX512910.2_EE.UU-SRV07

Figure 3. Phylogenetic tree of *Betaarterivirus europense* and *Betaarterivirus americense* species based on the glycoprotein 5 (GP5) sequence. 1. Purple branching: *Betaarterivirus americense;* 2. Turquoise branching (Lineage 1): sublineage 1.1 (98-3584), sublineage 1.5 (NADC34, study isolates and Peruvian reference strains), sublineage 1.6 (NADC31), sublineage 1.8(HENAN), sublineage 1.9 (MN184A); 3. Light green branching (lineage 5): sublineage 5.1 (Res-PRRS MLV and VR2332); 4. Dark blue branching (lineage 8): sublineage 8.7 (JXA1). Bootstrapping with 1,000 alignment replicates was used, and evolutionary analysis were conducted in DNAMAN 10.

Table 2. Results of the genetic diversity of 24 isolates of *Betaarterivirus americense* analyzed using the DnaSP and MEGA 6 programs from blood samples of pigs from the Lima region, Peru.

Strains	Sequence	ORF 5 (GP5)	C.S.	Variable sites	P. I. S	U. V. S	D Tajima
24	Nt (aa)	588 (196)	61	530	48	82	-0,78746

Nt: nucleotide; **aa**: amino acid; **C.S**.: conserved sites; **P.I.S**.: parsimony-informative sites; **U.V.S**: unique variable sites.

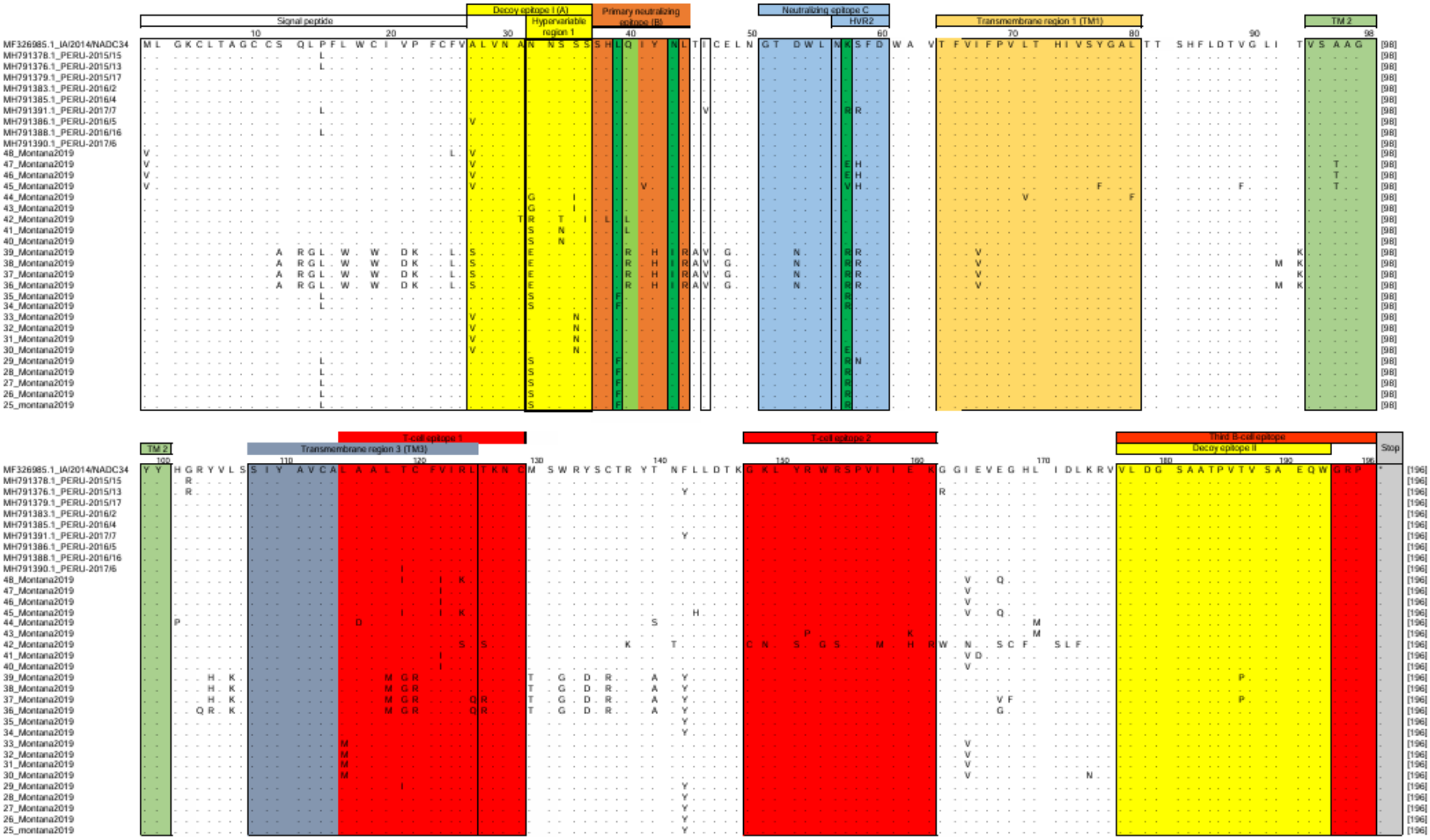
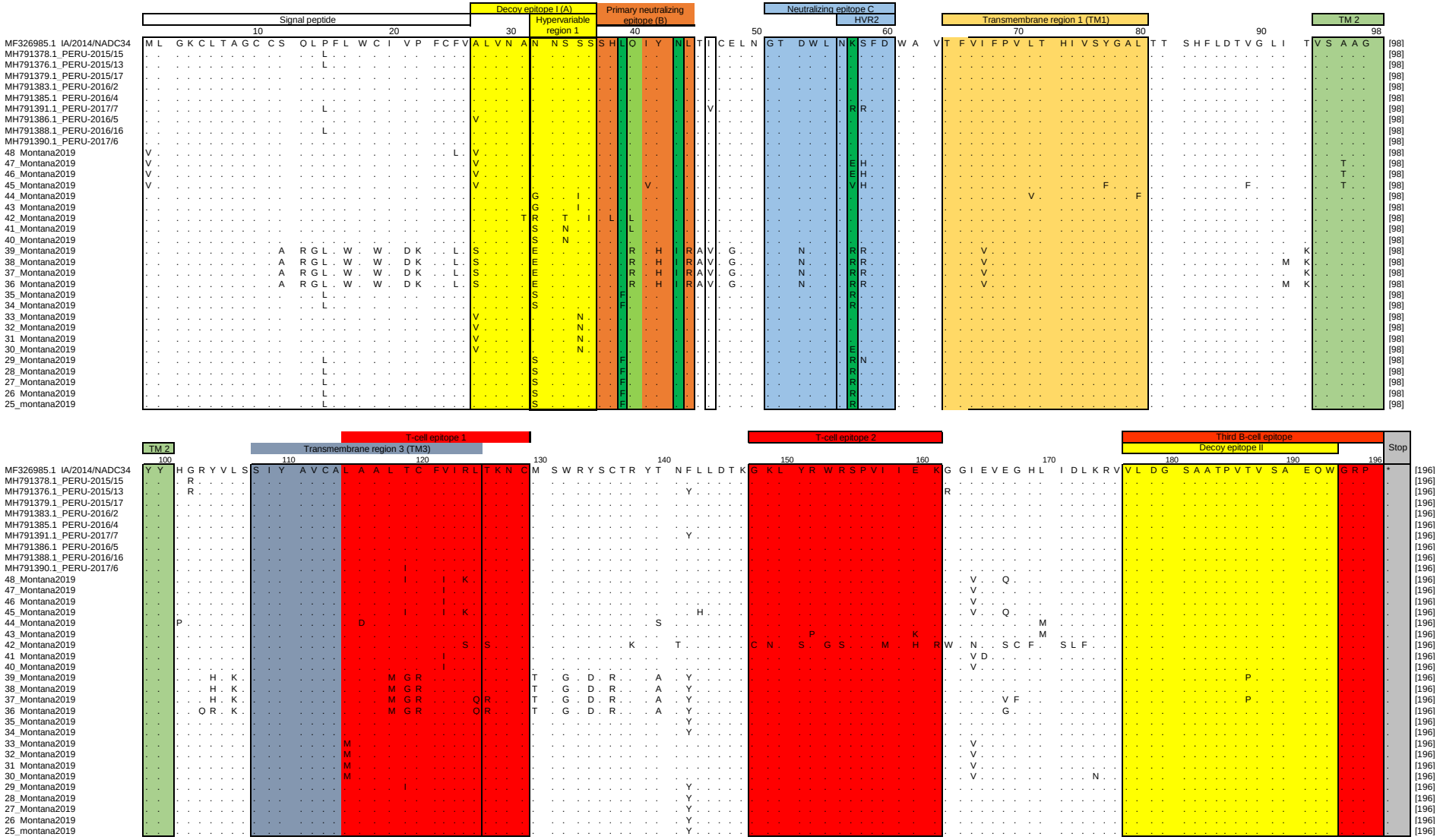

Signal peptide
Decoy epitope I (A)
Hypervariable region 1
Primary neutralizing epitope (B)
Neutralizing epitope C
HVR2
Transmembrane region 1 (TM1)
TM 2
T-cell epitope 1
Transmembrane region 3 (TM3)
T-cell epitope 2
Third B-cell epitope
Decoy epitope II
Stop
MF326985.1_IA/2014/NADC34
MH791378.1_PERU-2015/15
MH791376.1_PERU-2015/13
MH791379.1_PERU-2015/17
MH791383.1_PERU-2016/2
MH791385.1_PERU-2016/4
MH791391.1_PERU-2017/7
MH791386.1_PERU-2016/5
MH791388.1_PERU-2016/16
MH791390.1_PERU-2017/6
48_Montana2019
47_Montana2019
46_Montana2019
45_Montana2019
44_Montana2019
43_Montana2019
42_Montana2019
41_Montana2019
40_Montana2019
39_Montana2019
38_Montana2019
37_Montana2019
36_Montana2019
35_Montana2019
34_Montana2019
33_Montana2019
32_Montana2019
31_Montana2019
30_Montana2019
29_Montana2019
28_Montana2019
27_Montana2019
26_Montana2019
25_montana2019

Figure 4. Alignment of the 24 amino acid sequences corresponding to the antigenic regions of GP5 from *Betaarterivirus americense* sublineage 1.5 (NADC34). Boxed and color-coded regions depict the structural features of glycoprotein 5, highlighting antibody-neutralizing antigenic sites. The 24 field strains, the reference lineage MF326985.1_IA/2014/NADC34, and nine representative Peruvian variants of sublineage 1.5 (NADC34) were included as reference sequences in the alignment. Dots indicate conserved (homologous) amino acid positions, whereas letters denote amino acid substitutions.


Signal peptide
Decoy epitope I (A)
Hypervariable region 1
Primary neutralizing epitope (B)
Neutralizing epitope C
HVR2
Transmembrane region 1 (TM1)
TM 2
MF326985.1 IA/2014/NADC34
MH791378.1_PERU-2015/15
MH791376.1_PERU-2015/13
MH791379.1_PERU-2015/17
MH791383.1_PERU-2016/2
MH791385.1 PERU-2016/4
MH791391.1_PERU-2017/7
MH791386.1_PERU-2016/5
MH791388.1_PERU-2016/16
MH791390.1_PERU-2017/6
48_Montana2019
47_Montana2019
46_Montana2019
45_Montana2019
44_Montana2019
43_Montana2019
42_Montana2019
41_Montana2019
40_Montana2019
39_Montana2019
38_Montana2019
37_Montana2019
36_Montana2019
35_Montana2019
34_Montana2019
33_Montana2019
32_Montana2019
31_Montana2019
30_Montana2019
29_Montana2019
28_Montana2019
27_Montana2019
26_Montana2019
25_montana2019
TM 2
Transmembrane region 3 (TM3)
T-cell epitope 1
T-cell epitope 2
Third B-cell epitope
Decoy epitope II
Stop