

Complete sequence and variability of a new subgroup B nepovirus infecting potato in central Peru

Joao De Souza^{1,2} · Giovanna Müller¹ · Wilmer Perez¹ · Wilmer Cuellar^{1,3} · Jan Kreuze¹

Received: 8 August 2016 / Accepted: 31 October 2016 / Published online: 17 November 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract The complete bipartite genome (RNA1 and RNA2) of a new nepovirus infecting potato was obtained using small RNA sequencing and assembly complemented by Sanger sequencing. Each RNA encodes a single polyprotein, flanked by 5' and 3' untranslated regions (UTR) and followed by a poly (A) tail. The putative polyproteins encoded by RNA1 and RNA2 had sets of motifs which are characteristic of viruses in the genus *Nepovirus*. Sequence comparisons using the Pro-Pol region and the coat protein, including phylogenetic analysis of these regions, showed closest relationships with nepoviruses. The data obtained support the taxonomical status of this new virus (putative named Potato virus B, PVB) as a member of the genus *Nepovirus*, subgroup B.

The genus *Nepovirus* (order *Picornavirales*, family *Secoviridae*, sub-family *Comovirinae*) of plant viruses has been divided into three subgroups (A, B, C) based on the length and packaging of RNA2, serological properties, cleavage site specificity of the proteinase, and phylogenetic

analysis of the coat protein sequence [1]. There are few nepoviruses infecting potato. In the subgroup A, there are reports of a calico strain of *Tobacco ring spot virus* (TRSV) [2] and *Potato black ring spot virus* (PBRSV) [3]; however, a recent report [4] confirmed that the calico strain of TRSV, the only strain of TRSV that was reported to infect potato [2], is actually a strain of PBRSV. Accordingly, no reports exist of TRSV infecting potato. In the subgroup C, the only virus reported infecting potato is Potato virus U (PVU) [5], but there is no sequence information available yet that can confirm its classification in subgroup C. In subgroup B, *Tomato black ring virus* (TBRV) has been reported to infect potato, but only in Europe. Infection of TBRV in potatoes is normally only sporadic, and the virus is considered of only minor importance in this crop [6]. In this study, the complete genome of a new nepovirus from subgroup B infecting potato in Peru was determined from native potato plants showing calico-like symptoms and subsequently found to be relatively common in such plants in the central Peruvian highlands. This virus was tentatively named Potato virus B (PVB). Whereas the name PVB has been used before for a potato virus, the virus concerned was soon shown to be a strain of PVX [7, 8].

Leaves of native potato plants that showed calico symptoms were collected during the rainy season (March, 2010) at Quichas (Yanacancha, Daniel A. Carrión, Pasco, Peru) (latitude 10°38' 10.66" and longitude 76° 10'21.78"; at 4007 m.a.s.l.). Total RNA was extracted from these leaves using Trizol reagent (Invitrogen, Carlsbad, CA) following the manufacturer's instructions. Small RNAs were isolated of one plant as described previously [9], and then, sent to Fasteris Life Sciences SA (Plan-les-Ouates, Switzerland) for library preparation and sequencing on IlluminaHiSeq2000 platform in 50 bp single end read mode

Electronic supplementary material The online version of this article (doi:10.1007/s00705-016-3147-6) contains supplementary material, which is available to authorized users.

✉ Jan Kreuze
j.kreuze@cgiar.org

- ¹ International Potato Center (CIP), Apartado 1558, Lima 12, Peru
- ² Facultad de Ciencias Biológicas, Programa de Biología Molecular, Universidad Nacional Mayor de San Marcos, Lima 11, Peru
- ³ Present Address: International Center for Tropical Agriculture (CIAT), Apartado Aéreo 6713, Cali, Colombia

[10]. Reads were assembled de-novo using Velvet v0.6.04 [11] and the AssemblyAssembler 1.4 script (Jacob Crawford, Cornell University) with a range of hash lengths from 13 to 25. Contigs were identified by BLASTx against GenBank sequences and those corresponding to nepoviruses were extracted. Sequencing of the small RNA library produced a total of 15,871,227 reads between 21-24 nucleotides (nts). Six contigs with similarity to nepoviruses could be identified after assembly using Velvet. Five contigs corresponded to RNA1 of nepoviruses and three of these could be further assembled into a super-contig corresponding to the 3'-end and the other two contigs were close to the 5'-end of RNA1. The remaining contig corresponded to the almost complete RNA2 sequence. The missing intervening regions of RNA1 and terminal regions of both RNAs were successfully amplified by specifically designed primers and adapter ligation (Table 1), after which the corresponding regions were cloned into pGEM-T Easy vector (Promega, WI, USA) and sequencing (Macrogen-Korea). The 3' and 5' UTR were confirmed by RNA linker mediated rapid amplification of cDNA ends using the ModBan and BanTwo linkers, respectively (Table 1). Sequences were edited and annotated using Vector NTI v.9 software package (Invitrogen). Phylogenetic and molecular evolutionary analyses were conducted using MEGA5 [12]. The complete sequence of both RNAs were deposited in GenBank (accession no. KX656670 and KX656671).

The complete sequences of RNA1 and RNA2 of this putative new virus are 7148 and 4527 nts long excluding the poly(A) tail, respectively. Their predicted genomic

organization is similar to that of other members of the genus *Nepovirus*, subgroup B. Both RNAs (Fig. 1A) contained a single open reading frame (ORF). The putative 255.79-kDa (2264 aa) polyprotein (P1) from RNA1 was preceded by a 140 nts 5'-UTR and followed by a 216 nts 3'-UTR. P1 covers 95.01% of the RNA1 coding capacity, higher than other nepoviruses of subgroup B. Based on homology, P1 is predicted to be cleaved by its own viral proteinase into six mature proteins: X1 (48.6 kDa, unknown function), X2 (21.6 kDa, often referred to a putative cofactor by analogy to the related 32 kDa comovirus protein), NTB (66.6 kDa, nucleoside-triphosphate-binding protein or putative helicase), VPg (3.1 kDa, viral genome-linked protein), proteinase (Pro) (24.1 kDa) and RNA-dependent RNA polymerase – RdRp (Pol) (91.6 kDa). The putative dipeptides that serve as cleavage sites, deduced by the similarity of the dipeptide alongside the upstream sequence using nepoviruses from subgroup B previously identified, were R₄₂₇/A₄₂₈, R₆₁₇/S₆₁₈, K₁₂₁₂/S₁₂₁₃, R₁₂₃₉/S₁₂₄₀, K₁₄₄₉/S₁₄₅₀, producing the X1, X2, NTP, VPg, Pro and Pol proteins (Fig. 1A). Re-alignment of the reads of the RNA1 using MAQ confirmed the sequence and was able to call all nts with high quality scores at an average coverage of 1551x from 523,735 aligned reads (3.29% of the total siRNA reads) (Fig. S1a). RNA2 has a single ORF that encodes a polyprotein (P2) similar to subgroup A and B nepoviruses and is cleaved into three domains, whereas that of subgroup C nepoviruses contains four domains [1]. P2 has a predicted molecular weight of 153.31 kDa (1,371 aa), which is preceded by a 214 nts 5'-UTR and followed by a 200 nts 3'-UTR. P2 covers 90.85%

Table 1 Primers and linkers used in this study and the corresponding virus region amplified

Primer name	Sequence (5'-3')	Amplified region
NepoWP-5UTR-Rin1*	GGTAAAAAAACAAATTAAGACTATTATTATATTTGC	0 - 261
NepoWP1_Fn	TATGGATAGCGCCATAGTTTACCTTAT	197 - 486
NepoWP1_Rn	CTGTATAGGCAGAAATGGTGGTTATACC	
Nepo1-WP_Fa	TCTCCACTTATATCCAAGGTTCTTCC	402 - 1928
Nepo1-WP_Ra	CCGCCACTTTTTGCTGCTCTG	
NepoWP-RdRp-F1	CAGCCACCAAGACGGAACCAGCC	2671 - 4648
NepoWP-RdRp-R1	GGTTCTTAATTGGTACATTGAC	
NepoWP-F2	CAGATTATATGCTCTGAGTGGG	5384 - 6340
NepoWP-R2	GGCAAGCCACCGACACCAGCTCC	
NepoWP-RdRp-F3	GCAAAGCGTGTACAGGTATACG	6184 - 7082
NepoWP-3UTR-R3	GGACAAACTTATCTCATCGCCAC	
Nepo1-WP_3UTR-F*	CATTACAGTTCGAAGTTTACAGTTGCC	6797 - 7154
Nepo2-WP_5UTR-R*	CTACCAAGGGCTTCAGCTTTGGC	0 - 285
Nepo2-WP_3UTR-F*	GACGGCCGGACCTTCTCTCC	4204 - 4527
Modban linker	A5'pp CTG TAG GCA CCA TCA AT/ddC/	
Bantwo linker	ATC GTrA rGrGrC rArCrC rUrGrA rArA	

* Used together with RACE primer

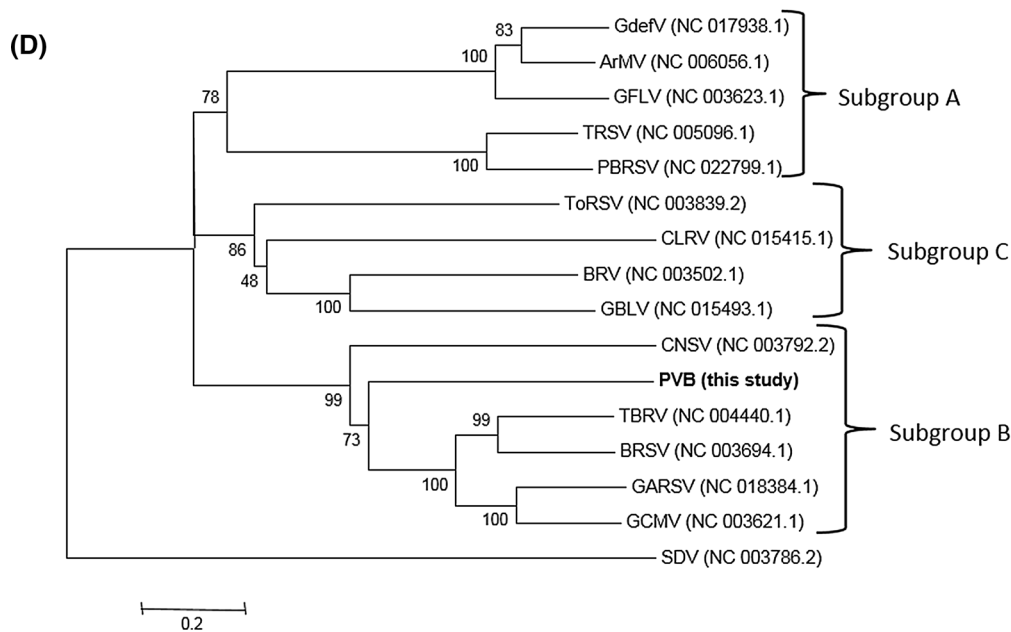
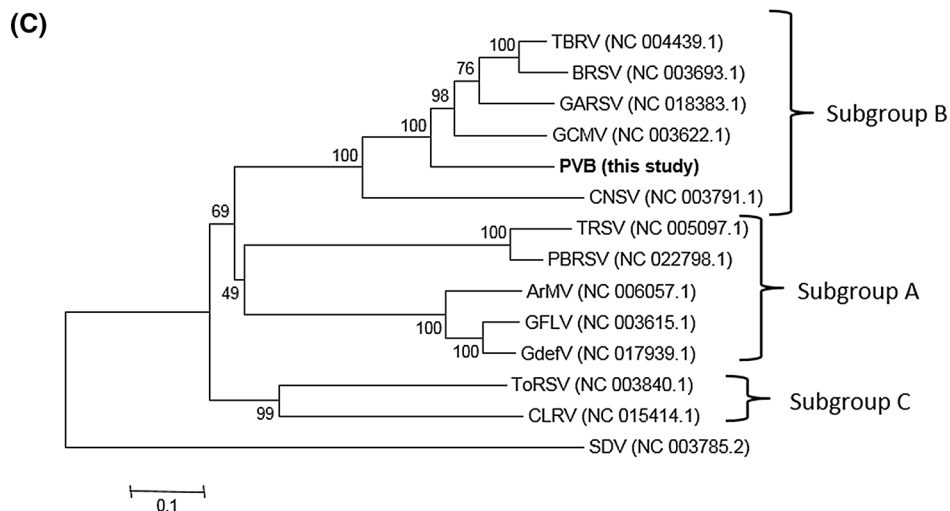
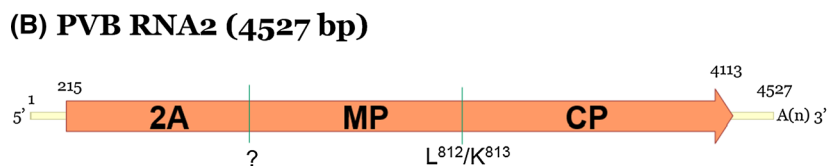
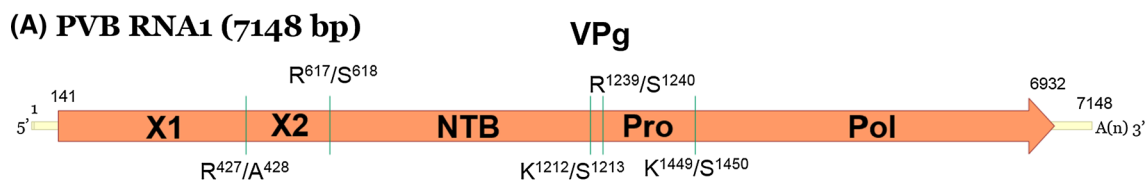


Fig. 1 A and B) Genome structure of Potato virus B. Open arrows indicate open reading frames, with the names of predicted gene products inside them. Polyprotein cleavage sites that release the specific proteins are indicated by vertical green lines. Question mark in RNA2 means that we were unable to identify the exact putative cleavage site to separate 2A from MP and its location is an estimation. C and D) Phylogenetic tree based on an alignment of the Pro-Pol and CP region respectively. Sequences were aligned using the ClustalW algorithm, and the tree was generated using the neighbour-joining method. Numbers on branches indicate percentage of bootstrap support out of 1000 bootstrap replications. *Satsuma dwarf virus* (SDV) was used as outgroup. TBRV (*Tomato black ring virus*), BRSV (*Beet ringspot virus*), GARSV (*Grapevine Anatolian ringspot virus*), GCMV (*Grapevine chrome mosaic virus*), CNSV (*Cycas necrotic stunt virus*), TRSV (*Tobacco ringspot virus*), PBRSV (*Potato black ringspot virus*), ArMV (*Arabidopsis mosaic virus*), GFLV (*Grapevine fanleaf virus*), GdefV (*Grapevine deformation virus*), ToRSV (*Tomato ringspot virus*), CLRV (*Cherry leaf roll virus*), BRV (*Blackcurrant reversion virus*), GBLV (*Grapevine Bulgarian Latent virus*)

of the RNA2 coding capacity, higher than RNA2 of other nepoviruses of subgroup B. They are predicted to be processed into mature protein 2A, the putative movement protein (MP), and the coat protein (CP, 61.4 kDa). Based on comparison to other subgroup B nepoviruses, the site containing the putative cleavage site to separate MP from CP is L₈₁₂/K₈₁₃ (Fig. 1B). Re-aligning siRNAs to the consensus sequence of the RNA2 using MAQ confirmed the sequence, calling all nts at an average coverage of 1,289x by 274,946 reads (1.73% of the total reads) (Fig S1b). The identity between the 3'-UTRs of both RNAs of PVB was 95% (typical for nepoviruses), but the identity between the 5'-UTRs was 38%, less than those of nepoviruses previously reported. Comparison among PVB and other members of the genus *Nepovirus*, show that, PVB has a similarity range with subgroup B nepoviruses reported from 53 - 71% in the Pro-Pol region, and from 25 - 34% in the CP region (Table S1). A phylogenetic tree generated using alignments of these regions located PVB in the subgroup B of the genus *Nepovirus* (Fig. 1C and D). Comparison of the sequence of PVB with the unpublished sequences of two previously reported potato infecting nepoviruses in Peru, Arracacha virus A and PVU, showed no significant homology (Ian Adams & Roger Jones, personal communication), confirming PVB is a newly discovered virus.

Primers, that amplify the polymerase region and were used to confirm the genome obtained, were also used to detect this virus in leaves of additional potato plants with calico symptoms collected from Pasco and Junín (departments in Peru) (Figure S2). From 151 samples from Pasco, 28 were positive (19%); however, out of 67 samples evaluated from Junín, only two were positive (3%) (Table S2). The amplified 1,977 bp RdRp fragment of some positive samples were also sequenced to evaluate the

PVB variability. In the phylogenetic tree produced with these sequences, two groups could be distinguished (86.8% average nt identity between the two groups and >98.7% or 95.7% identity within the two groups respectively), but there seemed to be little geographic relationship (Fig. S3)

The sequence similarity found, the organization genome and the phylogenetic analysis indicated that PVB corresponds to a new virus species belonging to the genus *Nepovirus* within subgroup B, according to [1]. Many nepoviruses are transmitted by soil-inhabiting nematodes belonging to three closely related genera *Xiphinema*, *Longidorus*, or *Paralongidorus* in the order *Dorylaimida*, family *Londigoridae* [13]. *Longidorus* is so far the only nematode genus reported to transmit subgroup B nepoviruses, whereas the genus *Xiphinema* and *Paralongidorus* were reported to transmit subgroup A and C nepoviruses. Genus *Longidorus* is considered a quarantine species by the National Agriculture Health Service (SENASA, its acronym Spanish) [14]. However Jones *et al.* [5] reported preliminary experiments with *Longidorus* spp., which had been collected in Peru (Cesar Fribourg, personal communication) and Ciancio *et al* [15, 16] also reported the presence of longidorid nematodes in Peru. It would be a priority to confirm what the vector of PVB is and confirm the presence of *Longidorus* spp. in Peru. The phylogenetic tree using the symptomatic plants sampled from Pasco and Huancayo show that variability is present in this virus, similar to GFLV as determined by Oliver *et al.* [17] where RdRp also was used as a source to analyze virus variability. On the other hand, whereas PVB was always found associated with calico like symptoms (all positive samples had calico symptoms), not all sampled plants with these symptoms tested positive for the virus. Thus it remains unclear to what extent PVB contributes to the observed symptoms. Future studies into this aspect, including tentative yield impacts should be an additional priority.

Acknowledgements We gratefully acknowledge PhD Dina Gutierrez, Biol. Rocío Silvestre, Carlos Chuquillanqui and Maria Scurrah for critical review of the manuscript and Eng. Ricardo Orrego for sample collection.

Compliance with ethical standards

This study was funded by the CGIAR research program on Roots, Tubers and Bananas (RTB). All authors declare that they have no conflict of interest. This article does not contain any studies with human participants or animals performed by any of the authors.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Sanfaçon H, Iwanami T, Karasev AV, van der Vlugt R, Wellink J, Wetzel T, Yoshikawa N (2012) Family Secoviridae. In: King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (eds) Virus taxonomy, IX Report of ICTV. Elsevier Academic Press, London, pp 881–899
2. Fribourg CE (1977) Andean potato calico strain of Tobacco ringspot virus. *Phytopathology* 67(2):174–178
3. Salazar LF, Harrison BD (1977) Two previously undescribed potato viruses from South America. *Nature* 264:337–338
4. Richards RS, Adams IP, Kreuze JF, De Souza J, Cuellar W et al (2014) The complete genome sequences of two isolates of potato black ringspot virus and their relationship to other isolates and nepoviruses. *Arch. Virol.* 159:811–815
5. Jones RAC, Fribourg CE, Koenig R (1983) A previously undescribed Nepovirus isolated from potato in Peru. *Phytopathology* 73:195–198
6. Jones RAC, Charkowski A, Fribourg CE, Stevenson WR, Slack SA (2009) Potato virus and viruslike diseases. In: *Virus diseases of plants: grape, potato, and wheat image collection and teaching resource CD-Rom*. APS Press, St. Paul, p 55121
7. Bawden FC, Sheffield FML (1944) The relationships of some viruses causing necrotic diseases of the potato. *Ann Appl Biol* 31:33–40
8. Matthews REF (1947) Status of potato virus B. *Nature* 159:713–714
9. Fuentes S, Heider B, Tasso RC, Romero E, ZumFelde T, Kreuze JF (2012) Complete genome sequence of a potyvirus infecting yam beans (*Pachyrhizus* spp.) in Peru. *Arch Virol* 157(4):773–776
10. Kreuze JF, Perez A, Untiveros M, Quispe D, Fuentes S, Barker I, Simon R (2009) Complete viral genome sequence and discovery of novel viruses by deep sequencing of small RNAs: a generic method for diagnosis, discovery and sequencing of viruses. *Virology* 388:1–7
11. Zerbino D, Birney E (2008) Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 18:821–829
12. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance and maximum parsimony methods. *Mol Biol Evol* 28:2731–2739
13. Sanfaçon H (2009) Nepovirus. In: Mahy B, Regenmortel M (eds) *Desk encyclopedia of plant and fungal virology*. Elsevier Ltd., pp 235–243
14. SENASA (2015) Lista de plagas cuarentenarias no presentes en el Perú. <http://www.senasa.gob.pe/senasa/wp-content/uploads/2014/12/Lista-de-plagas-cuarentenarias-agosto-2016.pdf>
15. Ciancio A, Mankau R (1989) Note on *Pasteuria* sp. parasitic in Longidorid nematodes. *Nematropica* 19(1):105–110
16. Ciancio A, Roca F, Carbonell Torres E (2011) New records of Longidorid nematodes in central Peru. Book of abstracts of XLIII meeting of Organization of Nematologists of Tropical America (ONTA), p 139
17. Oliver JE, Vigne E, Fuchs M (2010) Genetic structure and molecular variability of Grapevine fanleaf virus populations. *Virus Res* 152:30–40