Short Communication

Species Characterization in the Genus *Pestivirus* According to Palindromic Nucleotide Substitutions in the 5' Untranslated Region

Massimo Giangaspero^{1,2*} and Ryô Harasawa²

¹National Reference Centre for Wild Animal Diseases (CeRMAS),
Istituto Zooprofilattico Sperimentale del Piemonte, Liguria e Valle d'Aosta,
Via Regione Amerique No. 7/g, 11020 Quart (Aosta), Italy

²Veterinary Microbiology, School of Veterinary Medicine, Faculty of Agriculture, Iwate University,
18-8, Ueda 3 Chome, Morioka 020-8550, Iwate, Japan

*Corresponding author: Cermas@izsto.it (Tel: +39-0165-238558), Ryô Harasawa Veterinary Microbiology, School of Veterinary Medicine, Faculty of Agriculture, Iwate University, 18-8, Ueda 3 Chome, Morioka 020-8550, Iwate, Japan *Corresponding author: giangasp@iwate-u.ac.jp (Tel: +81-196216158)

Abstract

The palindromic nucleotide substitutions (PNS) at the three variable loci (V1, V2 and V3) in the 5' untranslated region (UTR) of the *Pestivirus* genome have been considered for taxonomical segregation of the species, through the evaluation of 533 strains. On the basis of qualitative and quantitative secondary structure characteristics, species have been identified within the genus, determining genetic distances between species isolates, clarifying borderline and multirelated sequences, and characterizing and clustering the *Pestivirus* strains showing unexpected genomic sequences. Nine genomic groups have been identified: the species *Bovine viral diarrhea virus 1* (BVDV-1), *Bovine viral diarrhea virus 2* (BVDV-2), *Border disease virus* (BDV) and *Classical swine fever virus* (CSFV) and the tentative species Pronghorn, Giraffe, *Bovine viral diarrhea virus 3* (BVDV-3) (HoBi group), *Border disease virus 2* (BDV-2) (Italian small ruminant isolates) and Bungowannah.

Palindromic positions have been characterized according to changes in nucleotide base-pairs identifying low variable positions (LVP) including base-pairs present in less than 80% of the genus. The determination of divergence between single strain sequences or genetic groups was obtained easily by comparing base-pairing combinations from aligned secondary structures. This provided clear information such as the level of heterogeneity within a species, the relatedness between species, or facilitating the characterization and clustering of specific strains. The BVDV-1 and BDV species resulted heterogeneous, showing isolates located on a borderline in the species. Within the BVDV-2 species, two main genogroups were identified, with strains showing common sequence characteristics to both groups (multirelated strains). They could be allocated correctly by quantitative analysis. Similarly, the relation between CSFV and BDV species appeared very clearly. Also in this case, ambiguous strain sequences could be clustered in the species showing the lowest divergence values.

In conclusion, the proposed taxonomical procedure is based on the evaluation of only the strategic and highly conserved genome regions in the 5' UTR. Furthermore, the application of quantitative analytical procedure allowed for a better determination of relation among species.

Key words: palindromic nucleotide substitutions, *Pestivirus*, taxonomy, 5'-UTR

The genus *Pestivirus* of the family *Flaviviridae* is represented by four established species *Bovine viral diarrhea virus 1* (BVDV-1), *Bovine viral diarrhea virus 2* (BVDV-2), *Border disease virus* (BDV) and *Classical swine fever virus* (CSFV) and a tentative “Giraffe” species (Faquet et al., 2005). These viruses infect a wide range of ungulate species like swine, cattle and sheep. They can cause prenatal and postnatal infections and heavy economical losses. Persistently infected animals are the main source of infection.

The pestiviruses were classified generally into different viral species according to their host origin: CSFV in pig, BVDV-1 and BVDV-2 in cattle, and BDV in sheep. However, cross infection may occur, thus, this approach resulted not sustainable. Therefore, isolates have to be classified according to their genetic sequence characteristics and the relatedness to viral strains that are used to define the type species. The Pestiviruses genome has a single-stranded, positive polarity RNA, composed by a sequence of about

12,500 nucleotides. It can be divided into three regions: a 5'-untranslated region (UTR), a single large open reading frame encoding a polyprotein, and a 3'-UTR. The virus-encoded polyprotein is cleaved into structural (C, Erns, E1 and E2) and non-structural proteins (Npro, P7, NS2-3, NS4A, NS4B, NS5A, NS5B). Previous reported phylogenetic analysis used different genomic regions, namely 5'-UTR, Npro or E2 genes, to distinguish genotypes within each viral species. Npro refers to an N-terminal autoprotease that has no counterpart in other flaviviruses, whereas the E2 protein plays a major role in virus attachment and entry. In addition, E2 is also important for the induction of neutralizing antibodies. So far, genotyping by using 5'-UTR, Npro or E2 sequences has given consistent results. As the 5'-UTR is relatively more conserved among all members within the genus, this genomic region was used to define pan-pestivirus reactive primers (Vilček et al., 1994), and was also more frequently utilised for the characterization of genotypes. Primary structure analysis, by sequence alignment and construction of phylogenetic trees, is the most common method for the classification of the virus isolates.

The nucleotide substitutions occurring at the level of the 5'-UTR genomic region are particularly important, since positive-sense RNA viruses include generally regulatory motifs, which are indispensable for viral survival. In Pestiviruses, the secondary structure of the 5'-UTR includes an internal ribosomal entry site (IRES), responsible for translational, transcriptional and replicational events (Deng and Brock, 1993). Thus, stable nucleotide variations at this level assume great importance in terms of virus evolutionary history. Nucleotide sequences at the variable loci, V1, V2 and V3, in the 5'-UTR of pestiviruses have been shown to be capable of forming a stable stem-loop structure peculiar to each *Pestivirus* species. The observation of nucleotide variations among virus strains at the level of the three specific variable loci in the secondary structure of the 5'-UTR has been conceived as a simple and practical procedure for genotyping (Harasawa and Giangaspero, 1998). The palindromic nucleotide substitutions (PNS) genotyping method has been further improved from the original concept limited to qualitative analysis (Giangaspero and Harasawa, 2007), allowing the taxonomical segregation of the genus *Pestivirus* into six species: BVDV-1, BVDV-2, CSFV, BDV, the tentative species Giraffe and a new proposed taxon named Pronghorn.

The relatively large number of new deposited sequences of isolates from domestic and wild animals and the recent evidence of novel “atypical” *Pestivirus* sequences, as for example the strains D32/00_‘HoBi’ (Schirrmeier et al., 2004) and Th/04_KhonKaen (Liu et al., 2009a) isolated in cattle infected naturally, in Brazil and Thailand, respectively, or the Bungowannah virus isolated from piglets in Australia (Kirkland et al., 2007), motivated the necessity for an updated application of the PNS method.

Qualitative and quantitative evaluation of genomic sequence divergence, in terms of palindromic nucleotide base pairings variations, has been applied for taxonomical segregation of species, through the evaluation of 533 genomic sequences. The nucleotide sequences in the 5'-UTR of *Pestivirus* strains, with different geographical origin, from different host species or contaminants of biological products, were obtained from the GenBank DNA database, provided by authors or obtained in our laboratories (Table 1) (detailed list of analysed strains available under request).

Nucleotide sequence secondary structures were predicted according to the algorithm of Zuker and Stiegler (1981) using the Genetyx-Mac version 10.1 program package (Software Development Co., Ltd., Tokyo, Japan). The minimum free energy was calculated by the method of Freier et al. (1986). Relevant secondary structure regions in the 5'-UTR were used for genotyping based on the palindromic nucleotide substitutions method (Harasawa and Giangaspero, 1998; Giangaspero and Harasawa, 2007). The identification of species within the genus has been achieved according to highly conserved base pairs in the three variable palindromic loci, based on a series of specific passages focused on the qualitative and quantitative analysis of the secondary structures of the strains in the genus. The first step was qualitative and consisted in the characterization of the different positions of the 3 stems and loops in the 5'-UTR sequences of all the considered strains belonging to the genus. Variable loci were positioned in the prototype *Pestivirus* strain BVDV-1 Osloss sequence (DeMoerlooze et al., 1993) as follows: V1 197-235, V2 267-289 and V3 293-311. Only V2 locus was composed by a constant number of nucleotides (n 23). The variation of loops in V1 and V3 loci determined difference in size of palindromes. Secondary structure sequences showing divergent base pair combinations have been aligned for comparison. For strain comparison, differences in number and type of nucleotides constituting the palindromes have been considered (Giangaspero and Harasawa, 2007). Palindromic positions have been characterized according to

changes in nucleotide base pairs identifying low variable positions (LVP) including base pairs present in less than 80% in the genus, thus, selecting specific base pairings suitable for genus and species determination. The second step was quantitative, allowing the identification of genomic groups among *Pestivirus* strains by clustering the base pair combinations according to LVP. Cross comparisons between types within the genus have been evaluated by computing the divergence percentage, identifying strains showing multirelation or borderlines, and quantifying the heterogeneity of a species and the genetic distance between species in terms of variation of base pairs in the secondary structure.

The observation made on the nucleotide sequences of the three variable loci at the level of the 5'-UTR genomic region of *Pestivirus* strains, according to the PNS genotyping method, allowed to the identification of consensus motifs shared by all the *Pestivirus* species, genus specific base-pairings positioned in the V1 and V2 loci, and characteristic species specific PNS, respectively (Table 3). On the base of qualitative and quantitative secondary structure characteristics, nine species have been identified: *Bovine viral diarrhea virus 1* (BVDV-1), *Bovine viral diarrhea virus 2* (BVDV-2), *Classical swine fever virus* (CSFV), *Border disease virus* (BDV), and the tentative species Pronghorn, Giraffe, *Bovine viral diarrhea virus 3* (BVDV-3) (HoBi group), *Border disease virus 2* (BDV-2) (Italian small ruminant isolates) and Bungowannah. The secondary structure variable loci in the 5'-UTR of the *Pestivirus* species are reported in Figure 1. The genetic markers useful for taxonomic purposes identified at genus level were represented by base pairings from highly conserved, homogeneous positions. At species level, characteristic base pairings determinative for species identification (species determinative at genus level – SDG, not shared with other species) were identified from conserved positions only in BVDV-3, Pronghorn, Giraffe and Bungowannah species. However, clear identification for BVDV-1, BVDV-2, BDV, BDV-2 and CSFV species was obtained by using specific combinations of base-pairings in the sequence (species determinative – SD, from conserved positions species specific, shared with other species). These base-pairings acquired significance only when combined, and they were non specific when considered separately.

The palindromic structures were identifiable in linear sequences. However, it was easier to find them observing the secondary structure. The evaluation of the secondary structure based on a qualitative

observation of the nucleotide variations was supported and confirmed through the application of a quantitative approach. The determination of divergence between single strain sequences or genetic groups was obtained easily by comparing base pairing from aligned secondary structure sequences. This provided clear information such as the level of heterogeneity within a species, the genetic distance between species in terms of variation of base pairs in the secondary structure, or helping for the characterization and clustering of specific strains (Table 4). The BVDV-1 and BDV species resulted heterogeneous, showing borderline strain sequences. Strains 3186V6, J, R, S, W (Vilček et al., 01), and MV98CB95 (Baule et al., 97) isolated from cattle, and the pig strain ZM-95 (Xu et al. 2006) were located on a borderline in the BVDV-1 species. Their sequences showed qualitative similarities with the BVDV-1 species, sharing partially the specific PNS species markers, but with high divergence values, thus, candidates for reclustering as separate groups in the genus. Similarly, the strains Aydın/04-TR and Burdur/05-TR (Oguzoglu et al., 2009), isolated from sheep in Turkey, showing partial genetic relatedness only with members of the BDV species, were clustered in the species as borderline. Within the BVDV-2 species, heterogeneity was particularly marked, represented graphically by a curve characterized by two distinct peaks (Figure 2). The left peak (low divergence values) was related to homogeneity within groups. Comparison of homogeneous groups resulted in higher divergence values (right peak), indicating clear divergence between the groups. However, mean divergence value (12.85) did not allowed their segregation into different species. This corresponded to two identifiable main genogroups: BVDV-2A, characterized by a cosmopolitan diffusion, and BVDV-2B, restricted to South America. In both groups, some strains showed common sequence characteristics to both groups (multirelated strains). They could be allocated correctly by quantitative analysis. Similarly, the relation between CSFV and BDV species appeared very clearly (Figure 3), as example of the genetic relatedness determination among *Pestivirus* species. The computing of the divergence by comparing sequences from both species showed very low values (mean 14.48) when compared to those obtained with other *Pestivirus* species (from 17 and 17.13 with Pronghorn and BDV-2, to 24.05 with Bungowannah) (Table 4). Also in this case, ambiguous strains, sharing common sequence characteristics to both species (multirelated strains), could be clustered in the species showing the lowest divergence values.

The classification among *Pestivirus* species strains according to PNS analysis based on changes in the secondary structure was compared with those based on the 5'-UTR primary structure, carried out through alignment and construction of phylogenetic trees. The results were generally comparable. In particular, new taxons have been defined due to specific base-pairings, despite the limited number of allocated strains. This corresponded to the observations made by different investigators, evaluating the 5'-UTR or other genomic regions (Schirrmeier et al., 2004; Vilček et al., 2005; Cortez et al., 2006; Kirkland et al., 2007). However, some atypical strains were related to controversial taxonomical clustering, in particular for sequences isolated from small ruminants. The sheep isolates from Tunisia (strains 33S, 35, 35T, Lot21, SN1T, SN3G, SN2T, 37A, RM and BM01 isolate 5), reported by Thabti et al. (2005), associated to clinical cases due to the use of a vaccine produced on contaminated cell lines derived from ovine, and natural infections, represented an interesting pestivirus intermediate group, genetically close to CSFV but antigenically related to BDV. They have been reported by the author as members of a new genotype of the BDV species. Two French ovine isolates, 91-F-6731 and 91-F-6732 (Dubois et al., 2008), showed high genetic similarity with the Tunisian strains. The author indicated that both group of strains belong to a novel species named Tunisian. Other studies reported the Tunisian strains as a separate species (Valdazo-Gonzalez et al., 2006, 2007, Liu et al., 2009b), or belonging to the BDV species (De Mia et al., 2005). The application of PNS segregated these isolates in the BDV species, based on the divergence values when compared to CSFV and BDV sequences, despite the sharing of CSFV specific base pairing U-A in position 2 in V3. In particular, the Tunisian and French isolates constituted a homogeneous group with divergence mean of 5.5 within the group. The strains showed lower divergence values with BDV than with CSFV, mean 12.07 and 12.72, respectively. Furthermore, the strains 712/02 (De Mia et al., 2005), LA/91/05 (Giammarioli et al., unpublished) and TO/121/04 (De Mia, unpublished) isolated from small ruminants in Italy, and reported as genotype of BDV species, have been clustered as new taxon in the genus *Pestivirus*, showing a divergence value mean of 16.30 with the BDV species. The ovine strains 098, 119 and 63 from Tunisia (Thabti et al., unpublished) have been clustered within the BVDV-2 species, group BVDV-2A, constituting a separate genotype in addition to the four genotypes, BVDV-2a, BVDV-2b, BVDV-2c, and BVDV-2d previously described (Giangaspero et al., 2008a). In the present study, the

nomenclature of HoBi group isolates as BVDV-3, proposed by Liu et al. (2009b), was adopted in reason to the close relation with the BVDV-1 species.

Problematic sequences showing difficulties for clustering based on primary structure analysis highlight the interest of applying PNS in order to clarify possible taxonomical discrepancies. However, as disadvantage, apart the lack of current informatics design of the method, only 5'-UTR is targeted and it is not applicable a combined study approach on different portions of the genome such as Npro, since no other genomic regions have been explored for the application of PNS.

The PNS demonstrated to offer a reliable alternative for viral investigations to classical taxonomical methods based on primary sequence structure. As secondary structures are essential for survival of the viral population, variations in the involved genomic regions might be related with biological characteristics, which in turn might correlate with classification. The relevant role of the 5' UTR in the translational, transcriptional and replicational mechanisms in pestiviruses implies that its secondary structures (e.g. palindromic structures) must be highly conserved, and that observed mutations are particularly meaningful. Since these mutations are observed as conserved nucleotide changes, they may be considered as the occurrence of a micro evolution step through selection in the phylogenetic history of the virus. On the base of the above mentioned considerations the nucleotide substitutions at the level of the palindromic structures in the genomic 5' UTR may represent a useful genetic marker for taxonomical procedures in pestiviruses and theoretically for other RNA positive strand viruses. Secondary structures predicted at the variable regions in the 5'-UTR showed typical PNS which were useful for classification or genotyping of hepatitis C virus (Giangaspero et al., 2008b).

In conclusion, the proposed method provides results comparable with other taxonomical procedures based on 5' UTR primary structure evaluation, but it differs from them in that only the strategic and highly conserved genome regions in the 5' UTR, and therefore the most meaningful nucleotide sequences at this level, are considered. Thus accurate parameters for species identification in terms of nucleotide sequence homology are made available with great advantage for simplification of virological investigations. Furthermore, the application of quantitative analytical procedure allowed to a better determination of relation among species and genotypes. *Pestivirus* strains showing unexpected genomic sequences such as

Giraffe (Plowright, 1969), Wisent Casimir (Becher et al., 1999), Chamois-1 (Arnal et al., 2004), Pronghorn (Vilček et al., 2005), HoBi group strains (Schirrmeyer et al., 2004; Cortez et al., 2006; Liu et al., 2009a) and Bungowannah (Kirkland et al., 2007) have been characterized easily and clustered within the genus by PNS method. The procedures require constant adaptability to eventual new and different observations. In addition, further investigations will possibly provide new variants in the species and other genotypes or subtypes might be characterized, requiring subsequent revision of identification markers. Nevertheless, the manual searching of relevant base-pairings and direct observation of the sequences still remain the method's main limitations. Further efforts are required in order to develop the PNS method in a fully computerized procedure for easy user-access and rapid testing with reliable results.

References

- Arnal, M., Fernandez-de-Luco, D., Riba, L., Maley, M., Gilray, J., Willoughby, K., Vilček, S., Nettleton, P., 2004. A novel pestivirus associated with deaths in Pyrenean chamois (*Rupicapra pyrenaica pyrenaica*). J. Gen. Virol. 85, 3653-3657.
- Baule, C., van Vuuren, M., Lowings, J.P., Belak, S., 1997. Genetic heterogeneity of bovine viral diarrhoea viruses isolated in Southern Africa. Virus Res. 52 (2), 205-220.
- Becher, P., Orlich, M., Kosmidou, M., Baroth, M., Thiel, J.-H., 1999. Genetic diversity of pestiviruses: identification of novel groups and implications for classification. Virology 262, 64-71.
- Cortez, A., Heinemann, M.B., de Castro, A.M.M.G., Soares, R.M., Pinto, A.M.V., Alfieri, A.A., Flores, E.F., Leite, R.C., Richtzenhain, L.J., 2006. Genetic characterization of Brazilian bovine viral diarrhea virus isolates by partial nucleotide sequencing of the 5'-UTR region. Pesqui Vet. Bras. 26 (4), 211-216.
- De Mia, G.M., Greiser-Wilke, I., Feliziani, F., Giammarioli, M., De Giuseppe, A., 2005. Genetic characterization of a caprine pestivirus as the first member of a putative novel pestivirus subgroup. J. Vet. Med. B Infect. Dis. Vet. Public Health 52 (5), 206-210.
- De Moerloose, L., Lecomte, C., Brown-Shimmer, S., Schmetz, D., Guiot, C., Vandenberg, D., Allaer, D., Rossius, M., Chappuis, G., Dina, D., Renard, A., Martial, J.A. 1993. Nucleotide sequence of the bovine viral

- diarrhoea virus Osloss strain: comparison with related viruses and identification of specific DNA probes in the 5' untranslated region. *J. Gen. Virol.* 74, 1433-1438.
- Deng, R., Brock, K.V., 1993. 5' and 3' untranslated regions of pestivirus genome: primary and secondary structure analyses. *Nucleic Acids Research* 21, 1949-1957.
- Dubois, E., Russo, P., Prigent, M., Thiéry, R., 2008. Genetic characterization of ovine pestiviruses isolated in France, between 1985 and 2006. *Vet. Microbiol.* 130, 69-79.
- Fauquet, C.M., Mayo, M.A., Maniiloff, J., Desselberger, U. and Ball, L.A. 2005. *Virus Taxonomy. Classification and Nomenclature of Viruses.* Elsevier, Academic Press, San Diego.
- Freier, S.M., Kierzek, R., Jaeger, J.A. *et al.* 1986. Improved free-energy parameters for predictions of RNA duplex stability. *Proc. Natl. Acad. Sci. USA* 83, 9373-9377.
- Giangaspero, M., Harasawa, R., 2007. Numerical taxonomy of genus *Pestivirus* based on palindromic nucleotide substitutions in the 5' untranslated region. *J. Virol. Methods* 146, 375-388.
- Giangaspero, M., Harasawa, R., Weber, E.L., Belloli, A., 2008a. Genoepidemiological evaluation of Bovine viral diarrhea virus 2 species based on secondary structures in the 5' genomic untranslated region. *J. Vet. Med. Sci.* 70 (6), 571-580.
- Giangaspero, M., Harasawa, R., Zanetti, A., 2008b. Taxonomy of Genus *Hepacivirus*. Application of Palindromic Nucleotide Substitutions for the Determination of Genotypes of Human Hepatitis C Virus Species. *J. Virol. Methods* 153, 280-299.
- Harasawa, R., Giangaspero, M., 1998. A novel method for pestivirus genotyping based on palindromic nucleotide substitutions in the 5'-untranslated region. *J. Virol. Methods* 70, 225-230.
- Kirkland, P.D., Frost, M.J., Finlaison, D.S., King, K.R., Ridpath, J.F., Gu, X., 2007. Identification of a novel virus in pigs-Bungowannah virus: A possible new species of pestivirus. *Virus Res.* 129 (1), 26-34.
- Liu, L., Kampa, J., Belák, S., Baule, C., 2009a Virus recovery and full-length sequence analysis of atypical bovine pestivirus Th/04_KhonKaen. *Veterinary Microbiology* **138**, 62-8.
- Liu, L., Xia, H., Wahlberg, N., Belák, S., Baule, C., 2009b Phylogeny, classification and evolutionary insights into pestiviruses. *Virology* **385**, 351-357.

- Oguzoglu, T.C., Tan, M.T., Toplu, N., Demir, A.B., Bilge-Dagalp, S., Karaoglu, T., Ozkul, A., Alkan, F., Burgu, I., Haas, L., Greiser-Wilke, I., 2009. Border disease virus (BDV) infections of small ruminants in Turkey: a new BDV subgroup?. *Vet. Microbiol.* 135, 374–379.
- Plowright, W., 1969. Other virus diseases in relation to the JP15 programme. In: Joint Campaign Against Rinderpest. Proceedings of the 1st technical review meeting, phase IV, Mogadiscio. Organization of African Unity, Kenya, 19-23.
- Schirrmeier, H., Strebelow, G., Depner, K., Hoffmann, B., Beer, M., 2004. Genetic and antigenic characterization of an atypical pestivirus isolate, a putative member of a novel pestivirus species. *J. Gen. Virol.* 85 (12), 3647-3652.
- Thabti, F., Letellier, C., Hammami, S., Pepin, M., Ribiere, M., Mesplede, A., Kerkhofs, P. and Russo, P., 2005. Detection of a novel border disease virus subgroup in Tunisian sheep. *Arch. Virol.* 150 (2), 215-229.
- Valdazo-Gonzalez, B., Alvarez-Martinez, M., Greiser-Wilke, I., 2006. Genetic typing and prevalence of Border disease virus (BDV) in small ruminant flocks in Spain. *Vet. Microbiol.* 117 (2-4), 141-153.
- Valdazo-Gonzalez, B., Alvarez-Martinez, M., Sandvik, T., 2007. Genetic and antigenic typing of border disease virus isolates in sheep from the Iberian Peninsula. *Vet. J.* 174 (2), 316-324.
- Vilček, S., Herring, J. A., Nettleton, P. F., Lowings, J. P., Paton, D. J., 1994. Pestiviruses isolated from pigs, cattle and sheep can be allocated into at least three genogroups using polymerase chain reaction and restriction endonuclease analysis. *Arch. Virol.* 136, 309-323.
- Vilček, S., Paton, D.J., Durkovic, B., Strojny, L., Ibata, G., Moussa, A., Loitsch, A., Rossmann, W., Vega, S., Scicluna, M., Palfi, V., 2001. Bovine viral diarrhoea virus genotype 1 can be separated into at least eleven genetic groups. *Arch. Virol.* 146, 99-115.
- Vilček, S., Ridpath, J.F., Van Campen, H., Cavender, J.L., Warg, J., 2005. Characterization of a novel pestivirus originating from a Pronghorn antelope. *Virus Res.* 108, 187-193.
- Xu, X., Zhang, Q., Yu, X., Liang, L., Xiao, C., Xiang, H., Tu, C., 2006. Sequencing and comparative analysis of a pig bovine viral diarrhoea virus genome. *Virus Res.* 122 (1-2), 164-170.
- Zuker, M., Stiegler, P., 1981. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary. *Nucleic Acids Res.* 9, 133-148.

Legend of figures

Figure 1: V1 - V3 palindromic loci in the 5' UTR of the genus *Pestivirus* species. Base pairings characteristic to the genus (PNS genus specific) are shown in bold. The characteristic base pairings of the species BVDV-1, BVDV 2, BDV, CSFV, and the new proposed taxons BVDV-3, BDV-2, Giraffe, Pronghorn and Bungowannah (PNS species specific) are represented in bold and italic. The position of base-pairings is defined by numbering from the bottom of the secondary structures. Watson-Crick base pairings are indicated by a dash (-); tolerated pairings in secondary structure are indicated by an asterisk (*); interchangeable base pairings are indicated by a column (:). M= A or C; R= A or G; W= A or U; S= C or G; Y= C or U; K= G or U; Z= A or C or G; H= A or C or U; D= A or G or U; B= C or G or U; N= A or C or G or U.

Figure 2: Determination of the level of heterogeneity within the single species in the genus *Pestivirus* (PNS method). The divergence values between single strain sequences were obtained by comparing base pairing from aligned secondary structure sequences, helping for the characterization and clustering of specific strains. Species divergence limit value was 13. Within the BVDV-2 species, heterogeneity was particularly marked, represented graphically by a curve characterized by two distinct peaks. The left peak (low divergence values) was related to homogeneity within groups. Comparison of homogeneous groups resulted in higher divergence values (right peak), indicating clear divergence between the groups. This corresponded to two identifiable main genogroups: BVDV-2A, characterized by a cosmopolitan diffusion, and BVDV-2B, restricted to South America. In both groups, some strains showed common sequence characteristics to both groups (multirelated strains). They could be allocated correctly by quantitative analysis.

Figure 3: Determination of the genetic relatedness among *Pestivirus* species (PNS method). The comparison of CSFV strain 5'UTR sequences with those from other species showed a very clear relation between CSFV and BDV species, with very low divergence values (mean 14.48). Ambiguous strains, sharing common sequence characteristics to both species (multirelated strains), could be clustered in the species showing the lowest divergence values.

Table 1. Summary of *Pestivirus* strains (n 533) evaluated according to the Palindromic nucleotide substitution (PNS) method at the 5' untranslated region of RNA.

Species	Number of strains	Host	Geographical origin
BVDV-1	274	Cattle, Sheep, Pig, Deer, Roe deer, Human, Contaminant	Argentina, Austria, Belgium, Brazil, Canada, China, Belgium, France, Germany, India, Ireland, Italy, Japan, New Zealand, Slovakia, South Africa, Spain, Sweden, Switzerland, UK, USA
BVDV-2	77	Cattle Sheep Contaminant	Argentina, Austria, Belgium, Brazil, Canada, France, Germany, Italy, Japan, Netherland, NewZealand, Slovakia, Tunisia, UK, USA
BVDV-3*	3	Cattle	Brazil, Thailand
BDV	131	Sheep, Pyrenean chamois, Cattle, Pig, Reindeer, Wisent	Australia, France, Germany, Japan, New Zealand, Spain, Switzerland, Tunisia, Turkey, UK, USA
BDV-2*	3	Sheep, Goat	Italy
CSFV	43	Pig, Sheep	China, France, Germany, Honduras, Italy, Japan, Malaysia, Netherlands, Poland, Russia, Spain, Switzerland, USA
Pronghorn*	1	Pronghorn	USA
Giraffe*	1	Giraffe	Kenya
Bungowannah*	1	Pig	Australia

* Tentative species.

Table 2: Palindromic nucleotide substitutions (PNS) characteristic to the genus *Pestivirus*. The position of base pairings is defined by numbering from the bottom of the variable locus.

Genus	Locus	Characteristic PNS markers
	V1	Absence in position 22 - size of V1 21 bp (exception U); C C bulge in position 11; A-U in position 10; C-G in position 8 (exceptions U*G, U-A and G G bulge); U-A in position 7 (exception G-C and A A bulge); A in position 6 (exception G); U*G in position 5; U in position 5 right nucleotide; G-C in position 4.
	V2	GGGGU loop (exception GGGGC); C-G in position 8 (exception U*G).
Species	Locus	Characteristic PNS markers
BVDV-1	V1	U-A in position 15 (exception U*G or C-G);
	V2	G-C in position 5 (exception A-U);
	V3	G-C in position 5; A in position 10 (exceptions R:H).
BVDV-2	V1	A-U or A C bulge in position 20 (exceptions Z:H); A,G or U in position 21 (exception GG);
	V2	U-A or U*G in position 6 (exception C A bulge);
	V3	A-U or A C bulge in position 7 (exception G-C).
BVDV-3 tentative species (HoBi group)	V1	U-A in position 15;
	V3	G-C or G*U in position 3; A-U or G-C in position 7; A in position 10.
BDV	V1	G-C or A-U in position 15 (exceptions C U and A C bulges);
	V3	U C and U U bulges or U*G in position 7 (exceptions A-U, U-A and C C, A C, C U and C A bulges).
BDV-2 tentative species (Italian ovine isolates)	V1	U-A or C A bulge in position 15;
	V3	G*U or G G bulge in position 8.
CSFV	V1	U-A in position 13 (exception U*G);
	V3	U-A in position 2; U or C in position 8 (exception A).
Pronghorn tentative species	V1	G-C in position 2; U-A in position 9; U-A in position 12; U-A in position 15;
	V2	G-C in position 4;
	V3	G A bulge in position 5.
Giraffe tentative species	V1	C-G in position 2; U*G in position 20;
	V2	C-G in position 7;
	V3	C-G in position 4; G*U in position 7.
Bungowannah tentative species	V1	A-U in position 2; G-C in position 7; U-A in position 9; U-A in position 12; G-C in position 13;
	V2	A-U in position 3; G-C in position 4;
	V3	U-A in position 4; G A bulge in position 10; A in position 11.

Table 3. Identified species within genus *Pestivirus*.

A) Relations within species.

Species	Divergence within genus	Divergence within species % / Mean value	Multirelated strains	Borderline strains
BVDV-1	0.33	6.43/9.54	None	3186V6, J, R, S, MV98CB95, W, ZM-95
BVDV-2	2.00	11.15/7.89	None	None
BVDV-2A	1.93	0.40/5.37	vs 2B 713-2, 5521-95, BSE1239, UVR420, 97/730	
BVDV-2B	2.25	0/7.28	vs 2A VS-123.4, VS-63	
BVDV-3	2.00	0/4	None	None
BDV	2.22	27.89/11.22	vs CSFV 91-F-6731, 91-F-6732	Aydin/04-TR, Burdur/05-TR
BDV-2	2.33	0/4.5	None	None
CSFV	2.33	0/5.52	vs BDV Kanagawa/74	None
Pronghorn	4	0/0	None	None
Giraffe	7	0/0	None	None
Bungowannah	7	0/0	None	None

B) Relations among species.

Species	Divergence % / Divergence Value Mean									
BVDV-2	100/19.53									
BVDV-2A	100/19.78									
BVDV-2B	100/18.54		33.20/12.85							
BVDV-3	100/16.94	100/17.10	100/17.02	100/17.41						
BDV	100/19.70	100/18.73	100/18.57	100/19.38	100/18.83					
BDV-2	99.62/17.86	100/17.79	100/17.71	100/18.08	100/19.33	97.74/16.30				
CSFV	100/19.29	100/20.68	100/20.63	100/20.88	100/19.51	69.30/14.48	100/17.13			
Pronghorn	100/19.43	100/21.02	100/20.64	100/22.50	100/19.00	100/20.39	100/17.00	100/21.83		
Giraffe	100/26.00	100/20.17	100/19.84	100/21.50	100/23.33	100/21.66	100/20.33	100/23.38	100/22.00	
Bungowannah	100/24.23	100/26.07	100/26.15	100/25.75	100/25.00	100/25.98	100/24.66	100/24.05	100/22.00	100/30.00
	BVDV-1	BVDV-2	BVDV-2A	BVDV-2B	BVDV-3	BDV	BDV-2	CSFV	Pronghorn	Giraffe

PNS pestivirus 9 species

V1

22	(U)							
21	<i>D(GG)</i>		(K,UU)				A	
20	(M, MR) <i>Z H</i>		(G,KB)	G	(UA)		<i>U G</i>	
19	N N	Y N	N N	A-U	N N		A G	GA
18	N N	B:R	AG	N Z	R: C	GU	A U	G U
17	N N	D:N	A G	N N	C-G	Y:G	A-U	A C
16	N N	Y:R	C A	N:N	G-C	G-C	A-U	A A
15	<i>U-A</i>	C-G	<i>U-A</i>	<i>Z:Y</i>	<i>Y:A</i>	R:Y	<i>U-A</i>	A C
14	Y:R	Y:R	A-U	N:N	G-C	R:Y	C-G	G-C
13	C-G	C-G	C-G	C-G	C-G	<i>U:R</i>	C-G	<i>G-C</i>
12	R:Y	R:Y	G-C	Z:Y	A-U	G:Y	<i>U-A</i>	<i>U-A</i>
11	C C	C C	C C	C C	C C	C C	C C	C C
10	A-U	A-U	A-U	A-U	A-U	A-U	A-U	A-U
9	R:Y	R:H	G-C (G)	Z:Y	G-C	R:Y	<i>U-A</i>	<i>U-A</i>
8	<i>C-G</i>	<i>B:R</i>	<i>C-G</i>	<i>C-G</i>	<i>C-G</i>	<i>C-G</i>	<i>C-G</i>	<i>C-G</i>
7	<i>U-AK</i>	<i>U-AK</i>	<i>U-AG</i>	<i>U-AD</i>	<i>U-AG</i>	<i>U-AK</i>	<i>U-AG</i>	<i>U-AU</i>
6	. R	. R	. A	. R	. (U)A	. A	. A	. A
5	<i>U*GU</i>	<i>U*GU</i>	<i>U*GU</i>	<i>U*GU</i>	<i>U*GU</i>	<i>U*GU</i>	<i>U*GU</i>	<i>U*GU</i>
4	<i>G-C</i>	<i>G-C</i>	<i>G-C</i>	<i>G-C</i>	<i>G-C</i>	<i>G-C</i>	<i>G-C</i>	<i>G-C</i>
3	R:Y	G-C	G-C	R:Y	A-U	A-U	G-C	A-U
2	U-A	U-A	U-A	U-A	U-A	U-A	<i>G-C</i>	<i>C-G</i>
1	5'-R:B-3'	5'-R:W-3'	5'-A-U-3'	5'-R:Y-3'	5'-A-U-3'	5'-G:Y-3'	5'-A-U-3'	5'-C-G-3'

V2

12	G	G	G	G	G	G	G	G
11	G G	G G	G G	G G	G G	G G	G G	G G
10	R U	G Y	G C	G U	G U	G U	G U	G U
9	R:Y	G*U	G-C	R: C	G: C	R: C	G*U	G-C
8	<i>C-G</i>	<i>C-G</i>	<i>C-G</i>	<i>C-G</i>	<i>C-G</i>	<i>C-G</i>	<i>C-G</i>	<i>C-G</i>
7	R:Y	G:Y	G-C	R:Y	G-C	R:Y	G*U	<i>C-G</i>
6	Z:H	<i>Y:R</i>	R:U	R:Y	G-C	R:Y	G*U	A-U
5	<i>G-C</i>	R:H	A-U	N:N	G-C	W:W	G-C	A-U
4	Y:R	Y:R	C-G	Y:R	U-A	C-G	<i>G-C</i>	U-A
3	Y:K	C-G	Y:G	H:N	C-G	C-G	C-G	<i>A-U</i>
2	Y:D	Y:G	C-G	B:D	U*G	M:G	C-G	C-G
1	5'-D:H-3'	5'-R:Y-3'	5'-A-U-3'	5'-Z:Y-3'	5'-G*U-3'	5'-A:Y-3'	5'-G*U-3'	5'-A C-3'

V3

11								A
10	A (R H)	(Y)	A	(YU)				<i>G A</i>
9	N H	YY	U U	(A,UU)	A			A-U
8	N N	Y D	U-A	NN	<i>G K</i>	<i>H</i>	UA	C-G
7	N:N	<i>R:Y</i>	<i>R:Y</i>	<i>H N</i>	U:W	Y A	<i>G U</i>	C-G
6	N:N	Y:G	Y:R	H:N	C-G	D:Y	CA	C-G
5	<i>G-C</i>	Y:R	Y:R	C-G	C-G	C-G	<i>G A</i>	G*U
4	R:Y	R:Y	G:S	R:Y	A-U	A-U	A-U	<i>C-G</i>
3	Y:R	C-G	<i>G:Y</i>	Y:R	C-G	C-G	U-A	C-G
2	D:Y	R:Y	G-C	D:H	G*U	<i>U-A</i>	G*U	G*U
1	5'-A:Y-3'	5'-A:Y-3'	5'-A-U-3'	5'-M:U-3'	5'-A-U-3'	5'-A:K-3'	5'-A-U-3'	5'-A-U-3'
	BVDV-1	BVDV-2	BVDV-3	BDV	BDV-2	CSFV	Pronghorn	Giraffe
								Bungwannah

Divergence between CSFV and other species