HAL
open science

Genetic Analysis and Phylogenetic Comparison of Black Queen Cell Virus Genotypes

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

To cite this version:

Zsuzsanna Tapaszti, Petra Forgách, Csaba Kővágó, 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-00526932v1

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 Tapaszti, 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: Tapaszti, 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.

ACCEPTED MANUSCRIPT

Genetic Analysis and Phylogenetic Comparison of Black Queen Cell Virus Genotypes

Zsuzsanna Tapaszti[1], Petra Forgách[1]*, Csaba Kővágó[1], Grażyna Topolska[2], Norbert Nowotny[3], Miklós Rusvai[4], Tamás Bakonyi[1,3]

[1]Department of Microbiology and Infectious Diseases, Faculty of Veterinary Science, Szent István University, Hungária krt. 23-25, H-1143 Budapest, Hungary

[2]Laboratory of Bee Diseases, Department of Clinical Sciences, Faculty of Veterinary Medicine, Warsaw University of Life Sciences, 8 Ciszewskiego str., 02-786 Warsaw, Poland

[3]Zoonoses and Emerging Infections Group, Clinical Virology, Department of Pathobiology, University of Veterinary Medicine, Vienna, Veterinärplatz 1, A-1210 Vienna, Austria

[4]Department of Pathology and Forensic Veterinary Medicine, Faculty of Veterinary Science, Szent István University, István u. 2, H-1078 Budapest, Hungary

Running title: Phylogeny of BQCV

*Corresponding author. Mailing address: Department of Microbiology and Infectious Diseases, Faculty of Veterinary Science, Szent István University, Hungária krt. 23-25, H-1143 Budapest, Hungary. Phone: 36 1 2519900. Fax: 36 1 2519260. E-mail: Forgach.Petra@aotk.szie.hu

Accepted Manuscript

ACCEPTED MANUSCRIPT

Abstract

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

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

Accepted Manuscript

ACCEPTED MANUSCRIPT

Introduction

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

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

ACCEPTED MANUSCRIPT

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

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

Accepted Manuscript

ACCEPTED MANUSCRIPT

Materials and Methods

Samples

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

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

RNA isolation and RT-PCR

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

ACCEPTED MANUSCRIPT

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

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

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

Nucleotide sequencing and computer analysis

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

ACCEPTED MANUSCRIPT

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

Nucleotide sequence accession numbers.

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

Accepted Manuscript

ACCEPTED MANUSCRIPT

Results

Phylogenetic analysis of central European Black queen cell virus genotypes

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

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

ACCEPTED MANUSCRIPT

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

Phylogenetic analyses were performed on the two selected genome regions (Figure 1. and 2.). The tree based on the partial helicase coding region is separated in three main groups. All Austrian and Hungarian genotypes and Polish genotypes 1 to 3 and 7 cluster within one branch of the tree. Three Polish genotypes (4 to 6) form a separate cluster, and the South-African strain represents the third main branch of the tree. The three unique Polish genotypes are closer to the South-African strain than to the other European genotypes (Figure 1.). However, the phylogenetic tree based on the partial structural polyprotein sequences exhibits partly a considerably different topology: the central European BQCVs form a common cluster, however it is sub-divided into two main groups. One group contains the Austrian and Hungarian genotypes, while the Polish viruses form the second group. The separation of the two groups is statistically supported (bootstrap of 1000 replicates is 944). Within the Austrian-Hungarian group, four Austrian genotypes form a separate subgroup together with two Hungarian genotypes, while the other eight Hungarian viruses and one Austrian genotype form a second, more diverse branch. In this tree, genotypes Poland 5 and 6 cluster together with the other Polish genotypes, with the exception of Poland 4, which has similar genetic distances from both the European genotypes and the South-African strain, the latter representing the third main branch of the tree (Figure 2.). Phylogenetic trees based on the putative amino acid sequences were also constructed, but the relatively short sequences and high similarity rates resulted in low bootstrap values, therefore these trees are not shown.

Genome comparisons of selected BQCV genotypes

The phylogenetic analysis of central European BQCV genotypes revealed that three viruses from Poland (named Poland 4, 5 and 6) represent a separate cluster of the investigated

ACCEPTED MANUSCRIPT

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

ACCEPTED MANUSCRIPT

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

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

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

ACCEPTED MANUSCRIPT

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

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

Accepted Manuscript

ACCEPTED MANUSCRIPT

Discussion

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

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

ACCEPTED MANUSCRIPT

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

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

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

ACCEPTED MANUSCRIPT

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

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

ACCEPTED MANUSCRIPT

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

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

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

ACCEPTED MANUSCRIPT

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

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

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

Accepted Manuscript

ACCEPTED MANUSCRIPT

Acknowledgements

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

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

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

Accepted Manuscript

ACCEPTED MANUSCRIPT

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J., 1990 Basic local alignment search tool. J. Mol. Biol. 215, 403-410.

Allen, M. and Ball, B.V. 1996. The incidence and world distribution of honey bee viruses. Bee World. 77, 141-162.

Anderson, D.L. 1993. Pathogens and queen bees. Australasian Beekeeper. 94, 292-296.

Antúnez, K., D'Alessandro, B., Corbella E., Ramallo G. and Zunino, P. 2006. Honeybee viruses in Uruguay. J. Invert. Pathol. 93, 67-70.

Bailey, L. and Woods, R.D. 1977. Two more small RNA viruses from honey bees and further observations on sacbrood and acute bee-paralysis viruses. J. Gen. Virol. 25, 175-186.

Bailey, L. 1982. Viruses of honeybees. Bee World. 63, 165-173.

Bakonyi, T., Grabensteiner, E., Kolodziejek, J., Rusvai, M., Topolska, G., Ritter, W. and Nowotny, N. 2002. Phylogenetic analysis of acute bee paralysis strains. Appl. Environ. Microbiol. 68, 6446-6450.

Bakonyi, T., Forgách, P., Topolska, G., Nowotny, N. and Rusvai, M. 2003. Nucleic acid investigations of two central European strains of acute bee paralysis virus isolated from honey bees (*Apis mellifera* L.), Apimondia 2003, 38th International Scientific Apicultural Congress, Ljubljana, Slovenia

Berényi, O., Bakonyi, T., Derakhshifar, I., Köglberger, H. and Nowotny, N. 2006. Occurrence of six honeybee viruses in diseased Austrian apiaries. Appl. Environ. Microbiol. 68, 6446-6450.

Berényi, O., Bakonyi, T., Derakhshifar, I., Köglberger, H., Topolska, G., Ritter W., Pechhacker, H. and Nowotny, N. 2007. Phylogenetic analysis of deformed wing virus genotypes from diverse geographic origins indicates recent global distribution of the virus.

ACCEPTED MANUSCRIPT

Appl Environ Microbiol 73, 3605-3611

Cherry, S. and Silverman, N. 2006. Host-pathogen interactions in drosophila: new tricks from an old friend. Nat. Immunol. 7, 911-917.

Felsenstein, J. 2004. PHYLIP Phylogeny Inference Package Version 3.6b (beta release) Department of Genome Sciences, University of Washington, Seattle, WA USA

Grabensteiner, E., Ritter, W., Carter, M.J., Davison, S., Pechhacker, H., Kolodziejek, J., Boecking, O., Derakhshifar, I., Moosbeckhofer, R., Licek, E. and Nowotny, N. 2001. Sacbrood virus of the honeybee (*Apis mellifera*): Rapid identification and phylogenetic analysis using reverse transcription-PCR. Clin Diagn. Lab. Immunol. 8, 93-104.

Heath, L., van der Walt, E., Varsani, A. and Martin, D.P. 2006. Recombination patterns in aphthoviruses mirror those found in other picornaviruses. J Virol. 80, 11827-11832.

Leat, N., Ball, B.V., Govan, V. and Davison, S. 2000. Analysis of the complete genome sequence of black queen-cell virus, a picorna-like virus from honey bees. J. Gen. Virol. 81, 2111-2119.

Lukashev, A.N. 2005. Role of recombination in evolution of enteroviruses. Rev. Med. Virol. 15, 157-167.

Ray, S. C. 2003. SimPlot for Windows 98/NT/2000/XP Version 3.5.1

Sasaki, J. and Nakashima, N. 1999. Translation initiation at the CUU codon is mediated by the internal ribosome entry site of an insect picorna-like virus *in vitro*. J. Virol. 73, 1219-1226.

Siede, R. and Buchler, R. 2003. Symptomatic Black Queen Cell Virus infection of drone brood in Hessian apiaries. Berl. Munch. Tierarztl. Wochenschr. 116, 130-133.

Accepted Manuscript

ACCEPTED MANUSCRIPT

Tentcheva, D., Gauthier, L., Zappulla, N., Dainat, B., Cousserans, F., Colin, M.E. and Bergoin, M. 2004. Prevalence and Seasonal Variations of Six Bee Viruses in *Apis mellifera* L. and *Varroa destructor* Mite Populations in France. Appl. Environ. Microbiol. 70, 7185-7191.

Tews, B.A., Schürmann, E.M. and Meyers, G. 2009. Mutation of cysteine 171 of pestivirus E rns RNase prevents homodimer formation and leads to attenuation of classical swine fever virus. J Virol. 83, 4823-4834.

Topolska, G. 2008. Zakażenia wirusowe czerwiu matecznego oraz matek pszczelich w dziesięciu pasiekach hodowlanych w Polsce [Virus infections of queen brood and queen bees in ten queen rearing apiaries in Poland] Warsaw University of Life Sciences, Warsaw, Poland

Wicker, J.A., Whiteman, M.C., Beasley, D.W., Davis, C.T., Zhang, S., Schneider, B.S., Higgs, S., Kinney, R.M. and Barrett, A.D. 2006. A single amino acid substitution in the central portion of the West Nile virus NS4B protein confers a highly attenuated phenotype in mice. Virology. 349, 245-253.

Zambon, R.A., Nandakumar, M., Vakharia, V.N. and Wu, L.P. 2005. The Toll pathway is important for an antiviral response in *Drosophila*. Proc. Nat. Acad. Sci. U S A. 102, 7257-7262.

Accepted Manuscript

ACCEPTED MANUSCRIPT

Legends of tables and figures

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

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

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

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

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

Accepted Manuscript

Figure 1

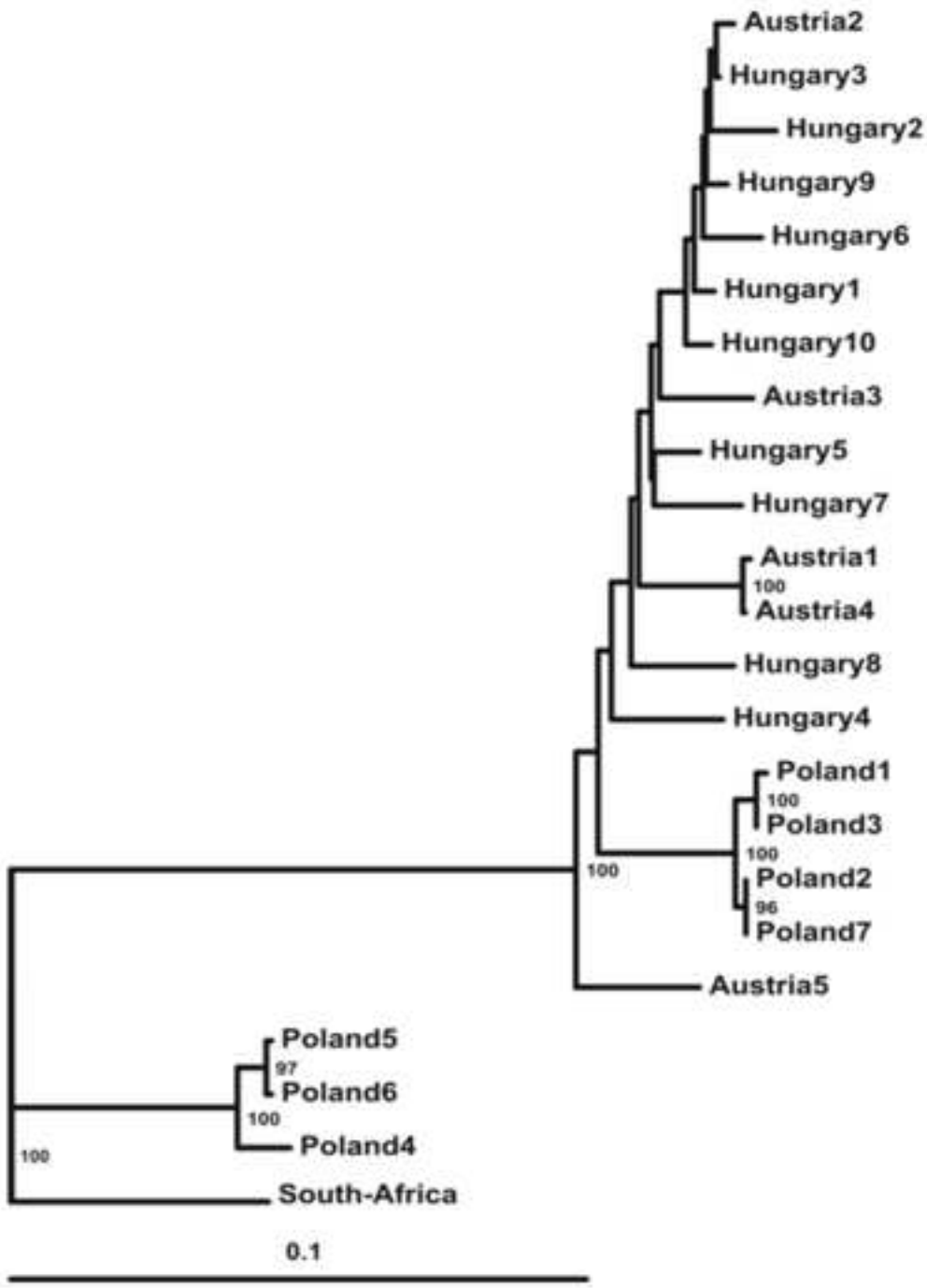
Austria2
Hungary3
Hungary2
Hungary9
Hungary6
Hungary1
Hungary10
Austria3
Hungary5
Hungary7
Austria1
100
Austria4
Hungary8
Hungary4
Poland1
100
Poland3
100
Poland2
96
Poland7
100
Austria5
Poland5
97
Poland6
100
Poland4
100
South-Africa
0.1

Figure 2

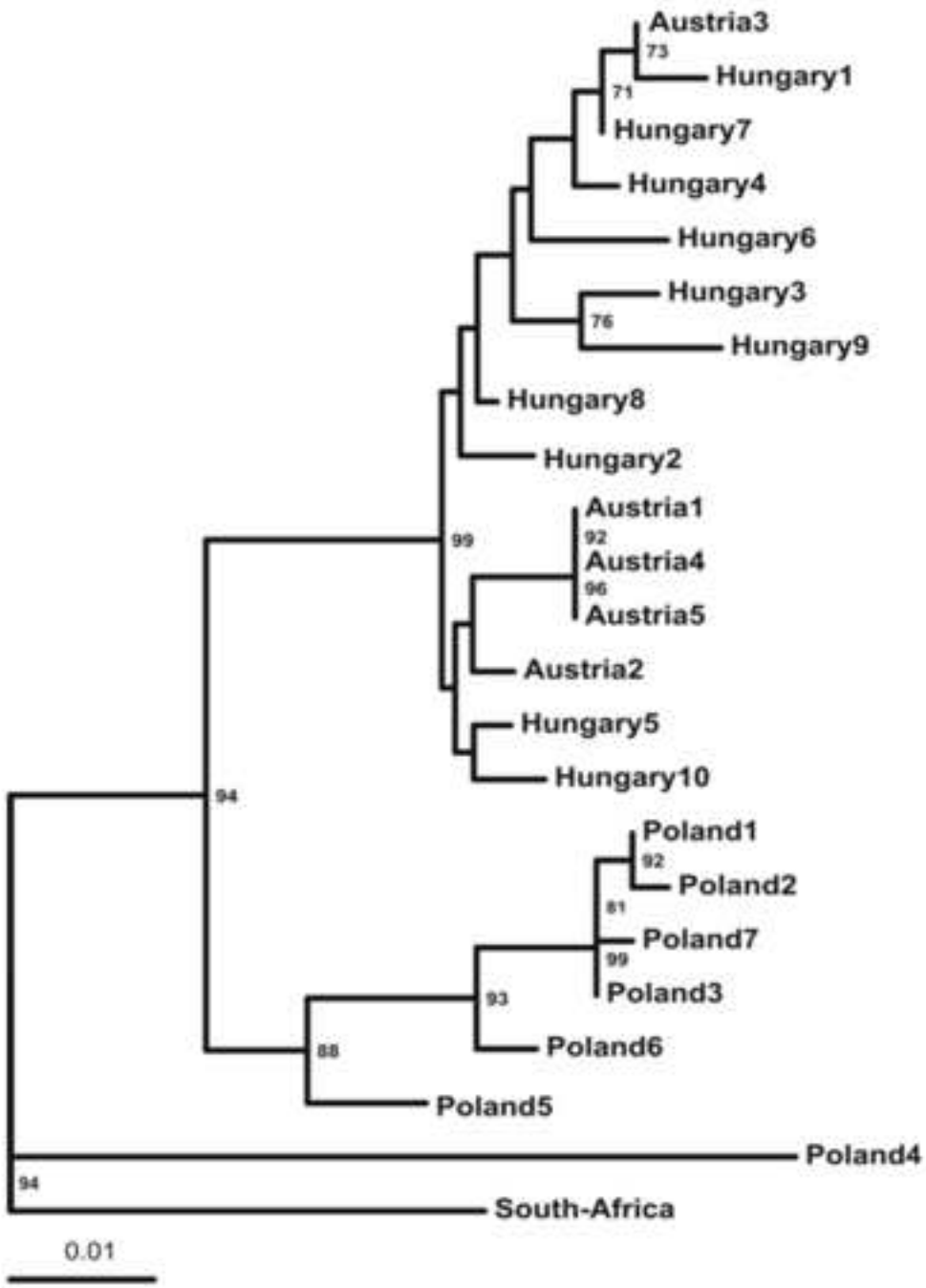
Austria3
73
Hungary1
71
Hungary7
Hungary4
Hungary6
Hungary3
76
Hungary9
Hungary8
Hungary2
Austria1
92
Austria4
96
Austria5
99
Austria2
Hungary5
Hungary10
94
Poland1
92
Poland2
81
Poland7
99
Poland3
93
Poland6
88
Poland5
Poland4
94
South-Africa
0.01

Figure 3

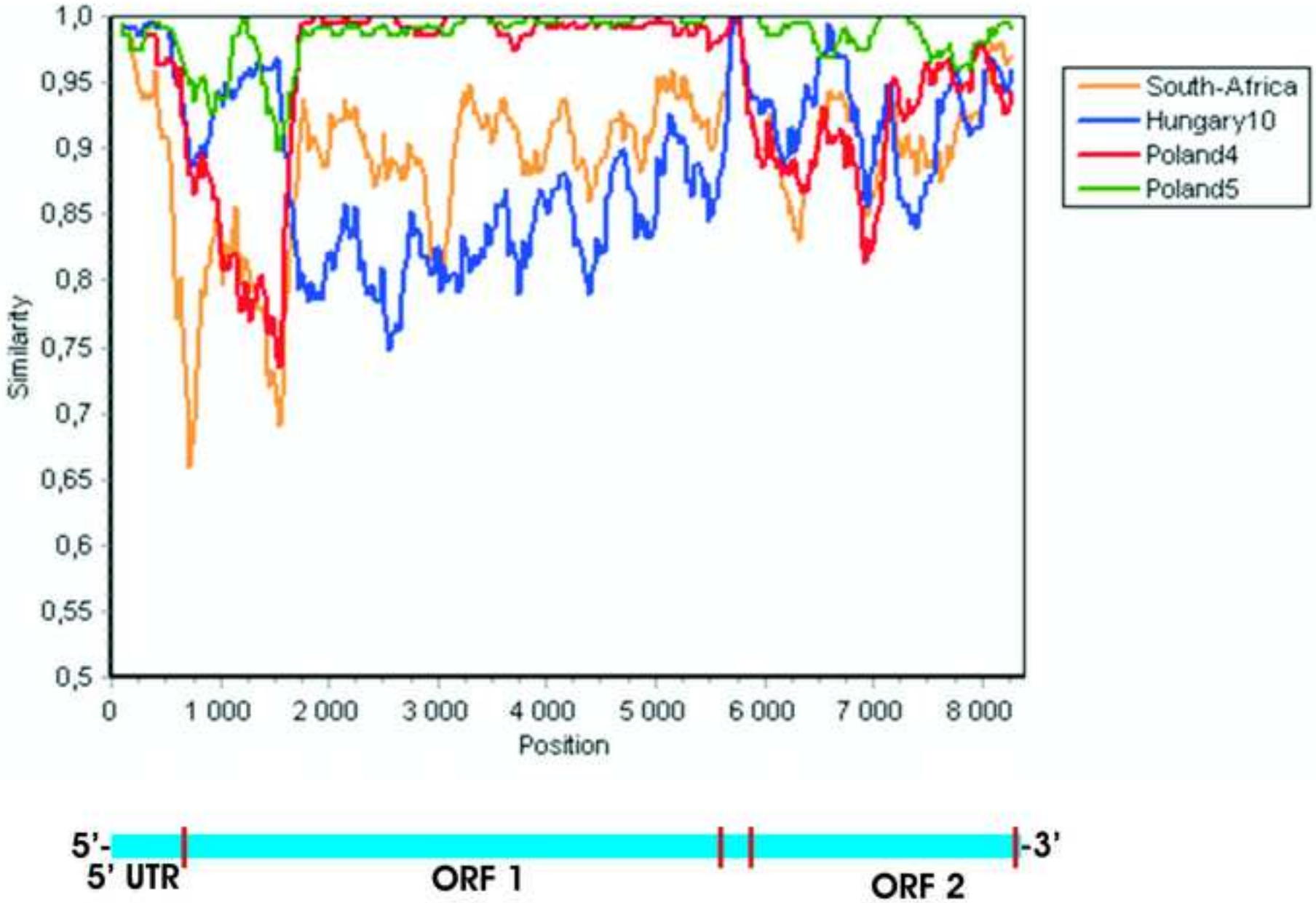
1,0
0,95
0,9
0,85
0,8
0,75
0,7
0,65
0,6
0,55
0,5
Similarity
0
1 000
2 000
3 000
4 000
5 000
6 000
7 000
8 000
Position
South-Africa
Hungary10
Poland4
Poland5
5'-
5' UTR
ORF 1
ORF 2
-3'

Table I

ACCEPTED MANUSCRIPT

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).

Accepted Manuscript

ACCEPTED MANUSCRIPT

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).

Accepted Manuscript