1 **Frequent intergenotypic recombination between the non-structural and structural**

2 **genes is a major driver of epidemiological fitness in caliciviruses**

3

4 Jackie E Mahar[1,2], Maria Jenckel[2], Nina Huang[2], Elena Smertina[2,4], Edward C Holmes[1],

5 Tanja Strive[2,3], Robyn N Hall[2,3*]

6   1. Marie Bashir Institute for Infectious Disease and Biosecurity, School of Life and

7   Environmental Sciences and School of Medical Sciences, The University of Sydney,

8   Sydney, NSW 2006, Australia.

9   2. Commonwealth Scientific and Industrial Research Organisation, Health and Biosecurity,

10   Black Mountain, ACT 2601, Australia.

11   3. Centre for Invasive Species Solutions, University of Canberra, Bruce, ACT 2617,

12   Australia.

13   4. Faculty of Science and Technology, University of Canberra, Bruce, ACT 2617, Australia.

14 *Corresponding author: robyn.hall@csiro.au; +61 2 6246 4245

15

16 Total word count: 6910

17

18 Keywords: calicivirus, recombination, genomic epidemiology, lagovirus, RHDV2

19 **Abstract**

20 The diversity of lagoviruses (*Caliciviridae*) in Australia has increased considerably. By

21 the end of 2017, five variants from three viral genotypes were present in populations of

22 Australian rabbits, while prior to 2014 only two variants were known. To understand the

23 interactions between these lagovirus variants we monitored their geographical

24 distribution and relative incidence over time through a landscape-scale competition

25 study, and from this, revealed potential drivers of epidemiological fitness. Within three

26 years of the arrival of GI.1bP-GI.2 (RHDV2) into Australia, we observed the emergence

27 of two novel recombinant lagovirus variants, GI.4eP-GI.2 (4e-recombinant) in New

28 South Wales and GI.4cP-GI.2 (4c-recombinant) in Victoria. Although both novel

29 recombinants contain the non-structural genes from benign, rabbit-specific, enterotropic

30 viruses, these variants were recovered from the livers of both rabbits and hares that had

31 died acutely. This suggests that determinants of host and tissue tropism for lagoviruses

32 are associated with the structural genes, and that tropism is intricately connected with

33 pathogenicity. Phylogenetic analyses demonstrated that the 4c-recombinant emerged

34 independently on multiple occasions, with five distinct lineages observed. Both new

35 recombinant variants replaced the previous dominant parental RHDV2 in their

36 respective geographical areas, despite sharing an identical or near-identical (i.e., single

37 amino acid change) major capsid protein with the parental virus. This suggests that

38 epidemiological fitness of these recombinants was not driven by antigenic variation in

39 the capsid, implicating the non-structural genes as key drivers of epidemiological

40 fitness. Molecular clock estimates place the GI4.e recombination event in early to mid-

41 2015, while the five GI.4c recombination events occurred from late 2015 through to

42 early 2017. The emergence of at least six viable recombinant variants within a two-year

43 period highlights an unprecedented frequency of these events, detectable only due to

44 intensive surveillance, and demonstrates the importance of recombination in lagovirus

45 evolution.

46

47 **Introduction**

48 Caliciviruses are an important group of vertebrate-infecting viruses. They include

49 noroviruses, the major cause of gastroenteritis in humans worldwide [1], and members

50 of the genus *Lagovirus*, which infect rabbits and hares. Some lagoviruses are

51 hepatotropic and cause an acute fulminant viral hepatitis with a case fatality rate

52 exceeding 90%, while others are enterotropic and are thought to be entirely benign [2].

53 These are referred to as rabbit haemorrhagic disease viruses (RHDVs) and rabbit

54 caliciviruses (RCVs), respectively.

55 Lagoviruses are hierarchically classified by their major capsid protein (VP60) type, and

56 less frequently by polymerase type, into genogroups (e.g. GI, GII), genotypes (e.g. GI.1,

57 GI.2, GI.4), and variants (e.g. GI.1a, GI.1b, GI.1c) [3]. The first lagoviruses identified,

58 from hares in Europe in the early 1980s, were those of the GII.1 genotype (EBHSV) [4].

59 Subsequently, mortality events in *Oryctolagus* rabbits in China in 1984 led to the

60 discovery of genotype GI.1 (RHDV1) viruses [5]. With an increasing diversity of GI.1

61 viruses, this genotype was further subdivided based on the VP60 phylogeny into several

62 variants. In 1997, the first antigenic variants of GI.1 viruses, now classified as GI.1a

63 (RHDVa), were described in Italy [6] and these spread throughout Europe and more

64 distantly. Then, in 2010, the GI.2 (RHDV2) variant first emerged in France [7-9]. RHDV2

65 rapidly spread globally, triggering epizootics worldwide in wild and domestic lagomorph

66 populations. As surveillance efforts and technologies improved, non-pathogenic GI.3

67 (RCV-E1) and GI.4 (RCV-A1, RCV-E2) lagoviruses were identified in the 1990s, first in

68 Europe [10-13] and then in Australia [14] and New Zealand [15]. Non-pathogenic hare

69 lagoviruses (GII.2, GII.3, GII.4, GII.5; HaCV) have also been reported more recently, in

70 2014 in Europe [16-18] and 2019 in Australia [19].

71 Lagoviruses, like other caliciviruses, are small, non-enveloped viruses containing a

72 monopartite positive-sense single-stranded RNA genome approximately 7.5 kb in length

73 (Figure 1) [2]. Virus particles contain both genomic RNA (gRNA) and 3' co-terminal sub-

74 genomic RNAs (sgRNA) approximately 2.5 kb in length [20]. In lagoviruses, the non-

75 structural (NS) genes are situated upstream of the major capsid protein, VP60, all of

76 which are encoded as a single large polyprotein (ORF1), while the 3' terminal ORF2

77   encodes a minor structural protein that is presumed to be important for genome release

78   during infection [21]. Lagoviruses encode seven NS proteins, including a helicase (2C-

79   like), a viral genome-linked protein (VpG), a protease (Pro), an RNA-dependent RNA

80   polymerase (RdRp), and three proteins of unknown function (p16, p23, p29). Both

81   gRNA and sgRNA are linked to VpG at their 5' end and are polyadenylated at their 3'

82   end [22]. There is considerable homology between the 5' terminal nucleotide sequences

83   of the gRNA and sgRNA (Figure 1) [22], facilitating recombination at the junction

84   between the RdRp and VP60; this region is highly conserved within and between

85   genogroups.

86

87
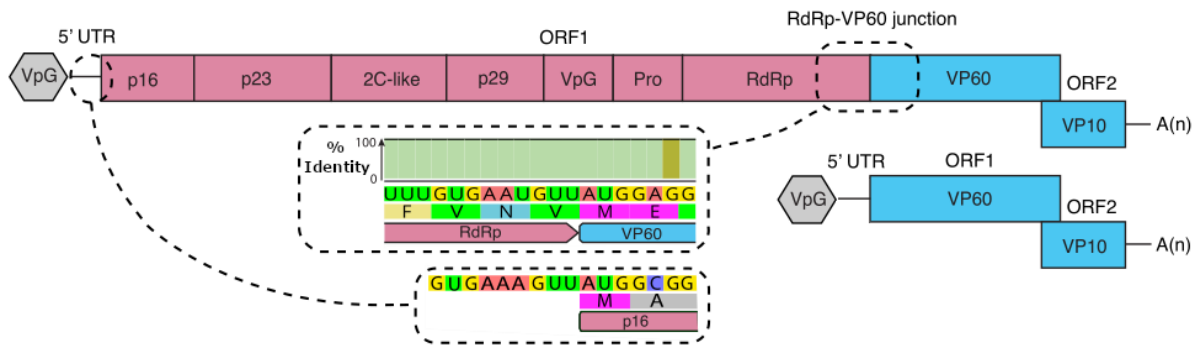
88   **Figure 1: Lagovirus genome organisation.** Genomic RNAs (gRNA) encode the non-

89   structural proteins (pink) and structural proteins (blue), while subgenomic RNAs

90   (sgRNA) encode the structural proteins only. Both RNAs are linked to VpG at their 5'

91   end and are polyadenylated at their 3' end. Inset boxes demonstrate the nucleotide

92   sequence identity between the 5' untranslated region (UTR) of the gRNA and the RdRp-

93   VP60 junction. An identity plot (based on n = 478 near-complete GI sequences from

94   GenBank; green indicates 100% identity) is also shown for the junction region.

95

96   Recombination is an important evolutionary mechanism in many RNA viruses [23]. In

97   the *Caliciviridae*, recombination frequently occurs at the junction between the NS and

98   structural genes, effectively mixing a set of structural genes with an entirely new set of

99   NS genes [24-26]. Recombinant lagoviruses are defined by the nomenclature [RdRp

100   genotype]P-[capsid genotype], and can include combinations of two pathogenic RHDVs

101   (e.g. GI.1bP-GI.2) as well as the seemingly more common combination of a benign

102   RCV and a pathogenic RHDV (e.g. GI.4eP-GI.1a, GI.4eP-GI.2, GI.3P-GI.2) [24, 27-32].

103   Retrospective phylogenetic analyses have demonstrated that all GI.2 viruses so far

104   described are recombinants (GI.3P-GI.2), implying that GI.2 is an orphan capsid-type

105   [24]. Intergenogroup recombinants between GI (rabbit and hare) and GII (hare) viruses

106   have also been reported, with two GII.1P-GI.2 viruses recovered from hares in Germany

107   in 2014 and 2019 [33].

108   Prior to 2016, three lagovirus genotypes (GI.1, GI.2, GI.4) from the G.I genogroup,

109   comprising five distinct variants — GI.4 (RCV-A1), GI.1c (RHDV1), GI.1a (RHDVa-K5),

110   GI.4eP-GI.1a (RHDVa-Aus), GI.1bP-GI.2 (RHDV2) — had been reported in Australia.

111   These variants differ in their host and tissue tropism and in pathogenicity [34]. RCV-A1

112   is a benign enterotropic virus that has circulated in wild and domestic rabbits, likely

113   since at least the 1950s [14, 35]. RHDV1 and RHDVa-K5 are pathogenic viruses, both

114   deliberately released [36]. RHDVa-Aus and RHDV2 were both exotic incursions, first

115   detected in January 2014 and May 2015, respectively [32, 37]. Phylogenetic analyses

116   suggest that both incursions arose from single point source introductions with

117   subsequent ongoing transmission in Australian rabbit populations [32, 34]. In mid-2017

118   a novel GI.4eP-GI.2 recombinant virus (4e-recombinant) was detected, comprising the

119   NS sequences of the RHDVa-Aus virus and RHDV2 VP60 sequences [31]. However,

120   the long-term epidemiological significance of this recombinant variant was unclear.

121   Australia is an ideal setting for understanding the evolution, recombination, and

122   epidemiological interactions of caliciviruses. Australia is an isolated landmass with a

123   large rabbit population that is distributed across most of the country. The genetic

124   diversity of Australian lagoviruses arose from a small number of introduction (or

125   incursion) events, and the parental sequence as well as the timing of the introduction is

126   known in several instances. Since there is little movement of rabbits into the country,

127   the study of lagoviruses can occur in a relatively controlled environment with a defined

128   number of pre-existing pathogenic viruses. Additionally, a robust surveillance system for

129   these viruses has been in place since 2014. The previous lack of diversity of the

130 Australian lagovirus population and sudden diversification of lagovirus variants in

131 2014/2015 created the perfect natural experimental conditions for monitoring the

132 frequency of successful recombination and drivers of viral evolution. In this study we

133 monitored the epidemiological and evolutionary dynamics of existing lagovirus variants

134 in Australia, while determining the frequency and importance of successful

135 recombination events between these viruses.

136

137 **Materials and Methods**

138 *Sample submission*

139 Samples from dead domestic and wild lagomorphs (i.e., rabbits and hares) were

140 submitted either directly or via RabbitScan (https://www.feralscan.org.au/rabbitscan/) to

141 the Commonwealth Scientific and Industrial Research Organisation (CSIRO) by

142 veterinarians, pet owners, landholders, and members of the public as part of ongoing

143 opportunistic lagovirus surveillance. No animal ethics approvals are required for

144 sampling rabbits that are found dead in Australia. Samples were provided either fresh-

145 frozen or stored in an RNA stabilization solution [31]. Additionally, 42 positive samples

146 submitted to the Elizabeth Macarthur Agricultural Institute (EMAI) for RHDV diagnostic

147 testing were forwarded to CSIRO, either as fresh-frozen tissue or as a swab of the cut

148 tissue stored in phosphate buffered gelatin saline (PBGS) [38]. Samples for this study

149 were restricted to those collected between 1 January 2014 (based on the estimated

150 time of incursion of GI.1bP-GI.2 into Australia [34, 39]) and 30 September 2020, and

151 were received from the Australian Capital Territory (ACT, n = 139), New South Wales

152 (NSW, n = 401), the Northern Territory (NT, n = 7), Queensland (QLD, n = 86), South

153 Australia (SA, n = 227), Tasmania (TAS, n = 221), Victoria (VIC, n = 649), and Western

154 Australia (WA, n = 194) (total n = 1,924).

155 *Initial testing*

156 RNA was extracted from 20 – 30 mg of tissue, predominantly liver or bone marrow, but

157 occasionally spleen, kidney, muscle, ear tip, eyelid, or blowfly maggots retrieved from

158 the carcass. After homogenisation with glass beads using a Precellys 24-dual tissue

159   homogenizer (Bertin Technologies), RNA was extracted using either the Maxwell 16

160   LEV simplyRNA tissue kit (Promega, Alexandria, NSW) or RNeasy mini kit (Qiagen,

161   Chadstone Centre, VIC) as per manufacturers' directions. For swab samples, RNA was

162   extracted from 200 µl PBGS using the Purelink viral RNA/DNA mini kit (ThermoFisher

163   Scientific). Samples were screened for known lagoviruses using a broadly-reactive

164   lagovirus SYBR-green based RT-qPCR targeting VP60 and sgRNA, as described

165   previously [31].

166   *Variant identification*

167   Of the samples that were positive by screening RT-qPCR (n = 1,209), 175 were

168   sequenced previously for other studies [31, 32, 34, 37, 40]. The remaining samples

169   were screened for recombination by RT-PCR and sequencing. RT-PCR primers were

170   designed that spanned the RdRp-VP60 junction, generating a 555 bp amplicon (Rec2

171   RT-PCR; Supplementary table 1). Primers were subsequently modified by the 5'

172   addition of Illumina pre-adapter and index primer binding sequences to enable high-

173   throughput Illumina sequencing (Rec2.tailed RT-PCR; Supplementary table 1). Assays

174   were experimentally validated to confirm reactivity with all known Australian lagoviruses.

175   For a subset of samples (prior to the development of the Rec2 RT-PCR), amplification

176   over the RdRp-VP60 junction was performed with alternative primer sets

177   (Supplementary table 1). Briefly, each 25 µl PCR reaction comprised 1x OneStep

178   Ahead RT-PCR master mix (Qiagen), 1x OneStep Ahead RT mix (Qiagen), 0.5 µM each

179   primer, and 1 µl of extracted RNA diluted 1/10 in nuclease-free water. Cycling

180   conditions were: 45 °C for 15 min, 95 °C for 5 min, followed by 40 cycles of 95 °C for 15

181   sec, 55 °C for 15s, 68 °C for 2 min, with a final extension of 68 °C for 5 min.

182   Amplicons were purified and either sent for Sanger sequencing at the Biomolecular

183   Resource Facility, The John Curtin School of Medical Research, Australian National

184   University, or were indexed, pooled, and sequenced on an Illumina MiSeq (300-cycle v2

185   chemistry) according to the manufacturer's directions. After quality trimming, sequences

186   were partitioned into regions upstream and downstream of the RdRp-VP60 junction.

187   Each partition was aligned with representative sequences of known Australian lagovirus

188   variants using MAFFTv7.450 [41] in Geneious Prime 2020.0.4, and an RdRp and VP60

189   variant type was assigned for each sample. Four isolates were identified as mixed

190   infections based on sequencing. Alignments of partial sequences are available at

191   https://doi.org/10.25919/758f-4t15.

Molecular epidemiological analyses

193   To investigate the geographical distribution and relative incidences of recombinant

194   variants over time, positive samples were restricted to those reported after 1 January

195   2016; i.e., 6 months prior to the first detection of a novel recombinant (n = 1,139).

196   Samples identified as RHDVa-K5 (n = 121) were excluded from further analyses since

197   these were mostly associated with release sites of this approved biocontrol virus; that is,

198   this variant was not transmitting through rabbit populations and was therefore not

199   competing epidemiologically with other virus variants [39]. Samples from the ACT were

200   grouped with those from NSW, since the ACT is a small jurisdiction (2,358 km^2) located

201   within NSW and was considered epidemiologically to be part of NSW.

202   For the 4c-recombinant lineage designation, 169 4c-recombinant sequences spanning

203   the RdRp-VP60 junction were aligned with 77 4c-recombinant full genomes and 26

204   publicly available GI.4 genomes using MAFFTv7.450 [41]. The taxa included in the final

205   data set are detailed in Supplementary table 2. The alignment was trimmed to

206   nucleotide positions 5,115 – 5,619, based on the M67473.1 DEU/FRG/1988.50

207   reference sequence (https://www.ncbi.nlm.nih.gov/genbank/). The remaining 109 4c-

208   recombinant sequences obtained in this study were not suitable for inclusion in the final

209   alignment because they were generated either using Illumina sequencing or an

210   alternative primer set and thus did not span the complete region of the alignment. Model

211   selection and maximum likelihood (ML) tree inference was conducted using IQ-

212   TREEv1.6.11 [42], and branch supported was estimated by 1,000 replicates of the SH-

213   like approximate likelihood ratio test [43]. The tree was rooted on the internal branch

214   leading to the RCV-A1 lineage.

Full genome sequencing

216   Of the 1,034 samples genotyped to the variant level, a subset of 224 were selected for

217   full genome sequencing (Supplementary table 3). These samples were selected to be

218 geographically and temporally representative, with a focus on the novel 4e- and 4c-

219 recombinants. Viral genomes were amplified in overlapping fragments, and DNA

220 libraries were prepared and sequenced using Illumina Miseq technology as described

221 previously [35, 40]. Primers used for amplification of the overlapping fragments are

222 detailed in Supplementary table 1. Consensus sequences were constructed by mapping

223 cleaned reads to the lagovirus GI reference sequence (GenBank accession M67473.1

224 DEU/FRG/1988.50) using Geneious Prime 2020.0.4. Sequences were deposited in

225 GenBank under accession numbers MW460205 – MW460242.

Recombination analysis

227 To further characterise the putative novel 4c-recombinant variant, and the newly

228 sequenced GI.4e-recombinants, recombination analyses were conducted on full

229 genome sequences using the RDP4 software [44]. The complete genome alignment

230 included lagovirus GI potential parent sequences from GenBank

231 (https://www.ncbi.nlm.nih.gov/genbank) (n = 384, 7,309 nt, Supplementary table 4).

232 Recombination was reported if detected by at least two of three primary scanning

233 methods (RDP, MaxChi and GENECONV), with a highest acceptable p-value of 0.05

234 with Bonferroni multiple comparison correction. The BootScan, 3Seq, CHIMAERA, and

235 SciScan methods were used to verify recombination events identified using the primary

236 methods.

Phylogenetic analyses

238 Newly sequenced lagovirus genomes were aligned with representative GI lagovirus

239 sequences from GenBank (https://www.ncbi.nlm.nih.gov/genbank) using MAFFTv7.271

240 [41]. The complete genome alignment was split into two data sets representing (i) the

241 NS genes (n = 240 sequences, 5,231 nt, Supplementary table 5), and (ii) the VP60

242 structural gene (n = 272, 1,740 nt, Supplementary table 6). An additional VP60

243 alignment was constructed that contained the VP60 of the newly sequenced viruses

244 along with all published Australian RHDV2 VP60 sequences (n = 332, 1,737 nt,

245 Supplementary table 7). Model selection and ML tree estimation was performed using

246 IQ-TREEv1.6.11 [42], as described above. Ancestral state reconstruction to identify

247 nonsynonymous changes along individual branches was carried out using PAMLv4.9 [45].

248   *Phylogeographic analyses*

249   A Bayesian Markov chain Monte Carlo (MCMC) approach was employed to infer time-

250   scaled phylogenies, and from this, to infer the temporal pattern and most likely

251   geographic location of internal nodes (i.e., phylogeography). All data sets were initially

252   screened using TempEst [46] to ensure that sufficient temporal signal was present,

253   using ML phylogenies as input to construct linear regressions of root-to-tip genetic

254   distances against sampling time. Due to recombination between the NS and capsid

255   genes, two data sets either side of the recombination breakpoint were analysed

256   separately: (i) a VP60 gene data set containing all published Australian GI.2 capsid

257   sequences together with the capsid genes of the recombinants sequenced here (n =

258   332, 1,737 nt, Supplementary table 7); and (ii) an NS genes data set of all published

259   Australian GI.4 sequences together with the NS genes of the recombinants sequenced

260   here (n = 188, 5,218 nt, Supplementary table 8). The Bayesian Evolutionary Analysis by

261   Sampling Trees (BEAST) software v1.8 [47] was used to conduct Bayesian MCMC

262   analysis of each data set, using substitution models inferred using ModelFinder as

263   implemented in IQ-TREEv1.6.11 [42] (VP60, SYM+G4; NS, SYM+I+G4) [48]. A discrete

264   traits partition indicating the sampling location (Australian state) was included to

265   facilitate ancestral state reconstruction of location (utilizing the symmetric substitution

266   model, inferring social network with Bayesian Stochastic Search Variable Selection, and

267   a strict clock model). All analyses were run twice to convergence (defined as an

268   effective sample size >200) to confirm consistency. Marginal likelihood estimation using

269   path sampling/stepping-stone sampling was used to assess the most appropriate clock

270   model prior (strict vs uncorrelated log-normally distributed [UCLD]) and tree prior

271   (Gaussian Markov random field Bayesian skyride model vs constant size coalescent vs

272   exponential coalescent) for the nucleotide partition.

273   The NS data set had a better temporal signal (R-squared of 0.86 on linear regression of

274   root-to-tip genetic distance against sampling time) and a larger sampling window

275   compared to the VP60 data set and was therefore considered the most informative. The

276   UCLD clock model and constant population size coalescent tree prior had the highest

277   marginal likelihood for this data set.

278   *Figures*

279   Figures were constructed using R v4.0.3 [49] using the following packages: 'ggtree

280   v2.2.1' [50], 'treeio v1.12.0' [51], 'scales v1.1.1' [52], 'tidyverse v1.3.0' [53], 'ape v5.4'

281   [54], 'phytools v0.7-47' [55], 'phangorn v2.5.5' [56], 'cowplot v1.0.0.9000' [57], 'lubridate

282   v1.7.8' [58], 'plyr v1.8.6' [59], and 'ozmaps v0.3.6' [60]. Figure 3B was generated in

283   GeneiousPrime 2020.2.4 using 3-D structure viewer.

284

285   **Results**

286   *Phylogenetic analyses reveal five independent recombination events in GI.4cP-GI.2*

287   *viruses*

288   A recombinant lagovirus variant (GI.4eP-GI.2; 4e-recombinant) was previously detected

289   in Australia using a multiplex RT-PCR assay. This assay is only able to detect specific

290   recombinant variants [31]. Therefore, to comprehensively screen for lagovirus

291   recombination events we sequenced the RdRp-VP60 junction region (i.e., the typical

292   recombination breakpoint) of 1,034 lagoviruses collected between January 2014 and

293   September 2020 (inclusive) from wild and domestic lagomorphs found dead across

294   Australia. This screen identified six distinct lagovirus variants circulating in Australia

295   during the study period: four known variants (GI.1cP-GI.1c (RHDV1), GI.1bP-GI.2

296   (RHDV2), GI.1aP-GI.1a (RHDVa-K5), GI.4eP-GI.1a (RHDVa-Aus)), the previously

297   reported 4e-recombinant (GI.4eP-GI.2), and a novel GI.4cP-GI.2 putative recombinant

298   (4c-recombinant).

299   We selected a subset of novel 4c-recombinant viruses, recently emerged 4e-

300   recombinant viruses, and parental RHDV2 viruses for full genome sequencing, selecting

301   isolates that were representative of the temporal and geographical spread of these

302   variants. Recombination analyses of newly sequenced 4e-recombinant viruses

303   confirmed these as recombinants with a breakpoint at the RdRp-VP60 junction,

304   supporting our previous finding (Table 1, Supplementary table 4). Recombination

305   analyses of the novel putative 4c-recombinant viruses also detected a recombination

306   breakpoint at the typical RdRp-VP60 junction (Table 1). The most likely parental

307   variants for the 4c-recombinant, as determined by the RDP4 software, were benign

308   Australian GI.4 (RCV-A1) lagoviruses in the NS region (GenBank accession KX357699)

309   and Australian GI.2 (RHDV2) viruses in the VP60 region (GenBank accession

310   MF421577), strongly suggesting that this recombination event occurred within Australia

311   after the arrival of GI.2 in late 2014.

312   Subsequently, the nucleotide alignment was partitioned either side of the putative

313   breakpoint and ML phylogenetic trees were inferred for the NS genes and the VP60

314   gene separately. There was clear incongruence between the NS and VP60

315   phylogenies, supporting the results of the RDP4 analysis (Figure 2). Strikingly, ML

316   phylogenies based on the VP60 region revealed five distinct lineages of the 4c-

317   recombinant variant, suggesting multiple independent emergences of this variant

318   (Figure 3A). In contrast, the NS gene sequences from this variant formed a

319   monophyletic group (collapsed in Figure 2A). All 4e-recombinant sequences fell within a

320   single clade in both trees (Figure 2).

321

322

323 **Table 1. Recombination analyses**

Recombinant details		Putative parent^		P-value for recombination event by detection method		
Representative Recombinant Sequence	Breakpoint Position*	NS Parent	S parent	RDP	Maxchi	GENE-CONV
GI.4eP-GI.2/O.cun*/AUS/ NSW/2017-03-17/CAR-8	5306	KY628315/ GI.4eP-GI.1a	MF421585/ GI.1bP-GI.2	4E-194	3E-43	3E-189
GI.4cP-GI.2/O.cun/AUS/ VIC/2017-02-05/ALV-1	5307	KX357699/ GI.4c	MF421577/ GI.1bP-GI.2	3E-148	2E-39	1E-154

324 A more detailed version of this table can is supplied as Supplementary table 4.

325 *Nucleotide position of putative breakpoint according to M67473.1/DEU/FRG/1988.50 reference sequence
326 numbering.

327 ^NS refers to putative parent providing the non-structural genes; S refers to putative parent providing the structural
328 genes.

329
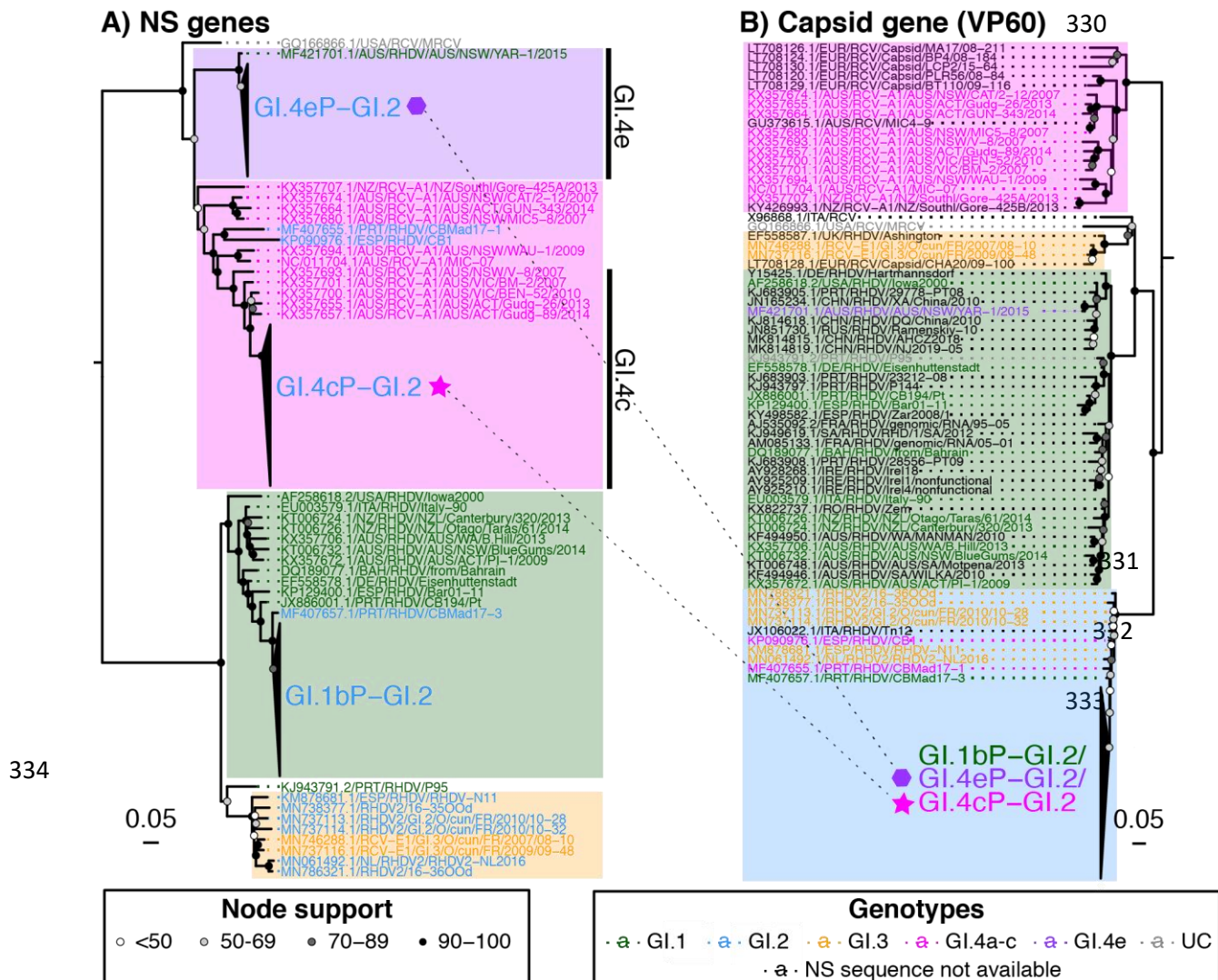
Figure 2. ML phylogenetic trees of the (A) NS genes and (B) the VP60 gene of representative lagovirus sequences and the Australian recombinant sequences. Discrepancies between the highlight and text colour indicate recombinant sequences. Clades are highlighted according to the genotype of the genetic segment analysed in the tree, while taxa names are coloured by the genotype of the alternative segment (i.e., taxa names in the NS gene tree are coloured by the genotype of the capsid gene, and vice versa). Taxa names coloured black indicate that the NS sequence data is unavailable. UC, unclassified. Major clades have been collapsed for visualization purposes. All Australian recombinant sequences from this study fall within the collapsed clades indicated by the star (GI.4cP-GI.2) or hexagon (GI.4eP-GI.2) symbols. The dotted lines between the two trees link the Australian recombinant clades. Phylogenies

346   were midpoint rooted and the scale bars represent the number of substitutions per site.

347   Bootstrap support values (1,000 replicates) are indicated by shaded circles at the

348   nodes.

349

350
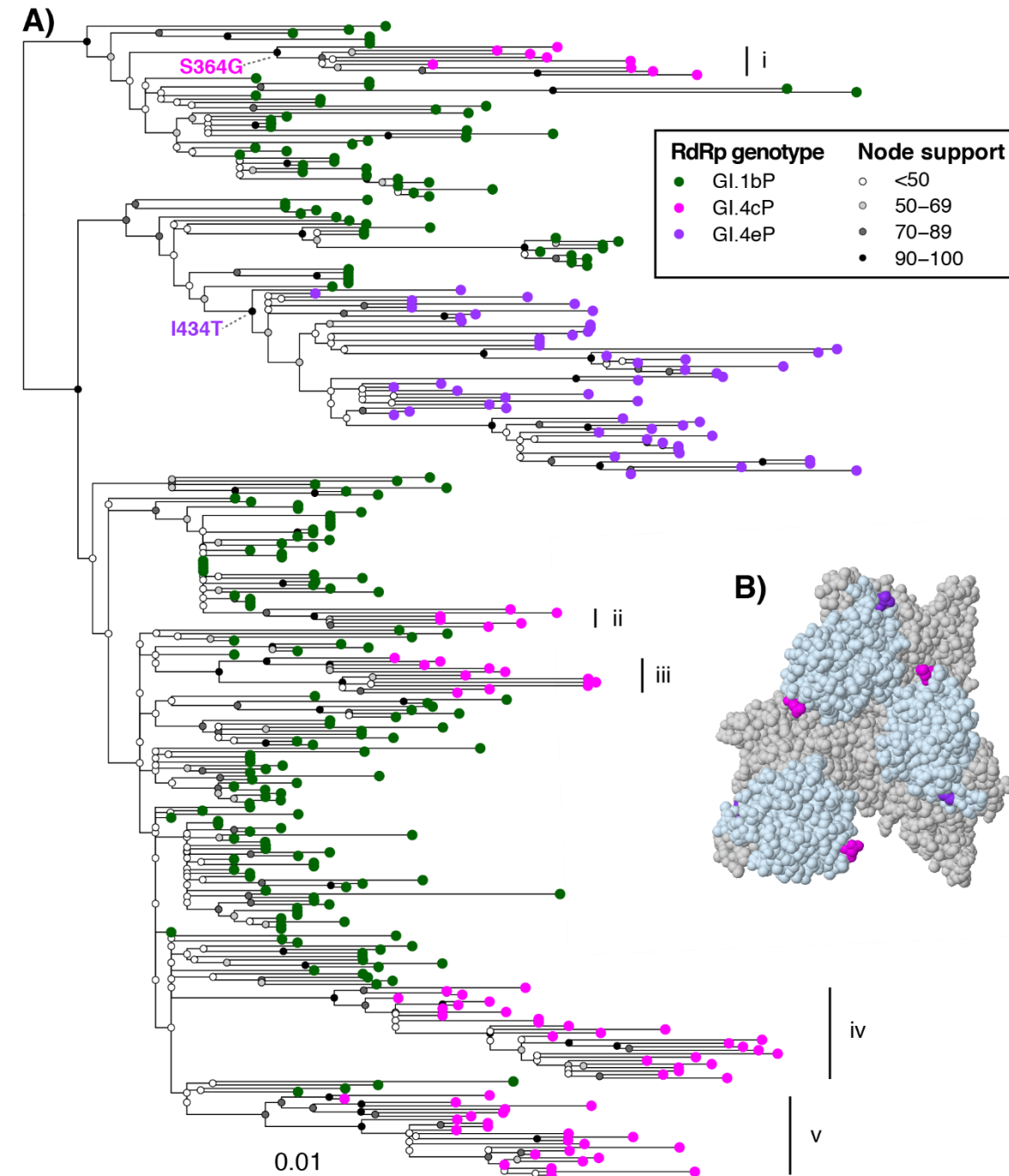
351 **Figure 3**: **ML phylogenetic tree of Australian GI.2 VP60 sequences and amino acid**

352 **changes associated with the emergence of 4c- and 4e-recombinant variants.** (A) A

353 phylogeny was inferred from the VP60 gene sequences (n = 332) of newly sequenced

354 Australian GI.2 lagoviruses (including recombinants) along with representative

355 published sequences. Coloured tip points represent the polymerase (RdRp) genotype:

356 GI.1bP, green; GI.4cP, pink; GI.4eP, purple. GI.4cP lineages are labelled (i through v).

357 The phylogeny was midpoint rooted and the scale bar represents nucleotide

358 substitutions per site. Non-parametric bootstrap support values (1,000 replicates) are

359 indicated by shaded circles at the nodes. Associated metadata including taxon names

360 are provided in Supplementary table 7. Ancestral state reconstruction to identify

361 nonsynonymous changes associated with internal divergence events was conducted

362 using PAML. All changes occurring on branches leading to the GI.4cP and GI.4eP

363 lineages are labelled at the relevant node. (B) These changes were mapped to an

364 asymmetric unit of the atomic model of lagovirus GI VP60 (PDB accession: 3J1P). The

365 P2 domain is coloured in light blue. Amino acid 364, associated with the change in 4c-

366 recombinant lineage i, is coloured pink and 434, associated with the 4e-recombinant, is

367 coloured purple.

368

369 *Rapid epidemiological replacement of parental RHDV2 by the newly emerged 4c- and*

370 *4e-recombinants*

371 Detections of the new 4c-recombinant variant consistently outnumbered those of the

372 parental RHDV2 in VIC by April 2018, 14 months after the first detection in Alvie, VIC,

373 on 5 February 2017, suggesting that this variant was outcompeting previously

374 circulating viruses (Figure 4). This pattern of emergence and replacement was also

375 observed in TAS (Figure 4). This variant was detected sporadically in NSW/ACT from

376 October 2017 where it cocirculated with the locally dominant 4e-recombinant; notably,

377 the proportional incidence of the 4c-recombinant variant has increased in this region

378 during 2020. Interestingly, a single detection of the 4c-recombinant was identified in WA

379 in November 2018 in two domestic rabbits recently imported from TAS as breeding

380 stock. The sequence of this 4c-recombinant virus detected in WA clearly nested within

381   the genetic diversity found in TAS (see Figure 4A, clade v), suggesting a single

382   incursion into WA with no ongoing local transmission. During the study period, the novel

383   4c-recombinant was recovered from several juvenile animals less than 800 g

384   bodyweight and as young as 4 weeks old, demonstrating pathogenicity in young rabbits,

385   like all other GI.2 variants characterized so far. It was also identified in one European

386   brown hare (*Lepus europaeus*).

387   We also examined the distribution of the previously reported 4e-recombinant variant

388   [31]. Most detections (n = 140) were identified in NSW/ACT. The earliest detection was

389   in Tubbul, NSW, on 13 July 2016, and by March 2017 it had mostly replaced the

390   parental RHDV2 to become the dominant lagovirus variant in this region (Figure 4B).

391   More recently, in late 2019 and throughout 2020, it has cocirculated with the 4c-

392   recombinant in NSW/ACT. Although most detections were in adult rabbits, the 4e-

393   recombinant was also recovered from three European brown hares and from several

394   juvenile rabbits down to 250 g bodyweight, suggesting a similar host tropism to the

395   parental RHDV2 virus and the novel 4c-recombinant.

396
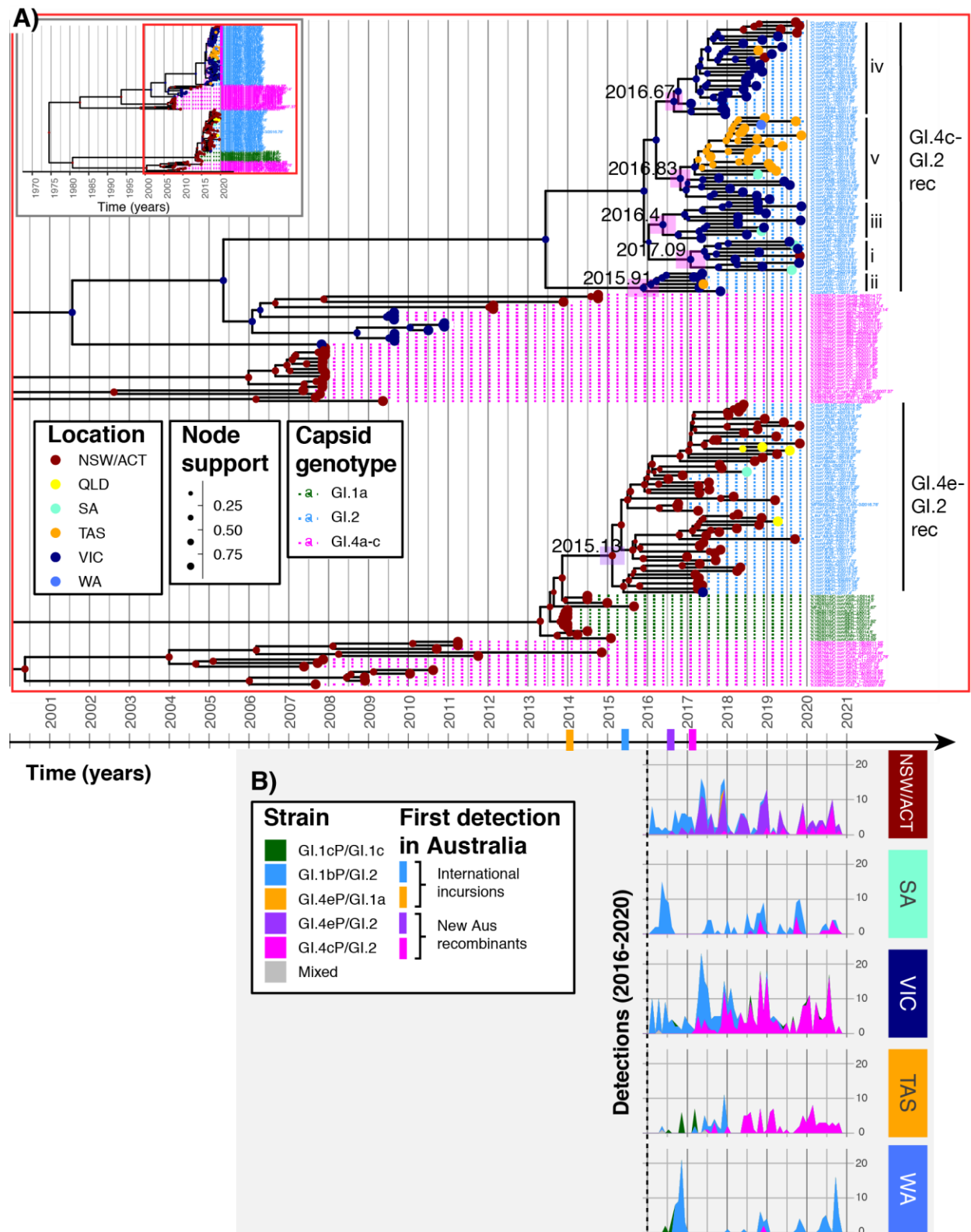
Figure 4: **Timing and location of the emergence of Australian lagovirus**

recombinants based on time-structured phylogeographic reconstruction and RT-

399 **qPCR detections.** Panels A and B are scaled to the same x-axis (shown between the

400 two panels), given as time in years. (A) Time-structured phylogeographic analysis of

401 Australian GI.4 lagovirus NS genes (uncorrelated log-normally distributed clock model,

402 constant tree prior). The inset indicates the section of the phylogeny shown. Node and

403 tip points are coloured according to location (Australian state) and node points are sized

404 according to posterior support for that clade. Taxon names are coloured by capsid

405 variant. The mean time to most recent common ancestor for the GI.4eP-GI.2 and each

406 GI.4cP-GI.2 recombinant lineage is indicated at the respective internal node, and

407 horizontal bars at these nodes represent the 95% highest posterior density (HPD). The

408 recombinant (rec) clades are labelled, including the GI.4cP-GI.2 lineages (i – v). (B)

409 Lagovirus positive samples collected in NSW/ACT, SA, VIC, TAS, and WA from 2016 to

410 2020 (n = 739) were genotyped to the variant level by sequencing either side of the

411 typical calicivirus recombination breakpoint. The number of detections of each variant

412 by month are shown for each geographical region as an area plot, with the plotted area

413 coloured by variant.

414

415 *All recombination events occurred within a two-year timeframe in eastern Australia*

416 To determine the temporal dynamics, the location of emergence, and the rate of viable

417 recombination events in Australian lagoviruses, we conducted time-structured

418 phylogenetic analyses for the NS and capsid genes of recombinant viruses and related

419 sequences (i.e., Australian GI.2 VP60 and GI.4 NS sequences). Based on regression

420 analysis, data from the GI.4 NS dataset was considered more accurate and is therefore

421 presented in Figure 4. Based on these analyses, the 4e-recombinant was the first

422 variant to emerge, with the most recent common ancestor of this lineage dated to early

423 to mid-2015 (95% highest posterior density [HPD] NS dataset, 2014.8 – 2015.43)

424 (Figure 4A). For the 4c-recombinants, the 95% HPD intervals of the time to most recent

425 common ancestor (TMRCA) overlapped between lineages, making it difficult to

426 confidently define the exact timings and order of emergence of each lineage (Figure

427 4A). However, the data clearly show that these five recombination events all occurred

428 within the space of two years. This two-year timeframe was consistently observed

429   regardless of whether the analysis was conducted on the NS genes GI.4 data set or the

430   VP60 gene GI.2 data set. Our phylogeographic analysis (Figure 4A) suggests that all

431   the 4c-recombinant lineages emerged in VIC (probability >0.99 for all lineages), with

432   subsequent spread to other states. Notably, this was observed even for lineage v, which

433   was first detected in Tasmania. In contrast, the 4e-recombinant clearly emerged in

434   NSW/ACT (probability >0.99, Figure 4A); this is further supported by our

435   epidemiological data, where we found only limited detections of this variant in other

436   states (QLD, VIC, SA).

437   *Continued cocirculation of 4c-recombinant lineages*

438   The 4c-recombinant viruses were initially assigned to lineages based on the GI.2 VP60

439   phylogeny (Figure 3A). For GI.4c-recombinant viruses where a full genome sequence

440   was not available, we first inferred a phylogeny from a 504 nt region spanning the

441   RdRp-VP60 junction (Figure 5); this phylogeny was annotated using taxa for which the

442   lineage was definitively known from the VP60 phylogeny. While sequences from each

443   4c lineage (as defined from the complete genome sequences) did not form clades with

444   individual common ancestors in the RdRp-VP60 junction phylogeny, they did form

445   distinct groups with visible genetic distance between them (Figure 5). These groups

446   were distinct enough to assign a 4c lineage type to those samples with partial sequence

447   only. After assigning lineages to all GI.4c-recombinants, we explored the interactions

448   between the different lineages by mapping their sampling locations over time. We found

449   that four of the five lineages were already circulating in 2017 (Figure 5). Lineage ii was

450   not detected after 2017 while the other four lineages continued to be detected

451   throughout 2018, 2019, and 2020. In VIC, where all lineages are circulating, there was

452   no clear indication of geographical clustering of the different lineages. In TAS, lineages

453   ii, iii, and iv were detected sporadically while lineage v has remained dominant over

454   time, suggestive of a founder effect in this region. Similarly, in NSW/ACT lineage iv has

455   remained dominant, also suggesting a regional founder effect, while lineages i, ii, and iii

456   were only detected intermittently.

457

458
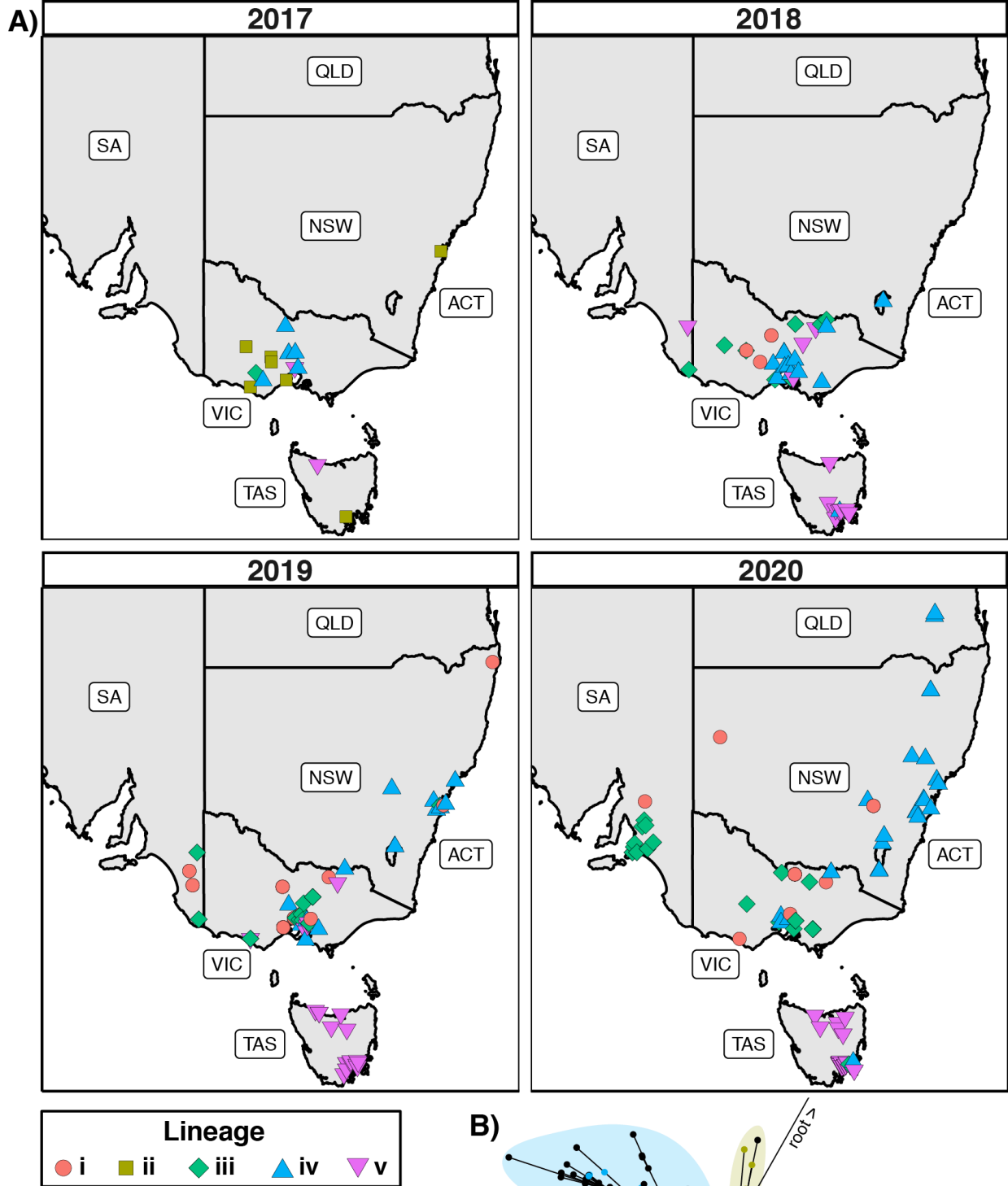
459 **Figure 5**: **Geographical distribution of 4c-recombinant lagoviruses in Australia,**

460 **2017 to 2020.** (A) Sampling locations for the 4c lineages were mapped for each year

461 following the emergence of this variant in 2017. (B) 4c-recombinant viruses were

462 allocated to a 4c lineage based on ML phylogenetic grouping of the partial RdRp-VP60

463 sequence (504 nt) with reference sequences. Reference sequences were annotated

464 based on full genome sequencing and are indicated in the tree by the coloured tip

465 points. The phylogeny was midpoint rooted and the scale bar represents nucleotide

466 substitutions per site.

467

468 *The emergence of 4c-recombinant variants was not associated with antigenic changes*

469 *in the VP60 capsid protein*

470 To determine whether the observed epidemiological replacement of 1) the parental

471 RHDV2 virus by the novel 4c-recombinant in VIC and TAS, and 2) the 4e-recombinant

472 in NSW/ACT were associated with nonsynonymous, potentially antigenic, changes in

473 the VP60 capsid protein we performed ancestral state reconstruction using our prior

474 Australian GI.2 VP60 phylogeny. Notably, four of the five 4c-recombinant lineages

475 showed no amino acid changes in the capsid protein relative to the inferred ancestral

476 RHDV2 node, suggesting that the emergence of these lineages was not associated with

477 changes in antigenicity (Figure 3A). However, nonsynonymous mutations were

478 identified in 4c-recombinant lineage i and in 4e-recombinant viruses (Figure 3A). When

479 we mapped these mutations back to an atomic model of lagovirus GI VP60, both

480 mutations were within the protruding domain, specifically the P2 subdomain, which is

481 known to contain immunodominant epitopes (Figure 3B) [61]. The 4c-recombinant

482 lineage i was associated with an S364G mutation, while an I434T mutation was present

483 in all lineage 4e-recombinants (Figure 3A). Thus, it is possible that the epidemiological

484 fitness of the 4e-recombinant variant in NSW/ACT may be attributable to antigenic

485 changes in the capsid; however, this cannot explain the replacement of parental RHDV2

486 by 4c-recombinant variants in VIC or TAS, or the continued co-circulation of 4c-

487 recombinant lineages without changes in VP60. Furthermore, in NSW/ACT the 4e-

488 recombinant variant, despite the I434T change in VP60, appears to be currently

489 undergoing replacement by the 4c-recombinant.

490

Discussion

492 In a global landscape where population immunity to GI.1 and GI.1a viruses was

493 widespread in lagomorphs, it is unsurprising that GI.2 viruses were so successful and

494 rapidly replaced GI.1 strains. We were interested to determine whether, after

495 establishment, GI.2 viruses would continue to evolve via immune selection and

496 antigenic drift or by other mechanisms. We were also interested to document whether

497 variants without the GI.2 capsid would emerge as population immunity increased

498 against this genotype. Instead, we found that the GI.2 capsid was retained and that the

499 emerging recombinant variants acquired alternative NS genes. It is becoming

500 increasingly clear that recombination is a major driver of calicivirus diversity, facilitating

501 the emergence of new variants including those with pandemic and panzootic potential

502 [24, 25, 62, 63]. However, the frequency at which viable recombination occurs and the

503 genetic drivers of epidemiological fitness are still poorly understood. In this study, a

504 continent-scale natural competition experiment of lagoviruses in Australia, we found that

505 viable recombination occurs at an extremely high frequency given the right agent, host,

506 and environmental circumstances.

Six viable recombination events in Australian lagoviruses between 2014 and 2018

508 Shortly after detecting two exotic lagovirus incursions in Australia, RHDVa-Aus (GI.4eP-

509 GI.1a) in 2014 and RHDV2 (GI.1bP-GI.2) in 2015 [32, 37], we reported the emergence

510 of a recombinant variant of these two viruses (4e-recombinant; GI.4eP-GI.2), detected

511 in July 2016 in NSW [31]. This finding prompted the development of improved molecular

512 tests for the detection of recombination and the subsequent retrospective and

513 prospective screening of lagovirus-positive samples for further recombination events.

514 This screen identified an additional novel recombinant lagovirus, a GI.4cP-GI.2 variant

515 (4c-recombinant), first detected in VIC in February 2017. Further analysis determined

516 that the VP60 genes of this 4c-recombinant separated into five distinct clades (Figure

517   3A), suggesting that this variant emerged on five separate occasions following at least

518   five independent recombination events. Recombination was already known to occur in

519   lagoviruses. For example, in the last decade, the lagovirus GI.2 capsid has been

520   detected in combination with five different sets of NS genes [24, 27, 31, 33]. Our study

521   demonstrates that viable recombination during natural infection occurs even more

522   frequently than previously appreciated but may often go undetected if different events

523   result in closely related progeny. Since the first detection of RHDV2 in Australia in mid-

524   2015 [37] this capsid gene has recombined and produced epidemiologically viable

525   variants at least six times in a two-year period.

526   The previously reported 4e-recombinant arose from a recombination event between

527   RHDVa-Aus (GI.1aP-GI.2) and RHDV2 (GI.1bP-GI.2), both recombinant viruses

528   themselves that were exotic incursions detected in Australia in January 2014 and May

529   2015, respectively, although the source of these viruses remains unknown [31, 32, 37].

530   Detections of the parental RHDVa-Aus were predominantly restricted to the Sydney

531   basin region of NSW, with two cases detected in QLD and one further case in a wild

532   rabbit in the ACT [32]. Therefore, it is not surprising that the 4e-recombinant emerged in

533   NSW/ACT, according to our epidemiological data and phylogeographic analysis (Figure

534   4). The 4c-recombinants arose from recombination events between RHDV2 (GI.1bP-

535   GI.1) and endemic RCV-A1 (GI.4c) viruses [14, 37]. Emergence of the 4c-recombinant

536   variants in the south-eastern states of Australia is consistent with high rabbit population

537   densities and a high seroprevalence to GI.4 viruses in these regions compared to more

538   arid regions of Australia [64].

539   Despite the VP60 phylogeny suggesting multiple recombination events between GI.4c

540   and GI.2 viruses, only a single unique recombination event was detected by the RDP4

541   program. This is not surprising given that the breakpoint locations are indistinguishable

542   and these recombinants (and their true parents) are very closely related. Indeed, it was

543   only possible to infer the true number of recombination events because of the high

544   sampling frequency of parental RHDV2 viruses, which has been maintained for this and

545   previous studies. Had RHDV2 sequences from intervening clades not been sampled,

546   the 4c-recombinants would have formed a monophyletic clade and a single

547   recombination event assumed. This is demonstrated by the NS genes phylogeny

548   (Figure 4A), in which the 4c-recombinant sequences form a single clade within the

549   greater GI.4c clade. Importantly, within this clade the sequences group into the same

550   five lineages as seen in the VP60 tree. Compared to pathogenic GI.2 viruses, benign

551   RCV-A1 (GI.4c) viruses are greatly under-sampled in Australia, with the most recent

552   RCV-A1 sequences (n = 6) derived from samples collected from the ACT in 2012 and

553   2014 [35]. Aside from three sequences sampled in 2014, no other RCV-A1 sequences

554   were available for the entirety of the study period. Consequently, a close relative of the

555   RCV-A1 parental virus has not been sequenced, suggesting a hidden diversity of

556   unsampled RCV-A1 viruses in Australian rabbits. This study provides a lower bound to

557   the rate of recombination in these viruses; it is possible that additional recombination

558   events have occurred with GI.4c viruses that are not observable against the current

559   background of available GI.2 VP60 sequences.

560   *Lagoviruses show frequent recombination between the GI.2 capsid and GI.4 NS*

561   *sequences*

562   Interestingly, all the recombination events detected during the study period involved the

563   pairing of a GI.2 capsid and the NS genes of a GI.4 virus. Despite screening over 800

564   lagovirus-positive samples, we found no evidence for recombinant variants with either

565   capsid or NS sequences of GI.1 viruses.

566   For recombination to occur several criteria must be met [23]. Firstly, there must be

567   coinfection of an individual host. This is influenced by host tropism, the prevalence of

568   each parental virus in the population, and the duration of infection. Secondly, there must

569   be coinfection of a single cell, both viruses must replicate within this cell, and precise

570   template switching must occur to generate viable gRNA. Finally, the resultant variant

571   must be epidemiologically competitive. That is, it must be able to successfully transmit,

572   to establish infection in new hosts, and to avoid being outcompeted by either the

573   parental or other circulating variants.

574   Following the epizootic incursion of an antigenically novel virus, RHDV2, into a naïve

575   population, the prevalence of this variant was extremely high in Australian rabbits,

576   reflected by the estimated population-wide mortality rate of 60% following the arrival of

577   RHDV2 [39]. The detection of four GI.1 and GI.2 co-infections in this study between

578   2015 and 2017 demonstrates that both viruses were circulating at sufficiently high

579   prevalence for mixed infections to occur. However, this doesn't explain the predilection

580   for GI.4 (RCV) NS sequences. The discovery of 4c-recombinants was surprising, since

581   the parental RCV-A1 is a benign enterotropic lagovirus [14], while RHDV2 (and RHDV1)

582   viruses are virulent and hepatotropic. This would seem to preclude coinfection of

583   individual cells with both viruses. Yet, several recombinants of RHDV2 and benign

584   enterotropic RCVs have also been reported from Europe [64]. This suggests that active

585   replication of both RCVs and RHDV2 must be occurring in the same target cell.

586   Extrapolating from our current understanding of human and murine norovirus tropism,

587   macrophages or other immune cells may be likely candidates [65]. This is further

588   supported by the detection of plus- and minus-strand viral RNAs in splenic and alveolar

589   macrophages of rabbits experimentally infected with RHDV1 [66]. Additionally, RCVs

590   typically infect young rabbits early in life [12, 14]. Although young rabbits can be

591   infected with RHDV1 viruses, robust innate immunity limits the extent of viral replication

592   in an age-dependent manner [34, 39]. Thus, mixed infections with RHDV1 and RCVs

593   are probably infrequent. This age-dependent resistance is not observed with RHDV2

594   infection [67]. Furthermore, the duration of infection is longer for benign RCVs

595   compared to pathogenic variants, where infected individuals typically die within 48 – 72

596   hours post-infection [2]. These factors may at least partly explain why RCV-RHDV2

597   recombinants appear to emerge more frequently than RHDV1-RHDV2 recombinants.

598   *Epidemiological drivers of lagovirus emergence and spread*

599   Both the 4e-recombinant and 4c-recombinant rapidly replaced the dominant circulating

600   parental RHDV2 in NSW/ACT and VIC/TAS, respectively. This replacement, at least for

601   several 4c-recombinant lineages, occurred without any associated antigenic changes in

602   the capsid protein, demonstrating that the replacement was not driven by antigenic

603   escape. This shows that NS sequence variation is an important driver of epidemiological

604   fitness in lagoviruses, complementing similar findings in human and murine noroviruses

605   [68, 69]. For example, the pandemic GII.P16/GII.4 Sydney 2012 norovirus, which does

606   not contain unique substitutions in the capsid, has substitutions within the RdRp that are

607   proposed to increase transmissibility [69]. Infection with this variant also resulted in

608   increased viral shedding compared to other norovirus genotypes, as measured by

609   higher faecal viral loads [70]. Previous studies have shown that RdRp fidelity and intra-

610   host viral diversity affect transmissibility of murine norovirus *in vivo*, with high-fidelity

611   variants being less efficiently transmitted than the wild-type variant [71]. Taken together,

612   these studies demonstrate that NS proteins, particularly the RdRp, are important drivers

613   of calicivirus fitness. Interestingly, the *in vitro* polymerase replication activity of the

614   cloned GI.4c RdRp was previously shown to be at least two times that of the RHDV1

615   (GI.1c) RdRp [72], although comparison to the GI.1b RdRp, the specific 'competing'

616   variant in the current study, was not reported. Analogous to the findings in murine

617   norovirus, we propose that a higher replication rate of GI.4 RdRps may lead to

618   increased intra-host viral diversity and transmissibility of GI.4 recombinants. However,

619   we cannot rule out that other NS proteins may also contribute to the observed high

620   fitness of the recombinant variants.

621   Within the 4c-recombinants, there was no evidence of dominance of any one lineage

622   over time. The apparent dominance of lineage iv in NSW/ACT and lineage v in TAS is

623   consistent with founder effects in both of those states. This demonstrates that different

624   lagoviruses can cocirculate at relative equilibrium over extended periods of time. This

625   further supports that the observed rapid replacement of parental RHDV2 (GI.1bP-GI.2)

626   by GI.4P recombinant variants in this study is due to a fitness advantage of these

627   variants, conferred by the NS proteins and possibly associated with the RdRp.

628   *The lagovirus capsid governs host and tissue tropism and is correlated with virulence*

629   Both the newly identified 4c-recombinant and the previously emerged 4e-recombinant

630   are virulent, hepatotropic viruses that were recovered from the livers of both rabbits and

631   hares and from rabbits of all ages in this study. This tropism mimics that seen with other

632   GI.2 viruses [73-80]. In stark contrast, the parental RCV-A1 is a benign, enterotropic

633   virus that has only been recovered from rabbits [14], while the RHDVa-Aus variant,

634   although virulent and hepatotropic, has only been found in adult rabbits. Our findings

635   suggest that it is the lagovirus capsid that confers both host and tissue tropism and that

636   tissue tropism is correlated with virulence.

637 With the broader host tropism conferred by the GI.2 capsid, there is increased potential

638 for the emergence of novel epizootic lagovirus variants through both intragenotypic and

639 intergenotypic recombination. Hares are known to carry their own, presumed benign,

640 caliciviruses [16-18, 31, 33, 73-79]. Indeed, the first intergenotypic lagovirus

641 recombinants were recently reported from Germany [33]. Since the incursion of RHDV2

642 into North America in 2020 this variant has also been reported to infect several

643 *Sylvilagus* species (cottontail rabbits) [81]. Although endemic *Sylvilagus* caliciviruses

644 have never been reported, very limited sampling has been conducted in this species. It

645 remains to be seen whether North American leporids may be a new reservoir for the

646 emergence of novel lagoviruses with panzootic potential. This highlights the need for

647 ongoing surveillance and full genetic characterization of lagoviruses and other

648 caliciviruses to facilitate detection of future emerging variants of significance to both

649 animal and human health.

650

Acknowledgements

666

Funding

Data availability

Full genome sequences are available in GenBank under accession numbers MW460205 – MW460242. All sequence alignments, tree files, and BEAST xml files are available at https://doi.org/10.25919/758f-4t15.

References

1. Parra, G.I., et al., *Static and Evolving Norovirus Genotypes: Implications for Epidemiology and Immunity.* PLoS Pathog, 2017. **13**(1): p. e1006136.
2. Abrantes, J., et al., *Rabbit haemorrhagic disease (RHD) and rabbit haemorrhagic disease virus (RHDV): a review.* Vet Res, 2012. **43**: p. 12.
3. Le Pendu, J., et al., *Proposal for a unified classification system and nomenclature of lagoviruses.* J Gen Virol, 2017. **98**(7): p. 1658-1666.
4. Gavier-Widen, D. and T. Morner, *Epidemiology and diagnosis of the European brown hare syndrome in Scandinavian countries: a review.* Rev Sci Tech, 1991. **10**(2): p. 453-8.
5. Liu, S.J., et al., *A new viral disease in rabbits.* J Vet Diagn Invest, 1984. **16**: p. 253-255.
6. Capucci, L., et al., *A further step in the evolution of rabbit hemorrhagic disease virus: the appearance of the first consistent antigenic variant.* Virus Res, 1998. **58**(1-2): p. 115-26.
7. Le Gall-Recule, G., et al., *Emergence of a new lagovirus related to Rabbit Haemorrhagic Disease Virus.* Vet Res, 2013. **44**: p. 81.
8. Le Gall-Recule, G., et al., *Detection of a new variant of rabbit haemorrhagic disease virus in France.* Vet Rec, 2011. **168**(5): p. 137-8.
9. Dalton, K.P., et al., *Variant rabbit hemorrhagic disease virus in young rabbits, Spain.* Emerg Infect Dis, 2012. **18**(12): p. 2009-12.

695   10.   Capucci, L., et al., *Detection and preliminary characterization of a new rabbit calicivirus related*
696   *to rabbit hemorrhagic disease virus but nonpathogenic.* J Virol, 1996. **70**(12): p. 8614-23.

697   11.   Lemaitre, E., et al., *First complete genome sequence of a European non-pathogenic rabbit*
698   *calicivirus (lagovirus GI.3).* Arch Virol, 2018. **163**(10): p. 2921-2924.

699   12.   Capucci, L., A. Nardin, and A. Lavazza, *Seroconversion in an industrial unit of rabbits infected*
700   *with a non-pathogenic rabbit haemorrhagic disease-like virus.* Vet Rec, 1997. **140**(25): p. 647-50.

701   13.   Le Gall-Recule, G., et al., *Characterisation of a non-pathogenic and non-protective infectious*
702   *rabbit lagovirus related to RHDV.* Virology, 2011. **410**(2): p. 395-402.

703   14.   Strive, T., J.D. Wright, and A.J. Robinson, *Identification and partial characterisation of a new*
704   *Lagovirus in Australian wild rabbits.* Virology, 2009. **384**(1): p. 97-105.

705   15.   Nicholson, L.J., et al., *Benign Rabbit Calicivirus in New Zealand.* Appl Environ Microbiol, 2017.
706   **83**(11).

707   16.   Cavadini, P., et al., *Widespread occurrence of the non-pathogenic hare calicivirus (HaCV*
708   *Lagovirus GII.2) in captive-reared and free-living wild hares in Europe.* Transbound Emerg Dis,
709   2020.

710   17.   Droillard, C., et al., *First complete genome sequence of a hare calicivirus strain isolated from*
711   *Lepus europaeus.* Microbiol Resour Announc, 2018. **7**(22).

712   18.   Droillard, C., et al., *Genetic diversity and evolution of Hare Calicivirus (HaCV), a recently*
713   *identified lagovirus from Lepus europaeus.* Infect Genet Evol, 2020. **82**: p. 104310.

714   19.   Mahar, J.E., et al., *The discovery of three new hare lagoviruses reveals unexplored viral diversity*
715   *in this genus.* Virus Evol, 2019. **5**(1): p. vez005.

716   20.   Ohlinger, V.F., et al., *Identification and characterization of the virus causing rabbit hemorrhagic*
717   *disease.* J Virol, 1990. **64**(7): p. 3331-6.

718   21.   Conley, M.J., et al., *Calicivirus VP2 forms a portal-like assembly following receptor engagement.*
719   Nature, 2019. **565**(7739): p. 377-381.

720   22.   Meyers, G., C. Wirblich, and H.J. Thiel, *Genomic and subgenomic RNAs of rabbit hemorrhagic*
721   *disease virus are both protein-linked and packaged into particles.* Virology, 1991. **184**(2): p. 677-
722   86.

723   23.   Worobey, M. and E.C. Holmes, *Evolutionary aspects of recombination in RNA viruses.* J Gen Virol,
724   1999. **80 (Pt 10)**: p. 2535-2543.

725   24.   Abrantes, J., et al., *Recombination at the emergence of the pathogenic rabbit haemorrhagic*
726   *disease virus Lagovirus europaeus/GI.2.* Sci Rep, 2020. **10**(1): p. 14502.

727   25.   Ludwig-Begall, L.F., A. Mauroy, and E. Thiry, *Norovirus recombinants: recurrent in the field,*
728   *recalcitrant in the lab - a scoping review of recombination and recombinant types of noroviruses.*
729   J Gen Virol, 2018. **99**(8): p. 970-988.

730   26.   Bull, R.A., et al., *Norovirus recombination in ORF1/ORF2 overlap.* Emerg Infect Dis, 2005. **11**(7):
731   p. 1079-85.

732   27.   Lopes, A.M., et al., *Full genomic analysis of new variant rabbit hemorrhagic disease virus*
733   *revealed multiple recombination events.* J Gen Virol, 2015. **96**(Pt 6): p. 1309-1319.

734   28.   Lopes, A.M., et al., *GI.1b/GI.1b/GI.2 recombinant rabbit hemorrhagic disease virus 2 (Lagovirus*
735   *europaeus/GI.2) in Morocco, Africa.* Arch Virol, 2019. **164**(1): p. 279-283.

736   29.   Silverio, D., et al., *Insights into the evolution of the new variant rabbit haemorrhagic disease*
737   *virus (GI.2) and the identification of novel recombinant strains.* Transbound Emerg Dis, 2018.
738   **65**(4): p. 983-992.

739   30.   Dalton, K.P., et al., *Conventional and real time RT-PCR assays for the detection and*
740   *differentiation of variant rabbit hemorrhagic disease virus (RHDVb) and its recombinants.* J Virol
741   Methods, 2018. **251**: p. 118-122.

742   31.   Hall, R.N., et al., *A strain-specific multiplex RT-PCR for Australian rabbit haemorrhagic disease*
743   *viruses uncovers a new recombinant virus variant in rabbits and hares.* Transbound Emerg Dis,
744   2018. **65**(2): p. e444-e456.

745   32.   Mahar, J.E., et al., *Detection and Circulation of a Novel Rabbit Hemorrhagic Disease Virus in*
746   *Australia.* Emerg Infect Dis, 2018. **24**(1): p. 22-31.

747   33.   Szillat, K.P., et al., *Full-genome sequencing of German rabbit haemorrhagic disease virus*
748   *uncovers recombination between RHDV (GI.2) and EBHSV (GII.1).* Virus Evolution, 2020. **6**(2).

749   34.   Mahar, J.E., et al., *Rabbit Hemorrhagic Disease Virus 2 (RHDV2; GI.2) Is Replacing Endemic*
750   *Strains of RHDV in the Australian Landscape within 18 Months of Its Arrival.* J Virol, 2018. **92**(2).

751   35.   Mahar, J.E., et al., *Benign Rabbit Caliciviruses Exhibit Evolutionary Dynamics Similar to Those of*
752   *Their Virulent Relatives.* J Virol, 2016. **90**(20): p. 9317-29.

753   36.   Kerr, P.J., R.N. Hall, and T. Strive, *Viruses for landscape-scale therapy: Biological control of*
754   *rabbits in Australia.* Methods Mol Biol, 2021. **2225**: p. 1-23.

755   37.   Hall, R.N., et al., *Emerging Rabbit Hemorrhagic Disease Virus 2 (RHDVb), Australia.* Emerg Infect
756   Dis, 2015. **21**(12): p. 2276-8.

757   38.   Kirkland, P.D. and M.J. Frost, *The impact of viral transport media on PCR assay results for the*
758   *detection of nucleic acid from SARS-CoV-2.* Pathology, 2020. **52**(7): p. 811-814.

759 39. Ramsey, D.S.L., et al., *Emerging RHDV2 suppresses the impact of endemic and novel strains of*
760   *RHDV on wild rabbit populations.* Journal of Applied Ecology, 2020. **57**(3): p. 630-641.

761 40. Eden, J.S., et al., *Comparative Phylodynamics of Rabbit Hemorrhagic Disease Virus in Australia*
762   *and New Zealand.* J Virol, 2015. **89**(18): p. 9548-58.

763 41. Katoh, K. and D.M. Standley, *MAFFT multiple sequence alignment software version 7:*
764   *improvements in performance and usability.* Mol Biol Evol, 2013. **30**(4): p. 772-80.

765 42. Minh, B.Q., et al., *IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the*
766   *Genomic Era.* Mol Biol Evol, 2020. **37**(5): p. 1530-1534.

767 43. Guindon, S., et al., *New algorithms and methods to estimate maximum-likelihood phylogenies:*
768   *assessing the performance of PhyML 3.0.* Syst Biol, 2010. **59**(3): p. 307-21.

769 44. Martin, D.P., et al., *RDP4: Detection and analysis of recombination patterns in virus genomes.*
770   Virus Evol, 2015. **1**(1): p. vev003.

771 45. Yang, Z., *PAML 4: phylogenetic analysis by maximum likelihood.* Mol Biol Evol, 2007. **24**(8): p.
772   1586-91.

773 46. Rambaut, A., et al., *Exploring the temporal structure of heterochronous sequences using TempEst*
774   *(formerly Path-O-Gen).* Virus Evol, 2016. **2**(1): p. vew007.

775 47. Drummond, A.J. and A. Rambaut, *BEAST: Bayesian evolutionary analysis by sampling trees.* BMC
776   Evol Biol, 2007. **7**: p. 214.

777 48. Zharkikh, A., *Estimation of evolutionary distances between nucleotide sequences.* J Mol Evol,
778   1994. **39**(3): p. 315-29.

779 49. R Core Team, *R: A language and environment for statistical computing*. 2020, R Foundation for
780   Statistical Computing: Vienna, Austria.

781 50. Yu, G., et al., *ggtree: an r package for visualization and annotation of phylogenetic trees with*
782   *their covariates and other associated data.* Methods in Ecology and Evolution, 2017. **8**(1): p. 28-
783   36.

784 51. Wang, L.G., et al., *Treeio: An R Package for Phylogenetic Tree Input and Output with Richly*
785   *Annotated and Associated Data.* Mol Biol Evol, 2020. **37**(2): p. 599-603.

786 52. Wickham, H. and D. Seidel, *scales: Scale Functions for Visualization*. 2020, R package version
787   1.1.1.

788 53. Wickham, H., et al., *Welcome to the Tidyverse.* Journal of Open Source Software, 2019. **4**(43): p.
789   1686.

790   54.   Paradis, E. and K. Schliep, *ape 5.0: an environment for modern phylogenetics and evolutionary*
791   *analyses in R.* Bioinformatics, 2019. **35**(3): p. 526-528.

792   55.   Revell, L.J., *phytools: an R package for phylogenetic comparative biology (and other things).*
793   Methods in Ecology and Evolution, 2012. **3**(2): p. 217-223.

794   56.   Schliep, K.P., *phangorn: phylogenetic analysis in R.* Bioinformatics, 2011. **27**(4): p. 592-3.

795   57.   Wilke, C., *cowplot: Streamlined Plot Theme and Plot Annotations for 'ggplot2'.* 2020.

796   58.   Grolemund, G. and H. Wickham, *Dates and times made easy with lubridate.* Journal of Statistical
797   Software, 2011. **1**(3).

798   59.   Wickham, H., *The split-apply-combine strategy for data analysis.* Journal of Statistical Software,
799   2011. **1**(1).

800   60.   Sumner, M., *ozmaps: Australia Maps*. 2020.

801   61.   Wang, X., et al., *Atomic model of rabbit hemorrhagic disease virus by cryo-electron microscopy*
802   *and crystallography.* PLoS Pathog, 2013. **9**(1): p. e1003132.

803   62.   Parra, G.I., *Emergence of norovirus strains: A tale of two genes.* Virus Evol, 2019. **5**(2): p. vez048.

804   63.   Eden, J.S., et al., *Recombination within the pandemic norovirus GII.4 lineage.* J Virol, 2013.
805   **87**(11): p. 6270-82.

806   64.   Liu, J., et al., *Distribution and prevalence of the Australian non-pathogenic rabbit calicivirus is*
807   *correlated with rainfall and temperature.* PLoS One, 2014. **9**(12): p. e113976.

808   65.   Wobus, C.E., et al., *Replication of Norovirus in cell culture reveals a tropism for dendritic cells and*
809   *macrophages.* PLoS Biol, 2004. **2**(12): p. e432.

810   66.   Kimura, T., et al., *Distribution of rabbit haemorrhagic disease virus RNA in experimentally*
811   *infected rabbits.* J Comp Pathol, 2001. **124**(2-3): p. 134-41.

812   67.   Neave, M.J., et al., *Robust Innate Immunity of Young Rabbits Mediates Resistance to Rabbit*
813   *Hemorrhagic Disease Caused by Lagovirus Europaeus GI.1 But Not GI.2.* Viruses, 2018. **10**(9).

814   68.   Tohma, K., et al., *Phylogenetic Analyses Suggest that Factors Other Than the Capsid Protein Play*
815   *a Role in the Epidemic Potential of GII.2 Norovirus.* mSphere, 2017. **2**(3).

816   69.   Ruis, C., et al., *The emerging GII.P16-GII.4 Sydney 2012 norovirus lineage is circulating*
817   *worldwide, arose by late-2014 and contains polymerase changes that may increase virus*
818   *transmission.* PLoS One, 2017. **12**(6): p. e0179572.

819   70.   Cheung, S.K.C., et al., *Higher Viral Load of Emerging Norovirus GII.P16-GII.2 than Pandemic GII.4*
820   *and Epidemic GII.17, Hong Kong, China.* Emerg Infect Dis, 2019. **25**(1): p. 119-122.

821   71.   Arias, A., et al., *Norovirus Polymerase Fidelity Contributes to Viral Transmission In Vivo.*

822   mSphere, 2016. **1**(5).

823   72.   Urakova, N., et al., *Purification and Biochemical Characterisation of Rabbit Calicivirus RNA-*

824   *Dependent RNA Polymerases and Identification of Non-Nucleoside Inhibitors.* Viruses, 2016. **8**(4):

825   p. 100.

826   73.   Puggioni, G., et al., *The new French 2010 Rabbit Hemorrhagic Disease Virus causes an RHD-like*

827   *disease in the Sardinian Cape hare (Lepus capensis mediterraneus).* Vet Res, 2013. **44**: p. 96.

828   74.   Camarda, A., et al., *Detection of the new emerging rabbit haemorrhagic disease type 2 virus*

829   *(RHDV2) in Sicily from rabbit (Oryctolagus cuniculus) and Italian hare (Lepus corsicanus).* Res Vet

830   Sci, 2014. **97**(3): p. 642-5.

831   75.   Hall, R.N., et al., *Detection of RHDV2 in European brown hares (Lepus europaeus) in Australia.*

832   Vet Rec, 2017. **180**(5): p. 121.

833   76.   Velarde, R., et al., *Spillover Events of Infection of Brown Hares (Lepus europaeus) with Rabbit*

834   *Haemorrhagic Disease Type 2 Virus (RHDV2) Caused Sporadic Cases of an European Brown Hare*

835   *Syndrome-Like Disease in Italy and Spain.* Transbound Emerg Dis, 2017. **64**(6): p. 1750-1761.

836   77.   Le Gall-Recule, G., et al., *Large-scale lagovirus disease outbreaks in European brown hares*

837   *(Lepus europaeus) in France caused by RHDV2 strains spatially shared with rabbits (Oryctolagus*

838   *cuniculus).* Vet Res, 2017. **48**(1): p. 70.

839   78.   Neimanis, A.S., et al., *Overcoming species barriers: an outbreak of Lagovirus europaeus*

840   *GI.2/RHDV2 in an isolated population of mountain hares (Lepus timidus).* BMC Vet Res, 2018.

841   **14**(1): p. 367.

842   79.   Buehler, M., et al., *Lagovirus europeus GI.2 (rabbit hemorrhagic disease virus 2) infection in*

843   *captive mountain hares (Lepus timidus) in Germany.* BMC Vet Res, 2020. **16**(1): p. 166.

844   80.   Dalton, K.P., et al., *Spread of new variant RHDV in domestic rabbits on the Iberian Peninsula.* Vet

845   Microbiol, 2014. **169**(1-2): p. 67-73.

846   81.   Sleeman, J., *National Wildlife Health Center Wildlife Health Bulletin 2020-06*. 2020, USGS

847   National Wildlife Health Center.

848