Full genomic analysis of new variant rabbit hemorrhagic disease virus revealed multiple recombination events

Ana M. Lopes, Kevin P. Dalton, Maria J. Magalhães, Francisco Parra, Pedro J. Esteves, Edward C. Holmes and Joana Abrantes

Correspondence
Joana Abrantes
jabrances@cibio.up.pt

1CIBIO, InBIO - Research Network in Biodiversity and Evolutionary Biology, Universidade do Porto, Campus de Vairão, Rua Padre Armando Quintas, 4485-661 Vairão, Portugal
2Departamento de Biologia, Faculdade de Ciências da Universidade do Porto, Porto, Portugal
3INSERM, UMR892, Université de Nantes, Nantes, France
4Instituto Universitário de Biotecnologia de Asturias, Departamento de Bioquímica y Biología Molecular, Universidad de Oviedo, Oviedo, Spain
5CITS, Centro de Investigação em Tecnologias da Saúde, IPSN, CESPU, Gandra, Portugal
6Marie Bashir Institute for Infectious Diseases and Biosecurity, Charles Perkins Centre, School of Biological Sciences and Sydney Medical School, The University of Sydney, Sydney, NSW, Australia

Rabbit hemorrhagic disease virus (RHDV), a Lagovirus of the family Caliciviridae, causes rabbit hemorrhagic disease (RHD) in the European rabbit (Oryctolagus cuniculus). The disease was first documented in 1984 in China and rapidly spread worldwide. In 2010, a new RHDV variant emerged, tentatively classified as ‘RHDVb’. RHDVb is characterized by affecting vaccinated rabbits and those <2 months old, and is genetically distinct (~20 %) from older strains. To determine the evolution of RHDV, including the new variant, we generated 28 full-genome sequences from samples collected between 1994 and 2014. Phylogenetic analysis of the gene encoding the major capsid protein, VP60, indicated that all viruses sampled from 2012 to 2014 were RHDVb. Multiple recombination events were detected in the more recent RHDVb genomes, with a single major breakpoint located in the 5′ region of VP60. This breakpoint divides the genome into two regions: one that encodes the non-structural proteins and another that encodes the major and minor structural proteins, VP60 and VP10, respectively. Additional phylogenetic analysis of each region revealed two types of recombinants with distinct genomic backgrounds. Recombinants always include the structural proteins of RHDVb, with non-structural proteins from non-pathogenic lagoviruses or from pathogenic genogroup 1 strains. Our results show that in contrast to the evolutionary history of older RHDV strains, recombination plays an important role in generating diversity in the newly emerged RHDVb.

INTRODUCTION

Rabbit hemorrhagic disease (RHD) is an acute fatal necrotizing hepatitis that affects wild and domestic European rabbits of both the Oryctolagus cuniculus cuniculus and Oryctolagus cuniculus algirus subspecies (Abrantes et al., 2012b). The disease was first detected in 1984 in China following the importation of commercially bred Angora rabbits from Germany (Liu et al., 1984). RHD rapidly spread worldwide and is currently endemic in several countries including Portugal, Spain and France. The economic burden of RHD on rabbit farming and hunting industries and the negative ecological impact on the wild rabbit populations and in their dependent predators, particularly in the Mediterranean ecosystem, are well known and of major concern (Delibes-Mateos et al., 2009). In contrast, in Australia, RHD has been successfully exploited as a form of biocontrol to reduce rabbit numbers, although...
there is mounting evidence for host resistance (Elsworth et al., 2012; Kovaliski, 1998; Kovaliski et al., 2014).

RHD is caused by rabbit hemorrhagic disease virus (RHDV), a Lagovirus of the family Caliciviridae. RHDV has a non-enveloped single-stranded (ss) RNA genome organized into two narrowly overlapping open reading frames (ORFs). ORF1 encompasses nucleotide residues 10–7044 and encodes a large polyprotein that is cleaved by a virus-encoded protease to generate several non-structural (NS) proteins and the major capsid protein, VP60 or VP1 (for convenience, we use VP60). ORF2 comprises nucleotide residues 7025–7378 and produces VP10 (or VP2), a minor structural protein (Wirblich et al., 1996). At the 5′ region, the RHDV genome presents a covalently linked protein, VPg, and is polyadenylated at the 3′ region (Gregg et al., 1991; Morales et al., 2004). Viral particles also package an abundant subgenomic RNA (sgRNA) that contains the coding sequences of VP60 and VP10. As is the case for the RHDV genomic RNA, the sgRNA is VPg-linked and polyadenylated (Meyers et al., 1991; Morales et al., 2004; Ohlinger et al., 1990).

Despite the importance of RHDV for animal health and ecosystem well-being, its origin and emergence as a pathogenic virus in the European rabbit is not well understood. On current data the most likely explanation is direct evolution from a non-pathogenic form of the virus (Kerr et al., 2009). Indeed, a number of non-pathogenic and a weakly pathogenic rabbit calciviruses have been described in Europe, America and Australia (Bergin et al., 2009; Capucci et al., 1996; Le Gall-Reculé et al., 2011b; Strive et al., 2009), and there is evidence for the existence of RHDV as a non-pathogenic form before the first documented outbreak (reviewed by Abrantes et al., 2012b). To date, however, the mutations (or recombination events) responsible for the profound change in virulence are unknown. It is also theoretically possible that a cross-species transmission event was central to the emergence of RHDV as a pathogenic form. Indeed, in the 1990s, RHDV affected a species other than the European rabbit (Lopes et al., 2014).

More recently, a new variant of RHDV was detected in France (Le Gall-Reculé et al., 2011a). As a definitive RHDV nomenclature has not yet been agreed (various terms such as ‘new variant’, ‘RHDVb’ and ‘RHDV2’ have been used), for simplicity we use the term ‘RHDVb’. RHDVb resulted in atypical RHD outbreaks in that it led to mortality in both vaccinated adult rabbits (Le Gall-Reculé et al., 2011a) and young rabbits (Dalton et al., 2014, 2012) that are typically resistant to RHDV. Phylogenetic relationships inferred from the VP60 gene showed that RHDVb formed a novel phylogenetic group that fell between the RCV-like viruses, i.e. European non-pathogenic viruses and the weakly pathogenic Michigan rabbit calcivirus (MRCV) and the Australian non-pathogenic rabbit calciviruses (RCV-A1) (Le Gall-Reculé et al., 2011a). Further work demonstrated that RHDVb is antigenically different from other RHDVs (Dalton et al., 2012; Le Gall-Reculé et al., 2013), although its genesis is uncertain. Importantly, the recent detection of RHDVb in leporid species other than the European rabbit (Camarda et al., 2014; Puggioni et al., 2013) indicates that it might infect a wider range of host species.

Both host-jumping and changes in virulence have been associated with the ability of RNA viruses to rapidly produce genetic diversity through mutation (Domingo & Holland, 1997; Holmes, 2010). It is also possible that the genotypes associated with successful emergence are generated by recombination (Simon-Loriere & Holmes, 2011). Because it requires template switching, (homologous) RNA recombination generally involves regions of high sequence similarity. Although homologous recombination has been detected in RHDV (Abrantes et al., 2008; Forrester et al., 2008), its impact on fitness and hence long-term importance remains uncertain. Notably, the RHDV strain isolated in 1984 in China had a recombinant origin (Forrester et al., 2008), which may have been in part responsible for the marked change in virulence observed in this virus.

RHDV has been divided into six genogroups, denoted G1–G6 (Le Gall-Reculé et al., 2003). In the Iberian Peninsula, RHDV is characterized by a high degree of virus isolation with only G1 detected in wild rabbit populations (Abrantes et al., 2012a; Alda et al., 2010; Muller et al., 2009), which differs from the epidemiological pattern seen in other European countries (Le Gall-Reculé et al., 2003). However, in 2011 and 2012, RHDVb was detected in both Spain and Portugal (Abrantes et al., 2013; Dalton et al., 2012). To better understand the evolution of RHDV and particularly the occurrence and impact of recombination, we determined the complete coding sequences of Iberian RHDV strains collected between 1994 and 2014.

RESULTS AND DISCUSSION

We determined the complete coding sequence (excluding the 5′ and 3′ UTRs) of 28 RHDV strains from rabbits found dead in the Iberian Peninsula. Based on their position in the maximum-likelihood (ML) tree of the capsid gene (VP60) (Fig. 1), 24 strains were identified as RHDVb and four as genogroup 1 (G1). Considering the distribution of the European rabbit subspecies in the Iberian Peninsula (Carneiro et al., 2011; Ferrand, 2008), these results further confirm that both O. cuniculus algirus and O. cuniculus cuniculus are susceptible to this new virus (Abrantes et al., 2013). Interestingly, the co-circulation of RHDVb and other RHDV genogroups, particularly G1, was not detected. This supports the notion that G1 has been replaced by RHDVb (Lopes et al., 2015) and is compatible with the rapid spread of this new variant through the Iberian Peninsula (Dalton et al., 2014).

The complete coding sequences of RHDVb exhibited some length variation due to deletions in the non-structural
Fig. 1. ML phylogenetic tree for 211 VP60 sequences of RHDV. Major genetic groups (genogroups) are indicated. For clarity of presentation, the sequences in genogroups G2–G5 and G6 (RHDVα) have been collapsed and have the following GenBank accession numbers: AB300693, AF231353, AF258618, AJ00619, AJ302016, AJ303106, AJ495586, AJ55092, AJ55094, AJ68694, AM085133, AY922826, AY922826, DQ069280–DQ069282, DQ189077, DQ189077, DG205345, DG841708, EF360335, EF58582, EF558574, EF558575, EF558577, EF558581–EF558584, EU003587–EU003582, EU250330, EU650619, EU650680, FJ12322, FJ12322, FJ74179, FJ74180, FN562800, FR823354, FR823355, GU392228, GU37617, GU37618, GU566448, HE963222, HM62309, HC917923, JF412629, JN165233–JN165236, JN581729–JN581735, JQ815931, JQ95154, JX393309–JX393312, KC345614, KC595207, KF70630, KF49490–KF494952, KF537692, KF537693, KF549474–KF549476, KF677011, KJ571917–KJ5719160, KJ606958, KJ60959, KJ63893, KJ683900, KJ683902, KJ683905–KJ683908, KJ814617–KJ814622, M67473, RHU49726, RHU54983, X875670, Y15424 and Y15427. The accession numbers are indicated for the remaining sequences. Viruses sequenced here (in bold) fall into two distinct clusters, G1 and RHDVβ. All horizontal branch lengths are drawn to a scale of nucleotide substitutions per site and the tree is mid-point rooted. Bootstrap values greater than 75% are shown for key nodes. *The main group of sequences in genogroups G2–G5 clustered together with 99% bootstrap support.
proteins p16/NS1 and p37/NS3: a deletion of codon 68 (p16) in CBAnd1/Spain/04-2012 and deletions of codons 134–135 (p16/NS1) and of codon 714 (p37/NS3) in several strains. While the biological role of p16/NS1 is unknown, a recent study of RHDV in Australia revealed a relatively high number of mutations in this protein, which were tentatively associated with increased virulence (Elsworth et al., 2014). Protein p37 has an ATPase activity and may be a member of the helicase superfamily III (Marín et al., 2000). Although helicase activity is crucial for replication hydrolysis (Marín et al., 2000). This indicates that the main biological functions of p37 are not compromised by the deletion. Additionally, the deletion is present in MRCV, RCV-A1 and in European brown hare syndrome virus (the other member of the genus Lagovirus), further suggesting that the replication activity of p37 is maintained. A mutation in the initiation codon (T to C in codon ATG, nucleotide positions 7025–7027) of ORF2 was observed in one strain, but expression of VP10 is not expected to be affected as translation of ORF2 relies on the presence of a sequence element at the 3′ end of ORF1 and not on a particular initiation codon (Meyers, 2003). Additionally, another ATG in this region (nucleotide positions 7037–7039) may act as the start site for translation of ORF2.

The most striking result of our study was the occurrence of recombination in a number of recent RHDVb viruses with strong statistical support (P values <0.001, Table 1). In these recombinant viruses, there was consistent evidence for a single recombination breakpoint located in a region close to the VP60 initiation codon (although two recombination breakpoints were detected they fall in very similar and overlapping locations, which is strongly suggestive of a single recombination event; Table 1). This was further confirmed with the SimPlot analysis (Fig. 2).

Notably, this recombination breakpoint is located within a region of high sequence similarity (data not shown) and divides the genome into two regions that include different protein subsets: one that includes the non-structural proteins and another that comprises the structural proteins VP60 and VP10. To confirm this recombinant history, ML phylogenetic trees were inferred for the different genome fragments on either side of the recombination breakpoint identified above. For simplicity, we basied this phylogenetic analysis on the recombinant breakpoint at nucleotide position 5305. For the phylogenetic tree of the genes encoding the structural proteins (VP60 and VP10 genes; Fig. 3a), all viruses previously identified as RHDVb (Fig. 1) clustered together in a single strongly supported monophyletic group (identified as RHDVb and shaded light grey). This cluster fell deep in the VP60 and VP10 phylogeny, close to the position of the mid-point of the tree and between the non-pathogenic Australian strain (RCV-A1) and the weakly pathogenic MRCV. Hence, RHDVb is phylogenetically distinct from the pathogenic RHDV sequences that form a separate highly supported cluster. In contrast, in the phylogeny inferred for the non-structural proteins (Fig. 3b), the RHDVb sequences fall into three distinct clusters: (i) one comprising likely non-recombinant RHDVb sequences (shaded light grey) that share a common ancestor with pathogenic RHDV strains, (ii) another composed of likely recombinant RHDVb sequences that cluster with MRCV and RCV-A1 (medium grey) and (iii) a group of RHDVb recombinant sequences that cluster closely with recently sampled pathogenic G1 viruses (dark grey).

The occurrence of multiple recombination events, i.e. RHDVb clusters (i) and (ii) mentioned above, makes it difficult to clearly distinguish recombinants from parents, but it is clearly necessary to invoke at least two independent recombination events involving RHDVb to explain these phylogenetic data. Moreover, the divergent position of the CBAnd1/Spain/04-2012 sequence within the RHDVb cluster (ii) may be indicative of an additional recombination event. It is also evident that one of the recombination events involved the combination of the RHDVb VP60 and VP10 coding sequences with a backbone composed of the coding sequences of the non-structural proteins of a G1-like strain, while in the other the RHDVb VP60 and VP10 coding sequences became linked with the coding sequences of the non-structural proteins of non-pathogenic lagoviruses (Fig. 3b).

Since some of the recombinant viruses possess non-structural genes that are closely related to the Australian non-pathogenic virus (as supported by the clustering of such recombinant viruses with RCV-A1 with 98 % bootstrap support), this evolutionary history implies that lineages related to the Australian viruses must circulate in the Iberian Peninsula but have not yet been detected [indeed, it is clear that RCV-A1 was imported into Australia (Jahnke et al., 2010)]. In addition, recombinant sequences were retrieved from animals found dead in the field with lesions compatible with RHD, and thus these strains are likely pathogenic despite their genome deriving mostly from non-pathogenic viruses.

Although circulation of ‘true’ G1 viruses has not been detected in this study from 2011 onwards, the co-circulation of G1-like and RHDVb strains within the same geographical region and time-frame must have occurred for recombination to take place. Interestingly, the two different types of recombinants were detected from the same geographical area only one year apart (e.g. CBAlgarve14-1/Portugal, detected in January 2014, and CBAlgarve1/Portugal/01-2013, detected in January 2013). These observations indicate that there is a relatively high diversity of RHDV strains circulating in this region, which clearly merits additional surveillance. In addition, the same type of recombinant was observed in several and distantly located populations and in different years, confirming its viability and capacity to spread at the epidemiological scale.

Two mechanisms of RNA recombination are possible for RNA viruses: replicative and non-replicative. The (copy-choice)
### Table 1. Results of the recombination analysis

<table>
<thead>
<tr>
<th>Recombinant strains</th>
<th>Most likely ‘parental’ strain</th>
<th>Genogroup</th>
<th>Breakpoint†</th>
<th>Average P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBAnd1/Spain/04-2012</td>
<td>EU871528/RCV-A1/ Australia/2007</td>
<td>NS†</td>
<td>RHDVb</td>
<td>5305</td>
</tr>
<tr>
<td>CBMestremoz14-1/Portugal/01-2014 CBMestremoz14-3/Portugal</td>
<td>Seg08-12/Spain/08-2012</td>
<td>VP60 and VP10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CBLavra10-13-1/Portugal/10-2013 CBAlgarve1/Portugal/01-2013 CBAlgarve3/Portugal/01-2013 Barrancos7-13/Portugal/01-2013 Barrancos10A-13/Portugal/01-2013 CBCoruche14-1/Portugal</td>
<td>G1</td>
<td>RHDVb</td>
<td>5336</td>
<td></td>
</tr>
<tr>
<td>CBMertola14-2/Portugal/01-2014 CBMertola14-1/Portugal/01-2014 CBAlgarve14-1/Portugal CBAlgarve14-3/Portugal CBAlgarve14-4/Portugal</td>
<td>JX886001/CB194/Chaves/Portugal/2006</td>
<td>Seg08-12/Spain/08-2012</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

†99 % confidence interval is indicated.
§Non-pathogenic strain.

*Non-structural proteins.
Replicative mechanism involves template switching by the viral polymerase during RNA synthesis. In other *Caliciviridae* genera, such as *Sapovirus*, *Vesivirus* and *Norovirus* (*NoV*), a recombination hotspot is similarly observed between the non-structural proteins and the capsid junction (Bull & White, 2010). In the case of *NoV*, a template switching mechanism has been proposed that reflects a combination of the copy-choice and the internal initiation model (Bull *et al.*, 2005). Accordingly, recombination occurs in the second round of replication when the polymerase starts the synthesis of a positive-sense RNA strand from a full-length negative strand. Due to the presence of complex secondary structures, such as stem–loops, at the ORF1/ORF2 overlap that coincides with the RNA promoter sequence, the polymerase loses processivity and switches to a negative strand of sgRNA synthesized by a co-infecting virus, thereby creating a hybrid virus (Bull *et al.*, 2005). Alternatively, the polymerase switches from a genomic RNA to another in regions with high sequence homology, such as the one found where ORF1 and ORF2 overlap (Bull *et al.*, 2005). Although the genomic organization of *NoV* differs from *RHDV* with the non-structural proteins and the capsid protein encoded by separate ORFs, the synthesis of the sgRNA by an internal initiation mechanism and the presence of a subgenomic

**Fig. 2.** Similarity plot for two types of RHDVb recombinants: (a) CBAlgare1/Portugal/01-2013 (non-pathogenic lagovirus/RHDVb) and (b) CBAlgare14-1/Portugal (G1/RHDVb). The vertical axis represents the percentage of sequence identity of each putative parental strain (grey and black lines) with the recombinant. The horizontal axis indicates the nucleotide positions. A window size of 200 bp with a step size of 20 bp was used. The site where the parental strains are identical in sequence to the recombinant is the predicted recombination breakpoint. A scheme of the genomic organization is shown above each plot with the encoded proteins indicated (pro, protease); arrowheads indicate cleavage sites.
promoter upstream of the sgRNA starting site in RHDV (Morales et al., 2004) suggests that a similar recombination mechanism may occur. In addition, the presence of a region of high sequence homology upstream of the initiation codon of VP60 in RHDV (Simmonds et al., 2008), which is compatible with the location of the sgRNA promoter (Boga et al., 1992; Miller & Koev, 2000; Morales et al., 2004; Sibilia et al., 1995), as well as the presence of a stem–loop (Simmonds et al., 2008), tentatively suggest that the recombination events observed could have been due to the template switching mechanism.

However, it is also possible that the recombinant RHDV genomes observed might have been generated by a non-replicative recombination mechanism. Such a mechanism, demonstrated for other positive-sense ssRNA viruses such as polioviruses and pestiviruses (Gallei et al., 2004; Gmyl...
et al., 1999), involves the cleavage of the recombining RNAs that are then ligated to form the recombinant molecule. In contrast to the template switching mechanism of recombination that is promoted by sequence conservation, the non-replicative recombination seems to be largely mediated by the presence of a RNA secondary structure (Chetverin et al., 1997) such as the stem–loop predicted for RHDV in the sgRNA promoter region (Simmonds et al., 2008).

While the impact of recombination on the fitness of RHDVb remains to be assessed, this process has clearly created substantial changes in the viral genome, although whether this has led to differences in virulence is unclear. Indeed, the recombination events described involve the combination of structural and non-structural protein subsets with different ancestries: protein coding sequences are combined as 'modules' with non-structural proteins originating either from non-pathogenic forms or GI viruses and structural proteins originating from RHDVb. A similar model of modular viral evolution has been described for enteroviruses (family Picornaviridae), in which the structural and non-structural regions of the genome evolve semi-independently (Lukashev, 2005; Lukashev et al., 2005; Oberste et al., 2004), giving rise to new virus variants. RNA signals and sequence similarity are among the factors that might impose such a recombination pattern (Runckel et al., 2013), although other constraints exist at several stages of viral infection and replication in the host. These constraints most likely reflect functional protein 'blocks' that work together for protein compatibility. Interestingly, previous studies revealed recombination breakpoints throughout the RHDV genome (Abrantes et al., 2008; Forrester et al., 2008), although most of these recombination events occurred between closely related RHDV strains. Hence, it is likely that the sequence similarity between the recombining strains allows the functional integrity to be maintained. As there is prior evidence that recombination may have been central to the emergence of RHDV as a pathogenic form (Forrester et al., 2008), we contend that more effort should
be directed toward revealing the frequency and fitness consequences of recombination in lagoviruses.

METHODS

Virus samples and genome amplification. Liver samples were collected from wild rabbits found dead in the field in different regions of Portugal and from domestic rabbits from Spanish rabbitries, and were kept frozen at −20 °C. No live animals were shot, trapped or handled to obtain tissues, as such, no animal ethics permit was required. Each frozen liver was thawed and up to 30 mg of the liver was collected. The tissue was homogenized in a rotor-stator homogenizer (Mixer Mill MM400; Retsch) at 30 Hz for 7 min. Total RNA was extracted using the RNeasy Mini Kit (Qiagen) according to the protocol provided by the manufacturer. Reverse transcription was performed using oligo(dT) as primers and SuperScript III Reverse Transcriptase (Invitrogen) according to the provided protocol. The protocol provided by the manufacturer. Reverse transcription was performed using oligo(dT) as primers and SuperScript III Reverse Transcriptase (Invitrogen) according to the provided protocol. The complete genome of each RHDV strain was obtained by the amplification of several overlapping fragments ranging from ~700 to 1700 bp of template and water to a final volume of 10 μl. The primers and PCR conditions used are summarized in Table 2. For recombinant sequences, each PCR was performed at least twice independently. After purification, PCR products were sequenced on an automatic sequencer ABI PRISM 310 Genetic Analyzer (Applied Biosystems). Sequences were determined using the amplification primers and are available in GenBank.

No incongruences were detected between the overlapping fragments that could indicate preferential amplification of templates in each PCR. While it is highly unlikely that the Phusion DNA Polymerase generates crossovers at the same position repeatedly, a long range PCR was also performed to obtain complete genomic sequences and exclude the possibility of PCR-mediated artefacts in generating the recombinant sequences. This PCR was conducted with the following conditions: 0.25 μl of Takara LA Taq (Clontech), 2.5 μl buffer, 2.5 mM MgCl₂, 2.5 mM of each dNTP, 0.4 μM of each primer, 1 μl of template and water to a final volume of 25 μl. Cycle conditions consisted of 2 min at 94 °C, 35 cycles of 30 s at 94 °C, 30 s at 63 °C and 8 min 30 s at 68 °C and a final extension of 15 min at 68 °C. The amplicons were sequenced as described above using amplification primers listed in Table 2 (primer 5′-GTGTATGCCATGACCTCGTGGT-3′ covers the regions located both upstream and downstream of the recombination breakpoint).

Recombination analysis. The complete genome sequences obtained (excluding the 5′ and 3′ UTRs) were aligned using the BioEdit software, version 7.0.9.0 (Hall, 1999). Full-length RHDV genome sequences available in public databases were retrieved and included in the alignment. This produced a final dataset of 68 sequences of 7569 nt, including representatives from all major groups of RHDV. This sequence alignment was screened for recombination using the RDP, GENECONV and BootScan methods implemented in the RDP software (version 4.26) (Martin et al., 2010) under the following parameters: sequences were set to linear, Bonferroni correction, highest acceptable P value 0.05 and 100 bootstraps. Only recombination events detected by all three methods were considered as viable and hence merited additional analysis (see Phylogenetic analysis section). Strains detected as recombinants were further analysed using the SimPlot software (version 3.5.1) (Lole et al., 1999) with the default parameters (window size, 200 bp; step, 20 bp; Ts/Tv ratio, 2:0).

Phylogenetic analysis. An initial reconstructed phylogenetic tree was inferred for the VP60 gene to provide a provisional picture of the evolutionary relationships of RHDVb to the remaining pathogenic and non-pathogenic lagoviruses. For this analysis, all available VP60 sequences were retrieved from public databases, producing a final dataset of 211 sequences of 1740 nt. A phylogenetic tree was then estimated using the ML method available in PhyML program (Guindon et al., 2010) utilizing the GTR+I+F model of nucleotide substitution and employing a combination of nearest-neighbour interchange (NNI) and subtree pruning and regrafting branch-swapping. The support for each node was determined from 1000 bootstrap replicate ML trees, employing the same substitution model and parameter values as above, and with NNI branch-swapping.

For the complete genome sequences, the dataset was partitioned at the putative recombination breakpoint(s) detected in the recombination analysis described in the Results and Discussion section and then subjected to phylogenetic analysis. This partitioning was as follows: (i) ORF1 except VP60 (nucleotide residue positions 10–5304) and (ii) VP60 and VP10 (nucleotide residue positions 5305–7378). Phylogenetic trees were then estimated for both partitions using the same phylogenetic procedure as described above (i.e. ML trees in PhyML).

ACKNOWLEDGEMENTS

We acknowledge all the collaborators who provided samples: Dr Paulo Célio Alves (CIBIO/InBIO-UP), Vítor Palmilha, José Manuel Maia and Jacinto Amaro. We thank INTERCUN and Spanish Ministerio de Economía y Competitividad (grant AGL2013–48550–C2–1-R co-financed by FEDER) for the logistical and financial support provided to the laboratory of F. P. This work was funded by National Funds through the FCT-Foundation for Science and Technology under the project ref. FCT-ANR/BIA-BIC/0043/2012. This work was also supported by ‘Genomics Applied to Genetic Resources’, co-financed by North Portugal Regional Operational
programme 2007/2013 (ON.2 - O Novo Norte), under the National Strategic Reference Framework (NSRF), through the European Regional Development Fund (ERDF). FCT supported the doctoral fellowship of A.M.L. (ref. SFHR/Bd/78738/2011) and the FCT Investigator grant of J.A. (ref. IF/01396/2013). E.C.H. is supported by an NHMRC Australia Fellowship (AF30) and by the ARC (grant no. DP140103362).

**REFERENCES**


Recombination in rabbit hemorrhagic disease virus


