n_1 and n_4 and between n_2 and n_3 were not considered. The first step realized by the software was the verification of the presence of a structure compatible with the V2 locus in a test sequence, giving that the determination of the following structures depended exclusively from this component. In case no V2 structure could be identified, the analyzed sequence had to be considered not corresponding to 5'-UTR of the genus *Pestivirus*, resulting in the termination of the execution of the program. Alternatively, if one or more V2 structures were identifiable in a given sequence, the program selected the appropriate one and successively continued the analysis.

The following palindrome to be identified was the V3 locus, according to the conserved separation by generally 3 nucleotides, and in some cases 2, as for strains VM, CP1874 and Marloie, from the V2 locus end and the starting of the V3 sequence. Consequently, the determination of the V3 was possible only when the V2 locus was identified. Similarly to V2, the construction was based on the search of a secondary structure with the highest number of Watson-Crick base pairings with strong bindings, taking into account the approximate length of the sequence, variable from 15 to 20 nucleotides, with a 6-7 base pairing composed stem and a 3-5 nucleotide loop. The construction of the stable palindrome was more problematic due to the variability of the sequence length.The V3 starting position was the third or the fourth nucleotide located after the end of the V2 locus sequence. In a first step, possible V3 palindromes were identified starting from the fourth nucleotide after the V2 end. From this nucleotide, n possible V3 structures named $v_{i=1,\ldots,n}$, were determined. Per each v_i structure was identified the one with the highest value of strong bindings. In case of two or more v_i had the same maximum value, the selection was done on the lowest index *i* among them, giving that the structure showed the minor number of nucleotides, thus, the more stable one. In a second step, the procedure is repeated starting from the third nucleotide after the V2 end identifying possible V3 structures identified as $v_{j=1\ldots,m}$. At the end of the two steps, the respective resulted structures v_i and v_j were compared in order to determine the correct V3. The structure showing the highest number of strong bindings and the minimum dimension (constructed starting from the fourth nucleotide after V2, irrespective of the number of nucleotides, in case of similarity) was selected as V3 sequence, resulting the output produced by the program.

The V1 was the last identified palindrome with the appropriate construction with a stem, composed by the highest number of strong bindings, interrupted by a lateral palindrome and the characteristic CC bulge and ending with a variable loop. Particular aspects had to be taken into account, as the variable number of nucleotides included in the palindrome, mainly 39 and up to 42 in BVDV 2 and Giraffe, and the variable starting point of the V1 in the genomic sequence, depending from the type of primers used for RT-PCR reaction. The parameters for the determination of the V1 locus were related to the position of V2 locus in the sequence, as per V3.

Starting from the beginning of the V2 locus and proceeding backward on the sequence, for a minimum of ten nucleotides, were identified G or A followed by a nucleotide and then by U and C (the description is following a backward search, in the sequence this appeared as 5'-nnnnCUnG/Ann..nnV2-3'). All the CUnG/A sequences were retained and recorded in a list. In case no CUnG/A sequences could be identified, the program concluded the procedure for the construction of the V1 locus. Furthermore, continuing backward, all the sequences C and A, followed by a nucleotide and a C or A or G (5'-n..nG/A/CnACnn..nnCUnG/An..nV2-3') were identified and recorded in a list. This search was related to the definition of the core of the V1 locus, identifying the two C composing the characteristic bulge located on the stem of the palindrome and considering two nucleotides more, one forward and the other backward (5'-n..n*nG/A/CnACnn..nnCUnG/An*n..nV2-3'). The V1 locus was completed considering backward five nucleotides, of which the first is usually U (5'-n..n*nnnnUnG/A/CnACnn...nnCUnG/An*n..nV2-3') and forward height nucleotides, of which the second, the third and the fourth were characteristic AUG (exception made for BVDV-1b strain Sanders where A is changed with G) (5'-n..n*nnnnUnG/A/CnACnn..nnCUnG/AnnAUGnnnn*n..nnV2-3'). The remaining part of the stem with the lateral palindrome was so constructed. At the level of the lateral palindrome, on the opposite was included a deletion point in order to equilibrate correctly the stem. Per each CUnG/A identified sequences was performed the construction of a palindromic locus according to each defined C/A/GnAC following sequences. All identified possible structure sequences were constructed (30-n+1) and included in a list of candidate V1 in case of the distance between the beginning of the sequence and the first element before a CUnG/A did not exceed the 30 nucleotides. V1 candidate structures were constructed and considered in the list according to the *m* number of C/A/GnAC identified sequences, with variable loop dimension not exceeding 30 nucleotides. The selected sequence showed the highest number of strong bindings in the stem, and in case of similarity, that with the minor number of nucleotides. The absence of the sequence C/A/GnAC indicated the incompleteness of the V1 locus therefore incomplete palindromes were constructed identifying the starting point of the V1, as at the possible level of nAC, AC or C from an incomplete C/A/GnAC sequence, and determining a series of possible V1 sequences, with different dimensions, with a variable loop

starting from the first, second or third nucleotide of the tested sequence, included in the V1 candidate list(the V1 were constructed only with variable loop dimension not exceeding 30 nucleotides).

Log file

The text file *"log.txt"* created during the execution of the program, is important for monitoring the execution of the program. Its content include all the different issues from all necessary steps for the determination of the final result. In case of unclear or incoherent results, the analysis of this file allowed to understand the cause of the occurred problem as malfunctioning of the program or erroneous rationale.

Application of *Pestivirus* identification keys and html file output

After the comparison between the identified structure and all genus, species and genotype parameters, taking also into account all the known exceptions of divergent base pairs in the genotypes, the final result consisted in a .html file, created by the program during the analytical processing and showed at the end of it. This document shows graphically the portions of the sequence reconstructed and identified as the palindromic V1, V2 and V3 variable loci. The corresponding of PNS *Pestivirus* identification keys in the tested sequence was highlighted at the level of genus and species specific PNS parameters indicating the matching with characteristic base pairings. At the genotype parameter level, the highlighting was applied only in relation to the identified species. Per each PNS pattner, the related control result was included in the output file, showing the expected base pairing and the observed one at each position in the structures, in order to evidence any relevant characteristic in the strategic region of the 5'-UTR.

The software was tested considering the nucleotide sequences in the 5'-UTR of five hundred-tirty-four *Pestivirus* strains of the species BVDV-1, BVDV-2, BDV, CSFV, and of the tentative species Giraffe BVDV-3, BDV-2, Pronghorn and Bungowannah. The sequences, with different geographical origin, from different host species or contaminants of biological products, were obtained from the the GenBank DNA database, provided by authors or obtained in our laboratories (Table 1) (detailed list of analysed strains available under request).

RESULTS

The realization of the PNS software resulted in satisfactory prototype, as demonstrated by the successful application of the testing on a large number of virus strains. The sequences were correctly displayed with their palindromes and the application of the keys for *Pestivirus* identification showed clear results presented in the output file. The identification step allowed to three distinct evaluations. The first was the comparison with genus specific PNS, identifying the appurtenance to the *Pestivirus* genus. The further evaluations were applied only in case of matching. The following comparison was performed with species specific PNS for BVDV 1, BVDV 2, BDV, CSFV and the new proposed taxons. The last comparison was performed for genotype determination within a selected species

A file html type was elaborated at the end of the procedure useful for stocking and printing the results, in which were shown the three palindromic fractions of the sequence, V1, V2 and V3, and the related parameters for genus, species and genotype characterization (Figures 2 and 3).

The secondary structure construction procedure showed results with slight differences and more precision at loop level from those obtained by Genetyx-Mac software, based on the algoritm of Zuker and Stiegler (1981) with minimum free energy calculated according Freier *et al.* (1986). For example, the BVDV-2 strain BS-95-II V1 locus presented a loop consisting of 9 nucleotides (5'-AUCAGUUGA-3'). This sequence was reorganized with a loop reduced to four nucleotides (5'-AGUU-3'), followed by two strong binding base pairs G-C and A-U and an unpaired adenine, resulting in an irregular palindromic shape, when calculated by Genetyx-Mac (not shown). However, this discrepancy was occasional and in general the two applications corresponded. In the Giraffe strain, the V1 locus obtained by Genetyx-Mac showed correct palindromic shape and the two potential strong bindings A-U at the level of the loop were not applied, avoiding a similar alteration as in BS-95-II.

The construction of the V2 palindrome did not presented particular difficulty due to the strictly conserved nucleotide sequences composing the locus. The variability of the nucleotide number composing the other two palindromic structures V1 and V3, was, at the contrary an element requiring particular attention. However, the resulted structures were correct and revealed relation with expected genus characteristics. Table 4 shows an example of construction of the V3 palindrome locus with determination of the length of the sequence. Only in four cases, the construction of V3 of the BVDV-1 strains zvr711, 1248/01, G and W was problematic due to specific and uncommon aspects of their sequences.

The V1 was the last identified palindrome. Particular aspects had to be taken into account. The number of nucleotides included in the palindrome was variable, mainly 39 and up to 42-44 in some BVDV 2 and BDV strains, and Giraffe. The starting point of the V1 in the

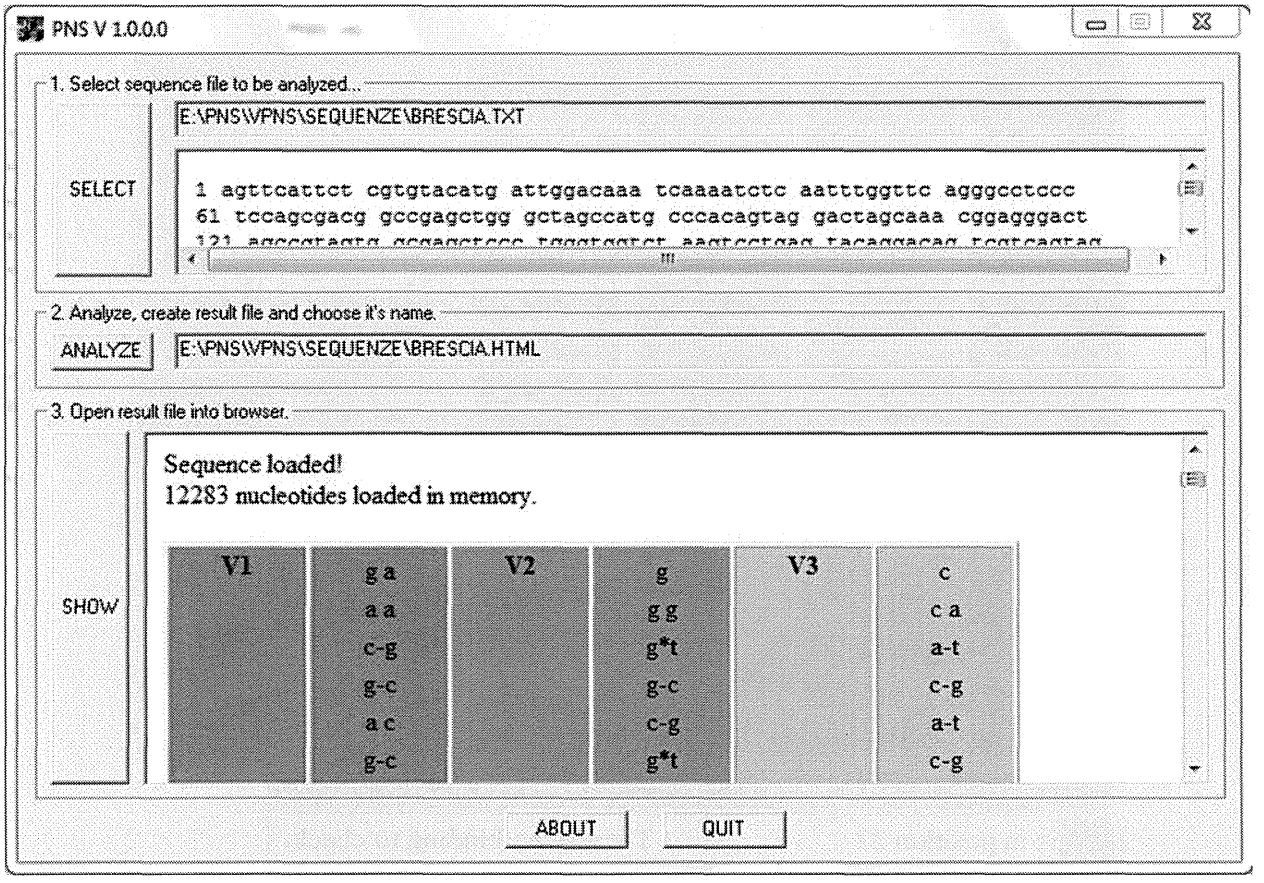


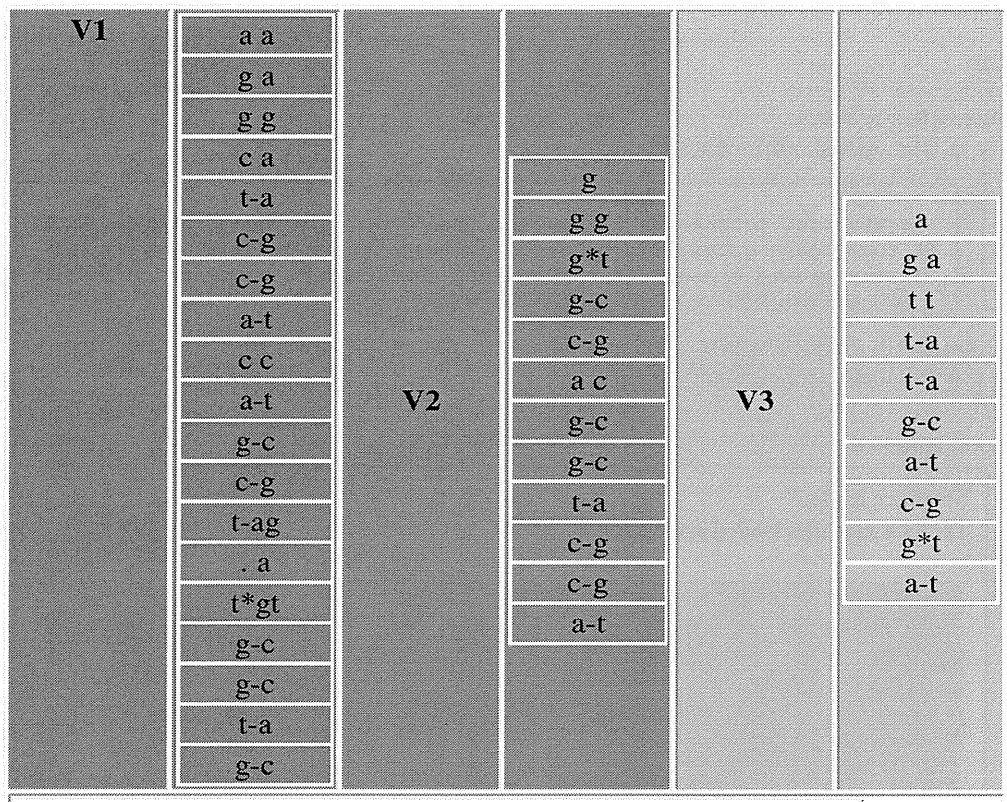
Figure 2. PNS software allowed a simple and intuitive utilization, through selection of sequence to be tested included in a text file format, and display of secondary structure relevant loci with subsequent analytical classification

genomic sequence was variable depending from the type of primers used for RT-PCR reaction. In order to construct the correct V1 palindrome, the specific condition of a minimum number of nucleotides between the C of the characteristic bulge (5'-nG/A/CnACn..nCUnG/An-3'), and between the V1 and V2, had to be applied. This parameter was necessary to avoid the construction of incoherent V1 palindromes, since some strains showed possible combination of G/A/CnAC and CUnG/A consecutive in the sequence. In CSFV strains Alfort and Brescia, starting from the V2, backward for V1 identification, with the first CUnG/A sequence G in position 217 and the following first G/A/CnAC sequence with C in position 210, the program, without correct parameters, constructed a very short V1 (not shown). The correct combination was obtained considering the following G/A/CnAC sequence in position 194. Similar aspects were observed in the BVD-2a strains 713/2, 11mi97 and 552195 and BVDV-1b Influenza2.

In case of incomplete V1 sequences, as for BVDV-1 strains Massimo 1, 2 and 4, the V1 palindrome was constructed according to a predicted structure corresponding to conserved locus in the genus. The program completed the construction of the V2 and V3 palindromes and the related genotyping procedure. The genus and species characteristics were determined. The identified genotype specific PNS were not sufficient to allocate these strains. In case of absence of the V1 sequence, as for BVDV-1 strain M98, BVDV-2 strains 59386 and Scp, and BDV strain L83/I84, the program was orientated to determine species and genotype specific PNS in remaining two loci, V2 and V3. For example, in strain M98 sequence the program could identify the BVDV-1 species specific PNS G-C in position 5 in V2 and in V3.

DISCUSSION

The PNS software represented the transposition of the *Pestivirus* genetic characterization method in a computerized procedure, a relevant improvement with a main advantage represented by the rapidity of the execution of the testing procedure providing data for accurate analyses. The prototype of the program



Genus characteristic PNS: PNS recognized.		
V1:	t in position 22	There is no binding to check.
	t in position 5 right nucleotide	YES (t*gt)
	c c in position 11	YES (c c)
	a t in position 10	YES (a-t)
	c g in position 8	YES (c-g)
	or t a in position 8	NO (c-g)
	or t g in position 8	NO (c-g)
	or g g in position 8	NO (c-g)
	t a in position 7	YES (t-a)
	or a a in position 7	NO (t-a)
	or g c in position 7	NO (t-a)
	. a in position 6	YES (. a)
	or . g in position 6	NO (. a)
	t g in position 5	YES (t*g)
	g c in position 4	YES (g-c)
V2:	g t in position 10	YES (g*t)

Figure 3. Result .hlmt type file of genotyping according to the PNS method of the Europa strain. The 5'-UTR sequence was compared to known *Pestivirus* species BVDV-1, BVDV-2, BDV, CSFV and the new proposed taxons. The three palindromic regions in sequences were identified in the sequence and shown in the first part of the file. Genus *Pestivirus* characteristic PNS and BVDV-1 species characteristic PNS were identified in the palindromes and highlighted in the following section of the file. Genotyping was completed by the evidencing of BVDV-1c specific PNS

successfully demonstrated to be a simple and useful tool for the sequence testing indicating clear results for the allocation of unknown isolates and providing support for research work trough identification of peculiar

Table 4. Palindromic nucleotide substitutions (PNS) genotyping method for genus *Pestivirus*. Construction of the V3 palindrome locus with determination of the length of the sequence. The example was applied on the V3 locus of strain Lees which includes 18 nucleotides. Until now, observations indicated V3 to change form 15 to 20 nucleotides, however, it not possible to exclude new and different data. a) Identification of the linear sequence. The first nucleotide of V3 had a highly conserved position as the fourth nucleotide after the end of V2, and as the third nucleotide only in some strains. Due to the variability of the V3 sequence, the last nucleotide was not determined. b) Determination of the most stable palindromic structure. The sequence was gradually replied onto itself, searching for a maximum number of strong bindings. I. A palindromic structure with 2 strong bindings and 1 weak binding was identified with a 13 nucleotide sequence. II. Sequence with 16 nucleotide showed 3 strong bindings. III. Sequence with 18 nucleotides showed 6 strong bindings and 1 weak binding. IV. A longer sequence showed *instable structure. The 18 nucleotide sequence resulted compatible with V3 locus*

<table>
<tr><td>a)</td><td colspan="2">5'-V2NNNAGCGCCAUUCGUGGCGUUNNNNNNNN-3'
\ last nucleotide of V2</td></tr>
<tr><td rowspan="2">b)</td><td>I.
A
C U
C U
G-C
C-G
G*U
5'-A GGCGUUNNNNNNNNNNN-3'</td><td>II.
UU
A C
C-G
C U
G G
C-G
G-C
5'-A GUUNNNNNNNNNN-3'</td></tr>
<tr><td>III.
UC
U G
A-U
C-G
C-G
G-C
C-G
G*U
5'-A-U-NNNNNN-3'</td><td>IV.
C
U G
U U
A G
C-G
C C
G G
C U
G*U
5'-A-U-3'</td></tr>
</table>

characteristics in strategic genomic regions. In addition to recognized PNS, were made available also all structures indicating similarity or divergence, in terms of specific nucleotide base pairings, among virus genomic sequences at the level of the 5'-UTR, possibly expression of evolutionary changes or virus biological activities, such as virulence (Topliff and Kelling, 1998).

The preparation of the software for the PNS method, presented in this study for the first time, named PNS, freely available at www.pns-software.com, with the full computerization of the procedure, eliminated the main limitation due to manual searching of relevant base pairings and direct observation of the sequence, simplified the genotyping procedure for an easy access of the users and a rapid testing with reliable results, allowing the consideration of secondary structures predicted at the three variable regions in the 5'-UTR for the classification of *Pestivirus*. Future improvement will be required to standardize the procedure and increase the performance of the software in order to eliminate any possible incoherence. This aspect could be important also for possible adaptation of the methodology to other positive polarity RNA virus species, as Poliovirus or Hepatitis C virus.

REFERENCES

Apicella C (1999). Una metodologia per la progettazione ad oggetti di applicazioni relazionali: moduli back-end. Thesis of Doctorate, University of Salerno, Italy.

Deng R, Brock KV (1993). 5´ and 3´ untranslated regions of pestivirus genome: primary and secondary structure analyses. Nucleic Acids Res. 21: 1949-1957.

Ellis MA, Stroustrup B (1990). The Annotated C++ Reference Manual. Editors Addison Wesley, London.

Freier SM, Kierzek R, Jaeger JA, *et al.*, (1986). Improved free-energy parameters for predictions of RNA duplex stability. Proc. Natl. Acad. Sci. USA 83: 9373-9377.

Giangaspero M (2011). Genetic variation of Border Disease Virus species strains. Vet. Ital. 47: 415 - 435.

Giangaspero M, Harasawa R (2011). Classification of Pestivirus species genotypes based on palindromic nucleotide substitutions, a genetic marker in the 5' untranslated region of genomic RNA. J. Virol. Meth. 174: 166-172.

Giangaspero M, Harasawa R (2008). Genetic variation of classical swine fever virus based on palindromic nucleotide substitutions, a genetic marker in the 5' untranslated region of RNA. Vet. Ital. 44: 305-318.

Giangaspero M, Harasawa R, Weber EL, Belloli A (2008). Genoepidemiological evaluation of Bovine viral diarrhea virus 2 species based on secondary structures in the 5' genomic untranslated region. J. Vet. Med. Sci. 70: 571-580.

Giangaspero M, Vacirca G, Harasawa R, Büttner M, Panuccio A, De Giuli Morghen C, Zanetti A, Belloli A, Verhulst A (2001). Genotypes of Pestivirus RNA detected in live virus vaccines for human use. J. Vet. Med. Sci. 63: 723-733.

Harasawa R (1994). Comparative analysis of the 5' non-coding region of pestivirus RNA detected from live virus vaccines. J. Vet. Med. Sci. 56: 961-964.

Harasawa R, Giangaspero M (1998). A novel method for pestivirus genotyping based on palindromic nucleotide substitutions in the 5'-untranslated region. J. Virol. Meth. 70: 225-230.

Harasawa R, Giangaspero M, Ibata G, Paton DJ (2000). Giraffe strain of pestivirus. Its taxonomic status based on the 5' untranslated region. Microbiol. Immunol. 44: 915-921.

King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (2012). Virus taxonomy. Ninth Report of the International Committee on Taxonomy of Viruses. Elsevier-Academic Press, Amsterdam.

Kirkland PD, Frost MJ, Finlaison DS, King KR, Ridpath JF, Gu X (2007). Identification of a novel virus in pigs-Bungowannah virus: a possible new species of pestivirus. Virus Res. 129: 26-3

Lodi E, Pacini G (1998). Introduzione alle strutture di dati. Editors Bollati Boringhieri, Bologna, Italy.

Thompson JD, Gibson TJ, Plewniak F, Jaenmougin F, Higgins DG (1997). The CLUSTAL X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25: 4867–4882.

Topliff CL, Kelling CL (1998). Virulence markers in the 5' untranslated region of genotype 2 bovine viral diarrhea virus isolates. Virology 250: 164-172.

Vilček S, Ridpath JF, Van Campen H, Cavender JL, Warg J (2005). Characterization of a Novel pestivirus originating from a Pronghorn antilope. Virus Res. 108:187-193.

Zuker M, Stiegler P (1981). Optimal computer folding of large RNA sequences using thermodynamics and auxiliary. Nucleic Acids Res. 9:133-148.

Copyright © 2013 American Scientific Publishers
All rights reserved
Printed in the United States of America

Journal of
Bioinformatics and Intelligent Control
Vol. 2, 40–64, 2013

RESEARCH ARTICLE

Palindromic Nucleotide Substitutions Software Version 2.0. Genotyping Based on the Secondary Structure Alignment in the 5' Untranslated Region of *Pestivirus* RNA

Massimo Giangaspero[1,*], Claudio Apicella[2], and Ryô Harasawa[1]

[1]*Department of Veterinary Microbiology, School of Veterinary Medicine, Faculty of Agriculture, Iwate University, Ueda 3-18-8, Morioka, Iwate 020-8550, Japan*
[2]*Directorate General of Veterinary Public Health, Food and Nutrition, Ministry of Health, Via Ribotta 5, 00166 Rome, Italy*

Pestivirus species are responsible for cosmopolitan diseases affecting cattle and other ruminants, presenting a wide range of clinical manifestations, with relevant impact on zootechnic production. Understanding of genomic characteristic and virus taxonomy is fundamental in order to sustain control measures and prophylactic strategies. The palindromic nucleotide substitutions (PNS) in the 5' untranslated region (UTR) of *Pestivirus* RNA have been described as a new, simple and practical method for genotyping. Specific software, named PNS, was made available for users easy access to the method, based on secondary structure sequence analysis. In the present study, the new PNS software version 2.0 was prepared and the application on the genotyping procedures with the keys for *Pestivirus* identification was evaluated on five hundred-forty-three genomic sequences. Using the C# programming language, the software allowed the construction of secondary structure sequence alignment, an additional tool for virus genome research purposes useful for computing genetic distance among strains and numerical taxonomy.

Keywords: Alignment, Genotypes, Palindromic Nucleotide Substitutions, *Pestivirus*, Secondary Structure, Software Version 2.0.

1. INTRODUCTION

Bovine viral diarrhea virus type 1 (BVDV-1), Bovine viral diarrhea virus type 2 (BVDV-2), Border disease virus (BDV) and Classical swine fever virus (CSFV) are established species of the genus *Pestivirus* of the family *Flaviviridae*,[1] responsible for cosmopolitan disease affecting cattle and other ruminants, presenting a wide range of clinical manifestations, with relevant impact on zootechnic production.

According to palindromic nucleotide substitutions (PNS) procedure,[2,3] other strains, characterized by atypical genomic sequences, have been proposed as tentative species in the genus *Pestivirus*. Harasawa et al.[4] characterized the taxonomic status of a giraffe strain, based on the 5' untranslated region (UTR), as a new cluster among *Pestivirus* species. The "Giraffe" strain of *Pestivirus* was isolated in 1967 from a captive giraffe in Kenya which was one of several suffering from a condition resembling mucosal disease.[5] A cytopathic agent was isolated at a post-mortem. This agent, initially named strain H138, was easily passaged in primary calf testis and calf kidney monolayers, and was neutralized by antiserum against the Oregon strain of BVDV. This is thus the only strain isolated so far from a giraffe, which brought up an important question in *Pestivirus* taxonomy. The "Giraffe" strain showed unexpected PNS, which suggested its genetic divergence from the previously described *Pestivirus*. In 2000, the International Committee on Taxonomy of Viruses included the "Giraffe" strain as tentative species in the genus *Pestivirus*.[6] Recently, other *Pestivirus* tentative species determined according to PNS have been proposed: Bovine viral diarrhea virus type 3 (BVDV-3) (HoBi group), Border disease virus type 2 (BDV-2) (Italian small ruminant isolates), Pronghorn and Bungowannah.[7]

Understanding of genomic characteristic and virus taxonomy is fundamental in order to sustain control measures and prophylactic strategies. The palindromic nucleotide

*Author to whom correspondence should be addressed.

 2326-7496/2013/2/040/025 doi:10.1166/jbic.2013.1027

substitutions (PNS) in the 5'-UTR of *Pestivirus* RNA have been described as a new, simple and practical method for genotyping, as an alternative to other analytical procedures based on the primary structure comparison. The PNS method showed comparable results with the primary structure comparison procedures, but it was based exclusively on well conserved and critical region containing an internal ribosomal entry site, responsible for translational, transcriptional and replicational events in pestiviruses,[8] resulting in an accurate determination of strain characteristics and clustering. The PNS genotyping method has been further improved from the original concept which was limited to qualitative analysis.[3] The PNS genotyping procedure was applied on identified *Pestivirus* species, showing heterogeneity among strains clustered within the same species. In the BVDV-2 species five genotypes were identified, BVDV-2a to BVDV-2e.[7,9] The CSFV species showed three genotypes.[10] Eight genotypes have been described in the BDV species, from BDV-a to BDV-h.[11] The nomenclature of genotypes was defined by alphabetical order to distinguish it from species definition by numerical order and was ranked according to increasing divergence in the species, with reference to prevalent base-pairs, thus providing an additional element of strain definition, possibly useful to researchers.

The PNS procedure's main limitation was due to the manual searching of relevant base-pairings and direct observation of the sequence, both complex and time-consuming activities. This refrained researchers from applying the method. With the aim of improving the method by fully computerizing the procedure, specific software was conceived for user easy access and rapid testing with reliable results applying the patterns for *Pestivirus* identification. The prototype of the new software, named PNS, was made available freely at www.pns-software.com.[12]

In the present study, the new PNS software version 2.0 was prepared and the application on the genotyping procedures with the keys for *Pestivirus* identification was preliminarily evaluated on five hundred-forty-three genomic sequences. Using the C# programming language, the software allowed the construction of secondary structure sequence alignment, an additional tool for virus genome research purposes useful for computing genetic distance among strains and numerical taxonomy.

2. EXPERIMENTAL DETAILS

The new PNS software version 2.0 was preliminarily evaluated for the application of the genotyping procedures with the keys for *Pestivirus* identification on genomic sequences. Qualitative and quantitative evaluation of genomic sequence divergence, in terms of palindromic nucleotide base-pairing variations, has been applied for taxonomical segregation of species, through the evaluation of five hundred-forty-three genomic sequences (281 BVDV-1, 77 BVDV-2, 3 BVDV-3, 131 BDV, 5 BDV-2, 43 CSFV, Pronghorn, Giraffe, Bungowannah). The nucleotide sequences in the 5'-UTR of *Pestivirus* strains, with different geographical origins, 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 I). Nucleotide sequence secondary structures were predicted using the PNS software version 2.0 and compared with those obtained with the PNS program package version 1.0.[12] The minimum free energy was calculated by the method described by Freier et al.[91] Relevant secondary structure regions in the 5'-UTR were used for genotyping based on the palindromic nucleotide substitution method.[2,3]

2.1. PNS Software Version 2.0

The development of PNS software version 2.0 was made by applying the C# programming language. The choice of this language was based on practical and organizational considerations. For all the most important functions, therefore a back-end of the previous version of the software an operation of porting in the new language has been undertaken.

The software version 1.0 has been developed in order to evaluate virus nucleotide sequences of a maximum length of 15,000 bases. In the new version, this limit has been eliminated allowing the elaboration of sequences of any dimension using dynamic memorization structures.

The primary objective of the software remained the identification of the three variable loci V1, V2 and V3 in the test sequence, characteristic of the genotyping procedure. The format of the sequences to be submitted in input for a subsequent analysis also remained identical. Therefore, text files (.txt) containing only the characters representing nucleotide bases were also suitable for the new version of the software.

2.2. PNS Class

The class named PNS remained the most important class of the software. Obviously, due to the porting phases from the C++ language to the C# language, the PNS class has been modified in some parts by using new data structures.

2.3. Loading Structure Key Data into Memory

The most important data structures used in the PNS software version 2.0 were as follows:

(1) The memorization of all possible pairs of nucleotides and the relative bindings was performed through the use of a matrix of characters;

(2) The nucleotides read in the input file were memorized in an ArrayList of characters;

Table I. *Pestivirus* strains (n 543) evaluated according to the Palindromic nucleotide substitution (PNS) method at the 5' untranslated region of RNA.

Species	Strain	Origin	Country	Accession	Reference
BVDV-1	0192	Contaminant	Japan	D31799	13
BVDV-1	1041/01	Sheep	Spain	AY159542	14
BVDV-1	107/01	Cattle	Spain	AY159523	14
BVDV-1	10-84	Cattle	France	AF298054	15
BVDV-1	10846/91	Cattle	Germany	AJ304389	16
BVDV-1	10A/LC/97	Cattle	Italy	AJ293602	17
BVDV-1	1103/88	Cattle	Germany	AJ304380	16
BVDV-1	11207/98	Cattle	Germany	AJ304390	16
BVDV-1	114 817	Sheep	UK	U65053	18
BVDV-1	1190V97	Cattle	Italy	not deposited	15
BVDV-1	12	Cattle	Japan	D26051	19
BVDV-1	1248/01	Cattle	Spain	AY159545	14
BVDV-1	125 85	Cattle	Ireland	AJ312909	20
BVDV-1	128/88	Cattle	Germany	AJ304379	16
BVDV-1	133/02	Cattle	Spain	AY159534	14
BVDV-1	1372/01	Cattle	Spain	AY159552	14
BVDV-1	14-102	Cattle	France	not deposited	15
BVDV-1	15-3	Cattle	UK	not deposited	15
BVDV-1	16-111	Cattle	France	AF298056	15
BVDV-1	16484/93	Cattle	Germany	AJ304383	16
BVDV-1	17-112	Cattle	France	AF298057	15
BVDV-1	1 77	Cattle	Ireland	AJ312927	20
BVDV-1	17P	Cattle	Argentina	AF244954	21
BVDV-1	1891/99	Cattle	Germany	AJ304384	16
BVDV-1	1946/01	Cattle	Spain	AY159539	14
BVDV-1	1/A/00	Cattle	Italy	not deposited	22
BVDV-1	1/B/01	Cattle	Italy	not deposited	22
BVDV-1	1/C/01	Cattle	Italy	not deposited	22
BVDV-1	1R	Cattle	Argentina	AF244955	21
BVDV-1	1R93	Contaminant	Argentina	AF244956	21
BVDV-1	2032/01	Cattle	Spain	AY159537	14
BVDV-1	20-V661-2	Cattle	France	AF298058	15
BVDV-1	2110C	Cattle	USA	L20921	23
BVDV-1	2204/82	Cattle	Germany	AJ304377	16
BVDV-1	22146/81	Cattle	Germany	AJ304376	16
BVDV-1	2218/01	Cattle	Spain	AY159525	14
BVDV-1	228/02	Cattle	Spain	AY159554	14
BVDV-1	23-13	Cattle	UK	not deposited	15
BVDV-1	23-15	Cattle	UK	AF298059	15
BVDV-1	2318/01	Cattle	Spain	AY159543	14
BVDV-1	2343/01	Cattle	Spain	AY159541	14
BVDV-1	24/15	Cattle	UK	AF298060	15
BVDV-1	2430 95	Cattle	Ireland	AJ312930	20
BVDV-1	252 84	Cattle	Ireland	AJ312925	20
BVDV-1	2555/01	Cattle	Spain	AY159553	14
BVDV-1	2586×99	Cattle	Ireland	AJ312915	20
BVDV-1	25H	Cattle	Argentina	AF244965	21
BVDV-1	26-V639	Cattle	France	not deposited	15
BVDV-1	2703D 99	Cattle	Ireland	AJ312913	20
BVDV-1	2708/01	Cattle	Spain	AY159536	14
BVDV-1	2750A 99	Cattle	Ireland	AJ312916	20
BVDV-1	28/1	Cattle	UK	AF298061	15
BVDV-1	2900/83	Cattle	Germany	AJ304375	16
BVDV-1	2/A/00	Cattle	Italy	not deposited	22
BVDV-1	2B	Cattle	Argentina	AF244957	21
BVDV-1	2/B/01	Cattle	Italy	not deposited	22
BVDV-1	2L91	Contaminant	Japan	D31800	13
BVDV-1	2/VR/95	Cattle	Italy	AJ293594	17
BVDV-1	3114 93	Cattle	Ireland	AJ312932	20
BVDV-1	318	Contaminant	Argentina	AF244958	21
BVDV-1	3186V6	Cattle	Italy	AF298062	15
BVDV-1	3187V	Cattle	Spain	not deposited	15
BVDV-1	3251/01	Cattle	Spain	AY159550	14
BVDV-1	3291-97-A	Cattle	Austria	not deposited	15

Table I. Continued.

Species	Strain	Origin	Country	Accession	Reference
BVDV-1	3310/01	Cattle	Spain	AY159532	14
BVDV-1	3336/00	Cattle	Spain	AY159549	14
BVDV-1	3340/01	Cattle	Spain	AY159547	14
BVDV-1	3417/00	Cattle	Spain	AY159538	14
BVDV-1	3425/01	Cattle	Spain	AY159521	14
BVDV-1	3478/00	Cattle	Spain	AY159548	14
BVDV-1	3479-97-I	Cattle	Italy	not deposited	15
BVDV-1	3499/00	Cattle	Spain	AY159546	14
BVDV-1	3596/86	Cattle	Germany	AJ304388	16
BVDV-1	368/02	Cattle	Spain	AY159530	14
BVDV-1	371 89	Cattle	Ireland	AJ312918	20
BVDV-1	383 76	Cattle	Ireland	AJ312926	20
BVDV-1	3/A/00	Cattle	Italy	not deposited	22
BVDV-1	3/B/01	Cattle	Italy	not deposited	22
BVDV-1	3P	Cattle	Argentina	AF244968	21
BVDV-1	3/VR/95	Cattle	Italy	AJ293595	17
BVDV-1	4050/00	Cattle	Spain	AY159524	14
BVDV-1	4092/00	Cattle	Spain	AY159533	14
BVDV-1	4163/00	Cattle	Spain	AY159551	14
BVDV-1	4171/00	Cattle	Spain	AY159519	14
BVDV-1	4283/00	Cattle	Spain	AY159527	14
BVDV-1	42M	Cattle	Argentina	AF417999	24
BVDV-1	4325/01	Cattle	Spain	AY159526	14
BVDV-1	438/02	Cattle	Spain	AY159540	14
BVDV-1	4382/01	Cattle	Spain	AY159529	14
BVDV-1	4629/01	Cattle	Spain	AY159528	14
BVDV-1	4771 94	Cattle	Ireland	AJ312920	20
BVDV-1	4796 94	Cattle	Ireland	AJ312910	20
BVDV-1	4898 94	Cattle	Ireland	AJ312931	20
BVDV-1	4998/89	Cattle	Germany	AJ304385	16
BVDV-1	4/B/01	Cattle	Italy	not deposited	22
BVDV-1	4H	Cattle	Argentina	AF244964	21
BVDV-1	4/VR/95	Cattle	Italy	AJ293596	17
BVDV-1	5284/00	Cattle	Spain	AY159544	14
BVDV-1	551/02	Cattle	Spain	AY159520	14
BVDV-1	5551/84	Cattle	Germany	AJ304378	16
BVDV-1	561/01	Cattle	Spain	AY159531	14
BVDV-1	5/B/01	Cattle	Italy	not deposited	22
BVDV-1	65.2	Cattle	Argentina	AF244960	21
BVDV-1	66.1	Cattle	Argentina	AF244953	21
BVDV-1	66.3	Cattle	Argentina	AF244967	21
BVDV-1	66.5	Cattle	Argentina	AF244961	21
BVDV-1	66.6	Cattle	Argentina	AF244966	21
BVDV-1	68.883	Cattle	Argentina	AF244962	21
BVDV-1	6/B/01	Cattle	Italy	not deposited	22
BVDV-1	720/02	Cattle	Spain	AY159535	14
BVDV-1	7535	Sheep	Sweden	U65060	18
BVDV-1	7546	Sheep	Sweden	U65061	18
BVDV-1	7548	Sheep	Sweden	U65062	18
BVDV-1	76865	Cattle	Argentina	AF244969	21
BVDV-1	7/B/01	Cattle	Italy	not deposited	22
BVDV-1	80/1 cp 82	Cattle	Ireland	AJ312929	20
BVDV-1	80/1 ncp 82	Cattle	Ireland	AJ312914	20
BVDV-1	8087 99	Cattle	Ireland	AJ312921	20
BVDV-1	819 85	Cattle	Ireland	AJ312924	20
BVDV-1	832/01	Cattle	Spain	AY159522	14
BVDV-1	86713	Cattle	Argentina	AF244970	21
BVDV-1	8/B/01	Cattle	Italy	not deposited	22
BVDV-1	9189	Cattle	Belgium	ALIGN_000012	25
BVDV-1	9466/91	Cattle	Germany	AJ304382	16
BVDV-1	95-4845A	Deer	New Zealand	U80903	26
BVDV-1	95-4845B	Deer	New Zealand	U80904	26

Table I. Continued.

Species	Strain	Origin	Country	Accession	Reference
BVDV-1	9-77	Cattle	France	not deposited	15
BVDV-1	985 84	Cattle	Ireland	AJ312917	20
BVDV-1	A	Cattle	Austria	AF298064	15
BVDV-1	A014	Contaminant	Japan	D31801	13
BVDV-1	A553	Sheep	UK	U65025	18
BVDV-1	akT1	Cattle	Argentina	not deposited	21
BVDV-1	B1056	Sheep	UK	U65029	18
BVDV-1	B551 98	Cattle	Ireland	AJ312911	20
BVDV-1	BO2340/01	Cattle	Spain	AY159518	14
BVDV-1	BR275	Cattle	Brazil	U94915	Canal et al., unpub.
BVDV-1	BRU*0615	Contaminant	Japan	AB008837	27
BVDV-1	BVR1199	Cattle	Belgium	ALIGN_000012	25
BVDV-1	C009T	Contaminant	Japan	D31802	13
BVDV-1	cb1	Cattle	Argentina	AF417998	Jimenez et al., unpublished
BVDV-1	CD89	Cattle	Belgium	ALIGN_000012	25
BVDV-1	CP1872	Cattle	Belgium	ALIGN_000012	25
BVDV-1	CP1874	Cattle	Belgium	ALIGN_000012	25
BVDV-1	CP1885	Cattle	Belgium	ALIGN_000012	25
BVDV-1	CP1887	Cattle	Belgium	ALIGN_000012	25
BVDV-1	CP1940	Cattle	Belgium	ALIGN_000012	25
BVDV-1	CP1945	Cattle	Belgium	ALIGN_000012	25
BVDV-1	CRFK	Contaminant	Japan	D50814	28
BVDV-1	Culi1	Cattle	Belgium	ALIGN_000012	25
BVDV-1	Culi4	Cattle	Belgium	ALIGN_000012	25
BVDV-1	Culi6	Cattle	Belgium	ALIGN_000012	25
BVDV-1	CV-1	Contaminant	Japan	D50815	29
BVDV-1	D	Cattle	Austria	not deposited	15
BVDV-1	D1120/1	Sheep	UK	U65032	18
BVDV-1	D1432/P	Sheep	UK	U65033	18
BVDV-1	D771/1	Sheep	UK	U65030	18
BVDV-1	D861	Sheep	UK	U65031	18
BVDV-1	Deer	Deer	UK	not deposited	26
BVDV-1	DeerGB1	Deer	New Zealand	U80902	30
BVDV-1	Draper	Cattle	USA	L32880	31
BVDV-1	Europa	Human	Belgium	AB000898	32
BVDV-1	F	Cattle	Austria	AF298065	15
BVDV-1	FLK	Contaminant	Belgium	ALIGN_000012	25
BVDV-1	FS720	Contaminant	Japan	D31803	13
BVDV-1	FU411	Contaminant	Japan	D31804	33
BVDV-1	G	Cattle	Austria	AF208066	15
BVDV-1	H	Cattle	Austria	not deposited	15
BVDV-1	H503	Contaminant	Japan	AB008841	27
BVDV-1	H686 98	Cattle	Ireland	AJ312923	20
BVDV-1	H851 98	Cattle	Ireland	AJ312908	20
BVDV-1	HC725	Contaminant	Japan	D31805	13
BVDV-1	HE726	Contaminant	Japan	D31806	13
BVDV-1	HE728	Contaminant	Japan	D31808	13
BVDV-1	HeLa	Contaminant	Japan	D50819	29
BVDV-1	HH	Contaminant	Japan	D50818	28
BVDV-1	i13	Cattle	Argentina	AF417989	Jimenez et al., unpublished
BVDV-1	i297	Cattle	Argentina	AF417997	24
BVDV-1	i36P	Cattle	Argentina	not deposited	21
BVDV-1	i393	Cattle	Argentina	AF417992	24
BVDV-1	i467	Cattle	Argentina	AF417994	24
BVDV-1	i53	Cattle	Argentina	AF417987	24
BVDV-1	i63	Cattle	Argentina	AF417990	Jimenez et al., unpublished
BVDV-1	i66.2	Cattle	Argentina	not deposited	21
BVDV-1	i6.89	Cattle	Argentina	not deposited	21
BVDV-1	i720	Cattle	Argentina	AF417988	24
BVDV-1	i736	Cattle	Argentina	AF417993	Jones et al., unpublished
BVDV-1	i89	Cattle	Argentina	AF417991	24
BVDV-1	Ind 446	Cattle	India	AY279087	34
BVDV-1	Ind S 1222	Cattle	India	AY278459	34
BVDV-1	Ind S 1166	Cattle	India	AY278460	34

Table I. Continued.

Species	Strain	Origin	Country	Accession	Reference
BVDV-1	Ind S 1168	Cattle	India	AY279086	34
BVDV-1	Ind S 1170	Cattle	India	AY279526	34
BVDV-1	Ind S 1171	Cattle	India	AY279527	34
BVDV-1	Ind S 1181	Cattle	India	AY279528	34
BVDV-1	Influenza2	Contaminant	Switzerland	AB010146	35
BVDV-1	IQ19A	Contaminant	Japan	D31812	13
BVDV-1	J	Cattle	Austria	AF298067	15
BVDV-1	JE	Contaminant	Japan	D26611	33
BVDV-1	K869 98	Cattle	Ireland	AJ312912	20
BVDV-1	KA-91	Cattle	Japan	AB019684	36
BVDV-1	KM	Cattle	Slovakia	AF298068	15
BVDV-1	KQ25A	Contaminant	Japan	D31809	13
BVDV-1	KQ25B	Contaminant	Japan	D31810	13
BVDV-1	KS86-1ncp	Cattle	Japan	AB042713	37
BVDV-1	Kyj	Cattle	Slovakia	not deposited	15
BVDV-1	L	Cattle	Austria	AF298069	15
BVDV-1	L1000 98	Cattle	Ireland	AJ312919	20
BVDV-1	L256	Cattle	France	ALIGN_000012	25
BVDV-1	L322 98	Cattle	Ireland	AJ312922	20
BVDV-1	Lamspringe735	Cattle	Germany	AJ304391	16
BVDV-1	Lamspringe738	Cattle	Germany	AJ304392	16
BVDV-1	Ln 68	Cattle	Ireland	AJ312928	20
BVDV-1	LQ28A	Contaminant	Japan	D31811	13
BVDV-1	M065B/93	Cattle	South Africa	U97409	38
BVDV-1	M1515A/90	Cattle	South Africa	U97429	38
BVDV-1	M169B/93	Cattle	South Africa	U97430	38
BVDV-1	M245A/91	Cattle	South Africa	U97436	38
BVDV-1	M346T/96	Cattle	South Africa	U97440	38
BVDV-1	M388A/90	Cattle	South Africa	U97442	38
BVDV-1	M557A/90	Cattle	South Africa	U97449	38
BVDV-1	M657GX/95	Cattle	South Africa	U97455	38
BVDV-1	M98	Contaminant	Switzerland	AB014339	27
BVDV-1	Marloie	Cattle	Belgium	ALIGN_000012	25
BVDV-1	Massimo1	Contaminant	Switzerland	AB008838	27
BVDV-1	Massimo2	Contaminant	Italy	AB008839	27
BVDV-1	Massimo4	Contaminant	Switzerland	AB008840	35
BVDV-1	MDBK	Contaminant	Japan	D50820	29
BVDV-1	MDCK	Contaminant	Japan	D50821	29
BVDV-1	MMR-K	Contaminant	Japan	D26050	19
BVDV-1	MOLT-4	Contaminant	Japan	D50822	29
BVDV-1	Mumps	Contaminant	Japan	D26049	19
BVDV-1	MV39CB/95	Cattle	South Africa	U97465	38
BVDV-1	MV98CB/95	Cattle	South Africa	U97467	38
BVDV-1	NADL	Cattle	USA	M31182	39
BVDV-1	ncp2	Cattle	Argentina	not deposited	21
BVDV-1	NY-1	Cattle	USA	L32879	31
BVDV-1	Oregon	Cattle	USA	L32876	31
BVDV-1	Osloss	Cattle	USA	M96687	40
BVDV-1	P	Cattle	Austria	AF298070	15
BVDV-1	PT810	Cattle	Germany	Z79766	41
BVDV-1	Q1161/1	Sheep	UK	U65040	18
BVDV-1	Q1161/2	Sheep	UK	U65041	18
BVDV-1	Q713	Cattle	Canada	L32882	31
BVDV-1	R	Cattle	Austria	AF298071	15
BVDV-1	R1935/72	Cattle	Brazil	U94916	Canal et al., unpub.
BVDV-1	Rebe	Cattle	Switzerland	AF299317	42
BVDV-1	Renindeer2	Pig	Germany	not deposited	Giangaspero et al., unpub.
BVDV-1	S	Cattle	Austria	not deposited	15
BVDV-1	S21	Cattle	Argentina	AF244963	21
BVDV-1	S-ALT5/K	Cattle	South Africa	U97474	38
BVDV-1	Sanders	Cattle	USA	L20928	23
BVDV-1	SD-1	Cattle	USA	M96751	43

Table I. Continued.

Species	Strain	Origin	Country	Accession	Reference
BVDV-1	SE1015	Cattle	Germany	Z79767	Wolfmeyer et al., unpub.
BVDV-1	SE5572	Cattle	Germany	Z79770	41
BVDV-1	SE5726	Cattle	Germany	Z79778	41
BVDV-1	SH9/11	Roe deer	Germany	not deposited	44
BVDV-1	Singer	Cattle	USA	L32875	31
BVDV-1	so CP/75	Cattle	Japan	AB042661	37
BVDV-1	SuwaCp	Cattle	Switzerland	AF117699	Schweizer et al., 45
BVDV-1	T	Cattle	Austria	AF298072	15
BVDV-1	TFB	Contaminant	Argentina	AF244971	21
BVDV-1	TFB2	Cattle	Argentina	AF418000	Jones et al., unpublished
BVDV-1	TGAC	Cattle	USA	not deposited	23
BVDV-1	TGAN	Cattle	USA	not deposited	23
BVDV-1	TK-87-2	Cattle	Japan	AB019669	36
BVDV-1	TY CP/91	Cattle	Japan	AB042670	37
BVDV-1	U	Cattle	Austria	not deposited	15
BVDV-1	U937	Contaminant	Japan	D50823	29
BVDV-1	Vero	Contaminant	Japan	D50824	28
BVDV-1	Vkl	Cattle	Slovakia	not deposited	15
BVDV-1	VM	Cattle	USA	L20933	23
BVDV-1	W	Cattle	Austria	AF298073	15
BVDV-1	Weybridge	Sheep	UK	U65024	18
BVDV-1	Wi-38	Contaminant	Japan	D50825	29
BVDV-1	WiDr	Contaminant	Japan	D50826	29
BVDV-1	YVD947	Cattle	Belgium	ALIGN_000012	25
BVDV-1	YVR2394	Cattle	Belgium	ALIGN_000012	25
BVDV-1	ZM-95	Pig	China	AF526381	46
BVDV-1	ZVD278	Cattle	Belgium	ALIGN_000012	25
BVDV-1	ZVR711	Cattle	Belgium	ALIGN_000012	25
BVDV-2	098	Sheep	Tunisia	AF462004	Thabti et al., unpublished
BVDV-2	104/98	Cattle	Germany	AJ304381	16
BVDV-2	11/Mi/97	Cattle	Italy	AJ293603	17
BVDV-2	119	Sheep	Tunisia	AF462003	Thabti et al., unpublished
BVDV-2	15-103	Cattle	France	AF298055	15
BVDV-2	167 237	Sheep	UK	U65055	18
BVDV-2	168 149	Sheep	UK	U65056	18
BVDV-2	17011-96	Cattle	USA	AF039179	47
BVDV-2	173 157	Sheep	UK	U65058	18
BVDV-2	175 375	Sheep	UK	U65059	18
BVDV-2	17583-97	Cattle	USA	AF039176	47
BVDV-2	23025	Cattle	USA	AF039172	47
BVDV-2	34b	Cattle	Argentina	AF244952	21
BVDV-2	354	Contaminant	Argentina	AF244959	21
BVDV-2	37Gr	Cattle	Austria	EU327594	48
BVDV-2	4-5174	Cattle	France	AF298063	15
BVDV-2	5521-95	Cattle	USA	AF039174	47
BVDV-2	59386	Sheep	UK	U17146	49
BVDV-2	63	Sheep	Tunisia	AF462005	Thabti et al., unpublished
BVDV-2	713-2	Cattle	USA	AF039177	47
BVDV-2	7937	Cattle	USA	AF039175	47
BVDV-2	890	Cattle	Canada	L32886	31
BVDV-2	97/730	Cattle	NewZealand	AF026770	Vilcek et al., unpublished
BVDV-2	AF112	Cattle	Germany	not deposited	50
BVDV-2	AZ Spl	Cattle	USA	not deposited	51
BVDV-2	B45-5	Cattle	Germany	not deposited	50
BVDV-2	B50-5	Cattle	Germany	not deposited	50
BVDV-2	B52-2	Cattle	Germany	not deposited	50
BVDV-2	B5-4	Cattle	Germany	not deposited	50
BVDV-2	B77-5	Cattle	Germany	not deposited	50
BVDV-2	BD-78	Sheep	USA	U18330	52
BVDV-2	BM01 isolate 11	Sheep	Tunisia	AF462006	Thabti et al., unpublished
BVDV-2	BS-95-II	Cattle	Italy	AJ288903	17
BVDV-2	BSE1239	Cattle	Belgium	ALIGN_000012	25

Table I. Continued.

Species	Strain	Origin	Country	Accession	Reference
BVDV-2	BSE341	Cattle	Belgium	ALIGN_000012	25
BVDV-2	BSE921	Cattle	Belgium	ALIGN_000012	25
BVDV-2	C413	Sheep	USA	AF002227	Chen & Berry, unpublished
BVDV-2	CD87	Cattle	Canada	L32887	31
BVDV-2	CPA	Contaminant	Japan	D50812	29
BVDV-2	CPAE	Contaminant	Japan	D50813	29
BVDV-2	EBTr	Contaminant	Japan	D50817	28
BVDV-2	Giessen-1	Cattle	Germany	AF104030	30
BVDV-2	HE727	Contaminant	Japan	D31807	29
BVDV-2	i33283	Cattle	Argentina	AF417996	24
BVDV-2	i4083	Cattle	Argentina	AF417995	24
BVDV-2	i61380	Cattle	Argentina	AF417986	24
BVDV-2	i628	Cattle	Argentina	AF417985	24
BVDV-2	IT-1732	Contaminant	Italy	AJ416018	Muscillo, unpublished
BVDV-2	Kosice	Cattle	Slovakia	EU360934	53
BVDV-2	Lees	Sheep	UK	U65051	18
BVDV-2	LV-96	Cattle	Brazil	AF410787	54
BVDV-2	MAD Spl	Cattle	USA	not deposited	51
BVDV-2	MMR-T	Contaminant	Japan	D26052	19
BVDV-2	MN Fetus	Cattle	USA	not deposited	51
BVDV-2	MP	Contaminant	Belgium	ALIGN_000012	25
BVDV-2	MS-1	Cattle	Japan	AB019688	36
BVDV-2	Munich 1	Cattle	Germany	not deposited	50
BVDV-2	Munich 2	Cattle	Germany	not deposited	50
BVDV-2	Munich 3	Cattle	Germany	not deposited	50
BVDV-2	ncp7	Cattle	Argentina	not deposited	21
BVDV-2	NY93	Cattle	USA	AF039173	47
BVDV-2	OY89	Cattle	Japan	AB003621	55
BVDV-2	Parvo	Contaminant	Japan	D26614	56
BVDV-2	Q126	Cattle	Canada	L32890	31
BVDV-2	Rubella	Contaminant	Japan	D26048	56
BVDV-2	SCP	Contaminant	UK	U17148	49
BVDV-2	Soldan	Cattle	Brazil	U94914	Canal et al., unpublished
BVDV-2	SW90	Cattle	Japan	AB003622	37
BVDV-2	SY-89	Cattle	Japan	AB019689	36
BVDV-2	TC Shinozaki	Cattle	Japan	AB04267	37
BVDV-2	UVR420	Cattle	Belgium	ALIGN_000012	25
BVDV-2	V-FLL	Cattle	Japan	AB019687	36
BVDV-2	VS-123.4	Cattle	Brazil	AF410790	54
BVDV-2	VS-260	Cattle	Brazil	AF410788	54
BVDV-2	VS-63	Cattle	Brazil	AF410789	54
BVDV-2	WG4622	Contaminant	Netherland	ALIGN_000012	25
BVDV-2	WVD829	Cattle	Belgium	ALIGN_000012	25
BVDV-3	D32/00_HoBi/Brazil/200/2002	Cattle	Brazil	EF683557	57
BVDV-3	D32/00_HoBi	Cattle	Brazil	AY489116	58
BVDV-3	Th/04_KhonKaen	Cattle	Thailand	DQ897641	59
BDV	0502234	Sheep	Spain	EU711348	60
BDV	0501209-052GI	Sheep	Spain	DQ679902	61
BDV	06-F-0083	Sheep	France	EF693999	62
BDV	06-F-0299/357	Sheep	France	EF694000	62
BDV	06-F-0299/369	Sheep	France	EF694001	62
BDV	06-F-0299/420	Sheep	France	EF694002	62
BDV	06-F-0299/477	Sheep	France	EF694003	62
BDV	135 661	Sheep	UK	U65054	18
BDV	137/4	Sheep	UK	U65052	18
BDV	170 337	Sheep	UK	U65057	18
BDV	2112/99	Sheep	Spain	AY159513	14
BDV	33S	Sheep	Tunisia	AF462002	63
BDV	35	Sheep	Tunisia	AF462001	63
BDV	35T	Sheep	Tunisia	AF462000	63
BDV	37A	Sheep	Tunisia	AF461999	63
BDV	79248/01	Sheep	Spain	AY159515	14

Table I. Continued.

Species	Strain	Origin	Country	Accession	Reference
BDV	80582/01	Sheep	Spain	AY159516	14
BDV	8320-22NZ	Sheep	New Zealand	U65063	18
BDV	8320-31NZ	Sheep	New Zealand	U65064	18
BDV	85-F-488	Sheep	France	EF693985	62
BDV	85-F-588	Sheep	France	EF693986	62
BDV	87877/01	Sheep	Spain	AY159517	14
BDV	89-F-5374	Sheep	France	EF693987	62
BDV	89-F-5415	Sheep	France	EF693988	62
BDV	90/8320/31	Sheep	UK	AF026769	Vilček et al., unpublished
BDV	90-F-6227	Sheep	France	EF693989	62
BDV	90-F-6335	Sheep	France	EF693990	62
BDV	90-F-6338	Sheep	France	EF693991	62
BDV	90-F-6339	Sheep	France	EF693992	62
BDV	91/5809	Sheep	UK	AF026768	Vilček et al., unpublished
BDV	91-F-6731	Sheep	France	EF988632	62
BDV	91-F-6732	Sheep	France	EF988633	62
BDV	91-F-7014	Sheep	France	EF693993	62
BDV	92-F-7119	Sheep	France	EF693994	62
BDV	93-F-7289	Sheep	France	EF693995	62
BDV	94-F-7446/1	Sheep	France	EF693996	62
BDV	94-F-7446/2	Sheep	France	EF693997	62
BDV	96-F-7624	Sheep	France	EF693998	62
BDV	A1263/2	Sheep	UK	U65027	18
BDV	A1870	Sheep	UK	U65028	18
BDV	A841/1	Sheep	UK	U65026	18
BDV	ARAN-1	Pyrenean chamois	Spain	AM765800	64
BDV	ARAN-2	Pyrenean chamois	Spain	AM765801	64
BDV	ARAN-3	Pyrenean chamois	Spain	AM765802	64
BDV	ARAN-4	Pyrenean chamois	Spain	AM765803	64
BDV	ARAN-5	Pyrenean chamois	Spain	AM765804	64
BDV	ARAN-6	Pyrenean chamois	Spain	AM765805	64
BDV	ARAN-7	Pyrenean chamois	Spain	AM765806	64
BDV	ARAN-8	Pyrenean chamois	Spain	AM765807	64
BDV	AV	Sheep	France	EF693984	62
BDV	BD31	Sheep	USA	U70263	65
BDV	BD ncp	Sheep	USA	not deposited	40
BDV	BDV/Aydin/04-TR	Sheep	Turkey	AM418427	66
BDV	BDV/Burdur/05-TR	Sheep	Turkey	AM418428	66
BDV	BM01 isolate 5	Sheep	Tunisia	AY453630	63
BDV	BT2305	Sheep	Germany	EU637004	Schirrmeier et al., unpub.
BDV	BU1-C3	Sheep	Spain	DQ361068	67
BDV	BU1-C4	Sheep	Spain	DQ361069	67
BDV	BU-1CRA22	Sheep	Spain	DQ275622	68
BDV	C121	Sheep	Spain	DQ275625	68
BDV	C27	Sheep	Spain	DQ275623	68
BDV	C290	Sheep	Spain	DQ275624	68
BDV	CADI-1	Pyrenean chamois	Spain	AM905918	69
BDV	CADI-2	Pyrenean chamois	Spain	AM905919	69
BDV	CADI-3	Pyrenean chamois	Spain	AM905920	69
BDV	CADI-4	Pyrenean chamois	Spain	AM905921	69
BDV	CADI-5	Pyrenean	chamois Spain	AM905922	69
BDV	CADI-6	Pyrenean chamois	Spain	AM905923	69
BDV	CADI-7	Pyrenean chamois	Spain	AM905924	69
BDV	CADI-8	Pyrenean chamois	Spain	AM905925	69
BDV	CADI-9	Pyrenean chamois	Spain	AM905926	69
BDV	CADI-10	Pyrenean chamois	Spain	AM905927	69
BDV	CADI-11	Pyrenean chamois	Spain	AM905928	69
BDV	CADI-12	Pyrenean chamois	Spain	AM905929	69
BDV	CERDANYA-1	Pyrenean chamois	Spain	AM905930	69
BDV	CERDANYA-2	Pyrenean chamois	Spain	AM905931	69
BDV	CERDANYA-3	Pyrenean chamois	Spain	AM905932	69
BDV	CERDANYA-4	Pyrenean chamois	Spain	AM905933	69

Table I. Continued.

Species	Strain	Origin	Country	Accession	Reference
BDV	Ch1Es	Sheep	Japan	D50816	29
BDV	Chamois1	Pyrenean chamois	Spain	AY738080	70
BDV	Chamois-Spain02	Pyrenean chamois	Spain	AY641529	71
BDV	Chemnitz	Sheep	Germany	EU637006	72
BDV	Colm24	Sheep	Spain	DQ361073	67
BDV	D1586/2	Sheep	UK	U65034	18
BDV	G1305	Sheep	UK	U65035	18
BDV	G2048	Sheep	UK	U65036	18
BDV	Genzkow 701	Sheep	Germany	EU636999	Schirrmeier et al., unpub.
BDV	Gifhorn	Pig	Germany	EU636997	Schirrmeier et al., unpub.
BDV	Gifhorn-sh	Sheep	Germany	EU637007	Schirrmeier et al., unpub.
BDV	isard4606	Pyrenean chamois	France	EU637005	Schirrmeier et al., unpub.
BDV	J1004	Sheep	Germany	EU637001	Schirrmeier et al., unpub.
BDV	JH2816	Sheep	UK	U65037	18
BDV	K1729/3	Sheep	UK	U65038	18
BDV	L83/L84	Sheep	Germany	U17144	49
BDV	L991	Sheep	UK	U65039	18
BDV	LA1108	Sheep	Germany	EU637000	Schirrmeier et al., unpub.
BDV	LE31C2	Sheep	Spain	DQ361072	67
BDV	Lot21	Sheep	Tunisia	AF461998	63
BDV	M3	Sheep	Spain	DQ275626	68
BDV	Moredun cp	Sheep	UK	U65022	18
BDV	Moredunncp	Sheep	UK	U65023	18
BDV	Orlu-Etagne	Pyrenean chamois	France	DQ898291	73
BDV	Orlu-ORL 2004 02 C	Pyrenean chamois	France	EU477593	Dubois et al., unpublished
BDV	Orlu-R36	Pyrenean chamois	France	DQ898294	73
BDV	Orlu-R41	Pyrenean chamois	France	DQ898295	73
BDV	Orlu-S24	Pyrenean chamois	France	DQ898292	73
BDV	Orlu-S36	Pyrenean chamois	France	DQ898293	73
BDV	Q1488/1	Sheep	UK	U66042	18
BDV	Q1488/6	Sheep	UK	U65043	18
BDV	Q1673/2	Sheep	UK	U65044	18
BDV	R1292/01	Sheep	Switzerland	AY081182	74
BDV	Rentier Rudolph	Reindeer	Germany	AB122086	75
BDV	RM	Sheep	Tunisia	AY583307	Thabti et al., unpublished
BDV	Rocco	Sheep	Spain	DQ361067	67
BDV	SN1T	Sheep	Tunisia	AF461997	63
BDV	SN2T	Sheep	Tunisia	AF461996	63
BDV	SN3G	Sheep	Tunisia	AY583306	Thabti et al., unpublished
BDV	ST1405	Sheep	Germany	EU637002	Schirrmeier et al., unpub.
BDV	ST1507	Sheep	Germany	EU637003	Schirrmeier et al., unpub.
BDV	Stolpe	Sheep	Germany	EU636998	Schirrmeier et al., unpub.
BDV	T1789/1	Sheep	UK	U65045	18
BDV	T1802/1	Sheep	UK	U65046	18
BDV	V1414	Sheep	UK	U65047	18
BDV	V2377/12	Sheep	UK	U65048	18
BDV	V2536/2	Sheep	UK	U65049	18
BDV	V3196/1	Sheep	UK	U65050	18
BDV	VFMIII	Sheep	Spain	DQ361071	67
BDV	V-TOB	Cattle	Australia	U80906	26
BDV	Wisent Casimir	Wisent	Germany	AB122085	75
BDV	X818	Sheep	Australia	AF037405	49
BDV	ZA1-1115	Sheep	Spain	DQ361070	67
BDV-2	712/02	Goat	Italy	AJ829444	76
BDV-2	LA/91/05	Sheep	Italy	FM163381	77
BDV-2	TO/121/04	Sheep	Italy	AM900848	77
BDV-2	LA/82/04	Sheep	Italy	FM163383	77
BDV-2	LA/26/04	Sheep	Italy	FM163382	77
CSFV	17-93	Pig	Poland	L42413	78
CSFV	39	Pig	China	AF407339	Wu et al., unpublished
CSFV	5440/99	Sheep	Spain	AY159514	14
CSFV	Alfort	Pig	France	J04358	79
CSFV	Alfort 187	Pig	France	X87939	80

Table I. Continued.

Species	Strain	Origin	Country	Accession	Reference
CSFV	Alfort A19	Pig	France	U90951	Smondack et al., unpubl.
CSFV	Brescia	Pig	Italy	M31768	81
CSFV	C strain	Pig	China	Z46258	82
CSFV	CAP	Pig	Switzerland	X96550	Tratschin et al., unpubl.
CSFV	cF114	Pig	China	AF333000	Mingxiao et al., unpubl.
CSFV	Chiba-80	Pig	Japan	AB019659	36
CSFV	Eystrup	Pig	Germany	AF326963	83
CSFV	Fukuoka/72	Pig	Japan	AB019150	84
CSFV	GPE (-)	Vaccine	Japan	AB019152	84
CSFV	HCLV	Pig	China	AF091507	Wang et al., unpublished
CSFV	Hokkaido/66	Pig	Japan	AB019154	84
CSFV	Honduras	Pig	Honduras	L42426	78
CSFV	Ibaraki/66	Pig	Japan	AB019156	84
CSFV	Ibaraki/81-115	Pig	Japan	AB019158	84
CSFV	Ibaraki/81-20	Pig	Japan	AB019160	84
CSFV	Ibaraki/81-38	Pig	Japan	AB019162	84
CSFV	Ibaraki/81-40	Pig	Japan	AB019164	84
CSFV	Kanagawa/74	Pig	Japan	AB019166	84
CSFV	KC	Vaccine	Russia	AF099102	85
CSFV	LOM	Pig	Japan	AB019655	36
CSFV	Miyazaki/81	Pig	Japan	AB019168	84
CSFV	Nakamura/66	Pig	Japan	AB019170	84
CSFV	Okinawa/86	Pig	Japan	AB019172	84
CSFV	Osaka/51	Pig	Japan	AB019174	84
CSFV	Osaka/71	Pig	Japan	AB019176	84
CSFV	Pader	Pig	Garmany	AY072924	86
CSFV	Saitama/81	Pig	Japan	AB019178	84
CSFV	Shimen	Pig	China	AF092448	Huang et al., unpublished
CSFV	Shizuoka/73	Pig	Japan	AB019180	84
CSFV	Switzerland 1/93	Pig	Switzerland	AF045068	87
CSFV	Switzerland 2/93'	Pig	Switzerland	AF045069	87
CSFV	Switzerland 3/93/1'	Pig	Switzerland	AF045070	87
CSFV	Switzerland 3/93/2'	Pig	Switzerland	AF045071	87
CSFV	Switzerland 4/93'	Pig	Switzerland	AF045072	87
CSFV	Vac A	Pig	USA	L42435	78
CSFV	Venhorst	Pig	Netherlands	AF084049	88
CSFV	VRI4762	Pig	Malaysia	L42437	78
CSFV	Yamanashi/69	Pig	Japan	AB019182	84
Giraffe	H138	Giraffe	Kenya	AB040131	5
Pronghorn	Pronghorn	Pronghhorn	USA	not deposited	89
Bungowannah	Bungowannah	Pig	Australia	DQ901402	90

Cattle (*Bos taurus*); Deer (*Cervus elaphus*); Giraffe (*Giraffa camelopardalis*); Goat (*Capra hircus*); Human (*Homo sapiens*); Pig (*Sus scrofadomesticus*); Pronghorn (*Antilocapra americana*); Pyrenean chamois (*Rupicapra pyrenaica*); Reindeer (*Rangifer tarandus*); Roe deer (*Capreolus capreolus*); Sheep (*Ovis aries*); Wisent (*Bison bonasus*).

(3) The parameters for the analysis for the identification of genus, species and genotype were memorized in a relational database and were loaded in memory at the start of the program in a DataSet type structure;
(4) The V1, V2 and V3 structures created during the analysis phase of the sequence in input were memorized in an ArrayList of defined structures (*struct*).

With regard to point 3, the choice of the use of a database external to the software was derived from the objective of obtaining a higher flexibility of the program in case parameters changed or required future updating. In fact, in version 1.0, in order to update parameters, it was necessary to modify the software source code. In contrast, with the solution chosen for version 2.0, the software will not require modifications. Only database updating will be necessary.

Figure 1 shows the schematic structure of tables and relations which constitute the database created for PNS software version 2.0.

2.4. *Log* File

Similarly to version 1.0, also in the version 2.0, the software, during the execution of the program, creates a text file named *"log.txt"* important for the monitoring of the procedure of the analysis. In this file are reported all the steps followed for the determination of the final result.

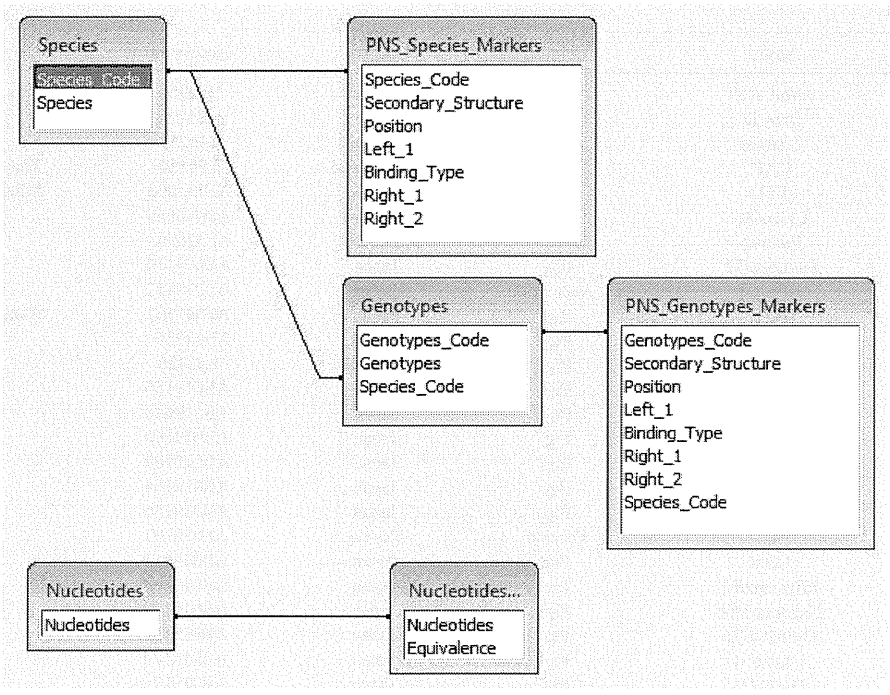


Fig. 1. Schematic structure of tables and relations constituting the database settled for PNS software version 2.0.

3. RESULTS AND DISCUSSION

The PNS software version 2.0 allowed for the characterization of the test sequences clustered in the different genogroups within the genus *Pestivirus* and constructed the secondary structures of the variable loci overlapping the results obtained with the PNS software version 1.0. Compared to version 1.0, the new PNS software version did not create any external file as result of the analysis procedure. The result and the analysis procedure output showing the three palindromic fractions of the sequence, V1, V2 and V3, and the related parameters for genus, species and genotype characterization were presented on the screen in a printable version by the user (Figs. 2 and 3).

The PNS software version 2.0, as the PNS version 1.0, demonstrated a reliable alternative for the construction of stable nucleotide sequence secondary structures previously predicted for the application of the PNS method according to the Zuker and Stiegler algorithm[92] using the Genetyx-Mac version 10.1 program package (Software Development Co., Ltd., Tokyo, Japan), or the DNASIS software package (Hitachi Software Engineering Co., Japan).

The new main characteristic of PNS 2.0 was the modified visualization of the V1, V2 and V3 palindromic structures, which were aligned horizontally. This method of visualization of the nucleotide base-pairings allowed an easier reading of the secondary structure. Furthermore, this was equivalent to the sequence representations used by most common software based on primary structure sequence analysis.

The construction of the secondary structures was displayed through aligned base-pairs from V1 locus on the top side, followed by V2 and ending with V3 at the bottom side. Table II shows the alignments of 5'-UTR RNA secondary structure variable loci sequences of prototype strains of genus *Pestivirus* species genotypes, segregated according to types of base-pair combinations. The different types are ranked according to increasing divergence in the genus, expressed in number of divergent base-pairs, with reference to the most common base-pairs in the prevalent positions.

The alignment for comparison of secondary structure sequences showing divergent base-pair combinations of the different positions of the three stems and loops in the 5'- UTR was the base for the quantitative step of the improved PNS genotyping procedure[3] allowing the identification of genomic groups among *Pestivirus* strains according to secondary sequence structure base-pair combinations in the three palindromes and identifying strains showing multi-relations (sequences showing low relatedness to different genomic groups with low divergence values) or borderlines (sequences showing qualitative similarities with a genomic group, but with high divergence values, candidates for reclustering as separate groups in the genus), and indicating divergence within groups and among groups quantifying the etherogeneity of a species and the genetic distance between species in terms of variation of base-pairs in the secondary structure. For species identification the divergence limit value

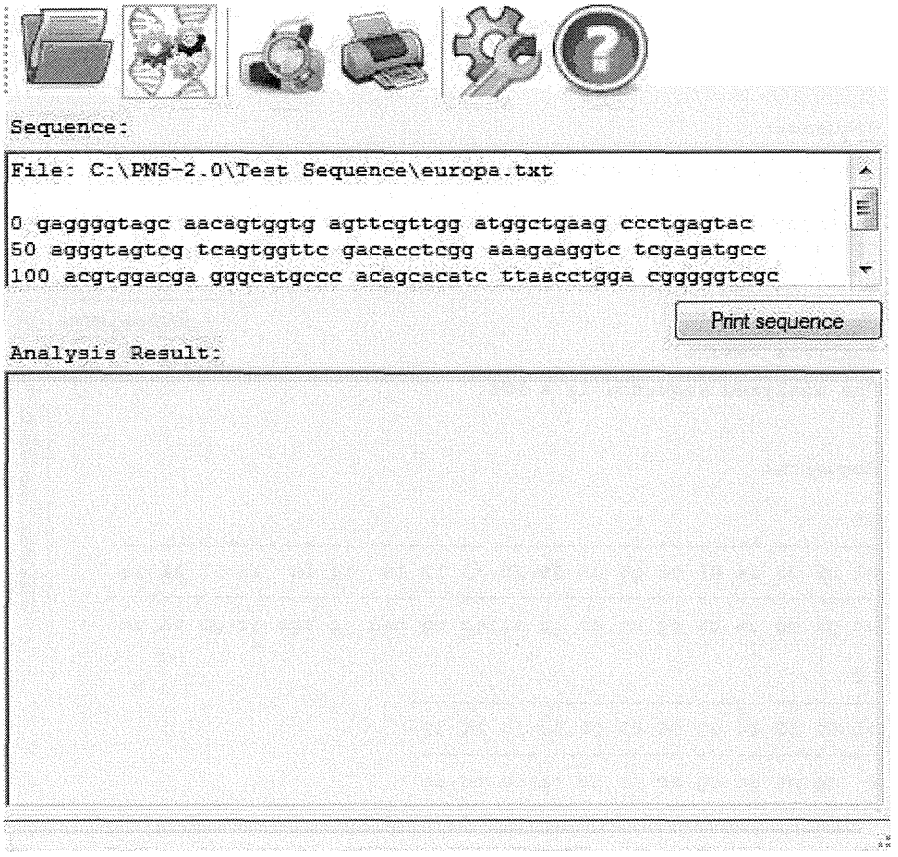


Fig. 2. PNS software version 2.0 allowed a simple and intuitive utilization, through selection of sequence to be tested included in a text file format, and display of aligned secondary structure relevant loci with subsequent analytical classification.

is 13, corresponding to rounded 30% of total number of compared positions (n 45). For genotype identification, the divergence limit value is 9. Base-pair divergence values not exceeding limit value indicate homogeneity between base-pair combinations. Higher values indicate heterogeneity. Therefore, through the provision of aligned secondary structure sequences, the new PNS software introduced an additional element essential for the application of numerical taxonomy applicable to virological research. Examples of sequence secondary structure alignment applied for the determination of the degree of genetic homogeneity among sequences clustered in a single group (PNS method numerical taxonomy qualitative verification - first step) as well as determination of genetic distance between genogroups (PNS method numerical taxonomy qualitative verification - second step) are reported in Tables III and IV.

Nucleotide sequences at the three variable loci, V1, V2 and V3, in the 5'-UTR of pestiviruses have been shown to be palindromic and capable of forming a stable stem-loop structure peculiar to each *Pestivirus* species. Nucleotide substitutions in the stem regions always occur to maintain a palindromic sequence and thereby form a stable stem-loop structure, and this type of mutation was referred to palindromic nucleotide substitutions (PNS). The observation of nucleotide variations among virus strains at the level of the three specific palindromes in the 5'-UTR has been conceived as simple and practical procedure for genotyping. The method, named PNS, proposed keys of identification at the genus, species and genotype levels.[2]

The keys of identification are characteristic and well conserved among all the different genotypes or specific to a single genotype. In terms of consensus motifs, shared by all the *Pestivirus* genotypes, it is possible to identify typical base-pairs suitable for a simple approach of genotyping procedure, given by the evaluation of the predicted secondary structure. The palindromic structures are identifiable in linear sequences. However, it is easier to find them observing the secondary structure.

The palindromic nucleotide substitutions (PNS) genotyping method has been further improved from the original concept limited to qualitative analysis,[3] allowing the taxomical segregation of the genus *Pestivirus*. Nine genomic groups have been identified: the species BVDV-1, BVDV-2, BDV and CSFV, and the tentative species Pronghorn, Giraffe, BVDV-3 (HoBi group), BDV-2 (Italian small ruminant isolates) and Bungowannah.[7]

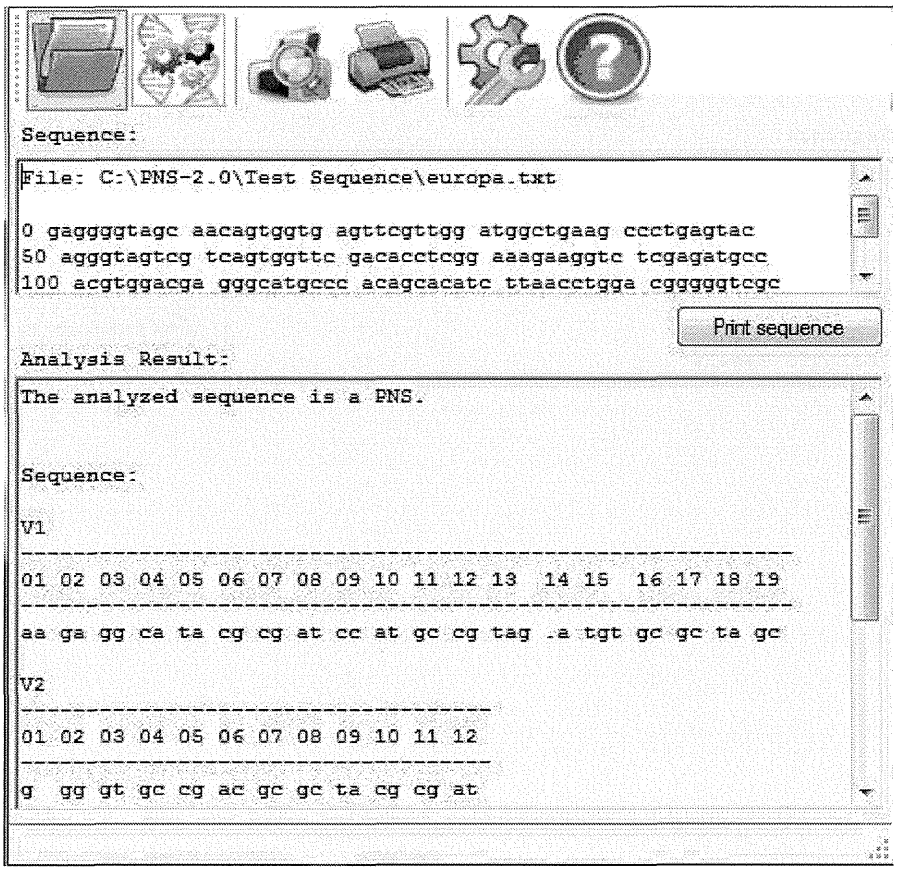


Fig. 3. Result of genotyping according to the PNS method of the Europa strain. The 5'-UTR sequence, presented as alignment format, was compared to known *Pestivirus* species BVDV-1, BVDV-2, BDV, CSFV and the new proposed taxons. The three palindromic regions in sequences were identified in the sequence and shown in the first part of the file. Genus *Pestivirus* characteristic PNS and BVDV-1 species characteristic PNS were identified in the palindromes and highlighted in the following section of the file. Genotyping was completed by the evidencing of BVDV-1c specific PNS.

The method provides results comparable with other genotyping 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 genotyping identification in terms of nucleotide sequence homology are made available with great advantage for simplification of virological investigations. *Pestivirus* strains showing unexpected genomic sequences such as Giraffe,[5] Wisent Casimir,[30] Chamois-1[70] or Pronghorn[89] have been easily characterized and clustered within the genus by PNS method. Furthermore, species could be segregated by the PNS method purely on the basis of genomic sequence characteristics, avoiding any reference to the animal host origin, and the application of quantitative analytical procedure allowed for a better determination of relation among species and genotypes.

The realization of the PNS software resulted in satisfactory prototypes, versions 1.0 and 2.0, as demonstrated by the successful application of the testing on a large number of virus strains. The sequences were correctly displayed with their palindromes and the application of the keys for *Pestivirus* identification showed clear results presented in the output result. The identification step allowed three distinct evaluations. The first was the comparison with genus specific PNS, identifying the appurtenance to the *Pestivirus* genus. The further evaluations were applied only in case of matching. The following comparison was performed with species specific PNS for BVDV 1, BVDV 2, BDV, CSFV and the new proposed taxons. The last comparison was performed for genotype determination within a selected species.

The PNS software represented the transposition of the *Pestivirus* genetic characterization method in a computerized procedure, a relevant improvement with a main advantage represented by the rapidity of the execution of the testing procedure providing data for accurate analyses. Both the prototypes successfully proved to be simple and useful tools for sequence testing indicating clear results for the allocation of unknown isolates and providing support