

1 **Widespread human exposure to ledanteviruses in Uganda – a population study**

2

3 **Author list**

4 James G. Shepherd<sup>\*1</sup>, Shirin Ashraf<sup>1</sup>, Jesus F. Salazar-Gonzalez<sup>2</sup>, Maria G. Salazar<sup>2</sup>, Robert  
5 Downing<sup>3</sup>, Henry Bukenya<sup>3</sup>, Hanna Jerome<sup>1</sup>, Joseph T. Mpanga<sup>3</sup>, Chris Davis<sup>1</sup>, Lily Tong<sup>1</sup>,  
6 Vattipally B. Sreenu<sup>1</sup>, Linda A. Atiku<sup>3</sup>, Nicola Logan<sup>1</sup>, Ezekiel Kajik<sup>3</sup>, Yafesi Mukobi<sup>3</sup>,  
7 Cyrus Mungujakisa<sup>3</sup>, Michael V. Olowo<sup>3</sup>, Emmanuel Tibo<sup>3</sup>, Fred Wunna<sup>3</sup>, Hollie Jackson  
8 Ireland<sup>1</sup>, Andrew E. Blunsum<sup>1</sup>, Iyanuoluwani Owolabi<sup>1</sup>, Ana da Silva Filipe<sup>1</sup>, Josephine  
9 Bwogi<sup>3</sup>, Brian J. Willett<sup>1</sup>, Julius J. Lutwama<sup>3</sup>, Daniel G. Streicker<sup>1</sup>, Pontiano Kaleebu<sup>2,3</sup>,  
10 Emma C. Thomson<sup>\*1,4</sup>

11

12 (1) MRC-University of Glasgow Centre for Virus Research, Glasgow, UK

13 (2) MRC/UVRI & LSHTM Uganda Research Unit, Entebbe, Uganda

14 (3) Uganda Virus Research Institute, Entebbe Uganda

15 (4) London School of Hygiene and Tropical Medicine, London, UK

16 \* Corresponding authors

17 **Abstract**

18 Le Dantec virus (LDV), type species of the genus *Ledantevirus* within the *Rhabdoviridae* has  
19 been associated with human disease but has gone undetected since the 1970s. We describe  
20 the detection of LDV in a human case of undifferentiated fever in Uganda by metagenomic  
21 sequencing and demonstrate a serological response using ELISA and pseudotype  
22 neutralisation. By screening a cohort of 997 individuals sampled in 2016, we show frequent  
23 exposure to ledanteviruses with 76% of individuals seropositive in Western Uganda, but  
24 lower seroprevalence in other areas. Serological cross-reactivity as measured by pseudotype-  
25 based neutralisation was confined to *Ledantevirus*, indicating population seropositivity may  
26 represent either exposure to LDV or related ledanteviruses. We also describe the discovery of  
27 a closely related ledantevirus, Odro virus, in the synanthropic rodent *Mastomys*  
28 *erythroleucus*. *Ledantevirus* infection is common in Uganda but is geographically  
29 heterogenous. Further surveys of patients presenting with acute fever are required to  
30 determine the contribution of these emerging viruses to febrile illness in Uganda.

31

## 32 **Introduction**

33 The epidemiology of acute febrile illness in sub-Saharan Africa is poorly characterised (1).  
34 Despite the presence of a variety of endemic human pathogens and zoonoses, the diagnostic  
35 infrastructure is limited in comparison to the extensive culture-based and molecular assays  
36 available to patients in high income countries (2). Accordingly, even in severe cases the  
37 overwhelming majority of non-malarial febrile illness in Africa is managed in line with  
38 syndromic guidelines in the absence of a microbiological diagnosis (3). Paradoxically, the  
39 existing diversity of human pathogens in sub-Saharan Africa is likely to expand further as  
40 populations encroach into novel ecosystems (4).

41 Although rapid diagnostic tests are widely available for malaria, diagnostics for the viral  
42 pathogens that cause febrile illness are largely unavailable outside tertiary hospitals or  
43 national referral centres. Viruses contribute heavily to common fever-related presentations in  
44 Africa (5), but many research studies perform limited viral diagnostics or omit them entirely  
45 (6). As such there is an incomplete understanding of the viruses causing febrile illness in  
46 Africa, limiting the potential for early epidemic and pandemic response. To address this,  
47 researchers have employed unbiased metagenomic Next Generation Sequencing (mNGS) for  
48 pathogen diagnosis in cohorts of patients with acute febrile illness. This approach involves  
49 amplification of all nucleic acids within a sample in a “pathogen agnostic” manner, and has  
50 discovered several novel viruses associated with human disease (7-9).

51 Here we describe detection of Le Dantec virus (LDV), a rhabdovirus first isolated in  
52 association with human febrile illness in Senegal in 1965, in the blood of a Ugandan patient  
53 with undifferentiated acute febrile illness in 2012 by mNGS. We demonstrate a serological  
54 response against the viral glycoprotein in patient serum by ELISA, immunocytochemistry  
55 and pseudotype neutralisation. By screening stored serum samples from a nationally

56 representative cohort by ELISA, we show evidence of significant but geographically  
57 heterogenous exposure to LDV across Uganda and investigate the ecological factors  
58 associated with seropositivity. Testing ELISA reactive sera against a panel of related  
59 pseudotype viruses reveals that in addition to LDV, other closely related viruses are likely to  
60 be causing human infection via zoonotic transmission. In support of this possibility, we report  
61 the detection of a novel *Ledantevirus*, closely related to LDV, in a rodent host derived from  
62 peri-domestic ecological sampling in Northern Uganda.

## 63 Results

### 64 Patient characteristics

65 The findings of the Acute Febrile Illness (AFI) study including the detection of LDV have  
66 been reported previously (10). Here we focus on the clinical presentation of the LDV patient.  
67 The patient was a male child (aged under 10 years) residing in Kasese district, Western  
68 Uganda, who was recruited to the AFI study, an observational study investigating the  
69 epidemiology of acute febrile illness in Uganda. He presented to his local health centre in  
70 May 2012 with a 4-day history of fever, chills, headache, arthralgia, abdominal pain, and  
71 vomiting. He had a fever of 38.5°C, but no other abnormal findings on physical examination.  
72 A presumptive diagnosis of typhoid was made by the treating physician based on clinical  
73 features and he was managed as an outpatient with oral ciprofloxacin. He had fully recovered  
74 at a follow-up appointment four weeks later and at further follow up in 2018 he remained  
75 well. After enrolment in the AFI study, blood samples were tested against multiple pathogens  
76 including leptospirosis, brucellosis, malaria, dengue, chikungunya, yellow fever, o'nyong  
77 nyong, typhoid and rickettsiosis, all of which were negative (Table S1).

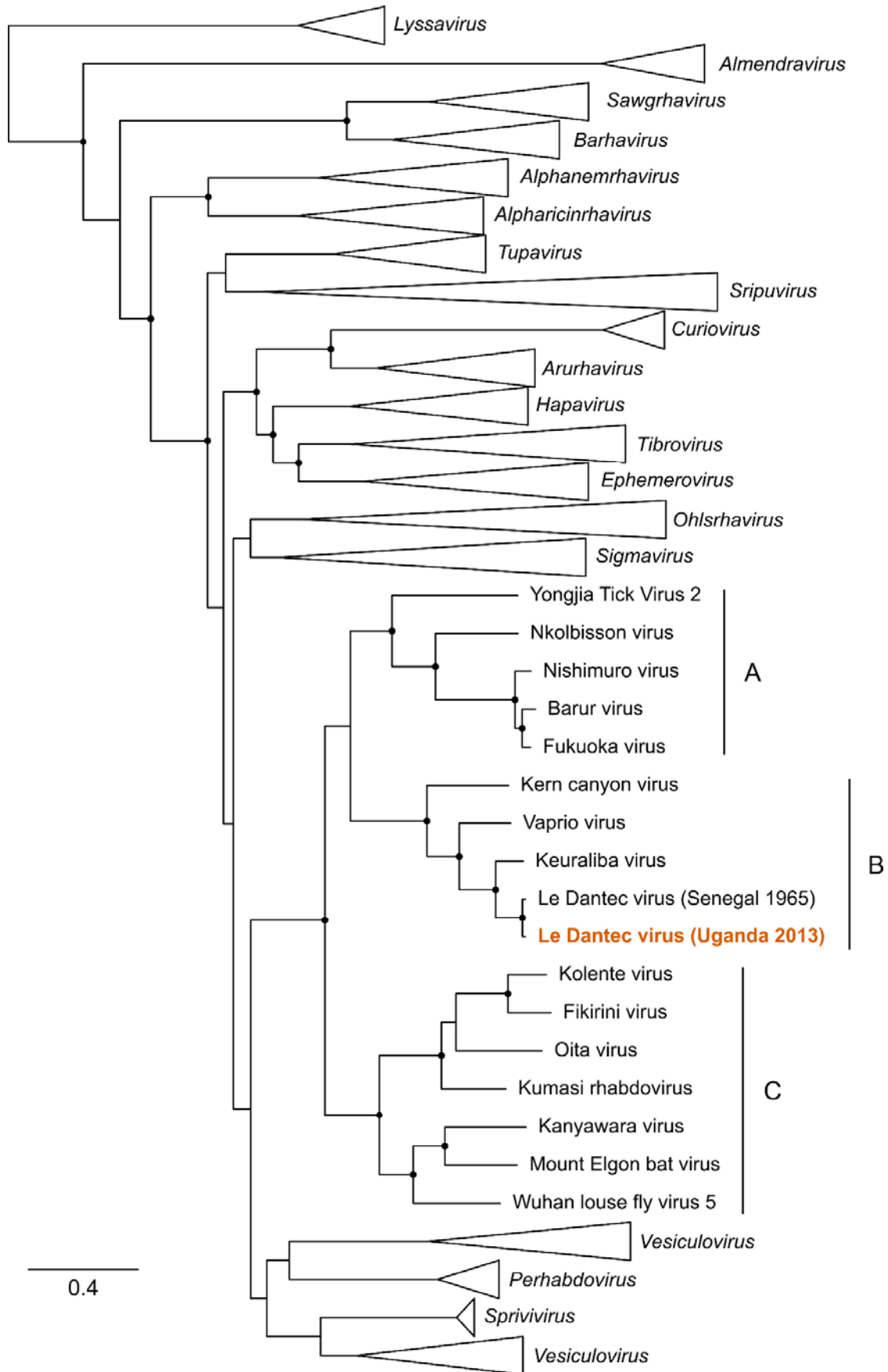
Table S1: Assays employed in the acute febrile illness study (AFI)	
Pathogen	Assay
Leptospirosis	Leptospirosis IgM lateral flow assay (Lifeassay diagnostics, Cape Town, SA)
Brucellosis	Brucella IgM lateral flow assay (Lifeassay diagnostics, Cape Town, SA)
Malaria	Thick and thin blood films Rapid 1-2-3 malaria HEMA Express rapid diagnostic kit (Miramar, Florida, USA.)
Dengue virus (acute samples)	Dengue NS1 Ag ELISA (Standard Diagnostics Inc. Kyonggido, Korea)
Dengue virus (convalescent samples)	Dengue IgM ELISA (Standard Diagnostics Inc. Kyonggido, Korea)
Chikungunya	Chikungunya IgM ELISA (Standard Diagnostics Inc. Kyonggido, Korea)
O'nyong nyong	In-house IgM ELISA developed at UVRI
West Nile Virus	West Nile virus IgM Capture ELISA (FOCUS Diagnostics Cypress, CA)
Yellow Fever Virus	In-house ELISA developed at UVRI
Typhoid Fever	Acute and convalescent IgM Tubex serology (IDL Biotech, Bromma, Sweden).
Rickettsioses	Rickettsia indirect fluorescent (IFA) IgG assay (Focus Diagnostics Cypress, CA)

78 **Detection of LDV by metagenomic sequencing of acute patient samples**

79 As initial diagnostic investigations were negative, a serum sample from the acute phase of the  
80 patient's illness was subjected to mNGS. Extracted RNA from serum was sequenced on an  
81 Illumina MiSeq instrument, resulting in 1677574 paired-end sequencing reads. *De novo*  
82 assembly yielded a 11,444 nucleotide contig that aligned to the genome of LDV on blastn  
83 query of the NCBI nucleotide database (isolate DakHD763, KM205006, 11450 base pairs).  
84 This contig was used as a reference sequence for alignment of raw sequencing reads,  
85 resulting in alignment of 8958 reads with 99.7% genome coverage at a minimum depth of 10  
86 reads. Pairwise alignment of the final genome showed greater than 90% nucleotide similarity  
87 to the original LDV isolate DakHD763 in all genomic regions (Table S2). Phylogenetic  
88 analysis confirmed the placement of the Ugandan LDV isolate within the genus *Ledantevirus*  
89 (Figure 1).

Region	N	P	M	G	U	L	Intergenic
Nucleotide p-distance	5.25	6.84	5.68	5.75	9.23	6.38	5.17
Amino acid p-distance	0.24	3.86	1.87	2.62	14.06	2.35	-

90



92 **Figure 1. Phylogenetic relationships of Le Dantec Virus identified in Uganda.**

93 Maximum likelihood phylogeny of the sub-family *Alpharhabdovirinae* based on full-length L protein amino  
94 acid sequences. The genus *Ledantivirus* is expanded with phylogroups A, B and C indicated to the right. Other  
95 genera are collapsed. The sequence derived from the febrile patient described in this study is indicated in orange  
96 text. The scale bar represents substitutions per site. Nodes with bootstrap support >70 are indicated with black  
97 circles.



98 **Confirmation of LDV infection by PCR and screening of acute fever samples**

99 To confirm mNGS discovery of LDV, a freshly prepared RNA extract from patient serum  
100 was assayed by RT-PCR with primers designed against conserved nucleotide positions in the  
101 original DakHD763 isolate and the sequence derived from Ugandan patient. The PCR assay  
102 yielded a 5.4 Kb fragment. A separate RT-PCR assay targeting a 2.4 Kb fragment in the  
103 glycoprotein gene was also positive. Sanger sequencing of the 5.4 Kb amplicon revealed all  
104 but one of the 5,406 nucleotides to be identical to the metagenomic sequence, validating the  
105 metagenomic sequence and confirming the presence of LDV RNA in the serum of the  
106 Ugandan patient.

107 The RNA was also tested with a separate LDV-specific real-time RT-PCR assay. Using serial  
108 10-fold dilutions of the DakHD763 isolate RNA as positive control we detected the 145-bp  
109 LDV sequence with a linear dynamic range of over four orders of magnitude. A fresh RNA  
110 extract from the source patient also showed a detectable signal; three µl of plasma-equivalent  
111 RNA yielded a Ct value of 30, consistent with a substantial viremia. Plasma samples (n=62)  
112 from a measles/rubella cohort and plasma pools (10 per pool) of unexplained acute febrile  
113 infections from Sudan tested all negative for LDV RNA in the real-time RT-PCR assay.

114

115 **Demonstration of a serological response to infection with LDV in the index patient**

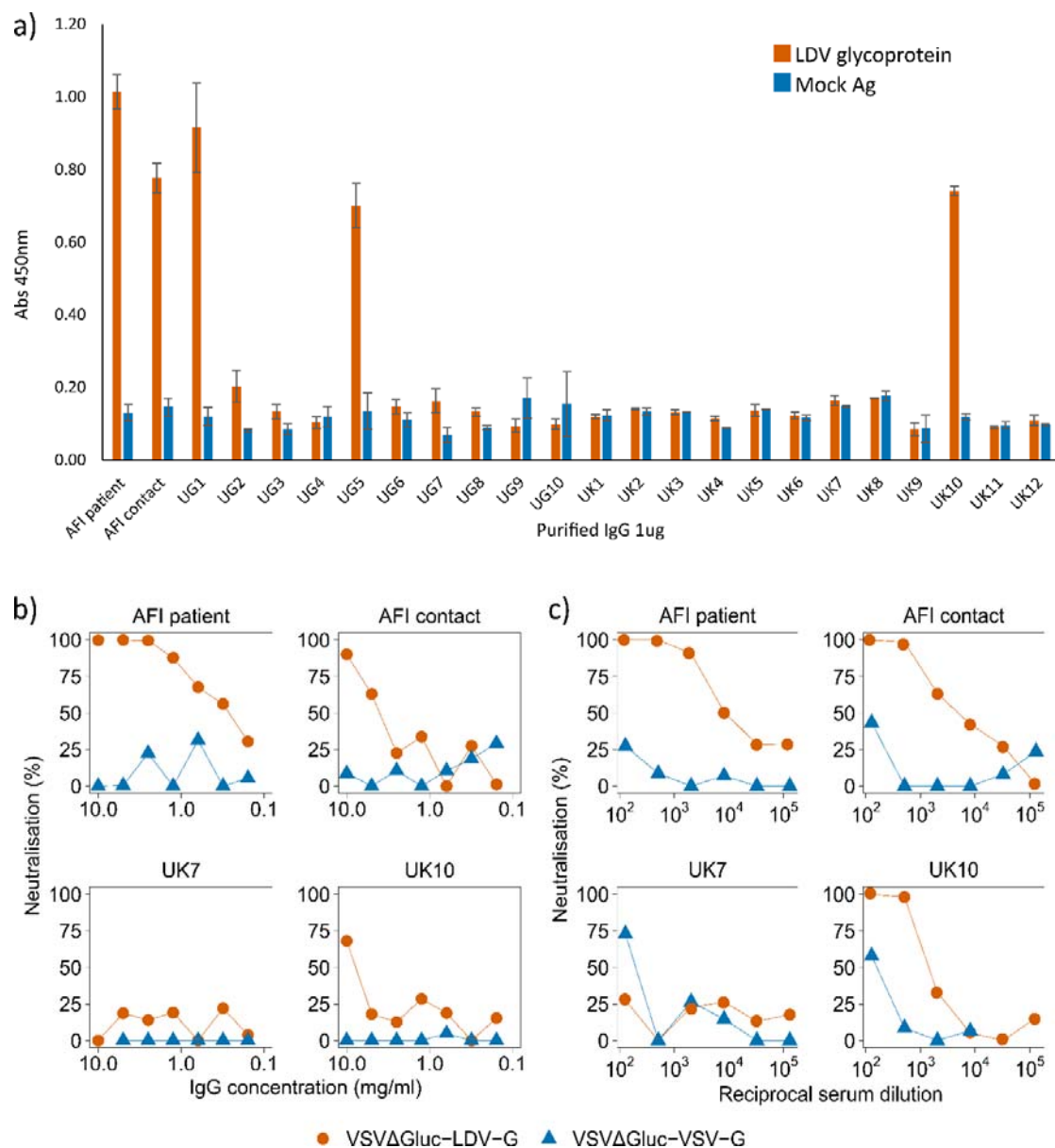
116 Using an in-house LDV glycoprotein (LDV-G) ELISA we demonstrated reactivity to LDV-G  
117 by both purified IgG and serum from blood samples collected from the patient at a follow-up  
118 visit in 2018 (Figure 2). Reactivity was also detected in a close relative who had experienced  
119 a contemporaneous illness but had not presented to hospital. A further 10 unrelated serum  
120 samples from Uganda and 12 from healthy UK controls were tested, with evidence of an  
121 antibody response to LDV-G detectable in both the index case and their close contact as well

122 as two individuals from Uganda and one sample from the UK (UK10). On further  
123 investigation it emerged that UK10 had previously resided in Africa.

124 A vesicular stomatitis virus (VSV) pseudotype virus with the glycoprotein gene replaced with  
125 the gene for firefly luciferase (VSV $\Delta$ Gluc) was created to express LDV-G as an outer  
126 membrane protein (VSV $\Delta$ Gluc-LDV-G). Both serum and purified IgG from the febrile  
127 patient specifically inhibited VSV $\Delta$ Gluc-LDV-G, without a similar effect on pseudotype  
128 viruses displaying the VSV glycoprotein (VSV $\Delta$ Gluc-VSV-G, Figure 2b). Consistent with  
129 the LDV-G ELISA, serum and purified IgG from the close family contact and UK10  
130 inhibited VSV $\Delta$ Gluc-LDV-G, although to a lesser degree than the patient.

131 To visualise the interactions between LDV-G and patient IgG, LDV-G engineered with a 6-  
132 histidine C-terminal tag (LDV-G<sub>his</sub>) was expressed in BHK-21 cells. Cells were incubated  
133 with a mixture of a rabbit anti-histidine primary antibody and patient sera. Co-localisation of  
134 the rabbit anti-histidine and IgG from patient samples was demonstrated for the patient and  
135 other samples found to be positive by ELISA, further demonstrating a serological response  
136 directed against LDV-G (Figure 3).

137

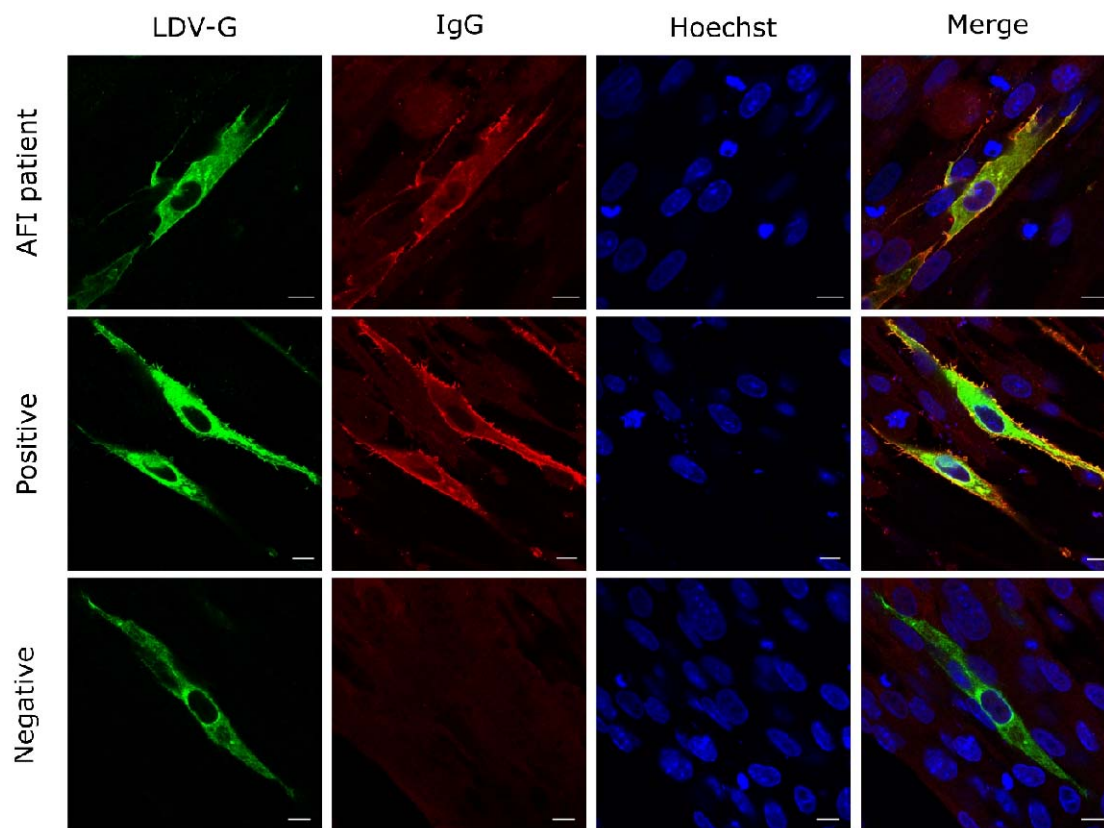


138

139 **Figure 2. Serological detection of LDV in patient serum by ELISA and pseudotype neutralisation**

140 a) Raw OD450 values for purified IgG from the index patient, a close contact and Ugandan and UK controls  
 141 tested by LDV-G ELISA. Bars represent the mean OD450 based on six technical replicates. Error bars represent  
 142 the standard error of the mean. b) Neutralisation of VSVΔGluc-LDV-G and VSVΔGluc-VSV-G by purified IgG  
 143 from human serum. Serial IgG dilutions were mixed with pseudotype preparations and added to HEK293T cells.  
 144 Luciferase activity was measured at 48 hours post infection. Datapoints represent the mean neutralisation  
 145 derived from three technical replicates. c) Neutralisation of VSVΔGluc-LDV-G and VSVΔGluc-VSV-G by  
 146 human serum. Serial serum dilutions were mixed with pseudotype preparations and added to HEK293T cells.  
 147 Luciferase activity was measured at 48 hours post infection. A representative experiment is shown. Datapoints  
 148 represent the percentage neutralisation derived from the mean of three technical replicates. Percentage  
 149 neutralisation represents luciferase readings relative to no-serum control wells.

150



151

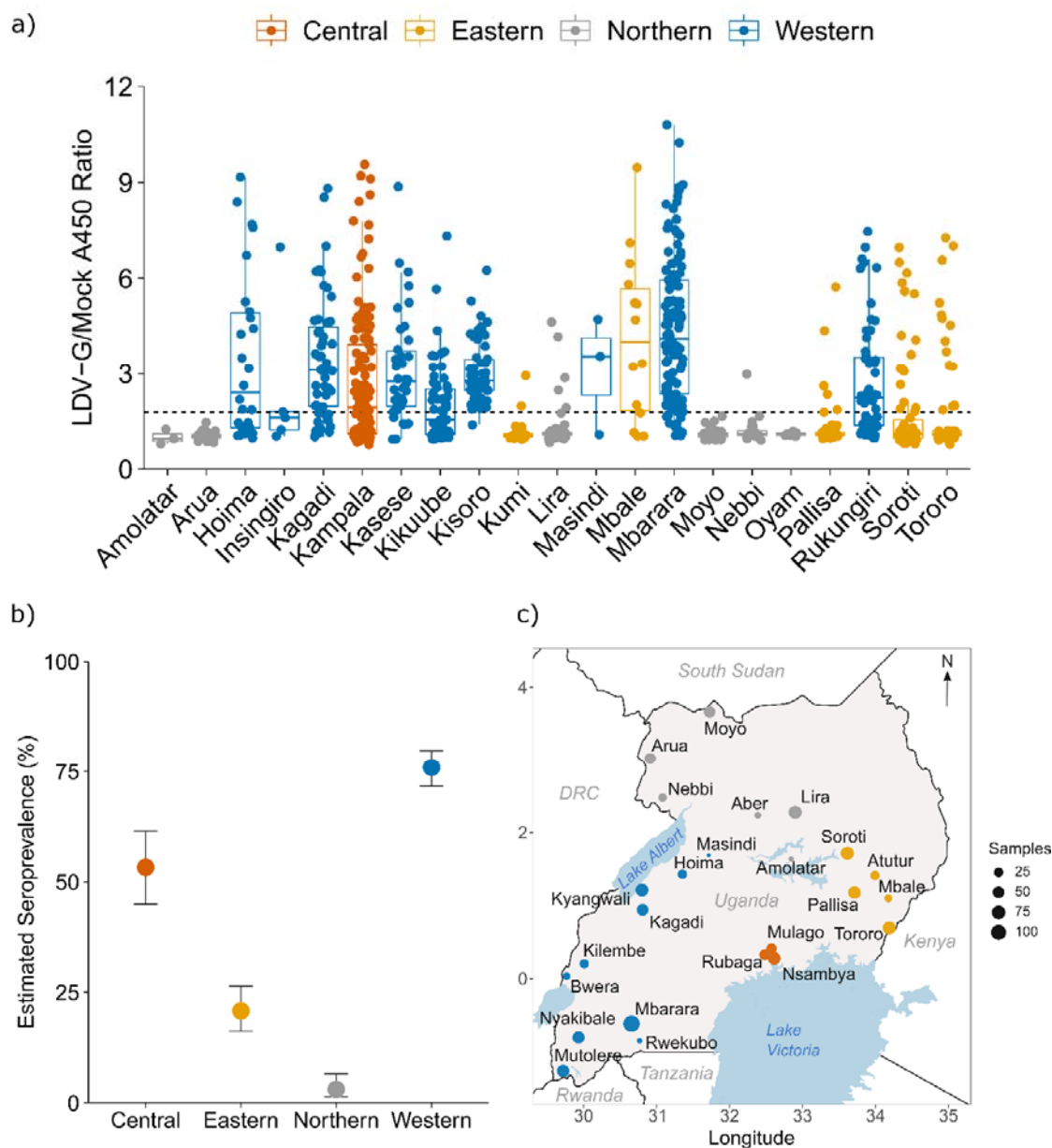
152 **Figure 3. Co-localisation of LDV-G and human IgG by immunocytochemistry**

153 BHK-21 cells were transfected with the mammalian expression plasmid VR1012-LDV-Ghis then fixed and  
154 permeabilised. Primary antibodies were a rabbit anti-his and human serum samples diluted 1:200. Secondary  
155 antibodies were an Alexa Fluor 488 conjugated anti-rabbit antibody (green) and anti-human IgG (red). Each  
156 sample was tested in duplicate. A representative experiment including the Ugandan patient serum and positive  
157 and negative samples as tested by LDV-G ELISA is shown. Three biological replicates were performed for each  
158 sample, with the exception of sera from the index patient where a single replicate was performed owing to  
159 limited sample availability.

160 **High population exposure to LDV in Uganda**

161 To investigate exposure to LDV in the Ugandan population, we screened a cohort of stored  
162 serum samples collected in 2016 as part of the Ugandan national HIV antenatal  
163 seroprevalence study (ANC-2016) by LDV-G ELISA. The cohort comprised a random  
164 selection of sera from 997 women of childbearing age (median age 25, IQR 21-29, range 15-  
165 48) from 24 locations across Uganda (Figure 4).

166 There was heterogeneity in reactivity of serum samples by LDV-G ELISA between the four  
167 administrative regions of Uganda (Figure 4). Notably, samples from Northern sites  
168 demonstrated low reactivity, in particular the sites in the West-Nile sub-region; Arua, Moyo  
169 and Nebbi. As these sites clustered spatially and shared a significantly lower reactivity when  
170 compared with other regions, they were considered to represent a low-prevalence population  
171 and were used to estimate a positive cut-off value for the assay (mean A450 ratio + 3  
172 standard deviations). Applying this cut-off to all samples we estimated seroprevalence for the  
173 different regions of Uganda, with exposure highest in the Western region (75.9%, 95% CI:  
174 71.7-79.8), followed by the Central region (53.3%, 95% CI: 44.9 - 61.4) and the Eastern  
175 region (20.8%, 95% CI: 16.2 - 26.4). Exposure was markedly lower in the Northern region  
176 (3.1%, 95% CI: 1.4 - 6.5, Figure 4b). The only patient-level metadata recorded for the ANC-  
177 2016 dataset was age, which was significantly associated with seropositivity on univariate  
178 analysis (Mann-Whitney U,  $p=0.016$ , Figure S1).



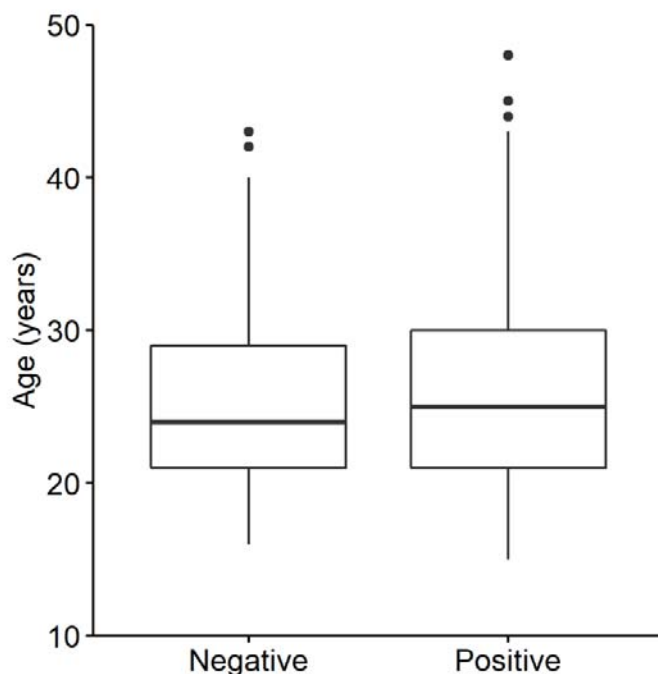
179  
180

181 **Figure 4. Seroprevalence of ledanteviruses in Uganda by LDV-G ELISA**

182 (a) Raw A450 index values by district of Uganda derived from in-house LDV-G capture ELISA. Colours  
 183 represent the regions of Uganda. A450 ratio values were derived by dividing the absorbance value of the test  
 184 well by that of the paired mock well for each sample. Data points represent the mean value for a single  
 185 individual derived from three independent experiments. Centre bars represent the median, box edges the IQR,  
 186 and vertical lines the range of values within 1.5 times the interquartile range. The dashed line represents the  
 187 positive cut-off derived from the mean + 3 × the standard deviation of samples derived from the West Nile  
 188 subregion. (b) Estimated seroprevalence of LDV across the major regions of Uganda based on in-house LDV-G  
 189 ELISA. Seroprevalence is displayed as the proportion of positive cases +/- 95% confidence intervals based on  
 190 the following sample sizes: Central n=137, Eastern n=240, Northern n=196, Western n=424. (c) map of Uganda  
 191 and surrounding countries and major waterbodies showing locations of study sites from the ANC-2016 study.

192 Coloured circles represent locations of study sites with the size of each circle corresponding to the number of  
193 participants from each site tested by the LDV-G ELISA.

194



195

196 **Figure S1. Age by LDV-G ELISA reactivity in the ANC-2016 cohort**

197 Age by seroprevalence as determined by LDV-G ELISA for individuals in the ANC-2016 cohort (n=997).  
198 Median age in the seropositive group was 25 years (n=451, IQR=21-30) compared to 24 in the seronegative  
199 group (n=546, IQR=21-29), Mann-Whitney U, p=0.016. Centre bars represent the median, box edges the IQR,  
200 vertical lines the range (1.5 times the IQR from the upper and lower quartile), and points the outliers.

201

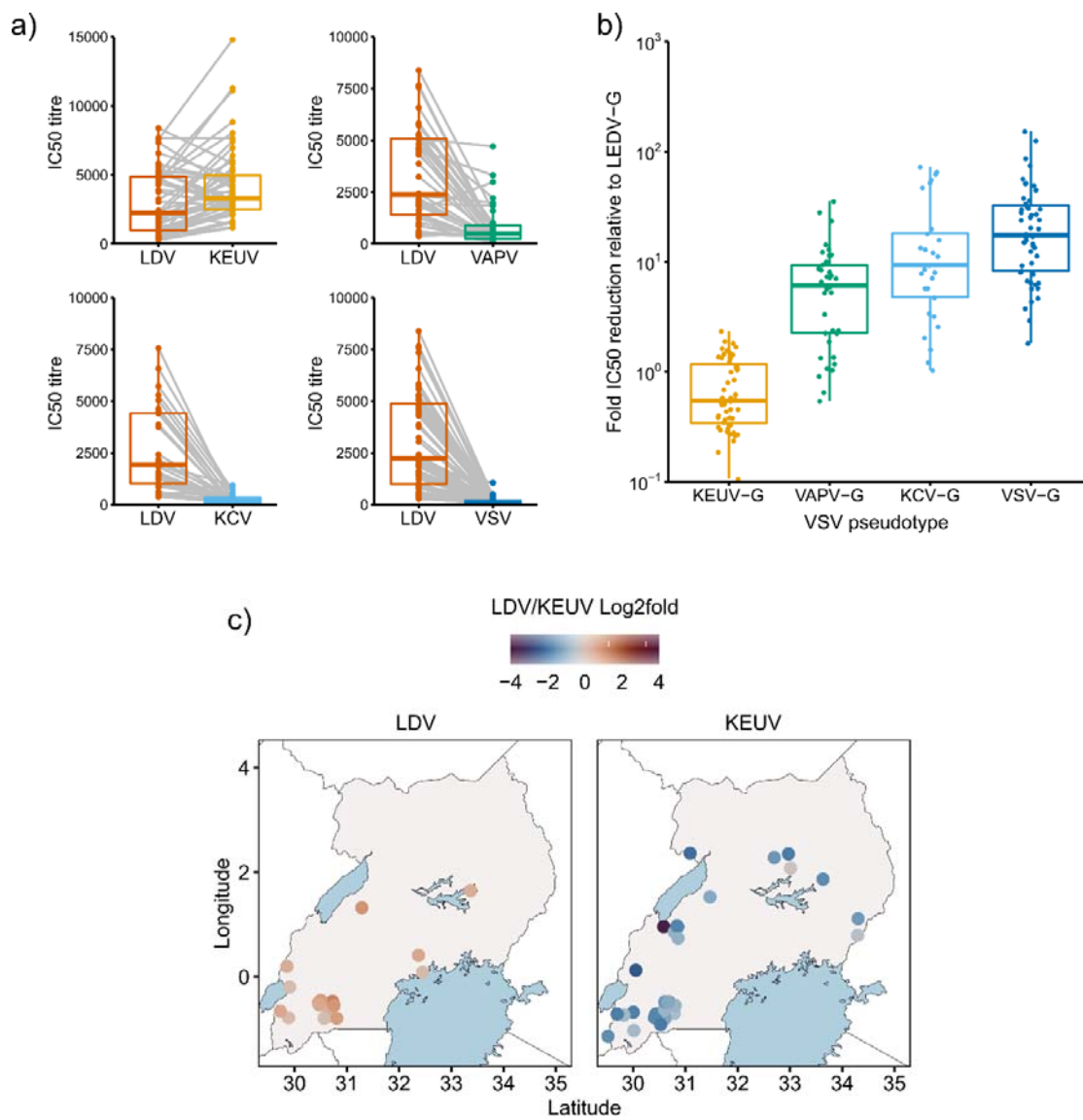
202 **Cross-reactivity of LDV-G reactive sera between phylogroup B ledanteviruses**

203 LDV antisera cross-reacts with Keuraliba virus (KEUV) in complement fixation assays (11,  
204 12). We therefore investigated whether reactivity to LDV-G by ELISA could be explained by  
205 serological cross-specificity to related viruses rather than exposure to LDV. We generated  
206 VSVΔ*Gluc* pseudotypes expressing the glycoproteins of three representative phylogroup B  
207 ledanteviruses; KEUV, Vaprio virus (VAPV) and Kern Canyon Virus (KCV), with  
208 pseudotypes expressing VSV glycoprotein as a negative control. First we determined IC50

209 values for sera from the index patient against each pseudotype. Consistent with prior  
210 infection serum titres were highest against VSVΔ*Gluc*-LDV-G (IC<sub>50</sub>=6864), followed by  
211 VSVΔ*Gluc*-KEUV-G (IC<sub>50</sub>=2981), VSVΔ*Gluc*-KCV-G (IC<sub>50</sub>=983) and VSVΔ*Gluc*-  
212 VAPV-G (IC<sub>50</sub>=141). There was minimal neutralising activity against VSVΔ*Gluc*-VSV-G  
213 (IC<sub>50</sub>=71).

214 Next, we randomly selected sera from individuals in the ANC-2016 cohort that were positive  
215 by LDV-ELISA and determined serum IC<sub>50</sub> values against the pseudotype viruses. There  
216 was significant cross-neutralisation of pseudotypes bearing LDV-G and KEUV-G by ELISA  
217 positive sera (Figure 5a, 5b). Notably, IC<sub>50</sub> titres against KEUV-G were in general higher  
218 than those for LDV-G (median LDV-G IC<sub>50</sub> 2121, median KEUV-G IC<sub>50</sub> 3166.4, median  
219 IC<sub>50</sub> fold change for KEUV-G relative to LDV-G 0.52, IQR: 0.24-1.14). There was cross-  
220 neutralisation to a lesser extent of VAPV-G (median IC<sub>50</sub> to VAPV-G 457.2, median fold  
221 change relative to LDV-G = 6.09, IQR: 2.26 – 9.27) and KCV-G (median IC<sub>50</sub> against KCV-  
222 G 252, median fold change relative to LDV-G = 9.32, IQR = 4.73-18.1). As expected,  
223 pseudotypes bearing VSV-G were neutralised to the lowest extent (median IC<sub>50</sub> 106.4, IQR:  
224 74.7-178.1, median fold change relative to LDV-G = 17.5, IQR: 8.21-32.5). To determine if  
225 there was an obvious spatial pattern differentiating the LDV-G ELISA positive sera based on  
226 their IC<sub>50</sub> against LDV-G or KEUV-G, the LDV/KEUV fold change in neutralisation was  
227 expressed as log<sub>2</sub> fold change (Figure 5c). There was no obvious geographical pattern  
228 consistent with spatial restriction of sera with higher IC<sub>50</sub> to either LDV or KEUV.





229

230 **Figure 5. LDV-G ELISA positive serum exhibits serological cross-reactivity within *Ledantevirus*.**

231 a) IC<sub>50</sub> titres by viral pseudotype for ANC-2016 serum samples positive by LDV-G ELISA. Between group  
 232 differences were tested by paired Wilcoxon test. a; n = 60, median LDV IC<sub>50</sub> 2121, median KEUV IC<sub>50</sub> 3166.4  
 233 , p < 0.001. b; n = 46, p < 0.001. c; n = 37, median LDV IC<sub>50</sub> 1863, median KCV IC<sub>50</sub> 252, p < 0.001. d; n = 60,  
 234 p < 0.001. Centre bars represent the median, box edges the IQR, and vertical lines the range of values within 1.5  
 235 times the interquartile range. Data points represent the IC<sub>50</sub> titre of each patient sample as determined by 4-  
 236 parameter logistic regression curves derived from 2 biological replicates. Grey lines link the titres for an  
 237 individual against different pseudotypes. B) Fold reduction in IC<sub>50</sub> titre of sera against each VSVΔ*Gluc*  
 238 pseudotype relative to VSV(LDV-G). VSV(KEUV-G) (n = 52, median = 0.54, IQR 0.33-1.14), VSV(VAPV-G)  
 239 (n = 38, median = 6.09, IQR = 2.26-9.27), VSV(KCV-G) (n = 29, median = 9.32, IQR = 4.73-18.1) and  
 240 VSV(VSV-G) (n = 52, median = 17.5, IQR = 8.21-32.5). Significance was determined by Kruskal-Wallis rank-  
 241 sum test (H = 111.31, df = 3, p < 0.001). c) Relative distribution of sera neutralising major *Ledantevirus* species  
 242 in Uganda. Colour intensity represents log<sub>2</sub> fold change in neutralising titre (LDV/KEUV IC<sub>50</sub>) for ELISA  
 243 positive samples. Panel A; Samples with LDV titre greater than KEUV (n = 18). Panel B; Samples with KEUV  
 244 titre greater than LDV (n=34). Data points represent individual patients. Exact locations are jittered to avoid  
 245 overlaps.

## 246 **Environmental correlates of LDV-G seropositivity in Uganda**

247 As there was evidence of significant regional variation in seropositivity by LDV-G ELISA  
248 within Uganda, the influence of environmental variation at the ANC-2016 study sites was  
249 investigated by general linear mixed modelling (GLMM). Climatic (relating to rainfall and  
250 temperature), geographical (elevation and tree cover) and livestock density variables were  
251 determined for each study site. Multicollinear variables were identified and excluded and a  
252 binomial GLMM constructed with the effect of recruitment site controlled by its inclusion as  
253 a random effect. The age of individual participants was included as an explanatory variable.  
254 Model selection based on AICc showed patient age (OR 1.04  $p = 0.004$ ), forest cover (OR  
255 1.89  $p = 0.016$ ) and the climatic variable isothermality, which describes daily and seasonal  
256 temperature fluctuations, (OR 4.81  $p < 0.001$ ) were positively associated with seroprevalence  
257 (Table 1). Notably, no effect of livestock density was observed.

258

	Estimate	Std. err	Z value	Pr(> z )	OR	95% conf int.
(Intercept)	-1.59483	0.45667	-3.492	<0.001	0.21	0.08 – 0.50
Age	0.04232	0.01472	2.875	0.004	1.04	1.01 – 1.07
Forest cover	0.63832	0.26530	2.406	0.016	1.89	1.10 – 3.33
Isothermality	1.57128	0.31730	4.952	<0.001	4.81	2.63 – 9.93
Marginal R <sup>2</sup> : 0.466, Conditional R <sup>2</sup> : 0.596						

259

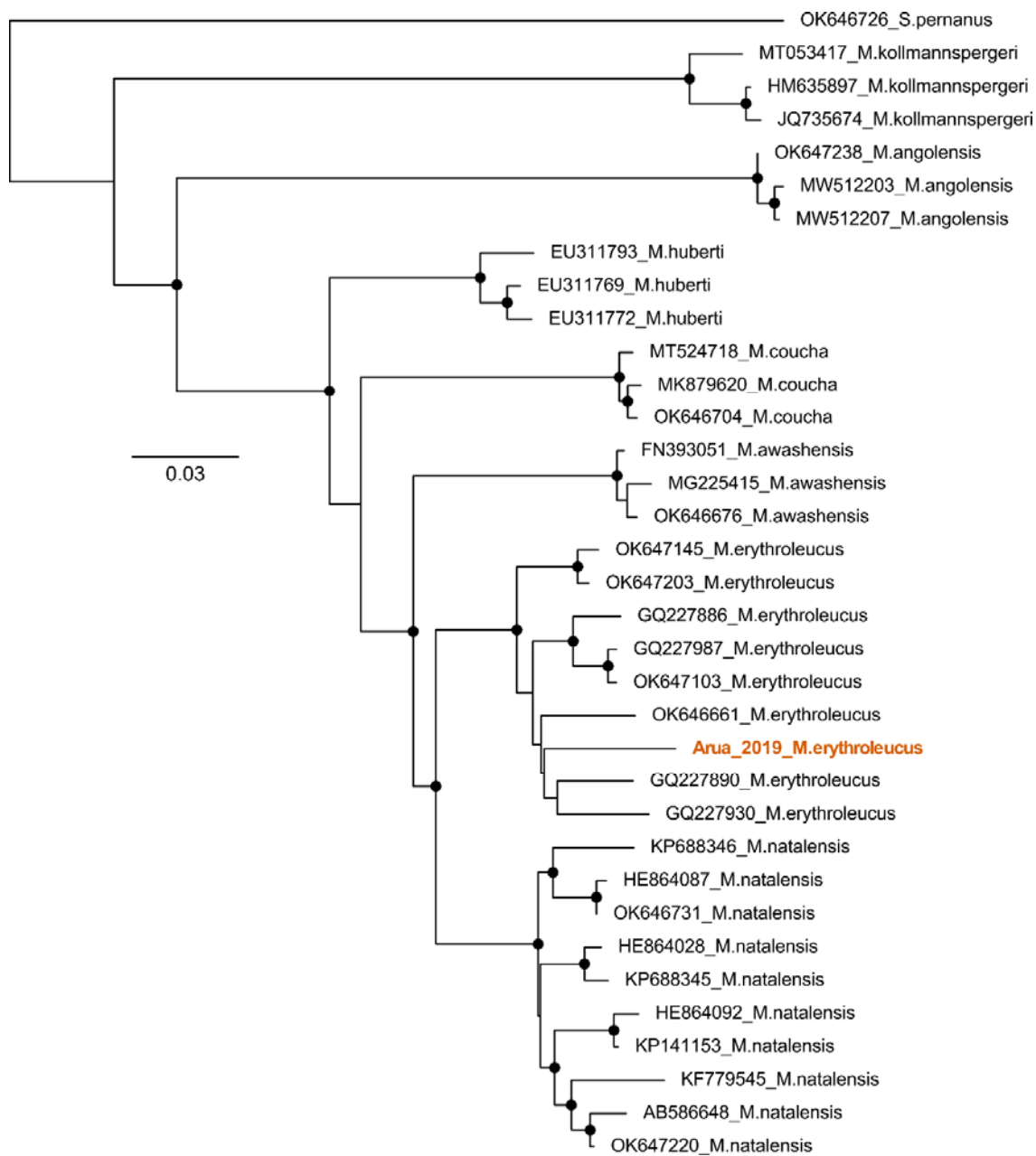
260

## 261 **Environmental evidence of novel ledanteviruses**

262 Whilst LDV has only been detected in humans, KEUV was originally isolated from various  
263 rodent species in Senegal (11). To investigate ecological reservoirs of viral pathogens in  
264 Uganda, we performed live rodent trapping in Arua district and used mNGS to investigate the  
265 blood virome of captured rodents. Over 2233 trap nights we captured 205 rodents (Table S3).

266 We detected genomic evidence of a novel rhabdovirus, which we have named Odro virus, in  
267 blood from a multimammate mouse, *Mastomys erythroleucus*, trapped in Arua district in  
268 2019. Host species was confirmed by sequence determination of the mitochondrial  
269 cytochrome B gene (Figure S2). Sequencing yielded 1383706 reads of which 584 mapped to  
270 viral species within *Ledantevirus* by blastx. Target enrichment sequencing of the NGS library  
271 yielded a further 1296041 reads, of which 11301 mapped to *Ledantevirus*. *De novo* assembly  
272 of the combined reads from metagenomic and target enriched sequencing yielded 13 contigs  
273 corresponding to fragments of the nucleoprotein, phosphoprotein, glycoprotein and RDRP of  
274 a rhabdovirus with high similarity to members of *Ledantevirus* (Table S4). Further iterative  
275 *de novo* assembly using these contigs as scaffolds yielded 88% coverage of the putative  
276 nucleoprotein, 75% of the phosphoprotein, 72% of the glycoprotein, and 89% of the RDRP  
277 with reference to the closest related sequence on the NCBI database. Phylogenetic analysis  
278 placed these fragments into phylogroup B of *Ledantevirus*, alongside LDV and KEUV  
279 (Figure 6). Pairwise amino acid distances in the G and L proteins suggest Odro virus is a  
280 novel species within *Ledantevirus* (Table S5). In addition to coding sequences, there was  
281 adequate sequencing coverage of gene junction sequences for the N-P, P-M and U-L  
282 intergenic regions, demonstrating the conserved rhabdovirus intergenic transcription  
283 termination-polyadenylation (3'-ACUUUUUUU-5'), and transcription initiation sequences  
284 (3'-UUGU~~n~~nUAG-5') (13). Consistent with other phylogroup B ledanteviruses, the  
285 polyadenylation signal at the P-M junction incorporates the termination codon of the  
286 upstream phosphoprotein ORF (Supplementary Figure S3).

287

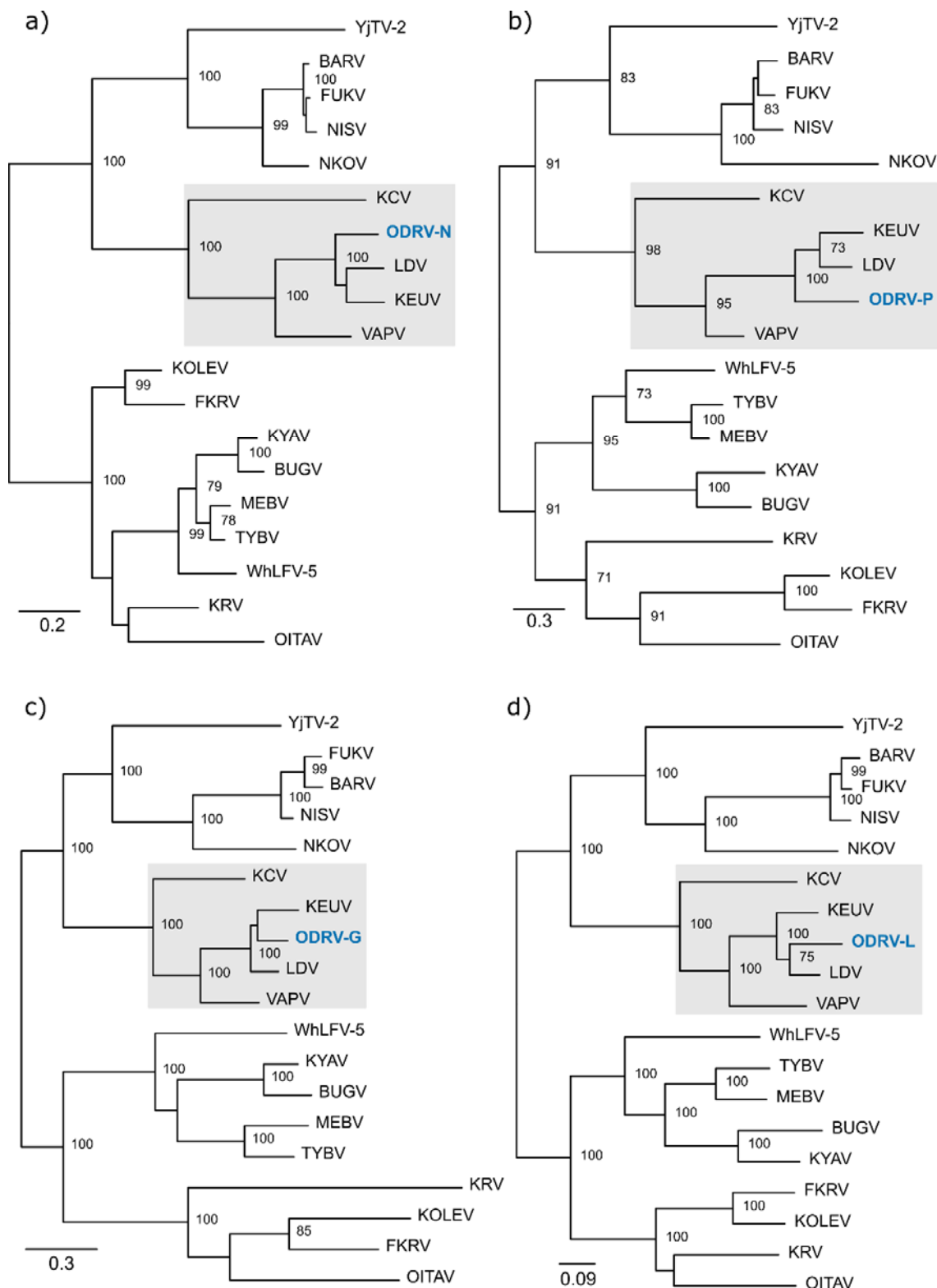


288

289 **Figure S2: Species identification of *Mastomys erythroleucus* by mitochondrial cytochrome B sequencing.**

290 Maximum likelihood phylogeny of mitochondrial cytochrome B sequences derived from representative  
291 *Mastomys* species. The sequence derived from the animal in which Odoro virus was detected is indicated in  
292 orange. *Serengetimys pernanus* is included as an outgroup. Circles indicate nodes with bootstrap support >70.

293



294

295 **Figure 6. Phylogenetic placement of Odro virus genomic fragments within *Ledantevirus***

296 (a) a continuous 373 amino acid fragment corresponding to the Odro virus nucleoprotein. (b) three continuous  
 297 fragments of 87aa, 72aa and 73aa corresponding to the Odro virus phosphoprotein. (c) a continuous 411 amino  
 298 acid fragment corresponding to the Odro virus glycoprotein. (d) five fragments of 65aa, 120aa, 529aa, 814aa

299 and 168aa corresponding to the Odro virus L protein. For each gene alignments were created based on complete  
 300 coding sequences of all members of *Ledantevirus* and gene fragments of Odro virus and maximum likelihood  
 301 phylogenies generated. Node labels represent bootstrap support values for nodes with support >70 based on  
 302 1000 replicates. The grey highlighted areas represent the phylogroup B viruses. ODRV; Odro virus. LDV; Le  
 303 Dantec virus, KEUV; Keuraliba virus, VAPV; Vaprio virus, KCV; Kern Canyon virus, YjTV-2; Yongjia tick  
 304 virus 2, BARV; Barur virus, FUKU; Fukuoka virus, NISV; Nishimuro virus, NKOV; Nkolbisson virus,  
 305 WhLFV-5; Wuhan louse fly virus 5, TYBV; Taiyi bat virus, MEBV; Mount Elgon bat virus, BUGV;  
 306 Bughendera virus, KYAV; Kanyawara virus, FKRV; Fikirini rhabdovirus, KOLEV; Kolente virus, KRV;  
 307 Kumasi rhabdovirus, OITAV; Oita virus.

308

**Table S3: number of rodents by species captured for serum mNGS in Adumi subcounty.**

<i>Aethomys kaiseri</i>	17
<i>Arvicanthus niloticus</i>	13
<i>Crocidura spp.</i>	57
<i>Lemniscomys striatus</i>	6
<i>Mus minutoides</i>	7
<i>Mastomys spp.</i>	26
<i>Rattus rattus</i>	64
<i>Gerbilliscus validus</i>	8
<i>Taterillus emini</i>	2
<i>Thamnomys spp.</i>	2
<i>Zelotomys hildegardeae</i>	3

309  
310

**Table S4: blastx results of de novo contigs from sequencing of blood from *Mastomys erythroleucus***

Contig name	Accession of blastx hit	Protein	Virus	Identity (%)	Length (AA)
merged_contig3927	YP_009361873	polymerase	LDV	78.2	271
merged_contig4193	YP_009361873	polymerase	LDV	87.9	471
merged_contig5160	YP_009361873	polymerase	LDV	67.9	153
merged_contig5201	YP_009362198	glycoprotein	KEUV	76.5	421
merged_contig5497	YP_009361873	polymerase	LDV	86.4	140
merged_contig6266	YP_009361873	polymerase	LDV	76.8	69
merged_contig7735	YP_009362195	nucleoprotein	KEUV	84.3	268
merged_contig8264	YP_009361873	polymerase	LDV	72.1	226
merged_contig10327	YP_009361873	polymerase	LDV	84.1	531
idba-121_14222	YP_009362196	phosphoprotein	KEUV	63.9	72
idba-121_8592	YP_009362194	polymerase	LDV	59.6	47
spades_37480	YP_009361868	nucleoprotein	LDV	70.8	96
spades_75880	YP_009361869	phosphoprotein	LDV	49.4	79

311

**Table S5: pairwise amino acid distance between Odro virus and other phylogroup B ledanteviruses**

Virus (accession)	Gene			
	N	P	G	L
KEUV (KM205021)	21.45	50.00	24.03	21.01
LDV (KM205006)	21.18	46.15	24.57	18.36
VAPV (MG021441)	36.46	58.19	37.71	29.37
KCV (KM204992)	50.40	70.39	47.79	36.46

312

313

a) N-P junction

ODRV 3' ACU - - - - -UUGUA - C UUUUUUUU CA UUG UAGUUAG UAC - - - 5'  
 LDV 3' AUU - - - - CAAUGUA - C UUUUUUUU GG UUG UAGUUAG UAC - - - 5'  
 KEUV 3' ACU - - - - GAGUGUA - C UUUUUUUU AA UUG UAGUUAG UAC - - - 5'  
 VAPV 3' ACUCUUAGAAUGUAC UUUUUUUU AG UUG UAGUUAGUUC UAC 5'  
 KCV 3' ACU - - - - -CGUA - C UUUUUUUU AG UUG UCAGUAG UAC - - - 5'

c) P-M junction

ODRV 3' ACU UUUUUUUU GA UUG UAGUUAGUUC UAC 5'  
 LDV 3' ACU UUUUUUUU GG UUG UAAUUAGUUU UAC 5'  
 KEUV 3' ACU UUUUUUUU AG UUG UAGCUAGUUG UAC 5'  
 VAPV 3' ACU UUUUUUUU AG UUG UAAACAGUUU UAC 5'  
 KCV 3' ACU UUUUUUUU AG UUG UCACUAGUUU UAC 5'



c) U-L junction

ODRV 3' AUU AACAGAACAGUAC UUUUUUUU GA - - - - UUG UUAUUAGUUC AAAGGU - - - UC UAC 5'  
 LDV 3' AUU - - - - CUCUGUAC UUUUUUUU GA - - - G UUG UCAAUAGAUCUUAGAUAGUUU UAC 5'  
 KEUV 3' AUU GGU - - - - -GUAC UUUUUUUU GA - - - - UUG UCCUAGGGU -GAGUU - - - UC UAC 5'  
 VAPV 3' ACU - AU - - - - -GUAC UUUUUUUU G - - - - UUG UAGUUAGUUU -UAACU - - - UU UAC 5'  
 KCV 3' ACU UAC - -GUUAGUAC UUUUUUUU ACAGAGAUGGUUCUCUA UUG UAGGU - - - AG UAC 5'

314

315 **Supplementary Figure S3. Odro virus shares conserved gene junction regions with other group B**  
 316 **ledanteviruses.**

317 Genomic gene junction regions of Odro virus compared with other group B ledanteviruses demonstrating the  
 318 presence of conserved rhabdovirus intergenic polyadenylation and transcription initiation signals. a)  
 319 nucleoprotein – phosphoprotein junction. b) phosphoprotein matrix protein junction. c) accessory protein –  
 320 RNA dependant RNA polymerase junction. ODRV; Odro virus, LDV; Le Dantec virus, KEUV, Keuraliba virus,  
 321 VAPV; Vaprio virus, KCV; Kern Canyon virus.

322

## 323 **Discussion**

324 We describe the second confirmed case of human infection with LDV and demonstrate  
325 evidence of high population exposure to LDV or closely related ledanteviruses in Western  
326 Uganda. In addition, we report genomic evidence of Odro virus, a novel ledantavirus, in  
327 blood from a *Mastomys erythroleucus*, suggesting a rodent reservoir of African  
328 ledanteviruses. *Ledantavirus* is an expanding genus within the family *Rhabdoviridae*, with 18  
329 members designated in the latest ICTV taxonomy update (14). The genus is divided into three  
330 phylogroups based on molecular similarity. In contrast to the related genera *Vesiculovirus*  
331 and *Sigmavirus*, whose members are predominantly associated with invertebrates, most  
332 ledanteviruses have been isolated from vertebrate hosts, or from ectoparasites such as bat  
333 flies (15, 16). LDV was originally isolated from a 10-year-old female who presented to  
334 hospital in Senegal in 1965 with fever (17). In addition to LDV, two other members of  
335 *Ledantavirus* have a reported association with human infection. Serological evidence of  
336 exposure to Kumasi virus was demonstrated in individuals residing in the area in which it  
337 was isolated from a bat in Ghana (18). Nkolbisson virus, originally isolated in 1964 from  
338 mosquitoes in Cameroon, was subsequently reported to have been isolated from human  
339 samples in the Central African Republic, although there are few details available of its  
340 association with human disease (19, 20).

341 Metagenomic sequencing has been increasingly applied in cases of unidentified acute febrile  
342 illness in Africa, leading to detection of several novel viruses including rhabdoviruses (9, 21).  
343 Determining the clinical significance of viral species detected by sequencing alone can  
344 however be challenging (22). Demonstration of seroconversion or detection of viral antigens  
345 in diseased tissue by histology can support the role of a novel agent as a pathogen (23, 24),  
346 but the clinical samples required for such analyses are often unavailable and can be  
347 challenging to access in low-resource settings. Tissue sampling for histological examination



348 is particularly difficult to obtain in non-fatal cases. Thus, it is often unclear if sporadically  
349 detected agents such as LDV represent pathogens or are incidental to an alternative disease  
350 process. The recent detection of two novel orthobunyaviruses in Uganda, both clearly  
351 associated with febrile illness but for which serological or other pathological parameters have  
352 not yet been reported, are illustrative of this challenge (7, 8). Systematic screening for novel  
353 or emerging pathogens by mNGS at sentinel clinical sites or through national surveillance  
354 systems may allow accurate estimation of their contribution to local infectious disease  
355 epidemiology, and inform the need for development of diagnostic platforms for use beyond  
356 research settings.

357 LDV has been exclusively detected in association with febrile illness and can be classed as a  
358 probable pathogen. To support this, we have shown a specific serological response by the  
359 patient directed against LDV-G, although owing to unavailability of samples from the acute  
360 phase of illness we were unable to demonstrate seroconversion. The patient tested negative  
361 against a panel of pathogens commonly causing febrile illness in this setting. Additionally,  
362 we report widespread population exposure to either LDV itself or other members of  
363 *Ledantevirus* in Uganda, particularly in the Western Region. The high seroprevalence in this  
364 area suggests ledanteviruses may predominantly cause subclinical or mild infection and that  
365 more severe cases such as that described here go unreported or undiagnosed. Notably the  
366 patient in this case made a full recovery. Such viruses are often able to spread undetected by  
367 surveillance systems in low and middle-income countries as demonstrated by the emergence  
368 and expansion of Zika virus (25). Even if resultant disease is in most cases non-severe,  
369 pathogens can exert significant pressure on health systems and cause substantial economic  
370 effects if transmission within human populations increases. Patients in Uganda are able to  
371 purchase antibiotics without a prescription and individuals may self-medicate at the onset of  
372 symptoms before presenting to healthcare if they fail to improve (26). Circulation of viral

373 pathogens causing even mild, self-limiting febrile illness thus serves to drive inappropriate  
374 antimicrobial use and antimicrobial resistance. As we detected LDV in a sample originally  
375 collected in 2012, and our serological data is based on samples from 2016, further  
376 prospective studies are required to determine the current contribution of ledanteviruses to AFI  
377 in Uganda.

378 Most of the LDV-G ELISA positive sera in the ANC-2016 cohort tested against a panel of  
379 ledantevirus glycoproteins had higher IC50 titres against KEUV-G than LDV-G, suggesting  
380 that in addition to LDV, human infection by closely related ledanteviruses may be occurring  
381 in Uganda. Serological cross-reactivity between LDV and KEUV was first noted in the 1980s  
382 (11), with the close serological relationship mirrored by relatively high similarity at the  
383 amino acid level (17). A reservoir species for LDV has not yet been identified, but KEUV  
384 was initially isolated from various rodent genera in Senegal in 1968 including *Mastomys* and  
385 gerbils of genus *Taterea* and *Taterillus*. Consistent with the serological data suggesting a  
386 further member of the LDV serogroup is causing human infection in Uganda, we detected the  
387 partial sequence of a close relative of KEUV and LDV by mNGS of blood sampled from a  
388 *Mastomys erythroleucus* in Arua district. The Odro virus glycoprotein displayed 76.7%  
389 amino acid identity with KEUV and 75.6% with the LDV glycoprotein. This putative novel  
390 *Ledantevirus* is therefore likely to exhibit cross-reactive behaviour in serological assays with  
391 both LDV and KEUV and its presence in a reservoir species known for its close interaction  
392 with humans potentially explains the high seroprevalence based on LDV-G ELISA observed  
393 for some regions in the ANC-2016 cohort. Unfortunately our study is limited by our inability  
394 to determine the complete Odro virus glycoprotein sequence, meaning we were unable to test  
395 for evidence of exposure in the ANC-2016 cohort by pseudotype-based neutralisation.

396 Ledantevirus seroprevalence as measured by LDV-G ELISA was geographically  
397 heterogenous within Uganda, suggesting either regional variation in the presence or

398 prevalence of infection within reservoir species or differences in the dynamics of human  
399 exposure. Notably, Odro virus, a potential driver of human LDV seroprevalence in Uganda,  
400 was discovered in Arua district where human LDV-G seroprevalence was low. One  
401 explanation is that other ledanteviruses, not present in Arua and other low seroprevalence  
402 regions, are contributing to human infection. However, if Odro virus is implicated in  
403 seropositive cases not attributable to LDV, factors other than the presence of the reservoir  
404 host are likely important in mediating the extent of human exposure. As an example, the  
405 practice of hunting of wild animals for food carries a risk of exposure to zoonotic pathogens.  
406 Regional variation in bushmeat consumption exists owing to local preferences, the  
407 availability of wildlife species, and poverty (27, 28). Rodents, particularly *Mastomys*, are  
408 often hunted by male children and adults (29, 30), highlighting a limitation of our use of an  
409 all-female adult cohort for seroprevalence.

410 The ecological sampling reported in this study was relatively limited in scale, and further  
411 study of wildlife in Uganda is required to accurately determine the host range of Odro virus.  
412 *Mastomys* species are important vectors of zoonoses in Africa, and are primarily found in  
413 peri-domestic or bush settings in Arua district rather than in huts or other human habitations,  
414 having been replaced as a domestic species by the expansion of *Rattus rattus* into Africa  
415 throughout the 20<sup>th</sup> century (29, 31). Our analysis of ecological conditions suggests forest  
416 cover and isothermality are associated with increased human exposure to ledanteviruses in  
417 Uganda. *Mastomys* species are known to breed and carry litters throughout the year in the  
418 presence of favourable ecological conditions and the absence of drought (29, 32). It may be  
419 hypothesised that the tendency for increased bushmeat consumption in areas adjacent to  
420 forested areas, combined with stable climatic conditions that permit higher rodent densities  
421 throughout the year could act to increase population exposure to zoonotic disease in these  
422 areas. Alternatively, isothermality and forest cover may simply be correlated with the

423 presence of a greater diversity of wildlife, and an as yet unsampled species may be driving  
424 the higher seropositivity in humans in these regions.

425 In summary we identified LDV as the likely causative agent in a case of acute febrile illness  
426 in Uganda and demonstrated widespread exposure to ledanteviruses in Western, Central and  
427 Eastern regions of the country with evidence of increased exposure with age and in regions  
428 with high isothermality and forest cover. A closely related novel phylogroup B *Ledantevirus*  
429 was detected in a common and geographically widespread rodent species. Identification of  
430 the infectious agents which routinely spill-over into human populations, providing  
431 opportunities for emergence and evolution within human hosts is important for pandemic  
432 preparedness. Further characterisation of the epidemiology of acute febrile illness and  
433 population level serosurveys are required to establish the precise contribution of  
434 ledanteviruses to human disease in Uganda. Our detection of a novel ledantevirus in a  
435 *Mastomys*, together with the earlier detection of KEUV in rodents, is supportive of a rodent  
436 reservoir of African ledanteviruses, which should be confirmed by further study of both  
437 rodent and non-rodent hosts in areas with evidence of human exposure.

438 **Methods**

439 **Ethical approvals**

440 Ethical approval for the AFI study, the use of stored serum samples from the ANC-2016  
441 cohort and for wild rodent collections in Arua district was granted by the Research Ethics  
442 Committee of the Uganda Virus Research Institute (reference numbers GC/127/10/02/19,  
443 GC/127/18/09/662 and GC/127/19/06/662). Research approvals were granted by the Uganda  
444 National Council for Science and Technology (UNCST, reference numbers HS2485 and  
445 HS767).

446

447 **Clinical samples**

448 Acute serum from the patient was collected as part of the acute febrile illness study (10).  
449 Recruitment to the study required written informed consent. A further episode of sampling  
450 for additional sera from this individual occurred in 2018. Samples from the ANC-2016 HIV  
451 surveillance study were collected in 2016 from individuals who had provided informed  
452 consent for the future use of their samples for medical research. ANC-2016 samples were  
453 supplied by Robert Downing (UVRI).

454

455 **Rodent sampling**

456 Rodent sampling and euthanasia was conducted in line with international best practice  
457 guidelines (33). Rodent trapping was performed in three villages in Adumi subcounty;  
458 Ombaci, Oniba, and Sua, in April 2019. Rodents were trapped in Sherman traps (H.B.  
459 Sherman Trap Company) and Tomahawk traps (Tomahawk Live Trap Company) baited with  
460 peanut butter mixed with smoked fish and a piece of sweet potato. Traps were set in four

461 locations in relation to homesteads: indoors (inside sleeping or cooking huts), the peri-  
462 domestic compound area (within five metres of the sleeping and cooking huts), the peri-  
463 domestic bush (within five metres of the homestead edge), and sylvatic (up to 300 metres into  
464 the bush from the centre of the village). Two Sherman and two Tomahawk traps were placed  
465 within the peri-domestic compound area. An additional three Sherman and three Tomahawk  
466 traps were placed in the peri-domestic bush setting. To sample the sylvatic environment one  
467 Tomahawk and one Sherman trap were placed every 20 metres for 300 metres from the  
468 centre of the village in each of the four cardinal directions (North, East, South, West).

469 Traps were set before dusk and opened shortly after dawn. Trapping was conducted over  
470 three nights in Ombaci and Oniba and for two nights in Sua for a total of 2233 trap nights.  
471 Seven traps were lost during the sampling period. Rodents were euthanised in the field by  
472 inhalation of halothane and cervical dislocation, followed immediately by blood sampling by  
473 heart puncture. Animals were speciated using taxonomic keys. Blood samples were  
474 centrifuged and sera heat inactivated and stored at -80 °C until further use.

475

#### 476 **Next generation sequencing**

477 For human and rodent serum samples RNA was extracted from 200 µl of plasma using the  
478 Agencourt RNAdvance Blood Kit (Beckman Coulter) according to the manufacturer's  
479 instructions and eluted in nuclease-free water. RNA was reverse transcribed using Superscript  
480 III (Invitrogen) followed by dsDNA synthesis with NEBNext (New England Biolabs). DNA  
481 libraries were prepared using LTP low-input Library preparation kit (KAPA Biosystems).  
482 Resulting libraries were quantified with the Qubit 3.0 fluorometer (Invitrogen) and their size  
483 determined using a 4200 TapeStation (Agilent). Libraries were pooled in an equimolar ratio  
484 and sequenced on the Illumina MiSeq platform using 150x2 v2 cartridges.

485 Target enrichment of RNA derived from rodent blood was performed using custom-designed  
486 biotinylated RNA probes targeting full genomes of all known arboviruses, including  
487 *Ledantevirus* (Agilent) using manufacturer's instructions. Briefly, dual-indexed DNA library  
488 was incubated overnight at 65°C. Captured DNA was recovered using Streptavidin T1  
489 Dynabeads, washed with provided buffers, and eluted in water. Captured library was  
490 amplified using Illumina primers and sequenced on an Illumina NextSeq 500.

491

#### 492 **Bioinformatic analysis**

493 Raw sequence reads were classified using diamond version 0.8.20.82 in blastx mode (34) and  
494 visualised using Krona (35). *De novo* assembly of the human serum sample was performed  
495 with SPAdes version 3.11.1 (36). The *de novo* contigs were screened using diamond blastx  
496 and confirmed by blastn against the NCBI nr database. The *de novo* assembly was then used  
497 as a reference to align raw sequencing reads using the short read aligner tanoti  
498 (<https://github.com/vbsreenu/Tanoti>). For sequencing of rodent serum *de novo* assembly was  
499 performed using the metavic pipeline (37) and contigs mapping to *Ledantevirus* species were  
500 determined by diamond blastx. The genome organisation of *Ledantevirus* contigs was  
501 manually determined by approximating their position based on their alignment by blastx  
502 against either the LDV or KEUV reference genomes. The resulting construct was used as a  
503 scaffold for further iterative *de novo* assembly with Gapfiller version 1-11 (38). Cytochrome  
504 B sequence barcoding for speciation of *Mastomys* was performed by aligning NGS  
505 sequencing reads to the optimal *Mastomys* mitochondrial reference sequences, with the  
506 optimal reference determined by blastn query of initial consensus sequence against the NCBI  
507 nucleotide database (MZ131552). The subsequent consensus was aligned with a panel of  
508 previously published reference sequences (39).

509

## 510 **Phylogenetic analysis**

511 To determine the phylogenetic placement of the LDV-Uganda L protein sequence recovered  
512 from Ugandan patient, full-length L protein amino acid sequences from representative  
513 rhabdovirus taxa within *Alpharhabdovirinae* were aligned. For phylogenetic placement of  
514 partial amino acid sequences of the putative N, P, G and L genes of the novel rhabdovirus  
515 derived from AV\_R184 the corresponding full-length amino acid sequences for each gene  
516 from all members of *Ledantevirus* were aligned. All alignments were created with mafft  
517 version 6.240 using the L-INS-i method (40). Poorly aligned regions were removed using  
518 trimAl version 1.2 (41). Phylogenetic trees were inferred in RAxML version 8.2.10 (42)  
519 using the LG substitution model (43) with an invariable site plus discrete Gamma model of  
520 rate heterogeneity across sites as determined by Modelfinder implemented in IQ-TREE  
521 version 1.6.12. Substitution models were selected based on the AIC criterion (44). Support  
522 for each node was determined using 1000 bootstrap replicates. Trees were visualised using  
523 figtree version 1.4.4 (<http://tree.bio.ed.ac.uk/software/figtree/>).

524

## 525 **Nested RT-PCR amplification**

526 Plasma RNA was extracted from 140 µl plasma aliquots using QIAamp Viral RNA Mini Kit  
527 (Qiagen Cat No 52904). Next, cDNA was generated in a 30 µl reaction containing RNA,  
528 reverse primer LEDPOL-51R (5'- CAACGCACATATCCTTCATCATCAGC-3') and master  
529 mix reagents (SuperScript IV (SSIV) Reverse Transcriptase Kit, ThermoFisher 18090050)  
530 following manufacturer's recommendations. Nested PCR amplification was carried out using  
531 Platinum™ Taq DNA Polymerase High Fidelity (ThermoFisher cat No 11304029) and the  
532 following primer sets aimed to amplify a 5'-half genome sequence of the metagenomic-



533 inferred LDV strain. This fragment contains nucleoprotein, phosphoprotein, matrix,  
534 glycoprotein genes, and a small polymerase subgenomic fragment .First round amplification  
535 performed with forward primer LEGA5-F (5'-  
536 TTTTCTGGTCTTCTCTTTTTCCTACTGAAA-3') and reverse primer LEDPOL-51R (see  
537 sequence above) generates a 5,706 bp fragment. Second round PCR amplification carried out  
538 with forward primer LDAN-F (5'- ATGGCTAACGAGACAATTTATCGTTTCTC-3') and  
539 reverse primer LEDPOL-52R (5'- GATGCATTCAACATCAACACCATGTCATG-3')  
540 generates a 5,464 bp fragment. A PCR positive control included an RNA extract from the Le  
541 Dantec viral isolate DakHD763 obtained from an infected Senegalese child in 1965.  
542 DakHD773, a BSL-2 rhabdovirus was imported as a lyophilized cultured supernatant derived  
543 from an experimentally infected mouse, as a kind donation from Dr. Thomas Ksiazek from  
544 the University of Texas Medical Branch, Galveston, TX 77555). Experimental positive PCR  
545 reactions visualized in agarose gel electrophoresis were sequenced by Sanger chemistry using  
546 ABI Big dye® Terminator v3.1 kit and ABI 3500 Genetic Analyser (ThermoFisher) per  
547 manufacturer's recommendations.

548

#### 549 **Le Dantec real-time RT-PCR**

550 RNA was extracted from 200µl of plasma using the Agencourt RNAdvance Blood Kit  
551 (Beckman Coulter) or the Qiagen viral minikit (see above) according to the manufacturer's  
552 instructions. Extracted RNA was tested directly for Le Dantec virus by real-time PCR or  
553 stored at -80°C until use. An in-house developed real-time RT-PCR assay targeting a 145 bp  
554 fragment of the LDV glycoprotein gene was develop as a highly sensitive detection assay of  
555 the Le Dantec virus in clinical samples. An aliquot of RNA was mixed in a 30 µl reaction  
556 volume containing reagents from SuperScript III Platinum One-Step Quantitative RT-PCR  
557 System (ThermoFisher Cat No. 11732-020) 0.2 mM forward primer LEDAG-F (5'-

558 GCTTGAAATGCCCTGAAGCT-3'), 0.2 mM reverse primer LEDAG-R (5'-  
559 TCACATCTRGTCAACCATCTTGA-3') and 0.1 mM TaqMan probes labelled with 6-  
560 carboxyfluorescein (FAM) LEDAG-FAM (5'-CCATCACCCCTATGTTTCATCAGGACC-3')  
561 in the presence of 0.05  $\mu$ M ROX<sup>TM</sup> Reference Dye. RT-PCR conditions were those  
562 recommended by the manufacturer for 50 cycles of amplification in an ABI 7500 Real-Time  
563 PCR System (Applied Biosystems).

564

### 565 **LDV-G ELISA**

566 The development of the LDV-G ELISA has been described previously (10). Briefly, the  
567 LDV-G transmembrane domain was identified using TMHMM (45). The predicted  
568 ectodomain of LDV-G was cloned into the secretory mammalian expression vector pHLSec  
569 containing a C-terminal 6xHistidine tag. Human embryonic kidney cells (HEK 293T),  
570 maintained in Dulbecco modified Eagle's medium supplemented with 100 IU/ml penicillin,  
571 100  $\mu$ g/ml streptomycin, 2 mM glutamine and 10% foetal bovine serum were transfected  
572 using Fugene 6 (Promega) and cell supernatant containing secreted LDV glycoprotein was  
573 harvested at 48 hours post transfection. Presence of His-tagged protein was confirmed by  
574 Western Blotting against 6xHis.

575 ELISA plates (Thermo Scientific) were coated with rabbit polyclonal anti-6xHis antibody  
576 (Abcam) diluted to 1:1000 in basic coating buffer and incubated overnight at 4 °C. The next  
577 day plates were blocked with 2.5% BSA in PBS at 37 °C for 1 hour. Cell supernatant  
578 containing LDV-G-His or mock protein was applied to the wells followed by human serum  
579 diluted 1:50 in blocking buffer or purified IgG for 1 hour at 37 °C. This was followed by goat  
580 anti-human IgG peroxidase antibody (Sigma) at room temperature for 1 hour. Wells were  
581 washed after each step with 0.1% Tween-20 in PBS. The ELISA reaction was developed

582 using TMB substrate (Thermo Scientific), stopped with 0.16M sulphuric acid and absorbance  
583 read at 450 nm using a Pherastar FS plate reader (BMG Labtech). Each sample was tested in  
584 a test well containing the truncated LDV-G-His protein as the capture antigen and, to identify  
585 nonspecific binding, a mock well with supernatant derived from mock-infected cells  
586 transfected with empty plasmid vector. The test result was reported as the OD450 reading  
587 from the sample test well divided by that of the sample mock well to give a test/mock OD450  
588 ratio for each sample. For each serum sample the ELISA was repeated on at least three  
589 occasions on separate plates. Each plate included known positive and negative control  
590 samples, as well as a blocking buffer only sample. All patient samples were heat-inactivated  
591 by incubation at 60 °C for 30 minutes before use in the assay.

592

### 593 **Production of VSV $\Delta$ Gluc pseudoparticles**

594 For expression of the KEUV, KCV and VAPV glycoproteins (KEUV-G, KCV-G and VAPV-  
595 G), complete coding sequences were derived from GenBank accession numbers  
596 NC\_034540.1, NC\_034451.1, and NC\_043538.1 and synthesised with the addition of a SalI  
597 restriction site and Kozak sequence immediately before the 5' initiation codon and a NotI  
598 restriction site immediately following the 3' stop codon (GeneWiz/BioBasic). For expression  
599 of LDV-G the glycoprotein sequence derived from the Ugandan patient was codon optimised  
600 and synthesised (Eurofins genomics). Coding sequences of LDV-G, KEUV-G, VAPV-G and  
601 KCV-G were then cloned into the multiple cloning site of mammalian expression vector  
602 VR1012 using the SalI and NotI restriction sites, or the NotI and BglIII restriction sites in the  
603 case of LDV-G, to create VR1012-LDV-G, VR1012-KEUV-G, VR1012-VAPV-G and  
604 VR1012-KCV-G. Plasmid pMDG was used to express VSV-G.

605 For each pseudotype virus  $2 \times 10^6$  HEK293T cells were seeded into 10 cm dishes in 10 ml  
606 DMEM supplemented with 10% FCS. The next day cells were transfected with the relevant  
607 glycoprotein expression plasmid using Fugene 6 (Promega) as per the manufacturer's  
608 instructions. One 10cm dish was not transfected to serve as a no glycoprotein control. Plates  
609 were incubated overnight at 37 °C in a 5% CO<sub>2</sub> atmosphere. The next day plates were  
610 removed from the incubator and infected with 15ul of  $1.15 \times 10^7$  TCID<sub>50</sub>/ml VSVΔ*Gluc*-  
611 VSV-G (corresponding to an approximate MOI of 0.03) and returned to the incubator for 3  
612 hours. Media was discarded and cells washed three times with warmed PBS before the  
613 addition of 10 mls DMEM with 10 % FCS. The cells were re-incubated for a further 72 hours  
614 after which supernatants containing pseudotype virus were harvested, passed through a 0.45  
615 μm filter (Starlab) and stored at -80 °C before further use.

616

#### 617 **Pseudotype-based neutralisation assay**

618 HEK293T cells were grown in a 75 cm<sup>2</sup> flask until 70% confluent. Media was removed and  
619 cells washed with 2 ml trypsin before being resuspended in 8 ml DMEM + 10% FBS. Cells  
620 were then plated at a density of  $5 \times 10^4$  cells per well in a volume of 50 μl per well in each  
621 well of a white 96-well plate.

622 Serum samples were serially diluted in DMEM supplemented with 10% FBS. Pseudotype  
623 preparations were mixed 1:1 with each dilution and incubated at 37 °C for 30 minutes. 50 μl  
624 of the pseudoparticle/sera mixture was added to the plated cells in triplicate for a final  
625 dilution series ranging from 1:32 to 1:131,072 and a “no sera” control well. Plates were  
626 incubated at 37 °C, 5 % CO<sub>2</sub> for 24 hours before addition of 75 μl Steadylite luciferase  
627 substrate solution (Perkin Elmer) to each well. Plates were incubated at room temperature for  
628 10 minutes before luminescence was read using a Chameleon V luminometer (Hidex).

629 Luciferase activity for each dilution was derived from the mean of three plate replicates  
630 normalised to the average for the no sera control wells. Percentage neutralisation for each  
631 dilution was calculated by subtracting the normalised luciferase activity from 1 and  
632 multiplying by 100. The IC50 value for each serum sample and pseudotype combination was  
633 determined by interpolation of a four-parameter logistic regression curve between percentage  
634 neutralisation and the reciprocal serum dilution. Regression curves were fitted in R version  
635 3.6.3/R Studio version 1.2.5033 using the package drc version 3.0-1 with an upper constraint  
636 of 100 using a minimum of 2 biological replicates to fit each curve.

637

### 638 **Immunocytochemistry**

639 The LDV-G sequence was cloned into the expression vector VR1012 with the addition of a  
640 C-terminal 6xHistidine tag to generate VR1012-LDV-G-His. Coverslips (Fisher Scientific)  
641 were added to each well of a 24-well plate (Corning) and 0.5 ml 70 % IMS added to each  
642 well for 20 minutes. IMS was removed and coverslips allowed to air dry before being washed  
643 3 times with 1ml PBS. BHK-21 cells were plated at a density of  $1 \times 10^5$  per well and incubated  
644 at 37 °C, 5% CO<sub>2</sub> until 70% confluent. Cells were transfected with VR1012-LDV-G-His  
645 using Fugene 6 transfection reagent (Promega) and returned to the incubator for 24-48 hours.  
646 Media was removed and cells fixed with 4% formaldehyde before permeabilization with 200  
647 µl 0.1% Triton-X-100 per well for 5 minutes. Cells were then washed three times with 1000  
648 µl PBS. The primary antibody mixture comprised a rabbit polyclonal to 6X His tag antibody  
649 (Abcam) at a dilution of 1:500 mixed with human serum at a dilution of 1:200 with BSA 1%  
650 as the diluent (ThermoFisher Scientific). The primary mixture was added to the cells at a  
651 volume of 200 µl per well and incubated overnight at 4 °C. The next morning cells were  
652 washed three times with PBS and blocked with 1 % BSA for 30 minutes. The secondary

653 antibody mixture was a combination of an Alexa Fluor 488-conjugated goat anti-rabbit IgG  
654 secondary antibody (Invitrogen) at a concentration of 4 µg/ml and an Alexa Fluor 594-  
655 conjugated goat anti-human IgG secondary antibody (Invitrogen) at a concentration of 5  
656 µg/ml diluted in 1% BSA. The secondary antibody mixture was added to wells at a volume of  
657 200 µl and the plate placed on a shaker at room temperature for 45 minutes before three  
658 further 5 minute washes with 1000 µl PBS. Coverslips were stained with Hoechst 33342, at a  
659 concentration of 1:2000 in PBS for 15 minutes. Coverslips were then washed three times with  
660 1000 µl PBS and once with ultrapure water before being mounted on glass slides with 2 µl  
661 CitiFluor AF1 Mountant solution (Electron Microscopy Sciences). Images were acquired  
662 using a LSM 710 confocal microscope (Carl Zeiss Microscopy), and images processed using  
663 ZEN Blue software (Carl Zeiss Microscopy).

664

## 665 **Environmental analysis**

666 The R package raster version 3.4-5 was used to extract bioclimatic, geographical and  
667 livestock data at the locations of the ANC-2016 study sites. Bioclimatic variables were  
668 derived from the Worldclim dataset (46). Tree cover data was extracted from the global forest  
669 change dataset (47). Livestock density data was extracted from the UN FAO dataset  
670 (<https://livestock.geo-wiki.org/home-2/>). Elevation was extracted from the STRM 90m DEM  
671 Digital Elevation Dataset (<https://srtm.csi.cgiar.org/>). Variables were averaged across a 10km  
672 radius around each study site. Multicollinearity between variables was tested by creating a  
673 correlation matrix by the method of Spearman and subsequently a hierarchal clustering  
674 dendrogram to identify highly correlated groups of variables. A correlation coefficient cut-off  
675 of 0.6 was used to group variables. One variable from each group was selected to be included  
676 in the global model.

677 Binomial generalised mixed models were constructed in R version 3.6.3/R Studio version  
678 1.2.5033 using package lme4 version 1.1.26. In addition to bioclimatic variables, patient level  
679 age was included in the analysis. To identify the optimal random effects structure an initial  
680 model was created combining all fixed effects terms and a random effects term of either site,  
681 region or a nested random effect (region/site), with the optimal random effect term selected  
682 based on the AICc criterion (48). The best fitting model was then identified using the dredge  
683 function package MuMIn version 1.43.17 (49), with the AICc criterion used to select the final  
684 model. The response variable was patient level positivity on LDV ELISA , with the final  
685 fixed effect explanatory variables being patient age, forest cover, and isothermality (BIO3),  
686 with a random effect of site (24 levels). Continuous response variables were normalised. The  
687 model was called as  $\text{glmer}(\text{seropositivity}) \sim \text{age} + \text{scale}(\text{forest\_cover}, \text{center}=(\text{mean forest}$   
688  $\text{cover})) + \text{scale}(\text{isothermality}, \text{center}=(\text{mean isothermality})) + (1|\text{site}), \text{family}=\text{binomial}$ .  
689 Model assumptions were tested using the R package DHARMA version 0.4.4 (50).  
690 Conditional and marginal R<sup>2</sup> measures were computed using R package MuMIn.

691

## 692 **Funding**

693 This work was funded by awards from the Wellcome Trust (102789/Z/13/Z and  
694 217221/Z/19/Z) and the MRC (MC\_UU\_12014/8 and MC\_ST\_U17020).

695

## 696 **Data availability**

697 Consensus viral genome and mitochondrial cytochrome B sequences have been submitted to  
698 GenBank under accession numbers OQ077988, OR497399 and OR607590. Raw sequencing  
699 read files derived from rodent blood have been submitted to NCBI sequence read archive  
700 (BioProject ID PRJNA1013481, Biosample accession SAMN37299165).

701

702 **Inclusion statement**

703 The research was conducted as part of a partnership between the MRC-University of  
704 Glasgow Centre for Virus Research, the MRC/UVRI & LSHTM Uganda Research Unit and  
705 the Uganda Virus Research Institute. The study was designed and conducted by researchers  
706 based in Uganda and the UK. Field and laboratory work were performed by trained  
707 individuals following a full health and safety risk assessment. Field PPE was employed when  
708 handling small mammals (rubber boots, coveralls, goggles and gloves). Laboratory samples  
709 were handled in a class 2 biological safety cabinet. The study was reviewed by the UVRI  
710 research ethics committee and the UNCST. All sampling and fieldwork protocols relating to  
711 wildlife sampling were based on established protocols in use at UVRI. All individuals  
712 meeting the criteria for authorship have been included as authors.



## 713 References

- 714 1. Maze MJ, Bassat Q, Feasey NA, Mandomando I, Musicha P, Crump JA. The epidemiology of  
715 febrile illness in sub-Saharan Africa: implications for diagnosis and management. *Clinical*  
716 *Microbiology and Infection*. 2018;24(8):808-14.
- 717 2. Fleming KA, Horton S, Wilson ML, Atun R, DeStigter K, Flanigan J, et al. The Lancet  
718 Commission on diagnostics: transforming access to diagnostics. *The Lancet*. 2021.
- 719 3. Crump JA, Morrissey AB, Nicholson WL, Massung RF, Stoddard RA, Galloway RL, et al.  
720 Etiology of Severe Non-malaria Febrile Illness in Northern Tanzania: A Prospective Cohort Study.  
721 *PLOS Neglected Tropical Diseases*. 2013;7(7):e2324.
- 722 4. Dunn RR, Davies TJ, Harris NC, Gavin MC. Global drivers of human pathogen richness and  
723 prevalence. *Proc Biol Sci*. 2010;277(1694):2587-95.
- 724 5. O'Brien KL, Baggett HC, Brooks WA, Feikin DR, Hammitt LL, Higdon MM, et al. Causes of  
725 severe pneumonia requiring hospital admission in children without HIV infection from Africa and  
726 Asia: the PERCH multi-country case-control study. *The Lancet*. 2019;394(10200):757-79.
- 727 6. Prasad N, Murdoch DR, Reyburn H, Crump JA. Etiology of Severe Febrile Illness in Low- and  
728 Middle-Income Countries: A Systematic Review. *PLOS ONE*. 2015;10(6):e0127962.
- 729 7. Ramesh A, Nakielny S, Hsu J, Kyohere M, Byaruhanga O, de Bourcy C, et al. Metagenomic  
730 next-generation sequencing of samples from pediatric febrile illness in Tororo, Uganda. *PLOS ONE*.  
731 2019;14(6):e0218318.
- 732 8. Edridge AWD, Deijs M, Namazzi R, Cristella C, Jebbink MF, Maurer I, et al. Novel  
733 Orthobunyavirus Identified in the Cerebrospinal Fluid of a Ugandan Child With Severe  
734 Encephalopathy. *Clinical Infectious Diseases*. 2019;68(1):139-42.
- 735 9. Grard G, Fair JN, Lee D, Slikas E, Steffen I, Muyembe J-J, et al. A Novel Rhabdovirus  
736 Associated with Acute Hemorrhagic Fever in Central Africa. *PLOS Pathogens*. 2012;8(9):e1002924.
- 737 10. Ashraf S, Jerome H, Bugembe DL, Ssemwanga D, Byaruhanga T, Kayiwa JT, et al. Emerging  
738 viruses are an underestimated cause of undiagnosed febrile illness in Uganda. *medRxiv*.  
739 2023:2023.04.27.23288465.
- 740 11. Cropp CB, Prange WC, Monath TP. LeDantec virus: identification as a rhabdovirus associated  
741 with human infection and formation of a new serogroup. *J Gen Virol*. 1985;66 ( Pt 12):2749-54.
- 742 12. Calisher CH, Karabatsos N, Zeller H, Digoutte JP, Tesh RB, Shope RE, et al. Antigenic  
743 Relationships among Rhabdoviruses from Vertebrates and Hematophagous Arthropods.  
744 *Intervirology*. 1989;30(5):241-57.
- 745 13. Hinzman EE, Barr JN, Wertz GW. Selection for gene junction sequences important for VSV  
746 transcription. *Virology*. 2008;380(2):379-87.
- 747 14. Walker PJ, Freitas-Astúa J, Bejerman N, Blasdel KR, Breyta R, Dietzgen RG, et al. ICTV Virus  
748 Taxonomy Profile: Rhabdoviridae 2022. *Journal of General Virology*. 2022;103(6).
- 749 15. Bennett AJ, Paskey AC, Kuhn JH, Bishop-Lilly KA, Goldberg TL. Diversity, Transmission, and  
750 Cophylogeny of Ledanteviruses (Rhabdoviridae: Ledantavirus) and Nycteribiid Bat Flies Parasitizing  
751 Angolan Soft-Furred Fruit Bats in Bundibugyo District, Uganda. *Microorganisms*. 2020;8(5):750.
- 752 16. Lelli D, Prosperi A, Moreno A, Chiapponi C, Gibellini AM, De Benedictis P, et al. Isolation of a  
753 novel Rhabdovirus from an insectivorous bat (*Pipistrellus kuhlii*) in Italy. *Virology Journal*.  
754 2018;15(1):37.
- 755 17. Blasdel KR, Guzman H, Widen SG, Firth C, Wood TG, Holmes EC, et al. Ledantavirus: A  
756 Proposed New Genus in the Rhabdoviridae has a Strong Ecological Association with Bats. *The*  
757 *American Society of Tropical Medicine and Hygiene*. 2015;92(2):405-10.
- 758 18. Binger T, Annan A, Drexler JF, Müller MA, Kallies R, Adankwah E, et al. A Novel Rhabdovirus  
759 Isolated from the Straw-Colored Fruit Bat *Eidolon helvum*, with Signs of Antibodies in Swine and  
760 Humans. *Journal of Virology*. 2015;89(8):4588-97.
- 761 19. Ndiaye M, Saluzzo JF, Digoutte JP, Mattei X. Identification du virus Nkolbisson par  
762 microscopie électronique. *Annales de l'Institut Pasteur / Virologie*. 1987;138(4):517-21.

- 763 20. Salaun JJ, Rickenbach A, Brès P, Brottes H, Germain M, Eouzan JP, et al. [The Nkolbisson virus  
764 (YM 31-65), a new prototype of arbovirus isolated in Cameroun]. *Ann Inst Pasteur (Paris)*.  
765 1969;116(2):254-60.
- 766 21. Stremlau MH, Andersen KG, Folarin OA, Grove JN, Odia I, Ehiane PE, et al. Discovery of Novel  
767 Rhabdoviruses in the Blood of Healthy Individuals from West Africa. *PLOS Neglected Tropical*  
768 *Diseases*. 2015;9(3):e0003631.
- 769 22. Lipkin WI. Microbe hunting. *Microbiol Mol Biol Rev*. 2010;74(3):363-77.
- 770 23. Bennett AJ, Paskey AC, Ebinger A, Pfaff F, Priemer G, Höper D, et al. Relatives of rubella virus  
771 in diverse mammals. *Nature*. 2020;586(7829):424-8.
- 772 24. Albariño CG, Foltzer M, Towner JS, Rowe LA, Campbello S, Jaramillo CM, et al. Novel  
773 paramyxovirus associated with severe acute febrile disease, South Sudan and Uganda, 2012.  
774 *Emerging infectious diseases*. 2014;20(2):211-6.
- 775 25. Faria NR, Quick J, Claro IM, Thézé J, de Jesus JG, Giovanetti M, et al. Establishment and  
776 cryptic transmission of Zika virus in Brazil and the Americas. *Nature*. 2017;546(7658):406-10.
- 777 26. Musoke D, Namata C, Lubega GB, Kitutu FE, Mugisha L, Amir S, et al. Access, use and  
778 disposal of antimicrobials among humans and animals in Wakiso district, Uganda: a qualitative study.  
779 *Journal of Pharmaceutical Policy and Practice*. 2021;14(1):69.
- 780 27. Meulen JT, Lukashevich I, Sidibe K, Inapogui A, Marx M, Dorlemann A, et al. Hunting of  
781 Peridomestic Rodents and Consumption of Their Meat as Possible Risk Factors for Rodent-to-Human  
782 Transmission of Lassa Virus in the Republic of Guinea. *The American Journal of Tropical Medicine*  
783 *and Hygiene*. 1996;55(6):661-6.
- 784 28. Friant S, Ayambem WA, Alobi AO, Ifebueke NM, Otukpa OM, Ogar DA, et al. Eating  
785 Bushmeat Improves Food Security in a Biodiversity and Infectious Disease “Hotspot”. *EcoHealth*.  
786 2020;17(1):125-38.
- 787 29. Isaäcson M. The ecology of *Praomys (Mastomys) natalensis* in southern Africa. *Bull World*  
788 *Health Organ*. 1975;52(4-6):629-36.
- 789 30. Douno M, Asampong E, Magassouba NF, Fichet-Calvet E, Almudena MS. Hunting and  
790 consumption of rodents by children in the Lassa fever endemic area of Faranah, Guinea. *PLOS*  
791 *Neglected Tropical Diseases*. 2021;15(3):e0009212.
- 792 31. Ensore RE, Babi N, Amatre G, Atiku L, Eisen RJ, Pepin KM, et al. The changing triad of plague  
793 in Uganda: invasive black rats (*Rattus rattus*), indigenous small mammals, and their fleas. *Journal of*  
794 *Vector Ecology*. 2020;45(2):333-55.
- 795 32. Leirs H, Verhagen R, Verheyen W. The basis of reproductive seasonality in *Mastomys* rats  
796 (Rodentia: Muridae) in Tanzania. *Journal of Tropical Ecology*. 1994;10(1):55-66.
- 797 33. AMVA. AVMA guidelines for the euthanasia of animals. Illinois: Schaumburg; 2020.
- 798 34. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nature*  
799 *Methods*. 2015;12(1):59-60.
- 800 35. Ondov BD, Bergman NH, Phillippy AM. Interactive metagenomic visualization in a Web  
801 browser. *BMC Bioinformatics*. 2011;12:385-.
- 802 36. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new  
803 genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*.  
804 2012;19(5):455-77.
- 805 37. Modha S, Hughes J, Bianco G, Ferguson HM, Helm B, Tong L, et al. Metaviromics Reveals  
806 Unknown Viral Diversity in the Biting Midge *Culicoides impunctatus*. *Viruses*. 2019;11(9):865.
- 807 38. Nadalin F, Vezzi F, Policriti A. GapFiller: a de novo assembly approach to fill the gap within  
808 paired reads. *BMC Bioinformatics*. 2012;13 Suppl 14(Suppl 14):S8-S.
- 809 39. Hánová A, Konečný A, Mikula O, Bryjová A, Šumbera R, Bryja J. Diversity, distribution, and  
810 evolutionary history of the most studied African rodents, multimammate mice of the genus  
811 *Mastomys*: An overview after a quarter of century of using DNA sequencing. *Journal of Zoological*  
812 *Systematics and Evolutionary Research*. 2021;59(8):2500-18.

- 813 40. Katoh K, Kuma K-i, Toh H, Miyata T. MAFFT version 5: improvement in accuracy of multiple  
814 sequence alignment. *Nucleic Acids Res.* 2005;33(2):511-8.
- 815 41. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment  
816 trimming in large-scale phylogenetic analyses. *Bioinformatics.* 2009;25(15):1972-3.
- 817 42. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large  
818 phylogenies. *Bioinformatics (Oxford, England).* 2014;30(9):1312-3.
- 819 43. Le SQ, Gascuel O. An Improved General Amino Acid Replacement Matrix. *Molecular Biology  
820 and Evolution.* 2008;25(7):1307-20.
- 821 44. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermini LS. ModelFinder: fast  
822 model selection for accurate phylogenetic estimates. *Nature Methods.* 2017;14(6):587-9.
- 823 45. Sonnhammer EL, von Heijne G, Krogh A. A hidden Markov model for predicting  
824 transmembrane helices in protein sequences. *Proc Int Conf Intell Syst Mol Biol.* 1998;6:175-82.
- 825 46. Fick SE, Hijmans RJ. WorldClim 2: new 1-km spatial resolution climate surfaces for global  
826 land areas. *International Journal of Climatology.* 2017;37(12):4302-15.
- 827 47. Hansen MC, Potapov PV, Moore R, Hancher M, Turubanova SA, Tyukavina A, et al. High-  
828 Resolution Global Maps of 21st-Century Forest Cover Change. *Science.* 2013;342(6160):850-3.
- 829 48. Akaike H. A new look at the statistical model identification. *IEEE Transactions on Automatic  
830 Control.* 1974;19(6):716-23.
- 831 49. Barton K. MuMIn: Multi-Model Inference (version 1.43.17). 1.43.17 ed2020.
- 832 50. Hartig F. DHARMA: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression  
833 Models. R package version 0.4.4. ed2021.

834