Proposal for a unified classification system and nomenclature of lagoviruses

Jacques Le Pendu,1,* Joana Abrantes,2† Stéphane Bertagnoli,3† Jean-Sébastien Guitton,4† Ghislaine Le Gall-Reculé,5† Ana Margarida Lopes,2† Stéphane Marchandeau,4† Fernando Alda,6 Tereza Almeida,2,4 Alves Paulo Célio,7,8 Juan Bárcena,9 Galina Burmakina,10 Esther Blanco,9 Carlos Calvete,11 Patrizia Cavadini,12 Brian Cooke,13 Kevin Dalton,14 Miguel Delibes Mateos,15 Wieslaw Deptula,16 John Sebastian Eden,17 Fang Wang,18 Catarina C Ferreira,19,20 Paula Ferreira,21 Pilar Foronda,22 David Gonçalves,27 Dolores Gavier-Widén,23,24 Robin Hall,25 Beata Hukowska-Szematowicz,26 Peter Kerr,27 John Kovalisli,28 Antonio Lavazza,12 Jackie Mahar,17,27 Alexander Malogolovkin,10 Raquel M. Marques,21 Sara Marques,2,29 Aaron Martin-Alonso,22 Pedro Monterroso,2 Sacramento Moreno,30 Greg Mutze,26 Aleksija Neimanis,23,31 Paulina Niedzwiedzka-Rystwej,26 David Peacock,28 Francisco Parra,14 Mara Rocchi,32 Carlos Rouco,33 Nathalie Ruvoën-Clouet,1 Eliane Silva,2,29 Diogo Silvério,2 Tanja Strive,25 Gertrudes Thompson,2,29 Beata Tokarz-Deptula16 and Pedro Esteves2,7,34,*

Abstract

Lagoviruses belong to the Caliciviridae family. They were first recognized as highly pathogenic viruses of the European rabbit (Oryctolagus cuniculus) and European brown hare (Lepus europaeus) that emerged in the 1970–1980s, namely, rabbit haemorrhagic disease virus (RHDV) and European brown hare syndrome virus (EBHSV), according to the host species from which they had been first detected. However, the diversity of lagoviruses has recently expanded to include new related viruses with varying pathogenicity, geographic distribution and host ranges. Together with the frequent recombination observed amongst circulating viruses, there is a clear need to establish precise guidelines for classifying and naming lagovirus strains. Therefore, here we propose a new nomenclature based on phylogenetic relationships. In this new nomenclature, a single species of lagovirus would be recognized and called Lagovirus europaeus. The species would be divided into two genogroups that correspond to RHDV- and EBHSV-related viruses, respectively. Genogroups could be subdivided into genotypes, which could themselves be subdivided into phylogenetically well-supported variants. Based on available sequences, pairwise distance cutoffs have been defined, but with the accumulation of new sequences these cutoffs may need to be revised. We propose that an international working group could coordinate the nomenclature of lagoviruses and any proposals for revision.
INTRODUCTION

The authors of this article constitute an international working group, that may be joined by other interested researchers, who, we suggest, should be in charge of defining the nomenclature of lagoviruses. Lagoviruses constitute a genus of the Caliciviridae family with viruses of this group causing severe diseases in the European rabbit (Oryctolagus cuniculus) and in several hare species (Lepus spp.). They emerged in the late 1970s to the early 1980s when a new disease epidemic, European brown hare syndrome (EBHS), began to cause death in the European brown hare population of Sweden [1]. Just a few years later, in 1984, a similar disease was recorded among farmed rabbits in China from where it rapidly spread worldwide [2]. This disease, termed rabbit haemorrhagic disease (RHD), is highly contagious, fulminant and usually fatal to infected rabbits over eight weeks old. Both diseases are very similar, characterized by acute necrotizing hepatitis and haemorrhages in many organs, particularly the lungs, heart and kidneys, generally associated with disseminated intravascular coagulation. Some cases also show tracheal congestion. The acute form of RHD involves depression, anorexia, apathy, rapid respiration, anemia, and some animals show signs of abdominal distress. Animals perish after 1–3 days [3, 4]. The etiological agents of RHD, rabbit haemorrhagic disease virus (RHDV), and of EBHS, European brown hare syndrome virus (EBHSV), were discovered in the late 1980s to the early 1990s. Both agents were identified as small positive-sense single-stranded RNA viruses that were quickly recognized as caliciviruses [5–11]. RHDV and EBHSV share typical features with other members of the Caliciviridae family. Viral particles are small (30–35 nm), andicosahedral. Their genomes of about 7.4 kb share the general organization of calicivirus genomes. They differ by having only two ORFs since the coding region for the capsid protein is included in ORF1 rather than constituting a separate ORF as in other members of the family [11, 12]. Many individual strains were later characterized from dead rabbits or hares. In 1996, a non-pathogenic virus related to RHDV was discovered on a farm in Italy [13]. This virus was called RCV for rabbit calicivirus, similar to the designation first suggested for the pathogenic viruses. Since then, several other non-pathogenic strains have been described in wild animals either in Australasia or in Europe [14–17]. In 1998, the first consistent antigenic variant of RHDV, called RHDVa, was detected in Italy [18]. Then, in 2010, another distinct pathogenic variant of RHDV was described in France and it soon appeared that this new virus was supplanting the ‘classical’ RHDV strains in France, Spain and Portugal [19–24].

A confused historical nomenclature of lagoviruses

The nomenclature of lagoviruses was initially based on the associated pathogenicity and the host of the originally identified strains (RHDV, EBHSV). As new strains were discovered, and based on their position in phylogenetic trees or antigenic properties, additional qualifiers were added such as RHDV G1, G2, G3, G4, G5 and G6 [25], the latter corresponding to the antigenic variant RHDVa, with the letter ‘a’ standing for antigenic variant [18]. The inclusion of additional strains in phylogenetic trees revealed that the G3, G4 and G5 subcategories could no longer be distinguished [26]. The discovery of other non-pathogenic strains that followed RCV generated the names RCV-A, with ‘A’ standing for Australia, and recently RCV-E, with ‘E’ standing for Europe [13, 27, 28]. Other names, such as Ashington virus or MRCV for Michigan rabbit calicivirus, have also been put forward for strains of unclear pathogenicity, regardless of their phylogenetic relationship with the other strains [29–31]. Moreover, the new pathogenic variant that emerged in 2010 has been given different names, which has led to some confusion in the field: at first it was named RHDV new variant or RHDVFrance2010 [19, 21, 24, 32]. Then, Dalton and colleagues considered it as a variant, which led them to propose the name RHDVb [24], while Le Gall-Reculé and colleagues considered this entity as a new virus [21], which led them to propose the name RHDV2. Regarding EBHSV, two major genetic groups were identified among the French strains collected between 1989 and 2003 (G1 and G2) and divided into some subgroups [33], whereas Swedish EBHSV strains collected between 1982 and 2008 were divided into two major groups (A and B), the second including the other...
European EBHSV collected since 1989 [34]. In addition, non-pathogenic strains of lagoviruses infecting hares have very recently been described and called HaCV for hare caliciviruses [35] (G. Le Gall-Reculé et al., unpublished results). To further complicate matters, recombination between strains occurs frequently, similar to what has been observed for noroviruses, another genus of the Caliciviridae family responsible for gastroenteritis in humans [36–39]. For all these reasons it is now clear that the historically rooted nomenclature can no longer accommodate new information, and that the use of different names to qualify the same entities is not tenable. A new nomenclature is therefore urgently required. Yet, there are a number of challenges that a sound nomenclature should be able to accommodate.

**Criteria for a new nomenclature**

Quite obviously, the new nomenclature should fit taxonomy. However, the criteria that may be used for classification are diverse. Broadly speaking, these criteria can be either biological or based on relationships between phylogenetically defined subgroups. Of note, in the Caliciviridae family, genera are named according to variable criteria, such as the location of first discovery (Norovirus, Sapovirus), the host species (Lagovirus, Recovirus, Becovirus) or associated symptoms (Vesivirus) [40]. Here we suggest that the new nomenclature of lagoviruses should be based on phylogenetic criteria since there are no sound biological criteria that can stably distinguish all taxa. Indeed, there is no reliable cell culture system available for any lagovirus, hampering the definition of serotypes based on neutralization assays. Cross-protection through in vivo assays can be tested but this is imprecise and only scant data is available [13, 14, 41]. Thus, although immune-related information is of paramount importance to understand virus evolution and patterns of virus circulation within and between species, it cannot be utilized in the nomenclature. Pathogenicity cannot be a reliable criterion either since it is sometimes difficult to define [30] and may be host-dependent [42]. Likewise, the virus names should no longer be based on the host species since species barriers appear more and more porous within the lagomorphs and the host species of origin have become obscure. It was originally thought that RHDV only infected the European rabbit, and that EBHSV only infected the European brown hare (Lepus europaeus). However, mounting evidence indicates that several lagoviruses can cross the host-species barrier, mainly within the Leporidae family (Table 1). Indeed, the ‘classical’ RHDV has been found in Lepus granatensis [43], and the new RHDV (RHDV2 or RHDVb) has been found in Lepus capensis, Lepus corsicanus and recently in Lepus europaeus and Lepus timidus [32, 44–48]. As for EBHSV, infections of other hare species, such as Lepus corsicanus [49] and Lepus timidus [50, 51], and also of a distantly related lagomorph species Sylvilagus floridanus have been described [52]. In addition, one-fifth of Eastern cottontails (Sylvilagus floridanus) freely living in Italy were found to have low anti-RHDV antibody titres, suggesting infection by a related lagovirus (non-pathogenic?).

<table>
<thead>
<tr>
<th>Table 1. Documented leporid host species of pathogenic lagoviruses</th>
</tr>
</thead>
<tbody>
<tr>
<td>RHDV (GL1*)</td>
</tr>
<tr>
<td>-------------</td>
</tr>
<tr>
<td>O. cuniculus</td>
</tr>
<tr>
<td>L. europaeus</td>
</tr>
<tr>
<td>L. granatensis</td>
</tr>
<tr>
<td>L. capensis</td>
</tr>
<tr>
<td>L. corsicanus</td>
</tr>
<tr>
<td>L. timidus</td>
</tr>
<tr>
<td>S. floridanus</td>
</tr>
</tbody>
</table>

*New suggested nomenclature of viruses.†[45–47].

†In Italy, 15–20% of wild animals were reported to have low anti-RHDV antibody titres, suggesting infection by a related lagovirus (non-pathogenic?).

A phylogenetically based new nomenclature

For all these reasons, biological criteria cannot be used as a basis for the new classification of lagoviruses. Therefore, we contend that a new nomenclature should be based on phylogeny and genetic distances, and should be consistent with previously established nomenclatures accepted for similar viruses. The new nomenclature should be robust to the discovery of new viruses, regardless of both their host species and their pathogenicity. In addition, any nomenclature should make suitable provision for recombinant viruses, which raises a major difficulty. Recombinants within the capsid gene have been described between RHDV strains [36, 37, 39]. Yet, multiple recombination events between the region coding for the structural and the nonstructural proteins were recently reported [17, 38]. This has also been observed frequently for noroviruses and the present nomenclature system used for these viruses accommodates this type of recombinant [57].

From the above considerations and the available phylogenetic data for lagoviruses, we propose the taxa organization described in Fig. 1. Since the vast majority of available sequences correspond to the major capsid protein (VP60), taxa organization will be based on major lineages identified in robust phylogenies obtained from alignments of full-
length VP60 gene sequences produced using a maximum-likelihood framework (Fig. 2). Phylogenetic groups on which nomenclature is based should be well-supported, with bootstrap values ≥70 (1000 bootstrap replicates). Suitable minimal and maximal distances used to support phylogeny-based classification will need to be defined, which should be compatible with distances defined for noroviruses. Genetic distances within both the noroviruses and lagoviruses indicate that the proposal should be robust enough to accommodate even genetically quite diverse putative new members of the *Caliciviridae* family.

We propose to classify the taxa according to four primary criteria:

(1) Within the *Caliciviridae* family, viruses that infect *lagomorph* species and belong to the *Lagovirus* genus, including all presently known lagoviruses should belong to a single species. The initial subdivision into two species was based on the clear-cut host-species demarcation between RHDV and EBHSV. As discussed above demarcation is largely blurred and co-infection of hares by the new RHDV variant and EBHSV is becoming likely, making the generation of recombinants plausible. In addition, two species (RHDV and EBHSV) would not be consistent with the criteria used within the *Caliciviridae* family as the level of sequence divergence between lagoviruses is far lower than that observed in either noroviruses or vesiviruses. Indeed, the level of genetic similarity between RHDV and EBHSV is ~70 % at the nucleotide level for the VP60 protein gene. Until recently, all noroviruses were considered as a single species, although Kroneman et al. recently suggested the existence of several species, since GI and GII noroviruses only have about 40 % sequence similarity [57]. To be consistent within the *Caliciviridae* family and to accommodate the now plausible event of recombinations between them, RHDV and EBHSV should therefore belong to the same virus species. Accordingly, the genus name should remain *Lagovirus*. The proposed species name is an historical reminder of the place of initial discovery of the virus. Although EBHS was...
described before RHD, RHDV was characterized before EBHSV. RHDV was correctly characterized as a calicivirus nearly simultaneously by several teams from Czech Republic, Spain, Italy and Germany between 1989 and 1990 [5, 6, 9, 58–60]. Thus, we propose that the species name should become Lagovirus europaeus. This will allow for accommodation of other species of lagoviruses that may exist either within other leporids or within other lagomorphs such as Ochotonidae. It should be clear that the species name europaeus is a reminder of the place where the presently known viruses were first correctly characterized as caliciviruses and not the geographic origin of the virus.

(2) Within this virus species (Lagovirus europaeus), there would be two genogroups recognized at this time. The first would correspond to those viruses related to RHDV, while the second would correspond to viruses related to EBHSV. The genogroups would be called GI and GII, respectively.

(3) Phylogenetic analyses strongly support a further subdivision within the genogroup 1 (RHDV-related viruses), defining genotypes. To define a genotype, we suggest that the genetic distance between two phylogenetic groups should be at least 15 %, with these genotypes characterized by a numeral such as GI.1, GI.2, etc. We are aware that cutoffs based on genetic distances may need revision upon addition of new sequences and can create incompatibilities with the phylogenetic tree because of potential differences within clusters arising from different rates of evolution. That should be controlled when additional phylogenetic groups (novel genotypes) are defined upon acquisition of new sequences. Based on this criterion, within GII (EBHSV-related viruses) there appears to be one valid genotype at present (GII.1) and a putative new genotype (GII.2) presently based on one sequence only. The initial name for this new virus was HaCV [35]. Kroneman et al. proposed a statistically based criterion to distinguish norovirus genotypes where the average distance between all sequences within a newly identified genotype and its nearest established genotype should not overlap within 2 standard deviations of each other [57]. As shown on Fig.

Fig. 2. Maximum-likelihood (ML) phylogenetic tree of the Lagovirus europaeus species. The tree was inferred using all published VP60 coding sequences (N=569, 1743 nucleotides) in MEGA6 [62] with the nucleotide substitution model GTR+G+I, partial deletion (95 % site coverage) for gaps/missing data, and 1000 bootstrap replicates (only bootstrap values≥70 are shown). Current names are shown beside each cluster in parentheses. The proposed new nomenclature is shown on the right-hand side and on each branch for the variants within genotypes. Due to a lack of sequences in some clusters, several strains remain unclassified (and are either noted with a question mark or unclassified; see Table 2). Accession numbers of all sequences used to build the tree are given in Table S1 (available in the online Supplementary Material).
Table 2. Nomenclature concordance

<table>
<thead>
<tr>
<th>Current names</th>
<th>New phylogenetically derived names*, †</th>
<th>Proposed common names</th>
</tr>
</thead>
<tbody>
<tr>
<td>RHDVa</td>
<td>Lagovirus europaeus/GI.1a/....</td>
<td>GI.1a/RHDVa</td>
</tr>
<tr>
<td>RHDV G6</td>
<td>Lagovirus europaeus/GI.1b/....</td>
<td>GI.1b/RHDV</td>
</tr>
<tr>
<td>RHDV G1</td>
<td>Lagovirus europaeus/GI.1c/....</td>
<td>GI.1c/RHDV</td>
</tr>
<tr>
<td>RHDV G2</td>
<td>Lagovirus europaeus/GI.1d/....</td>
<td>GI.1d/RHDV</td>
</tr>
<tr>
<td>RHDV G3/G4/G5</td>
<td>Lagovirus europaeus/GI.2/....</td>
<td>GI.2/RHDV2/b</td>
</tr>
<tr>
<td>RHDVb</td>
<td>Unclassified†</td>
<td>RCV</td>
</tr>
<tr>
<td>RCV</td>
<td></td>
<td>MRCV</td>
</tr>
<tr>
<td>Ashington</td>
<td></td>
<td>RCV</td>
</tr>
<tr>
<td>RCV-E1 or laboratory strain number i.e. 06–11</td>
<td>Lagovirus europaeus/GI.3/....</td>
<td>GI.3/RCV</td>
</tr>
<tr>
<td>RCV-A1(1,2)§</td>
<td>Lagovirus europaeus/GI.4/....</td>
<td>GI.4a/RCV</td>
</tr>
<tr>
<td>RCV-A1(3,4)</td>
<td></td>
<td>GI.4b/RCV</td>
</tr>
<tr>
<td>RCV-A1(5,6)</td>
<td></td>
<td>GI.4c/RCV</td>
</tr>
<tr>
<td>RCV-E2 or laboratory strain number i.e. B09</td>
<td>Lagovirus europaeus/GI.4d/....</td>
<td>GI.4d/RCV</td>
</tr>
<tr>
<td>EBHSV group A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EBHSV group B</td>
<td></td>
<td>.</td>
</tr>
<tr>
<td>EBHSV group B</td>
<td></td>
<td>.</td>
</tr>
<tr>
<td>HaCV</td>
<td>Unclassified†</td>
<td>GI.2/HaCV?</td>
</tr>
</tbody>
</table>

*Comprehensive names as described in the text and in Fig. 1.
†New comprehensive groups within each taxonomic subdivision will be incremented by order of discovery, i.e. GI.1, GI.6, GI.2, GI.1e, ...
‡This group of strains retains current names until acquisition of a sufficient number of strains that will allow inclusion in a phylogenetically supported new subgroup.
§RCV-A1 subtypes 1–6 refer to the variants described in Jahnke et al. [63].
||According to Lopes et al. [34].
¶Subgroup name proposed in Le Gall-Reculé et al. [33].

S1, all our proposed genotypes fulfilled that 2xSD criterion, indicating that they can be included in a consistent classification of the entire *Calicivirus* family. Well-supported subgroups within genotypes may be considered as variants defined by a letter, such as GL1a. Pairwise genetic distances of at least 6% between subgroups accounts for all presently known subgroups defined in Fig. 2.

(4) Finally, it should also be acknowledged that some individual sequences quite diverse from all other sequences in phylogenetic trees may correspond to new genogroups, new genotypes or new variants. However, these should be designated only when at least three independent strains that are not directly linked (i.e. through the same outbreak) will be available. Until then, they should remain unassigned.

With these taxonomic considerations, we propose a two-tier nomenclature. First, there is the need for comprehensive full-length virus names in order to avoid any confusion (complicated, but phylogenetically sound) and second, there remains a need for simpler common names for ease of communication.

The comprehensive names should be based on what has been implemented for noroviruses, picornaviruses and other viruses. It should not contain information regarding pathogenicity since that may not necessarily match phylogenetic relationships. In addition, we should anticipate species-barrier crossings that may result in widely different pathogenicity depending on the host.

GenBank records should have the genus and species name in the ‘Organism’ field:

**Lagovirus europaeus**

The strain name should be written as follows:

- Genogroup, genotype (i.e. GI.1).
- The species from which the virus was first detected should be given and it should be the Latin name to avoid species confusion. Thus *O cun* stands for *Oryctolagus cuniculus*, *L eur* for *Lepus europaeus*, *L gra* for *Lepus granatensis*, *S flo* for *Sylvilagus floridanus*, etc.
- The country where the viral sequence was obtained should be mentioned in an abbreviated form using ISO codes (www.iso.org/iso/home.html), i.e. FR for France, PT for Portugal, IT for Italy, ES for Spain, DE for Germany, AUS for Australia, etc.
- The year of isolation should follow.
- Finally, it would be terminated by the identification of the strain from the laboratory that submitted it.

An example of a full name for a lagovirus would thus be as follows:

*Lagovirus europaeus/GI.1d/O cun/FR/2003–24*
We propose that this comprehensive name should be given in the materials and methods sections of articles and for sequences deposited in data banks. Nevertheless, as noted by Kuhn et al. [61], a short common name could be used further in the text for convenience. In the above example, it would be GL.Id/03–24, indicating the genotype and strain. Here the variant denomination (d) has been introduced for convenience. Yet, as discussed above, this taxonomic level is strongly subject to modification upon discovery of new sequences. Taking into account non-virologist stakeholders, such as farmers, hunters and wildlife managers requires that common names relate as easily as possible with current names to facilitate transition. Thus, the common name may include a historical reminder. In the above example, it could be GL.Id/RHDV, corresponding to a generic name for a variant, not for a strain. For new variants that will not correspond to any of the historical names, the new short nomenclature followed by the first place of isolation could be used: for example GI.6/Porto. A table of correspondence between current names and the proposed new nomenclature is given in Table 2.

In addition, recombinants must be considered. Most recombination events between noroviruses occur between the polymerase and the capsid coding regions and this has also been observed in the Lagovirus genus [38]. Similar to the proposal of Kroneman et al. [57], identification of the recombinants may appear in the full name as follows for example:

- Lagovirus europaeus/GL.3P-GI.1a/O cun/FR/2005/05–12, indicating a recombinant between a GI.3 and a GI.1a strain; P standing for polymerase.

Since recombinants within the VP60 sequence have been described [36, 37, 39], it will have to be seen how this affects classification on a case-by-case basis before deciding on how that should be taken into account in the nomenclature.

In conclusion, we have defined here a new classification and nomenclature system for lagoviruses based on phylogeny, which sets guidelines for the field. The lagovirus working group will periodically monitor new sequences submitted to the public domain that may impose modifications of the minimal and maximal distances between phylogenetic levels or the addition (or removal) of taxa. The working group should also communicate the agreed nomenclature changes when necessary.

**Funding information**

This work was supported in part by a grant from the Agence Nationale de la Recherche (France). Cailloco and by a grant from the Région des Pays de la Loire (France) Armimo to JLP. It was performed within the framework of the ECALEP project selected during the 2nd joint call of the Animal Health and Welfare ERA-Net (Anihwa) initiative, a Coordination Action funded under the European Commission’s ERA-Net scheme within the Seventh Framework Programme (Contract No. 291815). The ECALEP project is funded by the Agence Nationale de la Recherche (France), the Ministry of Health, Depa for Veterinary Public Health, Nutrition and Food Safety (Italy) and the Research council FORMAS (Sweden). FCT-Foundation for Science and Technology, Portugal, supported the FCT Investigator grant of JA (ref. IF/01396/2013) and the Post-doc grant of AML (SFRH/BDP/115211/2016).

**Conflicts of interest**

The authors declare that there are no conflicts of interest.

**References**


---

Five reasons to publish your next article with a Microbiology Society journal

1. The Microbiology Society is a not-for-profit organization.
2. We offer fast and rigorous peer review – average time to first decision is 4–6 weeks.
3. Our journals have a global readership with subscriptions held in research institutions around the world.
4. 80% of our authors rate our submission process as ‘excellent’ or ‘very good’.
5. Your article will be published on an interactive journal platform with advanced metrics.

Find out more and submit your article at microbiologyresearch.org.