



HAL
open science

Genetic Analysis and Phylogenetic Comparison of Black Queen Cell Virus Genotypes

Zsuzsanna Tapaszti, Petra Forgách, Csaba Kóvágó, Grażyna Topolska,
Norbert Nowotny, Miklós Rusvai, Tamás Bakonyi

► **To cite this version:**

Zsuzsanna Tapaszti, Petra Forgách, Csaba Kóvágó, Grażyna Topolska, Norbert Nowotny, et al.. Genetic Analysis and Phylogenetic Comparison of Black Queen Cell Virus Genotypes. *Veterinary Microbiology*, 2009, 139 (3-4), pp.227. 10.1016/j.vetmic.2009.06.002 . hal-00526932

HAL Id: hal-00526932

<https://hal.science/hal-00526932>

Submitted on 17 Oct 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Accepted Manuscript

Title: Genetic Analysis and Phylogenetic Comparison of Black Queen Cell Virus Genotypes

Authors: Zsuzsanna Tapasztai, Petra Forgách, Csaba Kővágó, Grażyna Topolska, Norbert Nowotny, Miklós Rusvai, Tamás Bakonyi



PII: S0378-1135(09)00286-7
DOI: doi:10.1016/j.vetmic.2009.06.002
Reference: VETMIC 4456

To appear in: *VETMIC*

Received date: 31-10-2008
Revised date: 20-5-2009
Accepted date: 3-6-2009

Please cite this article as: Tapasztai, Z., Forgách, P., Kővágó, C., Topolska, G., Nowotny, N., Rusvai, M., Bakonyi, T., Genetic Analysis and Phylogenetic Comparison of Black Queen Cell Virus Genotypes, *Veterinary Microbiology* (2008), doi:10.1016/j.vetmic.2009.06.002

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1 **Abstract**

2 Phylogenetic analysis of 22 Black queen cell virus (BQCV) genotypes collected from
3 honeybee colonies in Poland, Austria and Hungary was performed on a partial helicase
4 enzyme coding region (ORF1) and on a partial structural polypeptide coding region (ORF2).
5 While the phylogeny based on the ORF2 region showed – with the exception of one strain
6 from Poland – clustering of the genotypes corresponding to their geographic origin, the
7 ORF1-based tree exhibited a completely different distribution of the Polish strains: three of
8 them clustered within a branch clearly separated from all other central European BQCVs,
9 while four other Polish strains remained well within the central European BQCV genotypes.
10 In order to investigate this discrepancy in more detail, the nearly complete genome sequences
11 of the 3 differing Polish strains were determined, together with one Hungarian sample. The
12 sequences were aligned to each other and to the reference strain from South-Africa.
13 Comparison of the different genome regions revealed that the 5'-UTR and the intergenic
14 regions of the BQCV genome are highly conserved with longer homologous sections. ORF1
15 (non-structural protein coding region) was found more variable compared to ORF2 (structural
16 protein coding region). The 5'-proximal third of ORF1 was particularly variable and contained
17 several deletions / insertions. The sudden changes in the similarity levels of BQCV strains in
18 different genomic regions are indicative of preceding recombination events.

19

20 Keywords: honey bee; Black queen cell virus; BQCV; RT-PCR; phylogenetic analysis

21

1 **Introduction**

2 Black queen cell virus (BQCV) was first isolated from queen prepupae and pupae, which
3 were found dead in their cells. The name of the virus was derived from the darkened areas on
4 the wall of queen cells containing infected pupae (Bailey and Woods, 1977). BQCV is a
5 widely prevalent virus; recent surveys in worker bees in France and Austria detected 86% and
6 30% infection rates, respectively (Tentcheva et al., 2004; Berényi et al., 2006). Infected
7 worker bees remain apparently healthy (Bailey, 1983b). It is believed that such bees transmit
8 the virus to bee larvae, especially queen bee larvae, with the secreted brood food - (Allen and
9 Ball, 1996); experiment showed that queen larvae can be successfully infected with BQCV
10 added into larval food (Topolska, 2008). In Australia and Poland BQCV was found to be the
11 main cause of death of the queen brood in queen rearing apiaries (Anderson, 1993; Topolska,
12 2008). Adult worker honeybees, worker or drone brood can also be infected by the virus, but
13 normally without apparent morphologic alterations, although symptomatic BQCV infection of
14 such broods was also reported (Bailey and Woods, 1977; Siede and Buchler, 2003).

15 BQCV particles contain a single stranded RNA genome of 8550 nt, consisting of two open
16 reading frames (ORFs): the 5'-proximal ORF (ORF1) encoding a putative replicase
17 polyprotein, and the 3'-proximal ORF (ORF2) encoding a capsid polyprotein (Leat et al.,
18 2000). The initiation of ORF1 is at nucleotide (nt) position 658, the termination at nt position
19 5625. ORF1 encodes the putative helicase, 3C-like cystein protease and RNA-dependent
20 RNA polymerase (RdRp) enzymes. ORF2 is located between nt pos. 5834 and 8395 (Leat et
21 al., 2000). The translation initiation of ORF2 is modulated by internal ribosomal entry site
22 (IRES), its translation initiation codon is CCU, and it encodes four capsid proteins, with
23 molecular masses of 34, 32, 29 and 6 kDa (Leat et al., 2000). BQCV is classified into the
24 *Cripavirus* genus within the *Dicistroviridae* family.

1 Insect viruses and vertebrate viruses differ significantly regarding their virus-host
2 interactions. In vertebrates, neutralizing antibodies represent an intensive selective pressure
3 on the surface antigens of the virions. Therefore, usually appreciable sequence diversity is
4 seen at the antigen-determinant regions of the structural protein coding genome sections of
5 different isolates of the same virus. In insects, however, the immune system does not produce
6 immunoglobulins, and the defence mechanisms are mainly based on innate immunity (Cherry
7 and Silverman 2006, Zambon et al., 2005). Hence, the evolution of insect viruses is probably
8 influenced by other intracellular factors targeting different regions of the viral genome.

9 While in many cases BQCV infections remain unapparent, considerable losses of queen
10 pupae are experienced in some colonies or apiaries. Besides host factors and beekeeping
11 conditions, differences in the virulence of the strains may also influence the development of
12 the disease. The aim of this study was to compare the nucleotide sequence diversity of BQCV
13 genotypes obtained from different geographic origins. Although our initial goal was a
14 phylogenetic analysis of central European BQCV strains, the results of the partial sequence
15 comparisons stimulated us to perform a comprehensive genome analysis of selected strains to
16 reveal the level of variance within the different genomic regions of the virus.

17

1 **Materials and Methods**

2 **Samples**

3 Honeybee samples were collected from three European countries: Poland, Hungary and
4 Austria. The samples from Poland (queen cells with dead larvae, prepupae and pupae)
5 originated from a queen rearing apiary, in which BQCV infection symptoms were observed,
6 The samples from Hungary (queen larvae, prepupae, pupae, and adult worker bees) and
7 Austria (adult worker bees) were collected from different apiaries without reported symptoms
8 of BQCV infection, since in these countries BQCV-associated losses were not observed and
9 reported by beekeepers.

10 Queen larvae, prepupae and pupae were processed individually, while the adult bees from the
11 same colony were pooled (50 bees/colony). Samples were homogenized in 10 ml sterile
12 phosphate buffered saline (PBS), and centrifuged at 1,500 g for 10 minutes to clear from
13 crude cell debris. The supernatants were centrifuged again at 12,000 g for 15 minutes to
14 pelletize fine cellular elements, bacterial and fungal spores. The supernatants were further
15 processed.

16

17 **RNA isolation and RT-PCR**

18 RNA was extracted from 140 μ l sample supernatant by using QIAamp viral RNA mini kit
19 (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Two pairs of
20 oligonucleotide primers were designed, based on the BQCV genome deposited in GenBank
21 database (accession number: AF183905), using Sci-Ed-Central primer designer program
22 (Scientific and Educational Software, version 3.0). Primers were designed to amplify partial
23 sequences of the structural polyprotein gene and the RNA helicase enzyme gene regions.
24 Further sets of primers were also designed to determine the complete genome sequence of
25 selected BQCV genotypes. The sequences, orientations and locations of the primers, as well

1 as product sizes are listed in Table I and Table II. The oligonucleotides were synthesized by
2 Creative Labor Ltd (Szeged, Hungary).

3 Reverse transcription and amplifications were carried out in a continuous RT-PCR method
4 using the QIAGEN OneStep RT-PCR Kit (Qiagen, Hilden, Germany). The reaction mixture
5 contained 10 μ l of 5 \times reaction buffer, 2 μ l of deoxy-nucleoside-triphosphate mix (dNTP), 50
6 pmol of the appropriate primers, 2 μ l of enzyme mix containing Omniscript and Sensiscript
7 reverse transcriptases and HotStarTaq DNA polymerase, 5 μ l template RNA, 10 μ l Q
8 solution, 40 U RiboLock RNase Inhibitor (Fermentas, Vilnius, Lithuania), and distilled water
9 was added to reach the final volume of 50 μ l.

10 The reverse transcription was carried out at 50°C for 30 minutes, followed by heat
11 denaturation at 95°C for 15 min. Thereafter samples were subjected to 40 cycles of
12 amplification: denaturation at 94°C for 45 sec, primer annealing at 55°C for 45 sec, and DNA
13 extension at 72°C for 1 min. Finally samples were maintained at 72°C for 10 min, and were
14 kept at 4°C until electrophoresis was carried out. The reactions were accomplished in a MJ
15 Research MiniCycler (MJ Research, Inc., Watertown, MA, USA) thermocycler.
16 Amplification products were electrophoresed in agarose gel containing ethidium bromide,
17 visualized by UV transillumination, and photographed by the Kodak DS Electrophoresis
18 Documentation and Analysis System. Product sizes were determined with reference to a 100
19 bp molecular weight ladder (GeneRuler, Fermentas, Vilnius, Lithuania).

20

21 **Nucleotide sequencing and computer analysis**

22 Specific amplification products were excised from the agarose gel and extracted using
23 QIAquick Gel Extraction Kit (Qiagen, Hilden, Germany), according to the manufacturer's
24 instructions. Fluorescence-based sequencing reactions were performed on the amplification
25 products using the same primers as in the RT-PCRs. The nucleotide sequences were

1 determined in at least two independent reactions using an ABI Prism 310 automated
2 sequencing system (Biological Research Centre of the Hungarian Academy of Sciences,
3 Szeged). The nucleotide sequences were identified by the Basic Local Alignment Search Tool
4 (BLAST, Altschul et al, 1990) at the National Center for Biotechnology Information (NCBI),
5 NIH, and they were compiled and aligned with using BioEdit 4.7.8 and Align Plus (Scientific
6 and Educational Software) programs. Multiple nucleotide alignments were created by the
7 ClustalX 1.8 program. The phylogenetic analysis was performed using the Phylogeny
8 Inference Program Package (PHYLIP, version 3.6b, Felsenstein, 2004). Bootstrap resampling
9 analysis of 1000 replicates was performed with the SEQBOOT program to prove the stability
10 of the trees. Distance matrices were generated by the DNADIST/Neighbor-Joining and Fitch
11 programs, using a transition/transversion ratio of 2.0. Phylogenetic trees were drawn using
12 TreeView (Win32, version 1.6.6.) software. The differences in the similarity levels of the
13 different genomic regions of selected BQCV genotypes were plotted by the SimPlot program
14 (version 3.5.1., Ray, 2003) using the 2-parameter (Kimura) distance model.

15

16 **Nucleotide sequence accession numbers.**

17 The BQCV sequences described in this paper were submitted to GenBank database under
18 accession numbers EF517501 - EF517522.

19

1 **Results**

2 **Phylogenetic analysis of central European Black queen cell virus genotypes**

3 Specific amplification products were generated from five Austrian, ten Hungarian and seven
4 Polish samples, from two regions of the BQCV genome (the helicase enzyme coding region,
5 between nt positions 2144 and 2726, and the structural polyprotein coding region between nt
6 positions 6540 and 7053, referring to the complete genome sequence, respectively). The
7 nucleotide sequences of the amplification products were determined, and the sequences were
8 aligned to each other. At the helicase domain region the central European genotypes showed
9 82 to 90% identity with the South-African reference strain. The Hungarian and the Austrian
10 BQCVs were more similar to each other (94-99% identity) than to the Polish genotypes (81-
11 96% identity). Genotypes Poland4, Poland5 and Poland6 were very similar to each other (98-
12 99% identity), and they were more similar to the South-African reference strain than to the
13 other investigated central European strains (exhibiting a similarity rate of 90% to the
14 reference strain, and 81-83% to the other central European genotypes). At the structural
15 polyprotein coding region generally higher identity rates (91 to 94%) were found between the
16 central European genotypes and the reference strain. In this region, the Hungarian and
17 Austrian genotypes were again more similar to each other (97-99%) than to the Polish strains
18 (93-96%). One genotype - Poland4 - showed particularly high divergence being 89-91%
19 identical to the other Polish, Austrian and Hungarian genotypes, and 91% identical to the
20 reference strain.

21 Nucleic acid sequences were translated to putative amino acid sequences, resulting in 186 and
22 132 amino acid (aa) long polypeptide sequences of the helicase and the structural protein
23 region, respectively. The aa sequences were also aligned, and compared to the South-African
24 reference strain, as well as to each other. In the helicase region 91 to 100%, and in the

1 structural polyprotein region 97 to 100% aa identities were found between the investigated
2 genotypes.

3 Phylogenetic analyses were performed on the two selected genome regions (Figure 1. and 2.).

4 The tree based on the partial helicase coding region is separated in three main groups. All

5 Austrian and Hungarian genotypes and Polish genotypes 1 to 3 and 7 cluster within one

6 branch of the tree. Three Polish genotypes (4 to 6) form a separate cluster, and the South-

7 African strain represents the third main branch of the tree. The three unique Polish genotypes

8 are closer to the South-African strain than to the other European genotypes (Figure 1.).

9 However, the phylogenetic tree based on the partial structural polyprotein sequences exhibits

10 partly a considerably different topology: the central European BQCVs form a common

11 cluster, however it is sub-divided into two main groups. One group contains the Austrian and

12 Hungarian genotypes, while the Polish viruses form the second group. The separation of the

13 two groups is statistically supported (bootstrap of 1000 replicates is 944). Within the

14 Austrian-Hungarian group, four Austrian genotypes form a separate subgroup together with

15 two Hungarian genotypes, while the other eight Hungarian viruses and one Austrian genotype

16 form a second, more diverse branch. In this tree, genotypes Poland 5 and 6 cluster together

17 with the other Polish genotypes, with the exception of Poland 4, which has similar genetic

18 distances from both the European genotypes and the South-African strain, the latter

19 representing the third main branch of the tree (Figure 2.). Phylogenetic trees based on the

20 putative amino acid sequences were also constructed, but the relatively short sequences and

21 high similarity rates resulted in low bootstrap values, therefore these trees are not shown.

22

23 **Genome comparisons of selected BQCV genotypes**

24 The phylogenetic analysis of central European BQCV genotypes revealed that three viruses

25 from Poland (named Poland 4, 5 and 6) represent a separate cluster of the investigated

1 viruses, when the partial ORF 1 (helicase coding region) was used for the analysis. Poland 5
2 and 6 genotypes, however, clustered together with the other Polish genotypes, when the
3 partial ORF 2 (structural polyprotein coding region) was used as basis of the analysis. This
4 phenomenon could be explained by intramolecular recombination between different BQCV
5 genotypes. To prove this theory, the genome sequence of strains Poland 4, 5, and 6 were
6 determined between nt positions 54 and 8431 (referring to the South-African strain's complete
7 genome record) using overlapping PCR amplification products (Table II). This region covers
8 the partial 5' UTR, the entire ORF 1, intergenic, and ORF 2 regions, and the partial 3' end of
9 the genome. As a representative of the "typical" central European genotypes, strain Hungary
10 10 was also sequenced in the same region. One Austrian genotype (Austria 5) was partially
11 sequenced between nt positions 59 and 2255, to investigate the diversity of the ORF 1 region.
12 Multiple sequence alignments were created and the similarity rates within the sequenced
13 regions were demonstrated by SimPlot graphs (Figure 3.). Strains Poland 5 and 6 shared a
14 high level of similarity (98%); they were 95% identical to the Poland 4 genotype, and 89%
15 identical to the Hungary 10 and the South-African reference strain. Strain Poland 4 showed 87
16 and 90% similarity to the Hungary 10 and South-African strains, respectively. The Hungarian
17 and South-African viruses shared 86% identity. The detailed analysis of the different genomic
18 regions, however, revealed interesting variance between the sequence identities. Within the 5'
19 UTR region, a 117 nucleotide long section was identified (between nt positions 69 and 235)
20 where all six investigated viruses (Poland 4 to 6, Hungary 10, Austria 5 and South-Africa)
21 were 100% identical to each other. ORF 1 (non-structural protein coding region) proved to be
22 highly interesting regarding sequence diversity in being split into two differing regions: the
23 substitutions and insertions/deletions are concentrated on the 5' proximal part of the gene,
24 within a ~1700 nucleotides long region. The European sequences are 84 nucleotides shorter
25 than the South-African strain's sequence. Between nt positions 1176 and 1244, the South-

1 African strain contains a 69 nucleotide-long continuous sequence, which is missing in the
2 investigated five European genotypes. There are no other BQCV sequences available in
3 GenBank database on this part of the genome, therefore, other virus strains could not be
4 included in the comparison. On the 3' side of ORF1 (approx. from nt pos. 1700 to 5625) the
5 similarity rates are higher than in the first third of this ORF, and the identity rates correspond
6 to the geographic origin of the viruses: the Polish genotypes are very similar to each other
7 (98-99%), and – surprisingly – they are more similar to the South-African strain (90-91%)
8 than to the Hungarian virus (84-85%). The Hungarian genotype differs widely from the
9 South-African strain as well (84-85%). In GenBank database partial sequence data had been
10 deposited on the 3' proximal side of ORF1 of one BQCV strain from Britain (helicase domain,
11 AF125252) and three genotypes from France (helicase AY669847, AY669848, Tentcheva et
12 al., 2004; and RdRp domain, AY230509). These western European genotypes shared 80-96%
13 identity with the investigated central European BQCVs, however the Poland 5 and 6 viruses
14 were 97% similar to the French viruses in the helicase domain region. The British strain,
15 however, showed similar and relatively low identities to the other investigated strains (82-
16 86%).

17 The two ORFs of BQCV are separated by a 208 nt long intergenic region between positions
18 5626 and 5833 (Leat et al., 2000). In the 3' proximal of this region a 155 nt long region was
19 found (between positions 5679 and 5833), in which all five investigated genotypes were
20 100% identical. This homologous region is extended to the first 56 nucleotides of ORF2. In
21 the GenBank database four further nucleotide sequences were found, which were also 100%
22 identical to the investigated viruses at the intergenic region of BQCV (BD173516,
23 BD177020, BD294724, and BD295735; sequences from Japan).

24 In ORF2 (structural polyprotein coding region) in general higher similarity rates (90 to 92%)
25 were found between the genotypes from different countries, compared to ORF 1 (82 to 88%).

1 Poland 5 and 6 genotypes shared 98% identity, while Poland 4 genotype was 92-93% similar
2 to them in ORF2. In GenBank database partial genome sequences were found in this region of
3 one genotype from Germany (AF521640, Siede and Buchler, 2003), one from the USA
4 (AY626246), and one from Uruguay (DQ364629, Antúnez et al., 2006). These viruses shared
5 93-98% identity to the central European BQCVs.

6 The distribution of the sequence differences within the BQCV genome are plotted on Figure
7 3. The conserved areas at the 5' UTR region and at the intergenic region are striking on the
8 plots, as well as the high level of diversity at the 5'-proximal side of ORF1. When Poland 4 is
9 used as reference, the Poland 5 and 6 sequences show high similarity to it at the 3'-proximal
10 side of the ORF1, while the South-African and the Hungarian genotypes exhibit higher
11 diversity. In the ORF2 region, however, all investigated strains differ in a similar level from
12 the Poland 4 genotype, especially at the 5'-proximal half of this ORF. The close genetic
13 relationship between the Poland 5 and Poland 6 strains is also indicated by Figure 3. The
14 greatest differences between these viruses are found at the ORF1 5' proximal region (~90%
15 identity), while at further genomic parts identities amount to 95 to 100%. Interestingly, in this
16 region the Poland 5 and 6 viruses are more similar to the Hungary 10 genotype than to the
17 Poland 4 strain, which suddenly turns to the opposite at approx. nt position 1700 nt. The
18 Hungarian genotype in general differs at the same level from the Polish and South-African
19 genotypes, however, within the ORF2, high similarity was found between the Poland 5 and
20 the Hungary 10 sequences.

21

22

1 Discussion

2 The scientific interest in insect viruses has significantly increased in the last ten years. On one
3 hand, insects represent the highest number of animal species on earth, and most of the viruses
4 infecting them presumably have not been discovered yet. Besides the possible use of viruses
5 in the control of insect pests and parasites, the knowledge on insect viruses is essential for the
6 prevention of mass mortality among economically important insects (such as honeybees) or
7 ecologically important and endangered species. On the other hand, because the physiology of
8 invertebrates and vertebrates differ in several aspects, the investigations of invertebrate
9 viruses may also provide information on differences in the pathophysiology, immunology,
10 and ecology of invertebrate and vertebrate infections.

11 Although virus infections in the honeybee are common, they are rarely associated with severe
12 symptoms and mortality. In certain cases, however, viruses are considered as the causative
13 agents of bee diseases. In vertebrate viruses it is well-known that different virus strains may
14 exhibit different virulence, but such alterations have not been described in any honeybee virus
15 yet. In our previous studies, the phylogenetic variability of two honeybee viruses, Sacbrood
16 virus (SBV; Grabensteiner et al., 2001) and Acute bee paralysis virus (ABPV, Bakonyi et al.,
17 2002) has been investigated. While SBV is member of the *Iflaviridae* family, *Iflavirus* genus,
18 ABPV is classified within the same family and genus as BQCV (*Dicistroviridae*, *Cripavirus*).
19 Although ABPV and BQCV share only 45% nucleotide sequence identity, their genome
20 organization is very similar. In general, phylogenetic clustering corresponding to the
21 geographic origin of the genotypes was observed both in SBV and in ABPV strains. Genetic
22 analysis of another iflavirus, the Deformed wing virus, however does not show any
23 phylogenetic clustering, based on its geographic origin (Berényi et al., 2007). Comparison of
24 the complete genome sequences of central European ABPV genotypes and the reference
25 strain did not result in the detection of significant differences in sequence divergences at any

1 regions of the investigated strains (Bakonyi et al., 2003). There were no significant
2 differences in the identity levels of the structural and non-structural protein coding regions of
3 SBV either (Grabensteiner et al., 2001).

4 In the study presented here, phylogenetic analyses were performed on two genomic regions of
5 22 central European BQCV strains and the South-African reference strain. The phylogenetic
6 trees indicated different clustering of the viruses. The tree based on the partial helicase coding
7 region could not differentiate the genotypes according to their origin, but three Polish
8 genotypes formed a separate cluster compared to the other central European BQCVs (Figure
9 1). Many of the "central European" subgroups were statistically poorly supported. The
10 phylogeny based on the partial structural protein coding region showed a much better
11 resolution: the viruses from Poland are separated from the Austrian and Hungarian genotypes,
12 and a closer genetic relatedness with limited mixing of the genotypes can be observed at the
13 Austrian–Hungarian clusters (Figure 2). The spatial distance, the limited trade of bee queens,
14 and the existence of ecological barriers (i.e. the Carpathian Mountains) between the Polish
15 and Austrian–Hungarian bee populations are plausible explanations for this sort of genetic
16 separation of the BQCV genotypes. The extreme geographical distance explains the
17 divergence of the South-African BQCV from the European ones. One Polish genotype,
18 however, cluster within a separate group, which is genetically distant from both the other
19 central European viruses, and from the South-African strain.

20 To estimate the reliability of the phylogenetic analyses, and to explain the unexpected
21 clustering of the three Polish genotypes, a more comprehensive genome analysis was
22 performed on the viruses concerned. Genetic recombination between virus genotypes could
23 explain the different clustering of the same virus strains on the two trees constructed using
24 sequences of the two different genomic regions. Recombination between foot-and-mouth
25 disease virus serotypes were described as frequent and common events, with typical

1 breakpoint distribution of "cold spots" and "hot spots" within the structural genes (Heath et
2 al., 2006). Similar recombination patterns were found in other picornaviruses too (Lukashev,
3 2005). The *Dicistroviridae* family is related to the *Picornaviridae* family; both of them are
4 due to be classified into the proposed order *Picornavirales*. Therefore, it would not be
5 surprising if recombinations were possible between BQCV genotypes as well. The
6 phylogenetic algorithms, however, calculate only independent sequence changes (mutations);
7 therefore recombinations may disturb the statistical analysis. Thus, such regions should be
8 selected for the phylogenetic comparisons which are not involved in recombination events.
9 However we can not exclude the possibility that the individual queen larvae (prepupae,
10 pupae) from Poland examined in the studies were infected by different virus genotypes, which
11 may enhance the opportunity for recombination events. The main reason for this is that each
12 queen larva is fed by many workers. In the queen rearing apiaries it is a common procedure to
13 pool the bees from different colonies together to obtain a rearing colony strong enough to rear
14 many queens and, in the apiary from which the samples for investigation originated, seven
15 BQCV genotypes (Poland 1-7) were found. On the other hand, the presence of different
16 BQCV genotypes in the apiary is not surprising because Polish native strains of bees are
17 rather aggressive and not very productive, so breeding material for queen rearing apiaries is
18 imported from different, sometimes distant regions.

19 By the comparison of the BQCV sequences, a typical pattern was found in the diversity of the
20 different genomic regions. Two genome sections (>100 nucleotides) were found where all
21 investigated viruses were homologous. Both areas are non-coding regions close to the
22 beginning of the ORFs. These areas usually play central roles in the transcription-initiation
23 and the expression regulation of the subsequent genes. The homologous area found in the
24 intergenic region covers the majority of the stem-loop structures predicted for the IRES of
25 BQCV and other dicistroviruses (Leat et al., 2000, Sasaki and Nakashima, 1999). Their

1 function might explain the high degree of sequence conservation; however, in other
2 dicistroviruses such 100% homologous areas were not found.

3 The 5'-proximal third of ORF1 proved to be the most variable region of the investigated
4 BQCVs. ORF1 encodes the non-structural proteins, but the exact role of this region is still
5 unknown (Leat et al., 2000). The domains of helicase, protease and RdRp are located in the
6 3'-proximal side of the ORF1. In the case of several vertebrate viruses non-structural proteins
7 are responsible for the differences in the virulence of the strains (Wicker et al., 2006; Tews et
8 al., 2009). Replication-associated genes, such as helicase and polymerase, are also likely to be
9 under some selective pressure for sequence conservation for protein function. Therefore, the
10 unique clustering of the Poland strains 4, 5 and 6 away from the other European strains in the
11 helicase phylogenetic tree is of interest as it may suggest that these are variants of other
12 BQCV genotypes and may have differing virulence. It is interesting to mention that the
13 Polish genotypes were all from symptomatic honeybees whereas the Austrian and Hungarian
14 genotypes were from asymptomatic honeybees. Further studies, involving experimental bee
15 (larvae) infections, are planned to investigate the possible connection between the sequence
16 divergence of this region and the virulence of BQCV strains. The latest investigations of
17 samples from queen rearing apiaries (Topolska, 2008) revealed a big diversity in the signs
18 which accompanied the death of a queen brood because of BQCV infection in Polish apiaries.
19 Although dead, decomposed black larvae and prepupae, known from literature, were present
20 (Bailey 1997), most often almost normal, pale yellow white-eyed pupae were found. Maybe
21 the lack of symptomatic BQCV infected queen brood in Hungary and Austria might be
22 explained by the presence of less virulent BQCV strains in these countries.

23 Another striking observation on the ORF1 is that the unique Polish genotypes are more
24 similar to the South-African strain than to the Hungarian genotype at the 3' side of the frame.
25 Sequences of two French BQCVs showed close relationship to these Polish genotypes.

1 Investigations of further western European strains could provide more information on the
2 possible origin of these particular viruses. The ORF2 in general exhibited a lower level of
3 sequence divergence than ORF1. However, the similarity rates between the closely related
4 viruses (i.e. the Polish genotypes) were lower in this region than in ORF1. Therefore, ORF2
5 seems to be a more suitable target region for phylogenetic analysis of BQCV, because the
6 mutation-derived individual sequence changes are more characteristic in this area.

7 Because the similarities between BQCV genotypes originating from the same geographic
8 regions (i.e. Poland 4 versus 5 and 6) suddenly changed within the genome, recombinations of
9 certain genome sections between genotypes is likely. In the current study the parties involved
10 in the recombination events were not accurately identified; therefore, further experiments are
11 necessary to document the exact process of a possible recombination event, and the most
12 frequently affected genomic regions. Poland 4 virus represents a unique genotype of the
13 central European BQCVs. Further investigations should reveal the pathological impact of this
14 virus, and identify closely related strains.

15 The RT-PCR-based diagnostic submissions should target the intergenic region of BQCV,
16 because due to the high level of conservation, the assays are likely to detect all genotypes,
17 regardless of their geographic origin and recombinations.

18

1 **Acknowledgements**

2 This study was supported by the Hungarian National Grants OTKA F043155, D048647, and
3 M027651; the Hungarian-Slovenian Intergovernmental S&T Cooperation Grant OMFB-
4 00482/07, and the Hungarian-Austrian Intergovernmental S&T Cooperation Grant OMFB-
5 00738/07. T. Bakonyi is grantee of the "Bolyai János" Postdoctoral Fellowship Grant of the
6 Hungarian Academy of Sciences.

7 The help of Eszter Roza (Faculty of Veterinary Science, Budapest) is greatly acknowledged.

8 The authors thank for the reviewers of Veterinary Microbiology for their valuable advices on
9 the manuscript.

10

Accepted Manuscript

1 **References**

- 2 Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J., 1990 Basic local
3 alignment search tool. *J. Mol. Biol.* 215, 403-410.
- 4 Allen, M. and Ball, B.V. 1996. The incidence and world distribution of honey bee viruses.
5 *Bee World.* 77, 141-162.
6
- 7 Anderson, D.L. 1993. Pathogens and queen bees. *Australasian Beekeeper.* 94, 292-296.
8
- 9 Antúnez, K., D'Alessandro, B., Corbella E., Ramallo G. and Zunino, P. 2006. Honeybee
10 viruses in Uruguay. *J. Invert. Pathol.* 93, 67-70.
11
- 12 Bailey, L. and Woods, R.D. 1977. Two more small RNA viruses from honey bees and further
13 observations on sacbrood and acute bee-paralysis viruses. *J. Gen. Virol.* 25, 175-186.
14
- 15 Bailey, L. 1982. Viruses of honeybees. *Bee World.* 63, 165-173.
16
- 17 Bakonyi, T., Grabensteiner, E., Kolodziejek, J., Rusvai, M., Topolska, G., Ritter, W. and
18 Nowotny, N. 2002. Phylogenetic analysis of acute bee paralysis strains. *Appl. Environ.*
19 *Microbiol.* 68, 6446-6450.
20
- 21 Bakonyi, T., Forgách, P., Topolska, G., Nowotny, N. and Rusvai, M. 2003. Nucleic acid
22 investigations of two central European strains of acute bee paralysis virus isolated from honey
23 bees (*Apis mellifera* L.), Apimondia 2003, 38th International Scientific Apicultural Congress,
24 Ljubljana, Slovenia
25
- 26 Berényi, O., Bakonyi, T., Derakhshifar, I., Köglberger, H. and Nowotny, N. 2006. Occurrence
27 of six honeybee viruses in diseased Austrian apiaries. *Appl. Environ. Microbiol.* 68, 6446-
28 6450.
29
- 30 Berényi, O., Bakonyi, T., Derakhshifar, I., Köglberger, H., Topolska, G., Ritter W.,
31 Pechhacker, H. and Nowotny, N. 2007. Phylogenetic analysis of deformed wing virus
32 genotypes from diverse geographic origins indicates recent global distribution of the virus.

- 1 Appl Environ Microbiol 73, 3605-3611
- 2
- 3 Cherry, S. and Silverman, N. 2006. Host-pathogen interactions in drosophila: new tricks from
4 an old friend. Nat. Immunol. 7, 911-917.
- 5
- 6 Felsenstein, J. 2004. PHYLIP Phylogeny Inference Package Version 3.6b (beta release)
7 Department of Genome Sciences, University of Washington, Seattle, WA USA
- 8
- 9 Grabensteiner, E., Ritter, W., Carter, M.J., Davison, S., Pechhacker, H., Kolodziejek, J.,
10 Boecking, O., Derakhshifar, I., Moosbeckhofer, R., Licek, E. and Nowotny, N. 2001.
11 Sacbrood virus of the honeybee (*Apis mellifera*): Rapid identification and phylogenetic
12 analysis using reverse transcription-PCR. Clin Diagn. Lab. Immunol. 8, 93-104.
- 13
- 14 Heath, L., van der Walt, E., Varsani, A. and Martin, D.P. 2006. Recombination patterns in
15 aphthoviruses mirror those found in other picornaviruses. J Virol. 80, 11827-11832.
- 16
- 17 Leat, N., Ball, B.V., Govan, V. and Davison, S. 2000. Analysis of the complete genome
18 sequence of black queen-cell virus, a picorna-like virus from honey bees. J. Gen. Virol. 81,
19 2111-2119.
- 20
- 21 Lukashev, A.N. 2005. Role of recombination in evolution of enteroviruses. Rev. Med. Virol.
22 15, 157-167.
- 23
- 24 Ray, S. C. 2003. SimPlot for Windows 98/NT/2000/XP Version 3.5.1
- 25
- 26 Sasaki, J. and Nakashima, N. 1999. Translation initiation at the CUU codon is mediated by
27 the internal ribosome entry site of an insect picorna-like virus *in vitro*. J. Virol. 73, 1219-
28 1226.
- 29
- 30 Siede, R. and Buchler, R. 2003. Symptomatic Black Queen Cell Virus infection of drone
31 brood in Hessian apiaries. Berl. Munch. Tierarztl. Wochenschr. 116, 130-133.
- 32

- 1 Tentcheva, D., Gauthier, L., Zappulla, N., Dainat, B., Cousserans, F., Colin, M.E. and
2 Bergoin, M. 2004. Prevalence and Seasonal Variations of Six Bee Viruses in *Apis mellifera* L.
3 and *Varroa destructor* Mite Populations in France. *Appl. Environ. Microbiol.* 70, 7185-7191.
4
- 5 Tews, B.A., Schürmann, E.M. and Meyers, G. 2009. Mutation of cysteine 171 of pestivirus E
6 ns RNase prevents homodimer formation and leads to attenuation of classical swine fever
7 virus. *J Virol.* 83, 4823-4834.
8
- 9 Topolska, G. 2008. Zakażenia wirusowe czerwiu matecznego oraz matek pszczelich w
10 dziesięciu pasiekach hodowlanych w Polsce [Virus infections of queen brood and queen bees
11 in ten queen rearing apiaries in Poland] Warsaw University of Life Sciences, Warsaw, Poland
12
- 13 Wicker, J.A., Whiteman, M.C., Beasley, D.W., Davis, C.T., Zhang, S., Schneider, B.S.,
14 Higgs, S., Kinney, R.M. and Barrett, A.D. 2006. A single amino acid substitution in the
15 central portion of the West Nile virus NS4B protein confers a highly attenuated phenotype in
16 mice. *Virology.* 349, 245-253.
17
- 18 Zambon, R.A., Nandakumar, M., Vakharia, V.N. and Wu, L.P. 2005. The Toll pathway is
19 important for an antiviral response in *Drosophila*. *Proc. Nat. Acad. Sci. U S A.* 102, 7257-
20 7262.
21
22

1 **Legends of tables and figures**

2

3 **Table I.** Oligonucleotide primer pairs selected for RT-PCR, specific for the helicase and the
4 structural protein region of BQCV, respectively. Amplicons were sequenced and used for
5 constructing phylogenetic trees (Figures 1. and 2.).

6

7 **Table II.** Oligonucleotide primer pairs designed to produce overlapping sequences for the
8 complete sequencing of selected BQCV genotypes.

9

10 **Figure 1.** Neighbor-joining tree based on the partial helicase enzyme region (ORF1) of
11 BQCV genotypes. Internal labels indicate the percentages of bootstrap values; only values
12 > 70% were indicated. Bar represents the genetic distance between the viruses.

13

14 **Figure 2.** Neighbor-joining tree based on the partial structural polypeptide coding region
15 (ORF2) of BQCV genotypes. Internal labels indicate the percentages of bootstrap values; only
16 values > 70% were indicated. Bar represents the genetic distance between the viruses.

17

18 **Figure 3.** Similarity plot of the Hungary 10, Poland 4, and Poland 5 genotypes and the South-
19 African reference strain compared to the Poland 6 genotype.

Figure 1

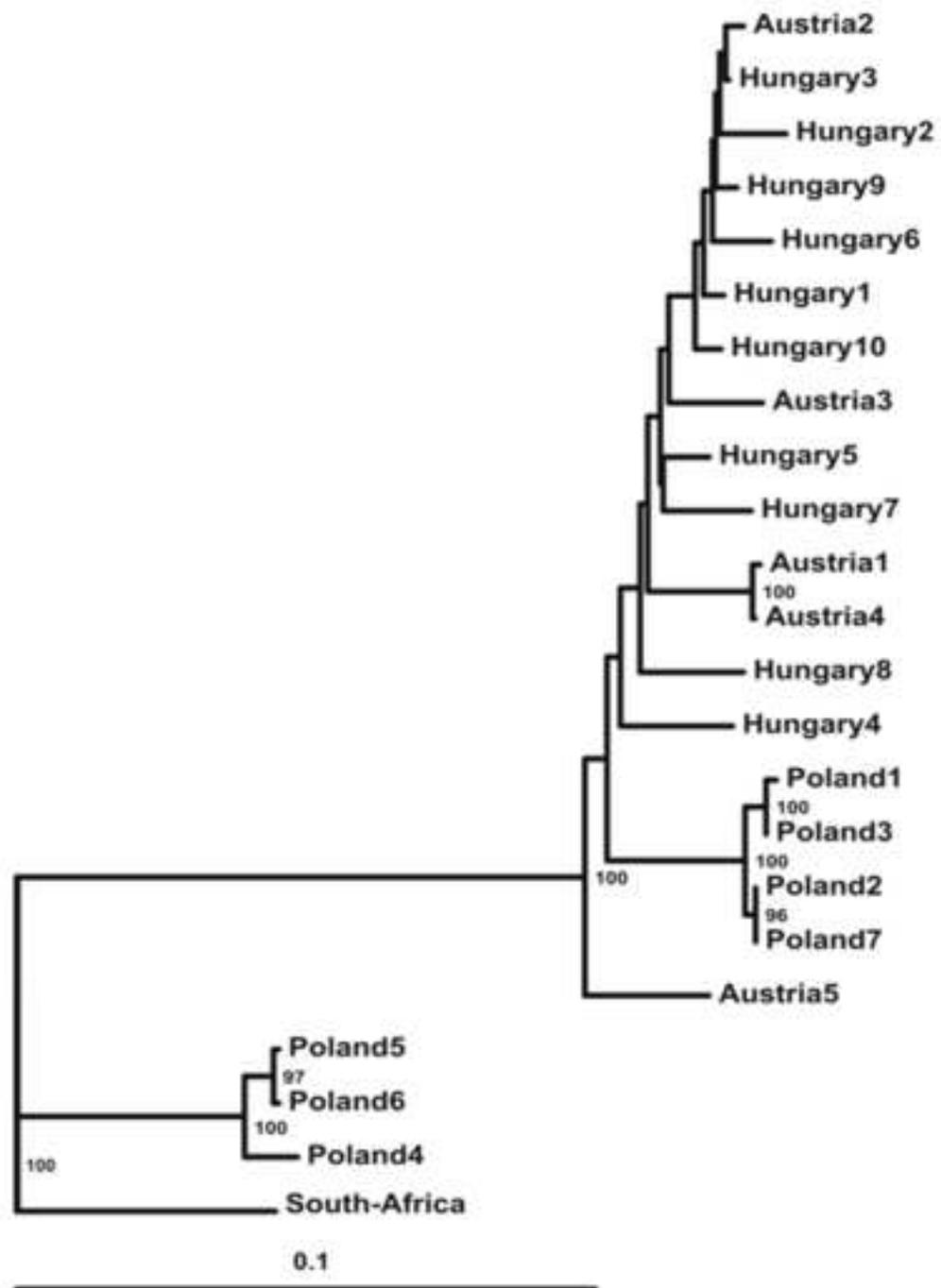


Figure 2

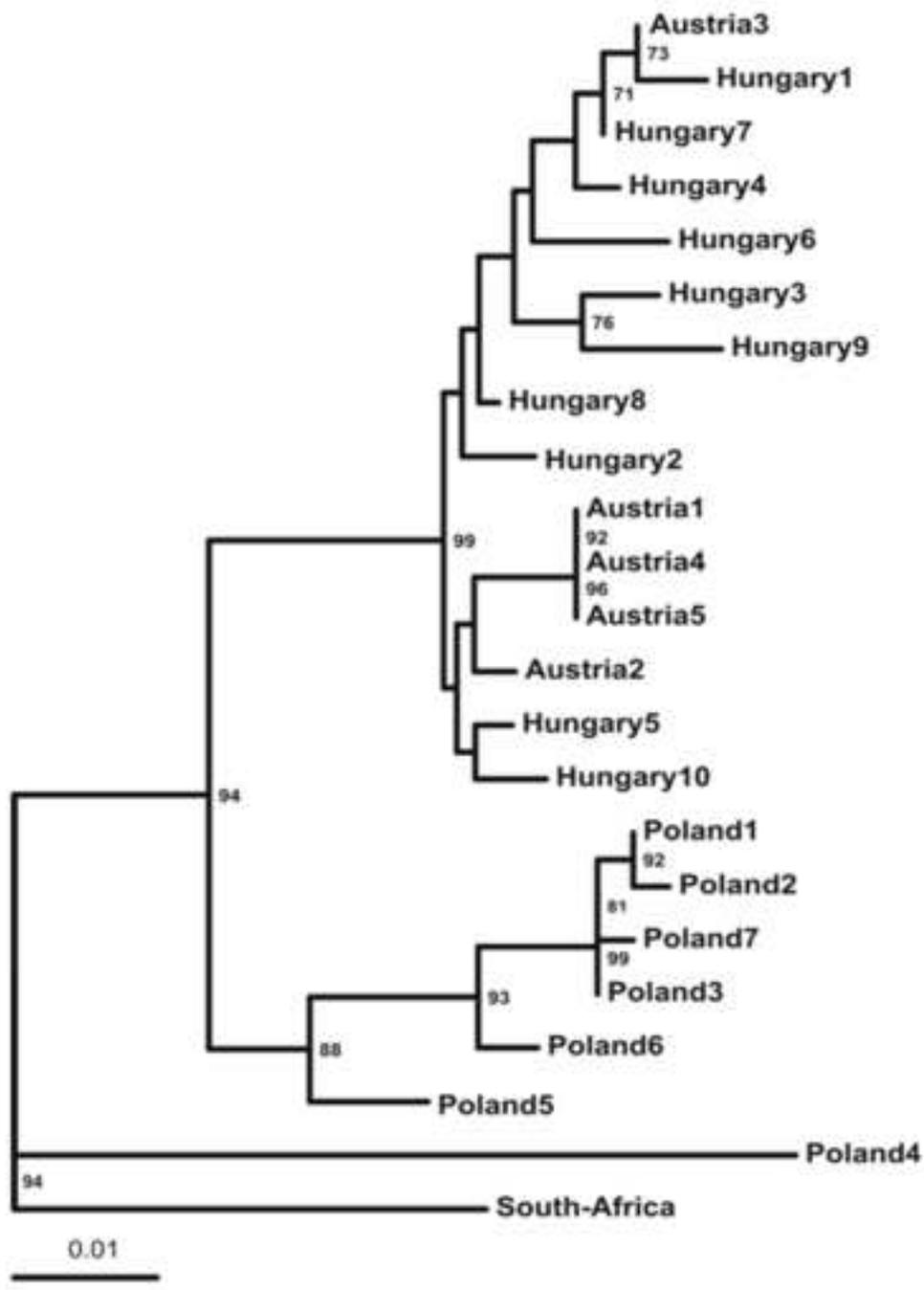
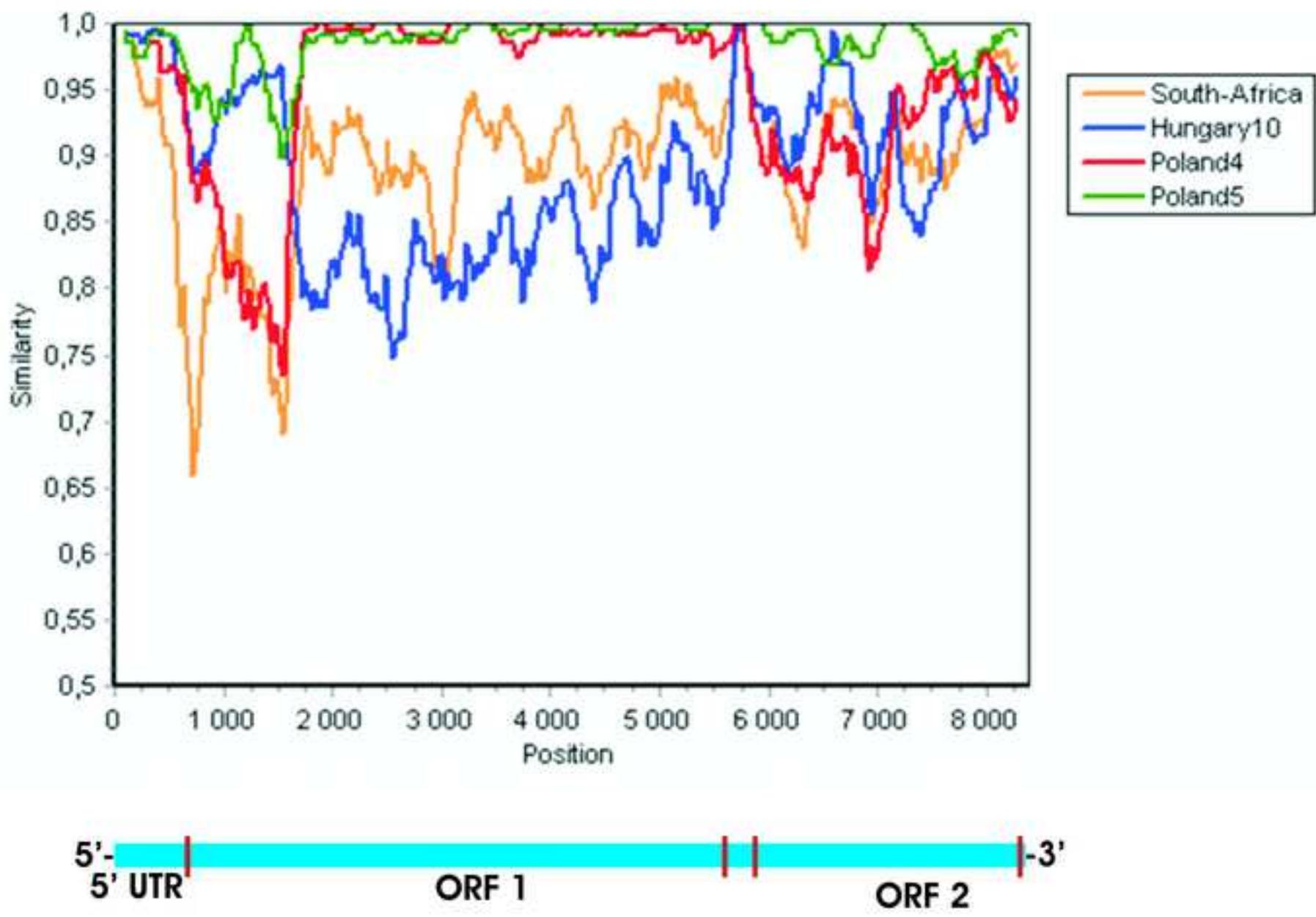


Figure 3

ripc



| Region | Primer ^(a) code | Sequence (5' to 3') | Nucleotide positions ^(b) | Size of the amplicon |
|--------------------|---------------------------------------|----------------------------|--|---------------------------------|
| Helicase | BQCV2144+ | TGA GAG CTG CAG AAC AAG AG | 2144-2163 | 583 bp |
| | BQCV2726- | TCA CGT GTC AAG GCT TCA TC | 2726-2707 | |
| Structural protein | BQCV6540+ | TGA GAG CTG CAG AAC AAG AG | 6540-6559 | 514 bp |
| | BQCV7053- | TCC ATG GCG ACA GTT ACA TC | 7053-7034 | |

^(a) +: forward, -: reverse primers

^(b) Nucleotide positions refer to the published complete BQCV sequence (GenBank accession number AF183905, Leat et al., 2000).

| Primer ^(a) code | Sequence (5' to 3') | Nucleotide positions ^(b) | Size of the amplicon (bp) |
|-------------------------------|----------------------------|--|------------------------------|
| BQ29+ | CTA TAC GCG CTT GGT TGT | 29-46 | 1807 |
| BQ1835- | TCC GCT AAC GTG GAC TCT GT | 1835-1816 | |
| BQ1645+ | ATA GGC ACG TGG TTG ACT AC | 1645-1664 | 1082 |
| BQ2726- | TCA CGT GTC AAG GCT TCA TC | 2726-2707 | |
| BQ2144+ | TGA GAG CTG CAG AAC AAG AG | 2144-2163 | 1033 |
| BQ3716- | CTA GAT TCC ACG CGC ACA GT | 3176-3157 | |
| BQ3016+ | GCG AGT AAG GCC AAG GTT GA | 3016-3035 | 1549 |
| BQ4564- | CCG CCT CTA TGC ATT CCT GT | 4564-4545 | |
| BQ4429+ | GAT ACC TCT TGC GGT TAT CC | 4429-4448 | 1510 |
| BQ5938- | CGG TCC ATC TTC AAG CAC TA | 5938-5919 | |
| BQ5744+ | CTC CAA GAT CGG TGG ATA GC | 5744-5763 | 1730 |
| BQ7473- | TCG TCG CCA GCA TTG ACT TC | 7473-7454 | |
| BQ7353+ | CCA AGT ACG CTC ATG CTA GT | 7353-7372 | 1140 |
| BQ8492- | TCA TGA GAA GAA CCG AGA AG | 8492-8473 | |
| BQ29+ | CTA TAC GCG CTT GGT TGT | 29-46 | 2275 |
| BQ2303- | GCC ATG TGT AAC GGA TAA GG | 2303-2284 | |
| BQ2298+ | CAT GGC GTC GAT AGA AGA GA | 2298-2317 | 989 |
| BQ3286- | TCGTTGTAACAGCCGTAAGC | 3286-3267 | |
| BQ7405+ | GTT GAA GGC GAC TAA TAC CG | 7405-7424 | 1088 |
| BQ8492- | TCA TGA GAA GAA CCG AGA AG | 8492-8473 | |
| BQ510+ | TGC TAC TGC GGT AGT GGA | 510-527 | Sequencing primer |
| BQ1062+ | CGT GGA CAC CTC TAT TGG TA | 1062-1081 | Sequencing primer |
| BQ506+ | ACT CTG CTA CTG CGG TAG CG | 506-525 | Sequencing primer |
| BQ1834- | CCG CTA ATG TGG ATT CTG | 1834-1817 | Sequencing primer |
| BQ276+ | GTA GCC TGT ACT ACC TGA AG | 276-295 | Sequencing primer |
| BQ951+ | GGT TGA CCG AAT GTA CCA | 951-968 | Sequencing primer |

^(a) +: forward, -: reverse primers

^(b) Nucleotide positions refer to the published complete BQCV sequence (GenBank accession number AF183905, Leat et al., 2000).