1 **Reconstruction of the cell entry pathway of an extinct virus**

2

3 Lindsey R. Robinson-McCarthy[1,2], Kevin R. McCarthy[3], Matthijs Raaben[4], Silvia Piccinotti[2#], Joppe

4 Nieuwenhuis[4], Sarah H. Stubbs[2], Mark J.G. Bakkers[2], and Sean P. J. Whelan[1,2]*

5 [1]Program in Virology and [2]Department of Microbiology and Immunobiology, Harvard Medical

6 School, Boston, Massachusetts, USA

7 [3]Laboratory of Molecular Medicine, Children's Hospital, Harvard Medical School, Boston,

8 Massachusetts, USA

9 [4]Division of Biochemistry, Netherlands Cancer Institute, Amsterdam, The Netherlands

10 [#]Current address: Department of Stem Cell and Regenerative Biology, Harvard University,

11 Cambridge, MA, USA

12

13 *To whom correspondence should be addressed: swhelan@hms.harvard.edu (SPJW)

14

Abstract

15

16   Endogenous retroviruses (ERVs), remnants of ancient germline infections, comprise 8%

17   of the human genome. The most recently integrated includes human ERV-K (HERV-K) where

18   several envelope (*env*) sequences remain intact. Viral pseudotypes decorated with one of those

19   Envs are infectious. Using a recombinant vesicular stomatitis virus encoding HERV-K Env as its

20   sole attachment and fusion protein (VSV-HERVK) we conducted a genome-wide haploid genetic

21   screen to interrogate the host requirements for infection. This screen identified 11 genes

22   involved in heparan sulfate biosynthesis. Genetic inhibition or chemical removal of heparan

23   sulfate and addition of excess soluble heparan sulfate inhibit infection. Direct binding of

24   heparin to soluble HERVK Env and purified VSV-HERVK defines it as critical for viral attachment.

25   Cell surface bound VSV-HERVK particles are triggered to infect on exposure to acidic-pH,

26   whereas acid pH pretreatment of virions blocks infection. Testing of additional endogenous

27   HERV-K *env* sequences reveals they bind heparin and mediate acid pH triggered fusion. This

28   work reconstructs and defines key steps in the infectious entry pathway of an extinct virus.

29

2

30 ## Author Summary

31   The genomes of all vertebrates are littered with the remains of once exogenous

32   retroviruses. The properties of these ancient viruses that fostered germline colonization and

33   their subsequent inheritance as genetic elements are largely unknown. The viral envelope

34   protein (Env) dictates the cell entry pathway. Here we define host factors involved in the cell-

35   entry of the youngest human ERV, HERV-K. Using a forward genetic screen, we identified

36   heparan sulfate as a critical mediator of productive cell-entry. The abundance of this

37   carbohydrate on almost all cells in the body suggests that HERV-K endogenization was a

38   consequence of a broad tropism and not a specific targeting of germ cells. We demonstrate

39   that multiple HERV-K Env protein encoded in the genome bind heparin. As HERV-K Envs are

40   expressed in some transformed and virus-infected cells as well as during inflammation, it is

41   tempting to speculate that this heparan sulfate binding property could be physiologically

42   relevant during disease.

43

Introduction

44

45   Endogenous retroviruses (ERVs) are remnants of ancient germline infections and

46   comprise approximately 8% of the human genome [1]. The degraded nature of ERV sequences

47   impedes investigation of the properties of the infectious progenitor viruses and the events that

48   led to their endogenization. During evolution, ERV sequences accumulate mutations,

49   consequently the most recently endogenized sequences are the most likely to reflect the

50   properties of the progenitor virus from which they were derived. The most recent human

51   endogenous retroviruses (HERVs) belong to the HERV-K (HML-2) group. Multiple

52   endogenization events resulted in approximately 90 proviral copies and 1,000 solo long

53   terminal repeats (LTRs) in the reference human genome [2]. The HERV-K (HML-2) group is

54   approximately 30-35 million years old [3], with evidence of endogenization as recently as

55   100,000-600,000 years ago [4, 5].

56   Many HERV-K sequences exist as largely intact proviral copies, some of which still

57   encode single functional proteins [6]. While no single locus has been demonstrated to produce

58   an infectious virus, many loci have retained the capacity to produce individual functional

59   proteins. For example, at least one copy, termed HERV-K 108, has retained the capacity to

60   produce an envelope (Env) that can mediate cellular attachment and entry [7]. Two replication-

61   competent infectious clones, Phoenix [8] and HERV-K$_{CON}$ [9] have been reconstructed from

62   consensus sequences comprising the most recently endogenized loci. The reconstructed viruses

63   grow poorly which has hampered efforts to study the biology of their envelope proteins.

64   The processes that govern endogenization are poorly defined. The first virus-cell

4

65   contacts are mediated through viral glycoproteins, which can dictate species, tissue and cellular

66   tropism. We have previously overcome some of the challenges imposed by viral titer by

67   generating an infectious vesicular stomatitis virus (VSV) in which the glycoprotein was replaced

68   by Phoenix Env (VSV-HERVK). Using this virus we determined that HERV-K Env imparts a broad

69   species and tissue tropism [10] and demonstrate that productive infection of mammalian cells

70   requires access to an acidified compartment that is accessed via a dynamin-dependent but

71   clathrin-independent pathway [10]. We also found that proteolytic processing and acid pH are

72   required for HERV-K Env to mediate membrane fusion. A broad species and cell-type tropism

73   was also described for a modified variant of a different ancestral sequence [11]. The broad host

74   range reported in those studies implies that host factors required for HERV-K entry are

75   evolutionarily conserved and ubiquitously expressed.

76   To identify such host factors we performed a genome-wide haploid genetic screen by

77   selecting cells resistant to VSV-HERVK infection. This approach has identified critical host

78   factors required for the entry of several extant viruses, including Ebola, Lassa, Lujo, Andes virus,

79   and Rift Valley fever virus [12-17]. We identify genes involved in heparan sulfate biosynthesis

80   and demonstrate a specific interaction between this glycosaminoglycan and multiple HERV-K

81   envelope proteins. We further show that acid pH is required to trigger membrane fusion by

82   these Envs and is sufficient to mediate infection of cell surface virus and to inactivate unbound

83   virions. Based on our findings we posit a model for the entry pathway of this extinct virus

84   where heparan sulfate binding followed by subsequent endosomal uptake and acidification

85   result in productive infection.

Results

86

87   To identify host factors required for HERV-K Env mediated entry, we performed a

88   haploid genetic screen [12] (Fig 1A-B). Briefly, HAP1 cells were mutagenized using a retroviral

89   gene-trap vector to generate a population with inactivating mutations across the genome and

90   infected with VSV-HERVK. Deep sequencing genomic DNA from cells that survived VSV-HERVK

91   infection identified sites of integration of the gene-trap retrovirus (Fig 1C). Among the genes

92   identified were 11 involved in the biosynthesis of heparan sulfate - a glycosaminoglycan (GAG)

93   ubiquitously expressed on the cell surface. Six of those genes (*GPC3, EXT1, EXT2, EXTL3,*

94   *HS2ST1,* and *NDST1*) are specific to heparan sulfate and heparin and not other GAGs (S1 Fig).

95   For follow up, we selected *EXT1* which encodes an enzyme that catalyzes the addition of a

96   glucaronic acid – N-aceytlglucosamine (GlcA-sGlcNAc) disaccharide onto the growing heparan

97   sulfate chain and *SLC35B2* which encodes the Golgi-resident transporter of the universal sulfate

98   donor 3'-phosphoadenosine-5'-phosphosulfate (PAPS) [18]. Three additional genes, myosin X

99   (*MYO10)*, sortilin (*SORT1*), and CREB binding protein (*CREBBP*) scored as significant and were

100   also selected for further follow up.

101   We independently generated single-cell clones of HAP1 cells lacking each of those 5

102   genes by gene editing and infected them with VSV or VSV-HERVK expressing eGFP as a marker

103   of infection (Fig 1D). This eliminated *MYO10, SORT1* and *CREBBP* from further analysis in viral

104   entry because the fraction of cells infected was only modestly changed (Fig 1D and S2 Fig). In

105   *MYO10*^{KO} cells we note however that the intensity of eGFP expression increased following

106   infection with VSV but slightly decreased following infection with VSV-HERVK (S2 Fig). This

6

107   result indicates that elimination of myosin X differentially impacts the kinetics of productive

108   infection perhaps reflective of the distinct uptake mechanisms of VSV compared to VSV-HERVK

109   (S2 Fig). By contrast, VSV-HERVK infection of $EXT1^{KO}$ and $SLC35B2^{KO}$ cells was reduced 4-fold

110   compared to VSV. VSV infection was unaffected in $EXT1^{KO}$ cells, but was diminished 3-fold in

111    $SLC35B2^{KO}$ cells. Those results demonstrate that elimination of cell surface heparan sulfate

112   reduces VSV-HERVK infection specifically and that suppression of sulfation also reduces VSV

113   infection in a manner that appears independent of heparan sulfate (S2 Fig). Flow cytometry

114   verified that cell surface expression of heparan sulfate was lost in both $EXT1^{KO}$ and $SLC35B2^{KO}$

115   cells and was restored following transduction with retroviruses expressing the corresponding

116   gene (S3 Fig). Restoration of cell surface heparan sulfate corresponded with an increase in VSV-

117   HERVK infection (Fig 2A). Loss of heparan sulfate did not completely block VSV-HERVK infection

118   as evident from the small fraction of infected cells. That small fraction, however, exhibits a 2-3

119   fold reduction in the intensity of eGFP, presumably reflecting a less efficient heparan sulfate

120   independent mechanism of viral entry (Fig 2B and S4 Fig).

121   As a complementary approach to genetic inactivation of heparan sulfate biosynthesis,

122   we employed a chemical approach. Sodium chlorate treatment of cells inhibits the synthesis of

123   PAPS and correspondingly reduces cell surface sulfation. Cells cultured in the presence of 50

124   mM sodium chlorate showed a 30-fold reduction in infectivity of VSV-HERVK compared to VSV

125   (Fig 2C). The fraction of cells that were infected by VSV-HERVK again showed a reduction in the

126   levels of eGFP expressed, following entry independent of heparan sulfate (Fig 2C and S4 Fig).

127   These results confirm the findings obtained following genetic inactivation of heparan sulfate

128   biosynthesis and support a role for heparan sulfate in HERV-K entry.

129   Heparan sulfate has been identified as a receptor for herpes simplex virus 1 (HSV1) [19]

130   and eastern equine encephalitis virus (EEEV) [20]. If heparan sulfate serves as a key entry factor

131   for HERV-K, VSV-HERVK infection should be sensitive to competition by excess soluble GAGs.

132   Incubation of purified virus with soluble heparin - a highly sulfated analog of heparan sulfate -

133   or with heparan sulfate, inhibits infection in a concentration dependent manner. The sulfated

134   GAGs chondroitin or dermatan sulfate had no effect on VSV or VSV-HERVK infection further

135   supporting a specific requirement for heparan sulfate in HERVK infection at the level of viral

136   attachment (Fig 2D). Consistent with this interpretation, attachment of single VSV-HERVK

137   particles to $SLC35B2^{KO}$ cells was reduced at least 2-fold compared to WT cells (Fig 2E-G) at both

138   4°C and 37°C. By contrast VSV particle binding was similar between both cell types at both

139   temperatures (Fig 2G).

140   Further evidence that heparan sulfate serves an entry factor was provided by the

141   demonstration that VSV-HERVK particles specifically associate with heparin but not protein A

142   beads (Fig 3A). This heparin bead binding was sensitive to inhibition by pre-incubation of virus

143   with soluble heparin. VSV did not bind either heparin or protein A beads, underscoring that

144   binding is dictated by the HERV-K glycoprotein. To test whether HERV-K Env directly interacts

145   with heparin, we generated a soluble, monomeric HERV-K SU subunit, which by extrapolation

146   from extant retroviruses would harbor the receptor-binding domain (Fig 3B, S5 Fig, and S6 Fig).

147   Soluble HERV-K SU specifically bound to heparin but not protein A beads, and this binding was

148   sensitive to pre-incubation of the protein with soluble heparin (Fig 3C). As expected a soluble

149   receptor-binding domain from Influenza A hemagglutinin (HA), which binds a different

150   carbohydrate receptor, sialic acid, failed to bind either the heparin or protein A beads, further

8

151 supporting the specificity of the HERVK-heparan sulfate interaction. Pre-incubation of HERV-K

152 SU with soluble GAGs prior to mixing with heparin beads demonstrates that binding is inhibited

153 by soluble heparin and heparan sulfate, but not chondroitin or dermatan sulfate (Fig 3D). These

154 data correlate with the suppression of infectivity, and provide further evidence that heparan

155 sulfate binding leads to productive infection by VSV-HERVK. We further found that HERV-K SU

156 binding to heparin beads was unaffected by pre-incubation with 2-O-desulfated heparin,

157 whereas 6-O-desulfated heparin showed partial inhibition of binding (Fig 3D). This result implies

158 that 2-O sulfation and not 6-O sulfation is important for HERV-K binding. Consistent with this,

159 our genetic screen identified the enzyme that catalyzes 2-O-sulfation (heparan sulfate 2-O-

160 sulfotransferase 1 (HS2ST1)) but not the enzymes that catalyze 6-O-sulfation (HS6ST1,2,3).

161   Acidic pH – such as that encountered in endocytic compartments – serves as the trigger

162 for conformational rearrangements in several viral envelope proteins necessary for membrane

163 fusion. In class I fusion proteins, such as influenza HA, those rearrangements are irreversible,

164 such that premature exposure to acidic pH inactivates the fusion machinery. Envelope proteins

165 from every extant betaretrovirus that has been tested, including Jaagsiekte sheep retrovirus

166 (JSRV), enzootic nasal tumor virus (ENTV), and mouse mammary tumor virus (MMTV), as well as

167 the alpharetrovirus avian leukosis virus (ALV) are not inactivated on exposure to mildly acidic

168 pH [21-24]. JSRV and ALV are only inactivated if first bound to their receptor, suggesting an

169 essential two-step fusion mechanism of receptor binding followed by exposure to low pH [22,

170 24]. To test whether exposure to acid pH inactivates HERVK Env, we exposed purified VSV-

171 HERVK to increasingly acidic pH for 30 minutes, neutralized the pH and then measured the

172 residual infectivity. Treatment of purified VSV-HERVK particles at pH <6.0 reduced infectivity

9

173   (Fig 4A). By contrast the infectivity of VSV was unaffected reflecting the reversibility of the

174   conformational changes in VSV G when exposed to acid pH [25-27]. We next examined whether

175   VSV-HERVK infection requires endocytosis beyond a need for acidic pH. For this purpose we

176   bound virus to the cell surface and exposed cells to a brief pulse of acidic pH (Fig 4B). Infection

177   was readily established demonstrating that endocytic uptake is not required and establishing a

178   minimal requirement and a necessary order of virus attachment and acidic pH for HERV-K Env

179   mediated entry

180   To determine whether heparin binding and acid pH triggered fusion are retained by

181   distinct HERVK Env sequences we compared the Phoenix Env sequence with that of two distinct

182   genomic copies, K108 and Xq21.33 [5, 28] (S7 Fig). All three Env sequences mediate acid pH

183   dependent cell-cell fusion, although the relative fusogenicity of Xq21.33 is reduced (Fig 4D).

184   Using lysates of cells overexpressing the individual envelope proteins we also demonstrate that

185   the 3 HERV K Envs, but not VSV G, are specifically bound by heparin beads (Fig 4D). These

186   results underscore that both heparin binding and acid pH triggered fusion are shared properties

187   of multiple HERV-K envelope sequences present in the genome.

188

Discussion

189

190   The major conclusion of this study is that heparan sulfate is a direct HERV-K Env

191   attachment factor. Binding of HERVK Env to heparan sulfate is most sensitive to the loss of 2-O

192   sulfation, implicating this modification in attachment. Combined with earlier work we posit the

193   following model for the entry of the retrovirus HERV-K (Fig 5). Following binding to cell surface

194   heparan sulfate, virus is taken up into cells in a dynamin-dependent, clathrin independent

195   manner with subsequent acidification of the endosome leading to membrane fusion and

196   productive infection. This model is reminiscent of the sialic-acid binding and acid pH

197   requirement for productive influenza virus entry. We cannot, however, rule out the possibility

198   that HERV-K entry may require additional host factors not identified through the haploid

199   genetic screening approach – such as essential host genes, or those with redundant function.

200   The low pH-mediated inactivation of VSV-HERVK, and lack of identification of endosomal

201   factors other than acidic pH, raise the possibility that heparan sulfate may act directly as a

202   receptor. Regardless of whether binding to heparan sulfate is sufficient to fulfill both

203   attachment and receptor functions, this study defines heparan sulfate as an important host

204   factor for HERV-K Env-mediated cell entry. The ability of HERVK Env to bind heparan sulfate

205   underscores that such binding is an ancient property of viruses.

206   For many viruses, heparan sulfate binding reflects an adaptation to growth in cell

207   culture. By contrast, heparan sulfate binding is an intrinsic property of the original HERV-K Env

208   because the *env* sequences we used are not derived from viruses grown in cell culture, and

209   such binding is apparent for multiple HERV-K Envs as they exist in the genome and at least one

210   putative ancestral sequence. Several extant retroviruses including the prototype foamy virus

211   (PFV), MMTV, and human T-lymphotropic virus 1 (HTLV-1) bind heparan sulfate [29-31]. MMTV

212   requires engagement of transferrin receptor [32] and HTLV-1 requires neuropilin 1 and glucose

213   transporter 1 [33, 34] for entry. Proteinaceous receptors for PFV have not been identified, and

214   like HERV-K it has proved difficult to identify cell types that are refractory to entry. Although we

215   cannot rule out the presence of an additional unidentified host factor for HERVK entry, the

216   demonstration that acid pH alone can trigger HERVK Env suggests that engagement of such a

217   second molecule may not be essential for infection. We do, however, observe some infectivity

218   in cells lacking heparan sulfate by its genetic or chemical inhibition demonstrating that

219   molecules other than heparan sulfate facilitate cell entry.

220   Our conclusions are based on the results from a genetic screen performed using VSV-

221   HERVK, combined with genetic, biochemical and cell biological follow up experiments to

222   validate the importance of heparan sulfate. In our prior work with VSV recombinants containing

223   heterologous envelope proteins we have always validated our findings using the respective wild

224   type virus. For HERV-K such validation experiments were not possible because the

225   reconstructed viruses replicate poorly. Lentiviral particles pseudotyped with HERV-K Env have

226   been described but they also produce low viral titers, ranging from approximately 60-1000

227   infectious units ml^{-1} as determined using a spinoculation based infectivity assay [7-9]. We

228   obtain similar titers of pseudotyped lentiviruses without such spinoculation - 179-517 infectious

229   units ml^{-1}. Those titers are substantially below the $3x10^{7}$ infectious units ml^{-1} of VSV-HERVK

230   [10], limiting the utility of the lentiviral pseudotypes in such validation experiments.

231   Nevertheless, we carried out experiments using such lentiviral pseudotypes using Env null

12

232 "bald" particles as a stringent background infectivity control. Using such lentiviral pseudotypes

233 we observe trends similar to those with VSV-HERVK when heparan sulfate biosynthesis

234 pathways are manipulated (S8 Fig). We therefore cannot rule out the possibility that

235 contributions of particle geometry and glycoprotein density might influence the entry of VSV-

236 HERVK into cells in a manner that does not fully recapitulate that of wild type HERV-K virus. We

237 cannot know the glycoprotein density on HERV-K – it may range from the low levels observed

238 for HIV to the high levels on MMTV [35]. Such considerations do not, however, affect the major

239 conclusions of this study as evidenced by the fact that biochemically pure wild type HERV-K Env

240 binds heparin, and heparin binding and acid pH triggered fusion are properties of three distinct

241 HERV-K Env sequences.

242 HERV-K(HML-2) proviruses are present in all humans and HERV-K Env is expressed

243 during a number of diseases, including viral infection, cancer, and autoimmune diseases [6, 36].

244 While these Envs are unlikely to be fusogenic at normal extracellular pH, they will likely act as

245 heparan sulfate binding proteins on the cell surface and would, in principle foster contacts with

246 the extra cellular matrix through heparan sulfate engagement. Heparan sulfate is involved in a

247 multitude of physiological functions, from cell adhesion and migration to cell signaling [37-40].

248 Heparan sulfate proteoglycans have been implicated in cancer invasion and metastasis, often

249 through dysregulation of cell signaling pathways [37]. Heparan sulfate binding by HERV-K Env

250 could thus play a role in these processes. Overexpression of HERV-K Env on cancer cells could

251 facilitate invasion and metastasis through binding heparan sulfate on the surface of neighboring

252 cells or the extracellular matrix. HERV-K Env binding heparan sulfate proteoglycans may also

253 disrupt normal signaling cascades in which these proteoglycans are involved.

254   As HERV-K is a relatively young group of ERVs, the ultimate fate of HERV-K Env is not yet

255   fixed. Several HERV envelopes have been coopted throughout evolution to perform important

256   functions for the host. These include the syncytins, which are essential for placentation [41],

257   HERV-T Env, which has antiviral properties and may have contributed to its own extinction [42],

258   and HEMO, a recently identified Env product that is shed in the blood of pregnant women [43].

259   HERV-K Env is known to be expressed in healthy tissues as well, including stem cells and during

260   early stages of embryogenesis [36, 44]. It remains to be determined whether there is any

261   physiological consequence of heparan sulfate binding by HERV-K Env in instances when it is

262   actively expressed.

263   The conservation of heparan sulfate throughout metazoans and its ubiquitous

264   expression presents no barrier for this ERV to enter into germ cells – a step essential for its

265   endogenization – and implies that other steps of the HERV-K replication cycle result in the

266   observed species tropism. The broad distribution of heparan sulfate is consistent with findings

267   for other endogenous retroviruses. These include murine leukemia virus and MMTV, which

268   utilize receptors that are broadly expressed in their respective host, and which exist as both

269   endogenous and exogenous viruses [45-51]. This also holds true for other extinct primate

270   endogenous retroviruses, chimp endogenous retrovirus 2 (CERV2) and human endogenous

271   retrovirus T (HERV-T) [42, 52]. Perhaps the great majority of endogenous retroviruses were

272   able to colonize the germline because their broad tropism allowed access to germ cells.

273   Germline integration and endogenization would become chance events by such "promiscuous"

274   viruses, rather than viruses that specifically target germ cells.

14

<div align="center">

Materials and Methods

</div>

275

276   **Cell lines, viruses, and plasmids.**

277   BSRT7 cells (a kind gift from U.J. Buchholz[53]) and 293T cells (ATCC CRL-3216; American

278   Type Culture Collection, Manassas, VA) were grown in Dulbecco's modified Eagle's medium

279   (DMEM) supplemented with 10% fetal bovine serum (FBS) and maintained at 37°C and 5% CO_2.

280   HAP1 cells (a kind gift from Thijn Brummelkamp [12]) were grown in Iscove's modified

281   Dulbecco's medium (IMDM) supplemented with 10% FBS and maintained at 37°C and 5% CO_2.

282   All cell lines were tested to be free of mycoplasma. VSV-HERVK[+], referred to throughout this

283   manuscript as VSV-HERVK, and VSV-eGFP were generated as previously described [10, 54]. VSV-

284   HERVK encodes the env from the Phoenix consensus sequence[8], where the cytoplasmic tail

285   has been replaced with that of VSV G. Both viruses express eGFP from the first position in the

286   genome. Viruses were grown on BSRT7 cells. Viruses were tittered by plaque assay on BSRT7

287   cells and flow cytometry on HAP1 cells.

288   pCAGGS-PhoenixEnv   and   pGEM-PhoenixEnv   were   previously   generated   [10].   We

289   synthesized   codon-optimized   versions   of   HERV-K   108   *env*   and   Xq21.33   *env*   (Genscript,

290   Piscataway, NJ) and subsequently cloned them into pCAGGS and pGEM to generate pCAGGS-

291   HERVK108Env,   pCAGGS-Xq21.33Env,   pGEM-HERVK108Env,   and   pGEM-Xq21.33Env,

292   respectively. Plasmids containing the cDNA for *EXT1* and *SLC35B2* were obtained from the

293   Dana-Farber plasmID repository (Dana-Farber/Harvard Cancer Center DNA Resource Core,

294   Boston, MA). These cDNAs were amplified with primers to add a C-terminal HA tag and were

295   cloned into pQCXIN (Clontech, Mountain View, CA). lentiCas9-Blast [55] (Addgene plasmid

<div align="center">

15

</div>

296   #52962), lentiGuide-Puro [55] (Addgene plasmid #52963), lentiCRISPRv2 [55] (Addgene plasmid

297   #52961) and pX330 [56] (Addgene plasmid #42230) were a gift from Feng Zhang.

Haploid genetic screen with VSV-HERVK.

299   HAP1 cells were mutagenized with a genetrap retrovirus as described [13].

300   Approximately 10^8 cells were infected with VSV-HERVK at a multiplicity of infection (MOI) of 3

301   infectious units (IU) per cell. Infection was allowed to proceed for several days, after which a

302   second round of infection was performed to kill remaining susceptible cells. Genomic DNA was

303   isolated from surviving cells and used to prepare a library for Illumina deep sequencing and

304   reads analyzed as described [13]. Inactivating insertion sites (mapping to exons or in sense

305   orientations in introns) in the VSV-HERVK selected cells (180,655 unique insertions) were

306   compared to that of a control data set from unselected cells (2,161,301 unique insertions). *P*

307   values for enrichment in the selected set versus control set were calculated using a Fisher's

308   exact test. Significance scores are reported as the inverse log of the *p* value. Genes with

309   insertions were also analyzed to identify bias in the direction of insertion of the genetrap

310   sequence within introns. Insertions in the forward direction are inactivating.

Generation of knockout cells.

312   To generate *CREBBP*[KO] cells, a guide RNA targeting the Histone Acetyl Transferase

313   domain of CREBBP (5'-GGAGGTTTTTGTCCGAGTGG-3') was cloned into pX330 creating pX330-

314   *CREBBP*. HAP1 cells were co-transfected with pX330-*CREBBP* and a plasmid containing an

315   expression cassette for a guide RNA targeting the zebrafish *TIA* gene (5'-

316   GGTATGTCGGGAACCTCTCC-3') followed by a CMV promotor sequence driving expression of a

16

317 blasticidin resistance gene flanked by two *TIA* target sites [57]. Co-transfection of these

318 plasmids resulted in the incorporation of the blasticidin resistance cassette at the site of the

319 targeted *CREBBP* locus. Four days after DNA transfection, the culture medium was

320 supplemented with blasticidin (30 μg/mL). A single cell, blasticidin-resistant clone (C4C2) was

321 expanded and disruption of *CREBBP* was verified by Sanger sequencing and Western blot for

322 protein expression (using anti-CREBBP; clone C-20, Santa Cruz Biotechnology, Santa Cruz, CA).

323   To generate *MYO10^{KO}* and *SORT1^{KO}* cells, we first generated HAP1 cells stably expressing

324 *Streptococcus pyogenes* Cas9 (HAP1-Cas9). Lentivirus was generated by transfecting 293T cells

325 with lentiCas9-Blast, pCD/NL-BH*DDD [58] and pCAGGS-VSVG. HAP1 cells were transduced

326 with the lentivirus and selected with 5μg/ml blasticidin. Guide RNAs targeting exon 20 of

327 *MYO10* (5'-CGCCTGGTCAAACTGGTC-3') and exon 3 of *SORT1* (5'-CAGTCCAAGCTATATCGA-3')

328 were individually cloned into lentiGuide-Puro. Lentivirus was generated by transfecting 293T

329 cells with lentiGuide-MYO10sgRNA or lentiGuide-SORT1sgRNA, pCD/NL-BH*DDD and pCAGGS-

330 VSVG. HAP1 cells were transduced with the lentivirus and selected with 2 μg/ml puromycin.

331 Single cell clones were screened by Western blot for expression of myosin-X and sortilin using

332 anti-MYO10 (HPA024223; Sigma-Aldrich, St. Louis, MO) and anti-sortilin (ab16640; Abcam,

333 Cambridge, MA) antibodies. To generate *EXT1^{KO}* cells, guide RNAs targeting two sequences in

334 exon 1 of *EXT1* (5'-GGCCAGAAATGATCCGGACT-3' and 5'-GCACAACGTCCTCCCCGTTA-3') were

335 individually cloned into pX330. The plasmids were cotransfected along with pLPCX-Puro

336 (Clontech) into HAP1 cells using lipofectamine 3000 transfection reagent (Life Technologies,

337 Grand Island, NY). Cells were selected with 2 μg/ml puromycin. Genomic DNA from single cell

338 clones was isolated using DirectPCR lysis reagent (Viagen Biotech, Los Angeles, CA) and

17

339   screened by PCR for an approximately 700 base pair deletion using the following primers: 5'-

340   GAGTTGAAGTTGCCTTCCCG-3' and 5'-AGGCTTTTCAGTTTGCCCGA-3'. To generate *SLC35B2*^{KO}

341   cells, a guide RNA sequence targeting exon 4 of *SLC35B2* (5'-GCTTTCCCATCAGCATGACA-3') was

342   cloned into the lentiCRISPRv2 plasmid [55]. Lentivirus was generated by transfecting 293T cells

343   with lentiCRISPRv2-SLC35B2sgRNA, pCD/NL-BH*DDD and pCAGGS-VSVG. HAP1 cells were

344   transduced with the lentivirus and selected with 2 µg/ml puromycin. Single cell clones were

345   screened by flow cytometry for reactivity with 10E4 (370255-1, AMS biotechnology, Cambridge,

346   MA), an antibody specific for sulfated heparan sulfate. After selection and single cell cloning, all

347   cells were maintained in IMDM+10% FBS without puromycin.

348   *B4GALT7*^{KO} HEK293T cells were generated by CRISPR/Cas9-mediated genome

349   editing essentially as described by Langereis *et al* [59] with two guide RNAs targeting

350   B4GALT7's active site encoded in exon 5 (gRNA 1: 5'-ATGGGATGTCCAACCGCTTC-3' and 5'-

351   GAGTTCTACCGGCGCATTAA-3'). Single cell clones were isolated by fluorescence activated

352   cell sorting (FACS) and genotyped by PCR and DNA sequencing. Loss of heparan sulfate

353   expression was confirmed by flow cytometry with antibody 10E4.

354
355

356   **Reconstitution of knockout cells with EXT1 or SLC35B2.**

357   Murine leukemia virus (MLV) carrying either EXT1-HA or SLC35B2-HA was generated by

358   transfecting 293T cells with pCS2-MGP (Moloney MLV gag/pol expression vector [60]), pCAGGS-

359   VSVG, and either pQCXIN-EXT1HA, pQCXIN-SLC35B2HA or pQCXIN empty vector. *EXT1*^{KO} cells

360   were transduced with MLV-EXT1HA or MLV-empty vector to generate *EXT1*^{KO}+EXT1-HA or

18

339   screened by PCR for an approximately 700 base pair deletion using the following primers: 5'-

340   GAGTTGAAGTTGCCTTCCCG-3' and 5'-AGGCTTTTCAGTTTGCCCGA-3'. To generate $SLC35B2^{KO}$

341   cells, a guide RNA sequence targeting exon 4 of *SLC35B2* (5'-GCTTTCCCATCAGCATGACA-3') was

342   cloned into the lentiCRISPRv2 plasmid [55]. Lentivirus was generated by transfecting 293T cells

343   with lentiCRISPRv2-SLC35B2sgRNA, pCD/NL-BH*DDD and pCAGGS-VSVG. HAP1 cells were

344   transduced with the lentivirus and selected with 2 µg/ml puromycin. Single cell clones were

345   screened by flow cytometry for reactivity with 10E4 (370255-1, AMS biotechnology, Cambridge,

346   MA), an antibody specific for sulfated heparan sulfate. After selection and single cell cloning, all

347   cells were maintained in IMDM+10% FBS without puromycin.

348    $B4GALT7^{KO}$ HEK293T cells were generated by CRISPR/Cas9-mediated genome

349   editing essentially as described by Langereis *et al* [59] with two guide RNAs targeting

350   B4GALT7's active site encoded in exon 5 (gRNA 1: 5'-ATGGGATGTCCAACCGCTTC-3' and 5'-

351   GAGTTCTACCGGCGCATTAA-3'). Single cell clones were isolated by fluorescence activated

352   cell sorting (FACS) and genotyped by PCR and DNA sequencing. Loss of heparan sulfate

353   expression was confirmed by flow cytometry with antibody 10E4.

354
355

356   **Reconstitution of knockout cells with EXT1 or SLC35B2.**

357   Murine leukemia virus (MLV) carrying either EXT1-HA or SLC35B2-HA was generated by

358   transfecting 293T cells with pCS2-MGP (Moloney MLV gag/pol expression vector [60]), pCAGGS-

359   VSVG, and either pQCXIN-EXT1HA, pQCXIN-SLC35B2HA or pQCXIN empty vector. $EXT1^{KO}$ cells

360   were transduced with MLV-EXT1HA or MLV-empty vector to generate $EXT1^{KO}$+EXT1-HA or

18

361   *EXT1*^KO Neo^r control, respectively. *SLC35B2*^KO cells were transduced with MLV-SLC35B2HA or

362   MLV-Empty vector to generate *SLC35B2*^KO+SLC35B2-HA or *SLC35B2*^KO Neo^r control, respectively.

363   Cells were selected with 1.5 mg/ml G418. *SLC35B2*^KO+SLC35B2-HA cells were further subcloned

364   and single cell clones were screened for heparan sulfate expression using the 10E4 antibody. All

365   reconstituted cells were maintained in IMDM with 10% FBS and 1.5 mg/ml G418.

Heparan sulfate staining.

367   Cells were collected with 10mM EDTA in PBS and subsequently resuspended in PBS with

368   1% BSA and BD human Fc block (1:25 dilution; 564220, BD Biosciences, San Jose, CA). Cells were

369   incubated with either the 10E4 monoclonal antibody to heparan sulfate (1:400), or mouse IgM

370   (1:200; ab18401, Abcam, Cambridge, MA) for 1 hour (h) on ice. Cells were washed twice with

371   PBS+1%BSA and incubated with goat anti-mouse IgG/IgM alexafluor 488 (1:500; A10680,

372   Molecular Probes, Eugene, OR) for 1h on ice. Cells were washed twice and fixed in 2%

373   paraformaldehyde (PFA). Fluorescence was measured using a FACSCalibur instrument (Cytek

374   Development, Freemont, CA) and analyzed using FlowJo software (FlowJo Inc, Ashland, OR).

Infectivity experiments.

376   WT HAP1, *MYO10*^KO, *CREBBP*^KO, *SORT1*^KO, *EXT1*^KO, *EXT1*^KO Neo^r control, *EXT1*^KO+EXT1-HA,

377   *SLC35B2*^KO, *SLC35B2*^KO Neo^r control, or *SLC35B2*^KO+SLC35B2-HA cells were infected with VSV-

378   HERVK or VSV-eGFP at an MOI of 1 IU/cell. Cells were collected 5h post-infection (pi) and fixed

379   in 2% PFA. eGFP fluorescence was measured using a FACSCalibur instrument and the

380   percentage of eGFP-positive cells, as well as mean fluorescence intensity was quantified using

381   FlowJo software. The number of individual clones tested for each cell line are as follows:

19

382   *MYO10*^{KO} (n=2), *CREBBP*^{KO} (n=1), *SORT1*^{KO} (n=3), *EXT1*^{KO} (n=3), *SLC35B2*^{KO} (n=1). Data are shown

383   from multiple independent replicates in a single clonal cell line for each gene. Data are

384   represented as fold difference in percent eGFP-positive cells or MFI compared to WT cells and

385   normalized to VSV infectivity in each cell type. Error bars represent standard error of the mean

386   from at least three independent biological replicates.

387   **Sodium chlorate treatment and infectivity**.

388   BSRT7 cells were passaged in sulfate-free Joklik modified minimal essential medium

389   (M8028, Sigma-Aldrich, St. Louis, MO) with 10% FBS with or without 50mM sodium chlorate for

390   at least two passages prior to seeding for infection. Cells were infected with either VSV-HERVK

391   or VSV at an MOI of 1 particle forming unit (PFU)/cell. Cells were collected 5hpi and fixed in 2%

392   PFA and fluorescence measured using a FACSCalibur instrument. The % eGFP-positive cells and

393   MFI were quantified using FlowJo software. Data are represented as fold difference in % eGFP-

394   positive cells or MFI normalized to VSV infected cells. Error bars represent standard error of the

395   mean from three independent biological replicates.

396   **Inhibition of infection by soluble glycosaminoglycans.**

397   Heparin (H3393, Sigma-Aldrich), heparan sulfate (AMS GAGHS01, AMS bioscience),

398   chondroitin sulfate A (C9819, Sigma-Aldrich) and dermatan sulfate (C3788, Sigma-Aldrich) were

399   reconstituted in PBS. 1 μg of purified VSV-HERVK or VSV was incubated with the compounds at

400   the indicated concentration in PBS for 1h at 37°C and then added to BSRT7 cells. Cells were

401   incubated with virus and compound for 1h at 37°C, washed with DMEM and incubated for 4h at

402   37°C. Cells were collected, fixed with 2% PFA and fluorescence measured using a FACSCalibur

20

403   instrument. The % eGFP-positive cells were quantified using FlowJo software and normalized to

404   infectivity with no compound. Error bars represent standard error of the mean from three

405   independent biological replicates.

406   **Imaging of HERV-K attachment to cells.**

407   Gradient purified VSV-HERVK particles were labeled with AlexaFluor 647 and VSV particles

408   were labeled with AlexaFluor 594, as described [61]. WT HAP1 or $SLC35B2^{KO}$ cells were pre-

409   stained with calcein (diluted 1:1000; C3099, Molecular Probes, Eugene, OR) and NucBlue live

410   cell stain (1:50, C34552, Molecular Probes) in IMDM for 30 minutes (min) at 37°C, followed by

411   blocking in 1% BSA in IMDM for 30 min at 37°C. Labeled VSV-HERVK and VSV were added

412   together to the cells. Cells were incubated with virus at either 37°C for 15 min or at 4°C for 1 h.

413   Samples were fixed in 2% PFA and mounted with ProLong Gold (P10144, Molecular Probes).

414   Samples were imaged using a Marianas system (Intelligent Imaging Innovations, Denver, CO)

415   based on a Zeiss observer microscope (Carl Zeiss Microimaging, Thornwood, NY) outfitted with

416   a 64 CSU-22 spinning-disk confocal unit (Yokogawa Electric Corporation, Musashino, Tokyo,

417   Japan) and a 63x (Plan-Apochromat, NA 1.4; Carl Zeiss Microimaging) objective lens. Excitation

418   wavelengths were 561 nm for AF594 and 660 nm for AF647. SlideBook 6.0 (Intelligent Imaging

419   Innovations) was used to command the hardware devices and visualize and export the acquired

420   data. Subsequent image analysis was conducted using ImageJ (National Institutes of Health).

421   Briefly, cellular cytoplasmic areas were approximated by manually tracing the 2D cellular

422   outline based on the calcein staining and determining its area. To simplify visualization, calcein

423   aggregates were eliminated using the Remove Outliers tool in ImageJ. Bound VSV and VSV-

424   HERVK particles were counted for each cell, excluding large aggregates. Particle binding per

425   area unit was calculated by dividing particle counts by the calculated areas. Data are

426   represented as box plots indicating the median values, first and third quartiles, minima and

427   maxima. Outliers were defined as those points 1.5 times the interquartile range, and severe

428   outliers as those 3 times the interquartile range. Data are from multiple images from a single

429   experiment. N values are as follows: for the 37°C experiment, 29 WT cells were counted, with

430   189 VSV particles and 492 VSV-HERVK particles; 62 $SLC35B2^{KO}$ cells were counted with 252 VSV

431   particles and 220 VSV-HERVK particles. For the 4°C experiment, 43 WT cells were counted, with

432   629 VSV particles and 793 VSV-HERVK particles; 72 $SLC35B2^{KO}$ cells were counted with 790 VSV

433   particles and 377 VSV-HERVK particles.

434   **Virus and cell lysate heparin pull-downs.**

435   For all pull-downs, heparin beads (H6508, Sigma-Aldrich) and protein A beads (P2545,

436   Sigma-Aldrich) were washed 3 times in the corresponding buffer and blocked for 1 h in PBS with

437   1% BSA. BSRT7 cells were transfected with pCAGGS-PhoenixEnv, pCAGGS-HERVK108Env, or

438   pCAGGS-Xq21.33Env. At 24h post-transfection, cells were lysed in TNE buffer (50mM Tris, 2mM

439   EDTA, 150mM NaCl,) supplemented with 1% tritonX-100 and 200 µl lysate was incubated with

440   50 µl of either heparin or control protein A beads at 4°C for 2 h. Beads were washed 5 times in

441   TNE+1% tritonX-100 and bound proteins eluted by boiling in 2X SDS loading buffer for 5 min.

442   Samples were loaded on a 10% (wt/vol) polyacrylamide gel and transferred to nitrocellulose

443   membrane which were blocked with 5% milk in PBS+0.1% Tween-20 and subsequently blotted

444   with anti-HERV-K Env antibody (1:2000; HERM-1811-5, Austral biological, San Ramon, CA)

22

445   followed by goat anti-mouse horseradish peroxidase (HRP) antibody (1:5000; Sigma-Aldrich).

446   Membranes were incubated with ECL reagent (Thermo-Fisher Scientific, Waltham, MA) and

447   signal was detected by film (Denville Scientific, Holliston, MA). Data shown are from a single

448   representative experiment from 3 biological replicates.

449   For virus pull-down we incubated 10 µg of purified virus +/- 50 µg/ml heparin in TNE buffer

450   with 1% BSA. Complexes were collected by incubation with 50 µl heparin or protein A beads

451   (4°C rotating for 2h), and washed 5 times in TNE buffer. Bound virions were eluted and

452   analyzed by Western blot as described above. Membranes were blotted with antibodies against

453   HERV-K Env, VSVG (1:10,000; V5507, Sigma-Aldrich), or VSVM (1:5000; 23H12, a kind gift from

454   Doug Lyles[62]) followed by goat anti-mouse HRP antibody (1:5000; Sigma-Aldrich).

455   Membranes were incubated with ECL reagent and signal was detected by film. Data shown are

456   from a single representative experiment from 3 biological replicates.

457   **Generation and characterization of HERV-K Env truncations.**

458   N-terminal truncations (N1-N7) were designed as outlined in S5 Fig to determine the

459   appropriate boundary between the signal peptide and SU domain. DNA fragments containing

460   the truncated versions of the envelope were cloned into a modified pVRC8400 expression

461   vector, which uses tissue plasminogen activator signal sequence [63]. These *env* sequences with

462   the new signal peptide were sub-cloned into pGEM3, under the control of a T7 polymerase

463   promoter. Env truncations were screened for expression, proteolytic processing, fusogenicity,

464   and pH dependency of fusion. Each screening experiment was a single replicate. Western blot

465   analysis and cell-cell fusion experiments were performed as described [10]. Briefly, BSRT7 cells

466   were infected with VVT7.3 [64], a vaccinia virus encoding the T7 RNA polymerase as source of

467   transcriptase. The cells were subsequently transfected with the HERV-K *env* expression

468   plasmids or an empty vector control. At 18h post-transfection, cells were either harvested for

469   Western blot analysis against HERV-K Env TM subunit (Austral biological) and Actin (Abcam), or

470   treated with phosphate-buffered saline (PBS) at the indicated pH for 20 min at 37°C, at which

471   point the cells were washed and standard growth medium was added. The cells were incubated

472   for 4 h at 37°C and subsequently imaged. Truncations N4 and N5 had similar expression,

473   processing, fusogenicity, and pH-dependency as WT.

474   C-terminal truncations based on N4 and N5 were tagged with a C-terminal HA tag, 3C

475   protease cleavage site, and a tandem His_{8X}-His_{6X} tag and cloned into pVRC8400. Recombinant

476   protein was produced by transient transfection of 293T cells using Lipofectamine 2000 (Life

477   Technologies), per manufacturer's protocol. Three days post-transfection supernatants were

478   harvested and clarified from cellular debris by low-speed centrifugation. HERV-K SU was

479   purified by passage over Co-NTA agarose (Clontech) and concentrated with an Amicon Ultra-4

480   filter (Millipore, Billerica, MA). Purified protein was run on both reducing and non-reducing

481   SDS-PAGE (4-20% polyacrylamide gel, 4561096; Bio-rad, Hercules, CA) followed by Coomassie

482   staining. Proteins were screened for the following criteria: 1. Expression: If a band of

483   appropriate size was observed on a reducing gel. 2. Solubility: the absence of major aggregate

484   bands under non-reducing conditions. 3. Monomeric: only proteins without evidence of major

485   aggregation were subject to size exclusion chromatography. Proteins for which a discrete peak

486   in the A280 trace corresponding to the approximate size of monomeric SU and produced a

487   single band of the appropriate size on a non-reducing SDS-PAGE were deemed to produce

488   monomeric species.

489   **Production of soluble HERV-K SU.**

490   A truncated version of the codon optimized HERV-K Phoenix SU domain, encoding

491   residues 96-433 (residue 1 being the initiating methionine) was synthesized by Integrated DNA

492   Technologies, Inc. to included a C-terminal HA tag, 3C protease cleavage site, and a tandem

493   His_{8X}-His_{6X} tag. This cDNA was cloned into the modified pVRC8400 expression vector [63].

494   Recombinant protein was produced by transient transfection of 293T cells using Lipofectamine

495   2000 (Life Technologies), per manufacturer's protocol. At 3 days post-transfection supernatants

496   were harvested and clarified from cellular debris by low-speed centrifugation and HERV-K SU

497   purified by passage over Co-NTA agarose (Clontech) followed by gel filtration chromatography

498   on Superdex 200 (General Electric Healthcare, Piscataway, NJ) in 10 mM Tris-HCl, 150 mM NaCl

499   at pH 7.5. Three major peaks were observed: an aggregate of SU that eluted in the void volume

500   of the column, a dimeric species that could be reduced into monomers by addition of a

501   reducing agent (likely the result of a disulfide bond linking two monomers), and a major peak

502   containing a homogeneous monomeric species. For binding assays, only gel filtration

503   chromatography fractions containing the monomeric species were used.

504   **Purification of recombinant HA.**

505   The hemagglutinin (HA) gene of Influenza A virus A/Leningrad/360/1986(H3N2) HA

506   (Accession number CY121277) was synthesized as a gBlock (Integrated DNA Technologies, Inc.,

507   Coralville, IA) and used as a template to amplify the globular head of HA, residues 37-319 (Hong

508   Kong 1968 H3N2 numbering). The resulting PCR product was cloned and expressed from a

25

509   baculovirus recombinant as previously described [63]. The HA head was purified by passage

510   over Co-NTA agarose (Clontech) followed by gel filtration chromatography on Superdex 200 (GE

511   Healthcare) in 10 mM Tris-HCl, 150 mM NaCl at pH 7.5.

Heparin pull downs with purified SU.

513   Briefly, 3 µg of HERV-K SU or HA was pre-incubated with 50 µg/ml heparin, heparan

514   sulfate, condriotin sulfate A, dermatan sulfate, 2-O-desulfated heparin (AMSDSH001-2; AMS

515   Bioscience), 6-O-desulfated heparin (AMSDSH002-6; AMS Bioscience), or no compound in 10

516   mMTris-HCl, 150 mM NaCl, 0.2% TritonX-100 for 1 h at 4°C, and subsequently mixed with 50 ul

517   of heparin, protein A or cobalt beads prior to incubation for 2 h at 4°C. Beads were washed 5

518   times, bound proteins eluted as above, separated on a 4-20% acrylamide gel (Bio-rad) and

519   transferred to nitrocellulose membranes. Membranes were blotted with an antibody against

520   the HA tag (1:5000; Abcam) followed by anti-rabbit HRP antibody (1:5000; Sigma-Aldrich). Data

521   shown are representative of 3 (Fig 3C) or 2 (Fig 3D) independent biological replicates.

Low pH inactivation of virions.

523   Virus was incubated in buffer (10mM Na_2HPO_4, 10mM HEPES, 10mM MES) at various

524   pH (7.0, 6.4, 6.0, 5.6, and 5.2) for 30 min at 37°C. pH was neutralized by adding an excess of

525   DMEM+10% FBS and residual viral infectivity determined by infection of BSRT7 cells. Cells were

526   collected 5 hpi, fixed in 2% PFA and eGFP fluorescence measured using a FACSCalibur

527   instrument. The % of eGFP-positive cells was quantified using FlowJo software and normalized

528   to pH7 treatment controls. Error bars represent standard error of the mean from 3

529   independent biological replicates.

26

Acid bypass of endocytosis.

BSRT7 cells were treated with 100nM bafilomycin A1 (Sigma-Aldrich; B1793) for 30 min at 37°C and VSV or VSV-HERVK subsequently bound by incubating cells with virus an MOI of 5 PFU/cell. for 1h at 4°C. Bound virus was then pulsed with buffer (10mM Na_2HPO_4, 10mM HEPES, 10mM MES) at either pH 7 or pH 5 for 10 min at 37°C, cells were washed twice and then incubated with DMEM (+/- 100nM bafilomycin A1). At 6 hpi cells were collected, fixed in 2% PFA and eGFP fluorescence measured as above. The % eGFP-positive cells in Bafilomycin treated cells is expressed relative to untreated cells. Error bars represent standard error of the mean from three independent biological replicates.

Cell-cell fusion experiments.

Cell-cell fusion experiments were performed as previously described [10]. Briefly, BSRT7 cells were infected with VTF7-3 [64], transfected with pGEM plasmids encoding the *env* of Phoenix, Xq21.33, or HERV-K 108 or empty vector and treated with a 20 min pulse of DMEM of varying pH at 18h post-transfection. Cells were washed, and incubated for 4h at 37°C in DMEM. Cells were fixed in cold methanol prior to Giemsa staining according to manufacturer's protocol (Sigma-Aldrich). Data shown are from a single representative experiment from three biological replicates.

Lentiviral pseudotypes infections.

Lentiviruses pseudotypes were generated by transfecting 293T cells with pCD/NL-BH*DDD, pNL-EGFP/CMV-WPREDU3 [65], and either pCAGGS-PhoenixEnv, pCAGGS-VSVG, or

27

550   pCAGGS empty vector (to generate bald particles). Supernatant was collected 48 hours post

551   transfection and particle concentration was determined using a p24 (HIV-1) antigen capture kit

552   from Advanced Bioscience Labs (Rockville, MD). Supernatant volumes were equilibrated to

553   equal particle amounts, based on p24 values, of each pseudotype virus were used to infect

554   293T, 293T-*B4GALT7*^{KO}, CRFK and CRFK cells treated with 50mM $NaClO_3$ for two passages prior

555   to infection. Supernatant was removed from cells 24 hours post infection and cells were

556   collected 48 hours post infection. For $NaClO_3$ treated cells, $NaClO_3$ was present during the

557   infection and subsequent incubation. eGFP fluorescence was measured using a FACSCalibur

558   instrument and the percentage of eGFP-positive cells, as well as mean fluorescence intensity

559   was quantified using FlowJo software. Error bars represent standard error of the mean from

560   three independent biological replicates.

561   **<u>Acknowledgments</u>**

References

564

565

566   1.   Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing
567 and analysis of the human genome. Nature. 2001;409(6822):860-921. doi: 10.1038/35057062.
568 PubMed PMID: 11237011.

569   2.   Subramanian RP, Wildschutte JH, Russo C, Coffin JM. Identification, characterization,
570 and comparative genomic distribution of the HERV-K (HML-2) group of human endogenous
571 retroviruses. Retrovirology. 2011;8:90. doi: 10.1186/1742-4690-8-90.

572   3.   Steinhuber S, Brack M, Hunsmann G, Schwelberger H, Dierich MP, Vogetseder W.
573 Distribution of human endogenous retrovirus HERV-K genomes in humans and different
574 primates. Human genetics. 1995;96(2):188-92. PubMed PMID: 7635468.

575   4.   Turner G, Barbulescu M, Su M, Jensen-Seaman MI, Kidd KK, Lenz J. Insertional
576 polymorphisms of full-length endogenous retroviruses in humans. Curr Biol. 2001;11(19):1531-
577 5. PubMed PMID: 11591322.

578   5.   Wildschutte JH, Williams ZH, Montesion M, Subramanian RP, Kidd JM, Coffin JM.
579 Discovery of unfixed endogenous retrovirus insertions in diverse human populations. Proc Natl
580 Acad Sci U S A. 2016. doi: 10.1073/pnas.1602336113. PubMed PMID: 27001843.

581   6.   Hanke K, Hohn O, Bannert N. HERV-K(HML-2), a seemingly silent subtenant - but still
582 waters run deep. APMIS : acta pathologica, microbiologica, et immunologica Scandinavica.
583 2016;124(1-2):67-87. doi: 10.1111/apm.12475. PubMed PMID: 26818263.

584   7.   Dewannieux M, Blaise S, Heidmann T. Identification of a functional envelope protein
585 from the HERV-K family of human endogenous retroviruses. J Virol. 2005;79(24):15573-7. Epub
586 2005/11/25. doi: 79/24/15573 [pii]
587 10.1128/JVI.79.24.15573-15577.2005. PubMed PMID: 16306628.

588   8.   Dewannieux M, Harper F, Richaud A, Letzelter C, Ribet D, Pierron G, et al. Identification
589 of an infectious progenitor for the multiple-copy HERV-K human endogenous retroelements.
590 Genome Res. 2006;16(12):1548-56. Epub 2006/11/02. doi: gr.5565706 [pii]
591 10.1101/gr.5565706. PubMed PMID: 17077319.

592   9.   Lee YN, Bieniasz PD. Reconstitution of an infectious human endogenous retrovirus. PLoS
593 Pathog. 2007;3(1):e10. Epub 2007/01/30. doi: 06-PLPA-RA-0426R2 [pii]
594 10.1371/journal.ppat.0030010. PubMed PMID: 17257061.

595   10.   Robinson LR, Whelan SP. Infectious Entry Pathway Mediated by the Human Endogenous
596 Retrovirus K Envelope Protein. J Virol. 2016;90(7):3640-9. doi: 10.1128/JVI.03136-15. PubMed
597 PMID: 26792739; PubMed Central PMCID: PMC4794671.

598   11.   Kramer P, Lausch V, Volkwein A, Hanke K, Hohn O, Bannert N. The human endogenous
599 retrovirus K(HML-2) has a broad envelope-mediated cellular tropism and is prone to inhibition
600 at a post-entry, pre-integration step. Virology. 2015;487:121-8. doi:
601 10.1016/j.virol.2015.10.014. PubMed PMID: 26517399.

602   12.   Carette JE, Raaben M, Wong AC, Herbert AS, Obernosterer G, Mulherkar N, et al. Ebola
603 virus entry requires the cholesterol transporter Niemann-Pick C1. Nature. 2011;477(7364):340-
604 3. Epub 2011/08/26. doi: nature10348 [pii]
605 10.1038/nature10348. PubMed PMID: 21866103.

606 13.   Jae LT, Raaben M, Riemersma M, Beusekom E, Blomen VA, Velds A, et al. Deciphering
607 the glycosylome of dystroglycanopathies using haploid screens for lassa virus entry. Science.
608 2013;340(6131):479-83. doi: 10.1126/science.1233675.
609 14.   Jae LT, Raaben M, Herbert AS, Kuehne AI, Wirchnianski AS, Soh TK, et al. Virus entry.
610 Lassa virus entry requires a trigger-induced receptor switch. Science (New York, NY).
611 2014;344(6191):1506-10. doi: 10.1126/science.1252480.
612 15.   Raaben M, Jae LT, Herbert AS, Kuehne AI, Stubbs SH, Chou YY, et al. NRP2 and CD63 Are
613 Host Factors for Lujo Virus Cell Entry. Cell host & microbe. 2017;22(5):688-96 e5. doi:
614 10.1016/j.chom.2017.10.002. PubMed PMID: 29120745.
615 16.   Kleinfelter LM, Jangra RK, Jae LT, Herbert AS, Mittler E, Stiles KM, et al. Haploid Genetic
616 Screen Reveals a Profound and Direct Dependence on Cholesterol for Hantavirus Membrane
617 Fusion. MBio. 2015;6(4):e00801. doi: 10.1128/mBio.00801-15. PubMed PMID: 26126854;
618 PubMed Central PMCID: PMCPMC4488941.
619 17.   Riblett AM, Blomen VA, Jae LT, Altamura LA, Doms RW, Brummelkamp TR, et al. A
620 Haploid Genetic Screen Identifies Heparan Sulfate Proteoglycans Supporting Rift Valley Fever
621 Virus Infection. J Virol. 2015;90(3):1414-23. doi: 10.1128/JVI.02055-15. PubMed PMID:
622 26581979; PubMed Central PMCID: PMCPMC4719632.
623 18.   Prydz K. Determinants of Glycosaminoglycan (GAG) Structure. Biomolecules.
624 2015;5(3):2003-22. doi: 10.3390/biom5032003. PubMed PMID: 26308067; PubMed Central
625 PMCID: PMC4598785.
626 19.   Shieh MT, WuDunn D, Montgomery RI, Esko JD, Spear PG. Cell surface receptors for
627 herpes simplex virus are heparan sulfate proteoglycans. J Cell Biol. 1992;116(5):1273-81.
628 PubMed PMID: 1310996; PubMed Central PMCID: PMC2289355.
629 20.   Gardner CL, Ebel GD, Ryman KD, Klimstra WB. Heparan sulfate binding by natural
630 eastern equine encephalitis viruses promotes neurovirulence. Proc Natl Acad Sci U S A.
631 2011;108(38):16026-31. doi: 10.1073/pnas.1110617108. PubMed PMID: 21896745; PubMed
632 Central PMCID: PMC3179095.
633 21.   Cote M, Zheng YM, Liu SL. Receptor binding and low pH coactivate oncogenic retrovirus
634 envelope-mediated fusion. J Virol. 2009;83(22):11447-55. doi: 10.1128/JVI.00748-09. PubMed
635 PMID: 19726505; PubMed Central PMCID: PMC2772678.
636 22.   Wang E, Obeng-Adjei N, Ying Q, Meertens L, Dragic T, Davey RA, et al. Mouse mammary
637 tumor virus uses mouse but not human transferrin receptor 1 to reach a low pH compartment
638 and infect cells. Virology. 2008;381(2):230-40. doi: 10.1016/j.virol.2008.08.013.
639 23.   Côté M, Kucharski TJ, Liu S-LL. Enzootic nasal tumor virus envelope requires a very acidic
640 pH for fusion activation and infection. Journal of virology. 2008;82(18):9023-34. doi:
641 10.1128/JVI.00648-08.
642 24.   Mothes W, Boerger AL, Narayan S, Cunningham JM, Young JA. Retroviral entry mediated
643 by receptor priming and low pH triggering of an envelope glycoprotein. Cell. 2000;103(4):679-
644 89. PubMed PMID: 11106737.
645 25.   Doms RW, Keller DS, Helenius A, Balch WE. Role for adenosine triphosphate in
646 regulating the assembly and transport of vesicular stomatitis virus G protein trimers. J Cell Biol.
647 1987;105(5):1957-69. PubMed PMID: 2824524; PubMed Central PMCID: PMCPMC2114842.

648   26.   Puri A, Winick J, Lowy RJ, Covell D, Eidelman O, Walter A, et al. Activation of vesicular
649   stomatitis virus fusion with cells by pretreatment at low pH. J Biol Chem. 1988;263(10):4749-53.
650   PubMed PMID: 2832405.
651   27.   Gaudin Y, Tuffereau C, Segretain D, Knossow M, Flamand A. Reversible conformational
652   changes and fusion activity of rabies virus glycoprotein. J Virol. 1991;65(9):4853-9. PubMed
653   PMID: 1870204; PubMed Central PMCID: PMCPMC248944.
654   28.   Mayer J, Sauter M, Racz A, Scherer D, Mueller-Lantzsch N, Meese E. An almost-intact
655   human endogenous retrovirus K on human chromosome 7. Nat Genet. 1999;21(3):257-8. doi:
656   10.1038/6766. PubMed PMID: 10080172.
657   29.   Nasimuzzaman M, Persons DA. Cell Membrane-associated heparan sulfate is a receptor
658   for prototype foamy virus in human, monkey, and rodent cells. Molecular therapy : the journal
659   of the American Society of Gene Therapy. 2012;20(6):1158-66. doi: 10.1038/mt.2012.41.
660   PubMed PMID: 22434139; PubMed Central PMCID: PMC3369305.
661   30.   Plochmann K, Horn A, Gschmack E, Armbruster N, Krieg J, Wiktorowicz T, et al. Heparan
662   sulfate is an attachment factor for foamy virus entry. J Virol. 2012;86(18):10028-35. doi:
663   10.1128/JVI.00051-12. PubMed PMID: 22787203; PubMed Central PMCID: PMC3446549.
664   31.   Picard-Maureau M, Jarmy G, Berg A, Rethwilm A, Lindemann D. Foamy virus envelope
665   glycoprotein-mediated entry involves a pH-dependent fusion process. J Virol. 2003;77(8):4722-
666   30. PubMed PMID: 12663779; PubMed Central PMCID: PMC152125.
667   32.   Zhang Y, Rassa JC, deObaldia ME, Albritton LM, Ross SR. Identification of the receptor
668   binding domain of the mouse mammary tumor virus envelope protein. Journal of virology.
669   2003;77(19):10468-78. doi: 10.1128/JVI.77.19.10468-10478.2003.
670   33.   Lambert S, Bouttier M, Vassy R, Seigneuret M, Petrow-Sadowski C, Janvier S, et al. HTLV-
671   1 uses HSPG and neuropilin-1 for entry by molecular mimicry of VEGF165. Blood.
672   2009;113(21):5176-85. doi: 10.1182/blood-2008-04-150342. PubMed PMID: 19270265;
673   PubMed Central PMCID: PMC2686187.
674   34.   Manel N, Kim FJ, Kinet S, Taylor N, Sitbon M, Battini JL. The ubiquitous glucose
675   transporter GLUT-1 is a receptor for HTLV. Cell. 2003;115(4):449-59. PubMed PMID: 14622599.
676   35.   Briggs JA, Watson BE, Gowen BE, Fuller SD. Cryoelectron microscopy of mouse
677   mammary tumor virus. J Virol. 2004;78(5):2606-8. Epub 2004/02/14. PubMed PMID: 14963166;
678   PubMed Central PMCID: PMCPMC369243.
679   36.   Grow EJ, Flynn RA, Chavez SL, Bayless NL, Wossidlo M, Wesche DJ, et al. Intrinsic
680   retroviral reactivation in human preimplantation embryos and pluripotent cells. Nature.
681   2015;522(7555):221-5. doi: 10.1038/nature14308. PubMed PMID: 25896322; PubMed Central
682   PMCID: PMC4503379.
683   37.   Iozzo RV, Sanderson RD. Proteoglycans in cancer biology, tumour microenvironment and
684   angiogenesis. Journal of cellular and molecular medicine. 2011;15(5):1013-31. doi:
685   10.1111/j.1582-4934.2010.01236.x. PubMed PMID: 21155971; PubMed Central PMCID:
686   PMC3633488.
687   38.   Whitelock JM, Melrose J, Iozzo RV. Diverse cell signaling events modulated by perlecan.
688   Biochemistry. 2008;47(43):11174-83. doi: 10.1021/bi8013938. PubMed PMID: 18826258;
689   PubMed Central PMCID: PMC2605657.

690   39.   Lindahl U, Li JP. Interactions between heparan sulfate and proteins-design and
691   functional implications. International review of cell and molecular biology. 2009;276:105-59.
692   doi: 10.1016/S1937-6448(09)76003-4. PubMed PMID: 19584012.
693   40.   Hacker U, Nybakken K, Perrimon N. Heparan sulphate proteoglycans: the sweet side of
694   development. Nature reviews Molecular cell biology. 2005;6(7):530-41. doi: 10.1038/nrm1681.
695   PubMed PMID: 16072037.
696   41.   Lavialle C, Cornelis G, Dupressoir A, Esnault C, Heidmann O, Vernochet C, et al.
697   Paleovirology of 'syncytins', retroviral env genes exapted for a role in placentation.
698   Philosophical transactions of the Royal Society of London Series B, Biological sciences.
699   2013;368(1626):20120507. doi: 10.1098/rstb.2012.0507.
700   42.   Blanco-Melo D, Gifford RJ, Bieniasz PD. Co-option of an endogenous retrovirus envelope
701   for host defense in hominid ancestors. eLife. 2017;6. doi: 10.7554/eLife.22519. PubMed PMID:
702   28397686; PubMed Central PMCID: PMCPMC5388530.
703   43.   Heidmann O, Beguin A, Paternina J, Berthier R, Deloger M, Bawa O, et al. HEMO, an
704   ancestral endogenous retroviral envelope protein shed in the blood of pregnant women and
705   expressed in pluripotent stem cells and tumors. Proc Natl Acad Sci U S A. 2017;114(32):E6642-
706   E51. doi: 10.1073/pnas.1702204114. PubMed PMID: 28739914; PubMed Central PMCID:
707   PMCPMC5559007.
708   44.   Fuchs NV, Loewer S, Daley GQ, Izsvak Z, Lower J, Lower R. Human endogenous retrovirus
709   K (HML-2) RNA and protein expression is a marker for human embryonic and induced
710   pluripotent stem cells. Retrovirology. 2013;10:115. doi: 10.1186/1742-4690-10-115. PubMed
711   PMID: 24156636; PubMed Central PMCID: PMC3819666.
712   45.   Ross SR, Schofield JJ, Farr CJ, Bucan M. Mouse transferrin receptor 1 is the cell entry
713   receptor for mouse mammary tumor virus. Proceedings of the National Academy of Sciences of
714   the United States of America. 2002;99(19):12386-90. doi: 10.1073/pnas.192360099.
715   46.   Miller AD, Bergholz U, Ziegler M, Stocking C. Identification of the myelin protein
716   plasmolipin as the cell entry receptor for Mus caroli endogenous retrovirus. J Virol.
717   2008;82(14):6862-8. doi: 10.1128/JVI.00397-08. PubMed PMID: 18463156; PubMed Central
718   PMCID: PMC2446966.
719   47.   Tailor CS, Nouri A, Lee CG, Kozak C, Kabat D. Cloning and characterization of a cell
720   surface receptor for xenotropic and polytropic murine leukemia viruses. Proc Natl Acad Sci U S
721   A. 1999;96(3):927-32. PubMed PMID: 9927670; PubMed Central PMCID: PMC15327.
722   48.   Albritton LM, Tseng L, Scadden D, Cunningham JM. A putative murine ecotropic
723   retrovirus receptor gene encodes a multiple membrane-spanning protein and confers
724   susceptibility to virus infection. Cell. 1989;57(4):659-66. PubMed PMID: 2541919.
725   49.   Miller DG, Miller AD. A family of retroviruses that utilize related phosphate transporters
726   for cell entry. J Virol. 1994;68(12):8270-6. PubMed PMID: 7966619; PubMed Central PMCID:
727   PMC237294.
728   50.   Miller DG, Edwards RH, Miller AD. Cloning of the cellular receptor for amphotropic
729   murine retroviruses reveals homology to that for gibbon ape leukemia virus. Proc Natl Acad Sci
730   U S A. 1994;91(1):78-82. PubMed PMID: 8278411; PubMed Central PMCID: PMC42889.
731   51.   Hein S, Prassolov V, Zhang Y, Ivanov D, Lohler J, Ross SR, et al. Sodium-dependent myo-
732   inositol transporter 1 is a cellular receptor for Mus cervicolor M813 murine leukemia virus. J
733   Virol. 2003;77(10):5926-32. PubMed PMID: 12719585; PubMed Central PMCID: PMC154034.

734 52.   Soll SJ, Neil SJ, Bieniasz PD. Identification of a receptor for an extinct virus. Proc Natl
735 Acad Sci U S A. 2010;107(45):19496-501. doi: 10.1073/pnas.1012344107. PubMed PMID:
736 20974973; PubMed Central PMCID: PMC2984158.
737 53.   Buchholz UJ, Finke S, Conzelmann KK. Generation of bovine respiratory syncytial virus
738 (BRSV) from cDNA: BRSV NS2 is not essential for virus replication in tissue culture, and the
739 human RSV leader region acts as a functional BRSV genome promoter. J Virol. 1999;73(1):251-9.
740 PubMed PMID: 9847328; PubMed Central PMCID: PMC103829.
741 54.   Whelan SP, Barr JN, Wertz GW. Identification of a minimal size requirement for
742 termination of vesicular stomatitis virus mRNA: implications for the mechanism of transcription.
743 Journal of virology. 2000;74(18):8268-76.
744 55.   Sanjana NE, Shalem O, Zhang F. Improved vectors and genome-wide libraries for CRISPR
745 screening. Nature methods. 2014;11(8):783-4. doi: 10.1038/nmeth.3047. PubMed PMID:
746 25075903; PubMed Central PMCID: PMC4486245.
747 56.   Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering
748 using CRISPR/Cas systems. Science. 2013;339(6121):819-23. doi: 10.1126/science.1231143.
749 PubMed PMID: 23287718; PubMed Central PMCID: PMC3795411.
750 57.   Lackner DH, Carre A, Guzzardo PM, Banning C, Mangena R, Henley T, et al. A generic
751 strategy for CRISPR-Cas9-mediated gene tagging. Nature communications. 2015;6:10237. doi:
752 10.1038/ncomms10237. PubMed PMID: 26674669; PubMed Central PMCID: PMC4703899.
753 58.   Zhang XY, La Russa VF, Reiser J. Transduction of bone-marrow-derived mesenchymal
754 stem cells by using lentivirus vectors pseudotyped with modified RD114 envelope
755 glycoproteins. J Virol. 2004;78(3):1219-29. PubMed PMID: 14722277; PubMed Central PMCID:
756 PMC321376.
757 59.   Langereis MA, Rabouw HH, Holwerda M, Visser LJ, van Kuppeveld FJ. Knockout of cGAS
758 and STING Rescues Virus Infection of Plasmid DNA-Transfected Cells. J Virol.
759 2015;89(21):11169-73. doi: 10.1128/JVI.01781-15. PubMed PMID: 26311870; PubMed Central
760 PMCID: PMCPMC4621133.
761 60.   Yamashita M, Emerman M. Capsid is a dominant determinant of retrovirus infectivity in
762 nondividing cells. J Virol. 2004;78(11):5670-8. doi: 10.1128/JVI.78.11.5670-5678.2004. PubMed
763 PMID: 15140964; PubMed Central PMCID: PMC415837.
764 61.   Cureton DK, Massol RH, Saffarian S, Kirchhausen TL, Whelan SP. Vesicular stomatitis
765 virus enters cells through vesicles incompletely coated with clathrin that depend upon actin for
766 internalization. PLoS pathogens. 2009;5(4). doi: 10.1371/journal.ppat.1000394.
767 62.   Lyles DS, Puddington L, McCreedy BJ, Jr. Vesicular stomatitis virus M protein in the
768 nuclei of infected cells. J Virol. 1988;62(11):4387-92. PubMed PMID: 2845149; PubMed Central
769 PMCID: PMC253880.
770 63.   Schmidt AG, Xu H, Khan AR, O'Donnell T, Khurana S, King LR, et al. Preconfiguration of
771 the antigen-binding site during affinity maturation of a broadly neutralizing influenza virus
772 antibody. Proc Natl Acad Sci U S A. 2013;110(1):264-9. doi: 10.1073/pnas.1218256109. PubMed
773 PMID: 23175789; PubMed Central PMCID: PMC3538208.
774 64.   Fuerst TR, Niles EG, Studier FW, Moss B. Eukaryotic transient-expression system based
775 on recombinant vaccinia virus that synthesizes bacteriophage T7 RNA polymerase. Proc Natl
776 Acad Sci U S A. 1986;83(21):8122-6. PubMed PMID: 3095828; PubMed Central PMCID:
777 PMC386879.

33

778   65.   Ricks DM, Kutner R, Zhang XY, Welsh DA, Reiser J. Optimized lentiviral transduction of
779   mouse bone marrow-derived mesenchymal stem cells. Stem cells and development.
780   2008;17(3):441-50. doi: 10.1089/scd.2007.0194. PubMed PMID: 18513160; PubMed Central
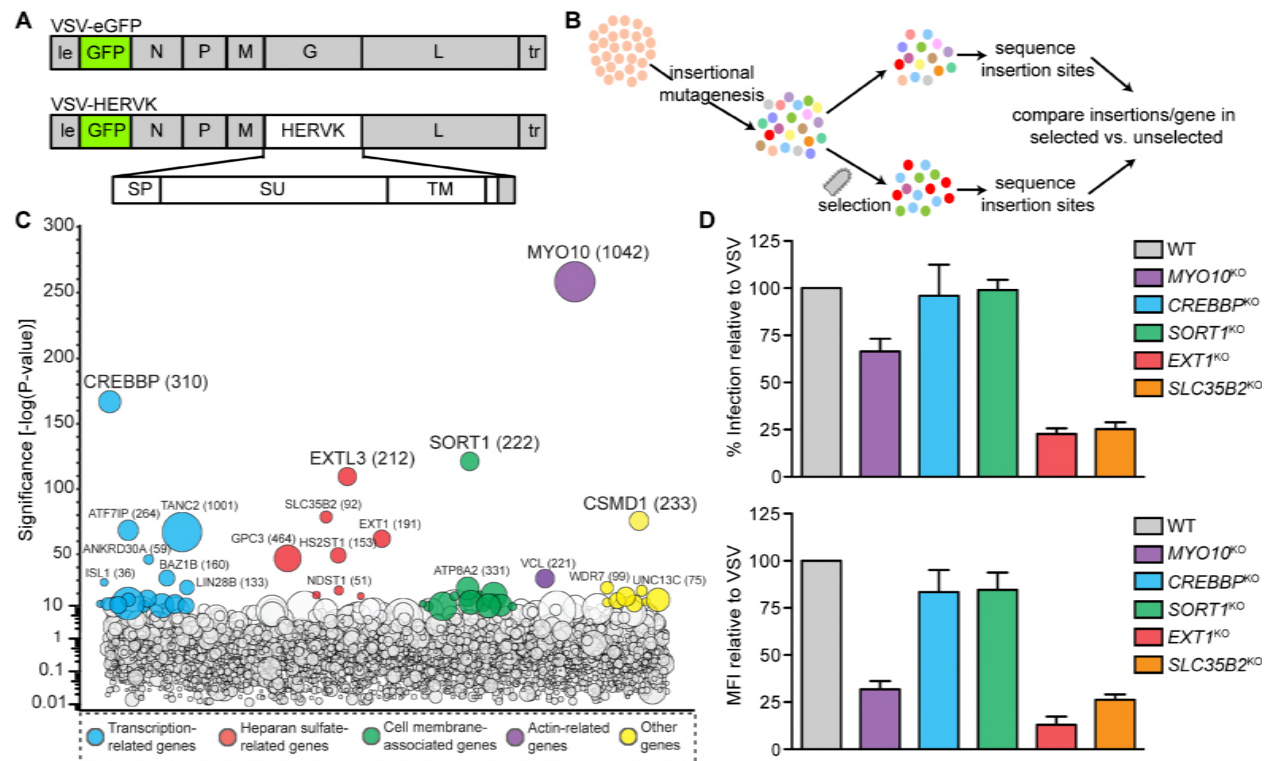781   PMCID: PMC2996877.
782

783

Fig 1. Haploid genetic screen identifies host factors required for infection by VSV-HERVK. (A)
Viral genome structures. Sequences from VSV are shown in grey: N, nucleocapsid; P, phosphoprotein; M, matrix; G, glycoprotein; L, large polymerase. le: leader. tr: trailer. Viruses also encode an eGFP reporter gene. VSV-HERVK encodes the HERVK glycoprotein, which contains the signal peptide (SP), surface (SU) subunit, transmembrane subunit (TM), and membrane-spanning domain of HERV-K env, and cytoplasmic tail of VSV G. **(B)** Schematic of haploid genetic screen. HAP1 cells were subjected to insertional mutagenesis, followed by selection with VSV-HERVK. Surviving cells were deep sequenced to identify the position of insertion sites. The number of insertions per gene in the selected set was compared to that of an unselected set to identify genes that were associated with survival of infection. **(C)** Screen results. The y-axis indicates the significance of enrichment of gene-trap insertions compared with unselected control cells. Circles represent individual genes and their size corresponds to

797   the number of unique insertion sites in the selected population. Genes with significance scores

798   above 10 are colored according to function and grouped horizontally. Genes with significance

799   score above 25 are labeled. **(D)** HAP1 cells were gene edited to lack the indicated genes and

800   infected with VSV or VSV-HERVK. The fold difference in percent infected cells (top) and mean

801   fluorescence intensity (MFI, bottom) of VSV-HERVK infected cells normalized to that of VSV is

802   shown. Error bars represent standard error of the mean (SEM) for at least three independent

803   experiments.

804

Fig 2. Heparan sulfate facilitates HERV-K Env-mediated entry and attachment. (A) WT HAP1,

EXT1^{KO} Neo^r Ctrl (transduced with a control retrovirus), *EXT1^{KO}* EXT1-HA (+ addback), *SLC35B2^{KO}*

Neo^r Ctrl and *SLC35B2^{KO}* SLC25B2-HA (+ addback) cells were infected with VSV-HERVK or VSV

and infectivity analyzed by flow cytometry. Fold difference in percent infected cells compared

to WT for VSV-HERVK was normalized to that of VSV for each condition. Error bars represent

810   SEM for four independent experiments. **(B)** Relative MFI of cells from (A). Data were normalized

811   as in (A). **(C)** BSRT7 cells were treated with 50mM sodium chlorate and infected with either VSV

812   or VSV-HERVK. Fold difference in both percent infected cells (left) and MFI (right) compared to

813   untreated cells for VSV-HERVK was normalized to that of VSV. Error bars represent SEM for

814   three independent experiments. **(D)** VSV or VSV-HERVK was incubated with the indicated

815   soluble glycosaminoglycans prior to infecting BSRT7 cells. Percent infected cells was normalized

816   to untreated virus controls. Error bars represent SEM for three independent experiments. **(E)**

817   Schematic of virus attachment experiment. Cells were blocked with BSA then incubated with

818   both fluorescently labeled VSV-HERVK and VSV at either 37°C or 4°C. **(F)** Representative images

819   from 4°C attachment experiment. Red: VSV. Green: VSV-HERVK. Blue: DAPI. Grey: calcein. **(G)**

820   Results of attachment experiment. Numbers of particles/μm^2 are plotted. Grey circles indicate

821   outliers. Total number of particles counted per condition is indicated above each box.
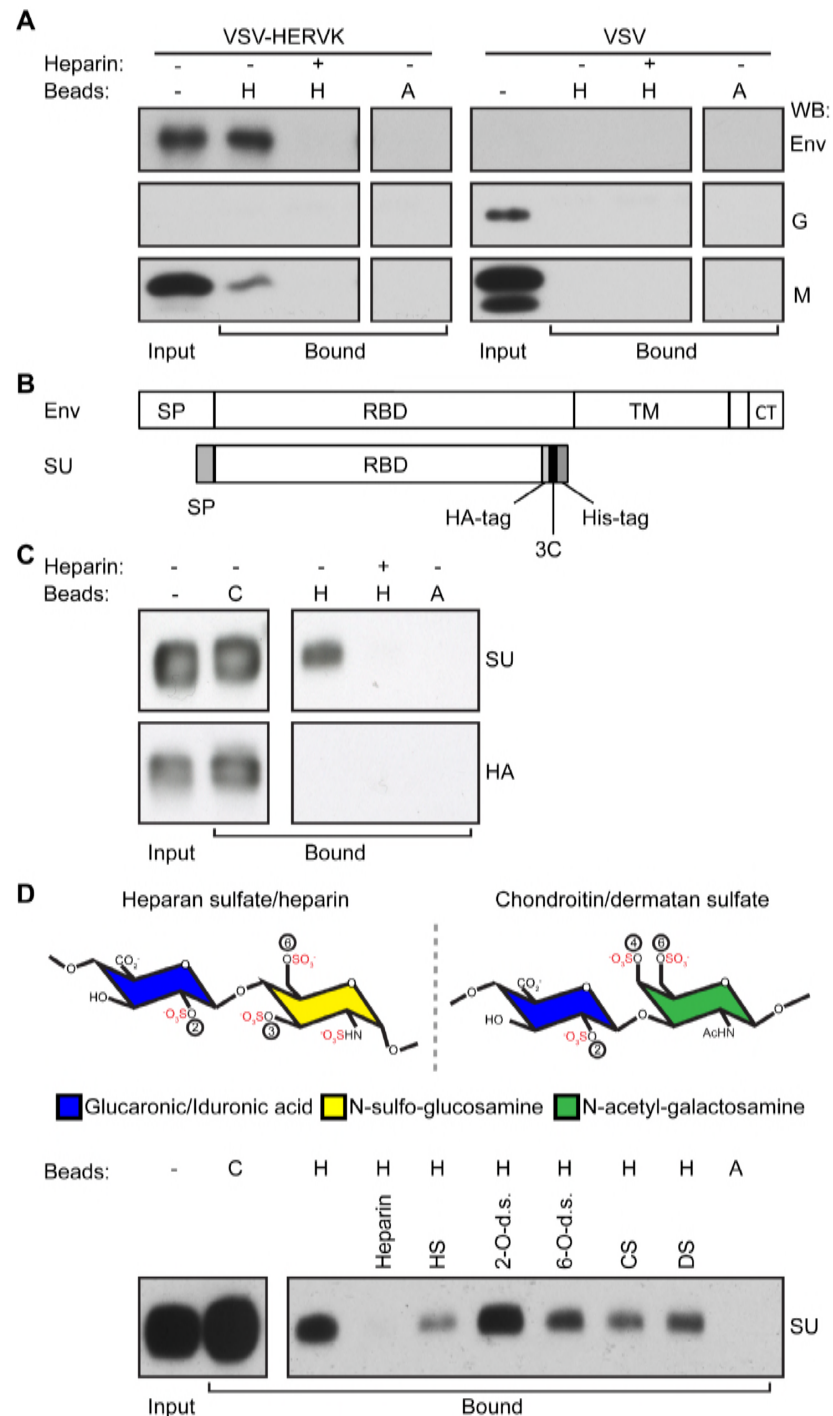
822

Fig 3. HERV-K Env binds heparin. (A) Purified VSV-HERVK or VSV were incubated with heparin

(H) or protein A (A) beads, with (+) or without (-) soluble heparin added as a competitor. Bound

virions were analyzed by Western blot against HERV-K Env, VSV G, and VSV M. Input: total input

virus. **(B)** Schematic of HERV-K Env and HERV-K SU used in this study. SP: signal peptide. RBD:

827   receptor binding domain. TM: transmembrane subunit. CT: cytoplasmic tail. 3C: 3C protease

828   cleavage site. **(C)** HERV-K SU or Influenza A HA receptor binding domain (HA) were pre-

829   incubated with or without soluble heparin prior to incubation with either cobalt (C, maximum

830   pull-down control), heparin (H), or protein A (A) agarose beads. Bound protein was eluted from

831   the beads and subjected to SDS-PAGE followed by Western blot against the HA tag. Input: 10%

832   of total input protein. **(D)** Top: Structure of glycosaminoglycans. The repeating disaccharides of

833   heparan sulfate/heparin (left) and chondroitin/dermatan sulfate (right) are shown. Sulfates are

834   highlighted in red. Positions of O-sulfations are indicated with circled numbers. Disaccharides

835   are shown as fully sulfated, however individual sugars will not always be sulfated at each

836   position. Bottom: HERV-K SU was pre-incubated with soluble competitor compounds (heparin,

837   heparan sulfate, 2-O-desulfated heparin, 6-O-desulfated heparin, chondroitin sulfate A, and

838   dermatan sulfate) prior to incubation with either cobalt, heparin, or protein A agarose beads.

839   Bound protein was eluted from the beads and subjected to SDS-PAGE followed by Western blot
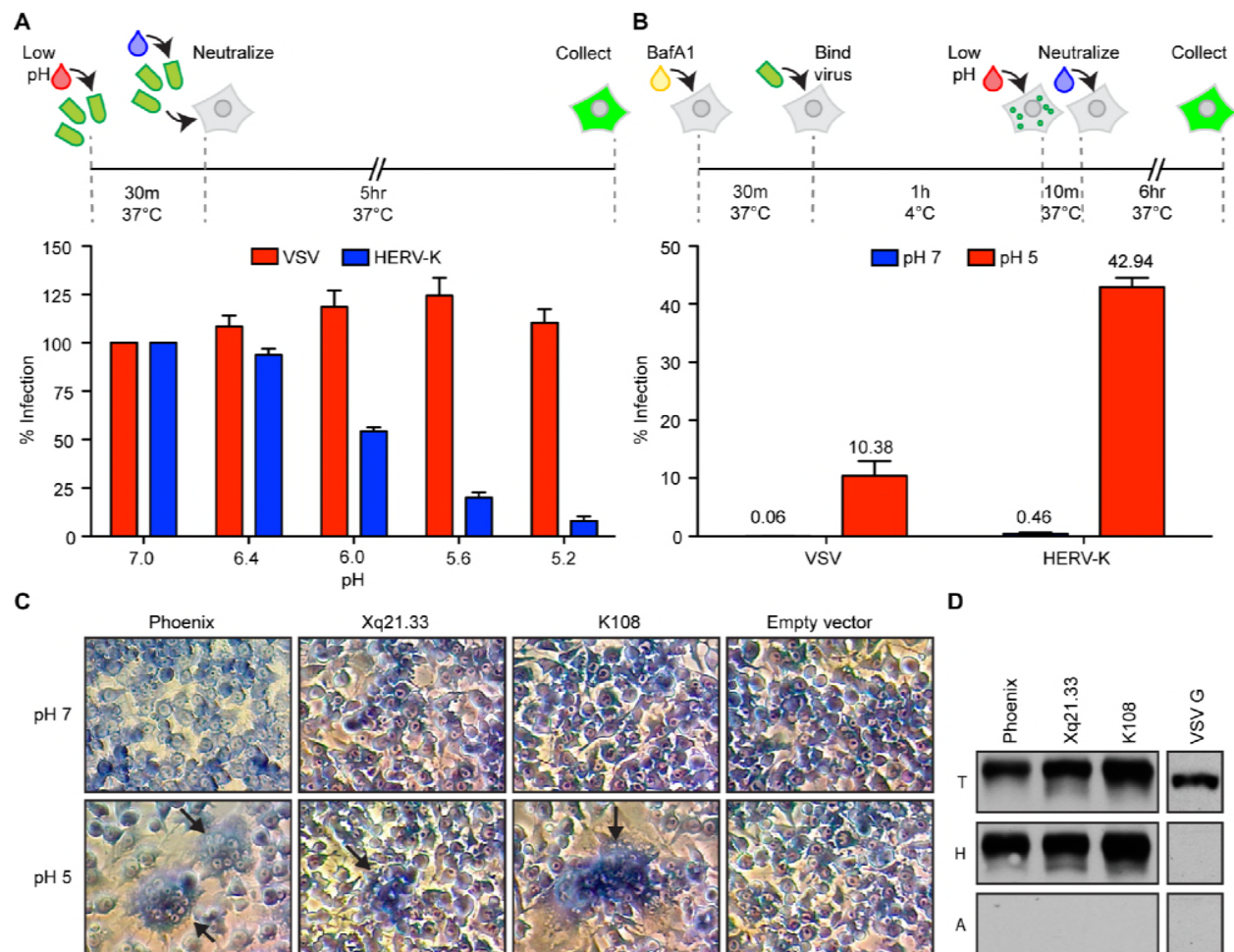
840   against the HA tag. Input: 10% of total input.

841

Fig 4. Acidic pH is sufficient to trigger HERV-K Env. (A) HERV-K Env is inactivated by exposure to acidic pH. VSV-HERVK and VSV were incubated in buffer at the indicated pH for 30 minutes at 37°C. Samples were returned to neutral pH before infecting BSRT7 cells. Cells were collected 5 hours post infection and percent GFP-expressing cells quantified by flow cytometry. Values are normalized to the pH 7 condition. Error bars represent SEM from three independent experiments. **(B)** VSV-HERVK fuses at the plasma membrane when treated with acidic pH. BSRT7 cells were pre-treated with bafilomycinA1 prior to binding virus at 4°C. Cells were treated with buffer at pH 7 or 5. Unbound virus was washed off and cells were collected 6 hours post-infection. Percent infected cells was normalized to cells not treated with bafilomycinA1.

41

851   Error bars represent SEM for three independent experiments. **(C)** Endogenous HERV-K Envs are

852   fusogenic at acidic pH. BSRT7 cells were transfected with *envs* from Phoenix, Xq21.33, and

853   HERV-K 108 and subsequently exposed to buffer at the indicated pH. Syncytia are highlighted

854   with arrows. Data are from a single representative experiment. **(D)** Endogenous HERV-K Envs

855   bind heparin. 293T cells were transfected with the *envs* from Phoenix, Xq21.33, HERV-K 108, or

856   VSV-G. Cell lysates were incubated with either heparin (H) or protein A (A) beads and bound

857   protein analyzed by Western blot against HERV-K Env and VSV-G. T: 10% of total input. Data are

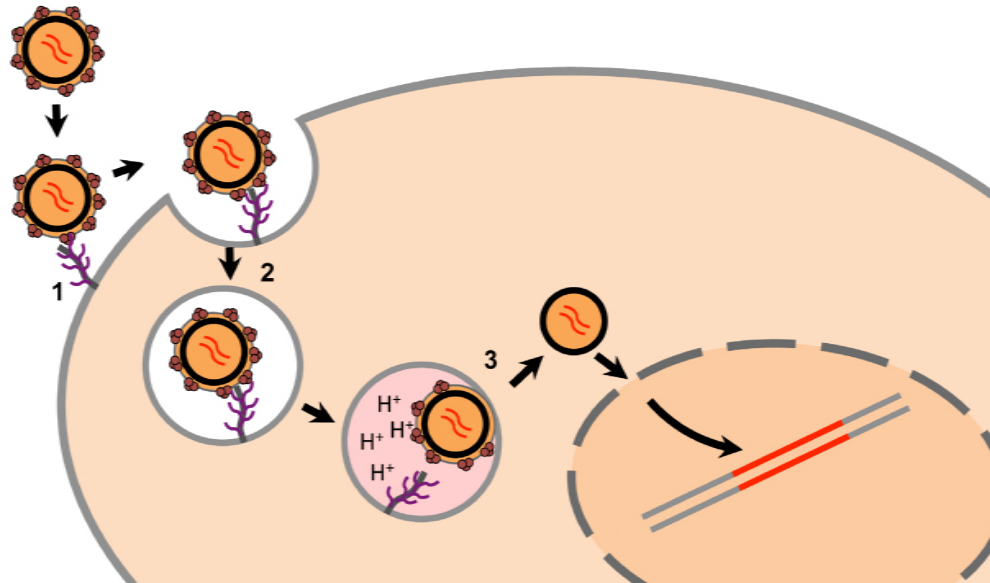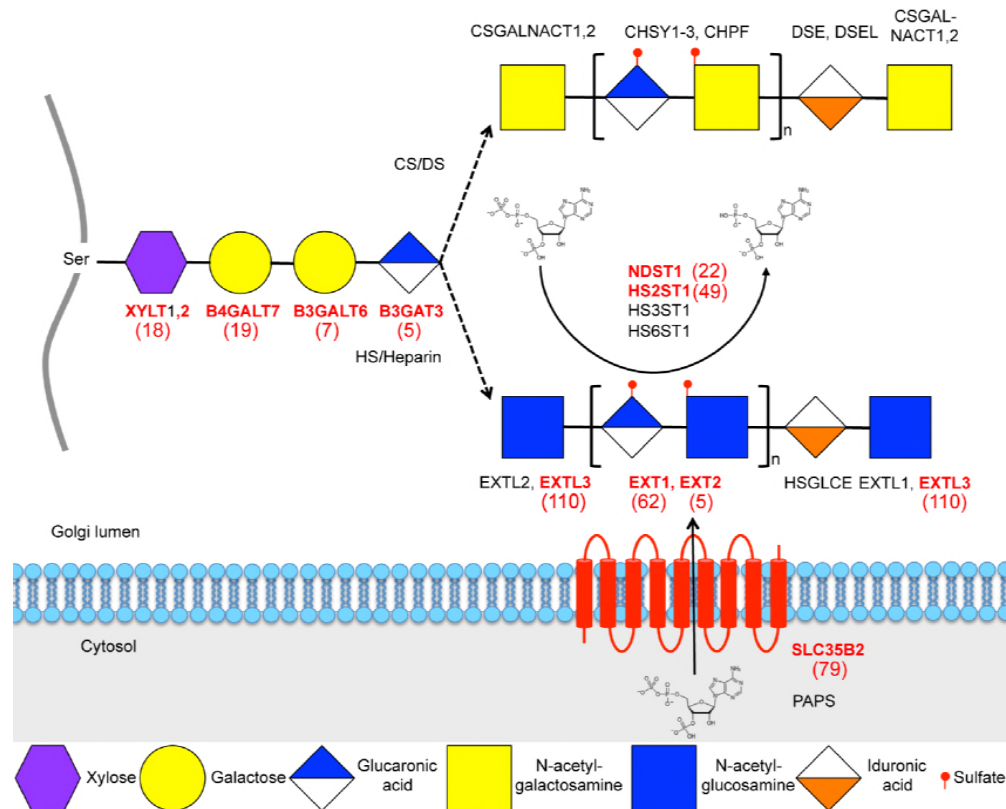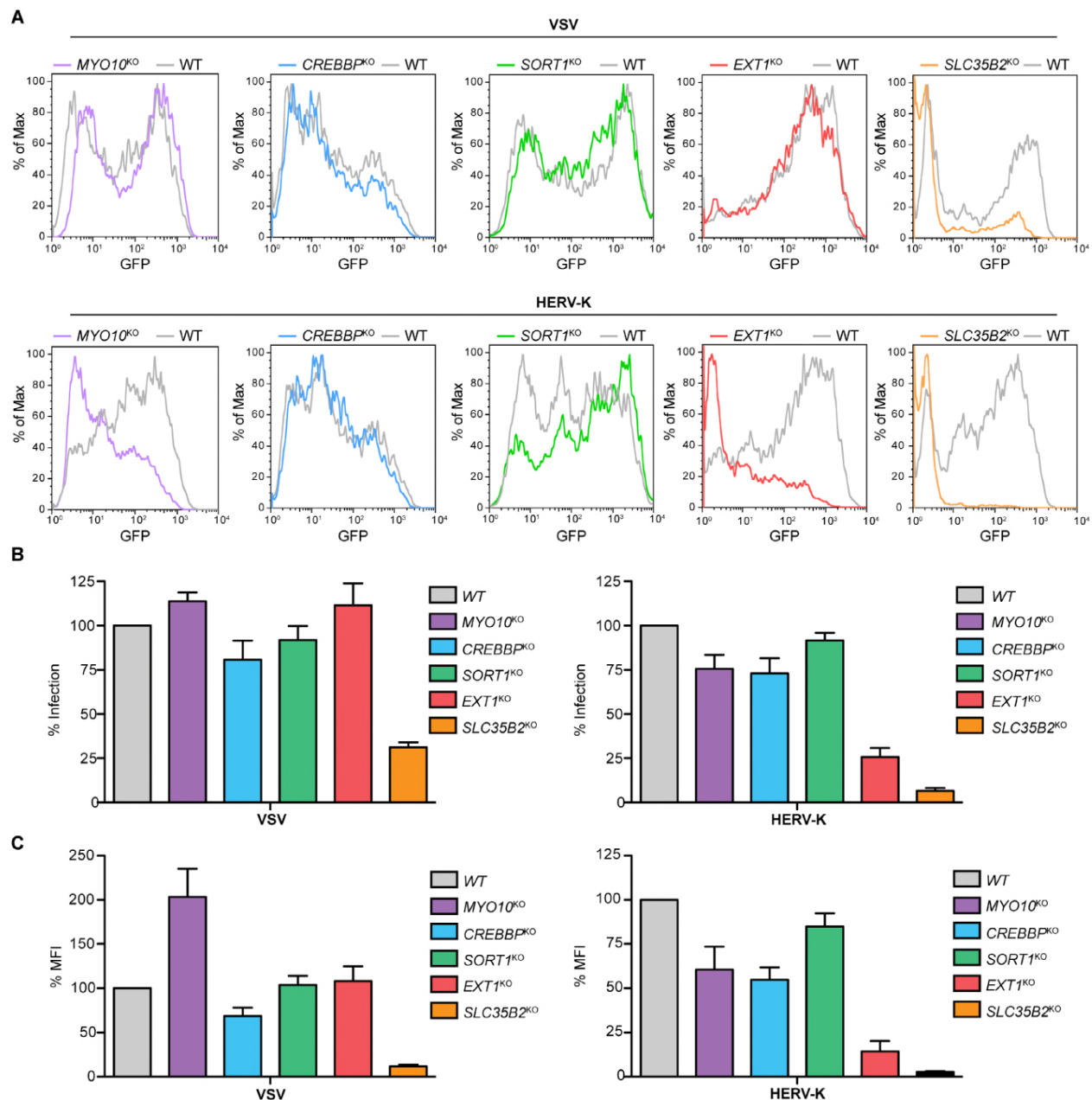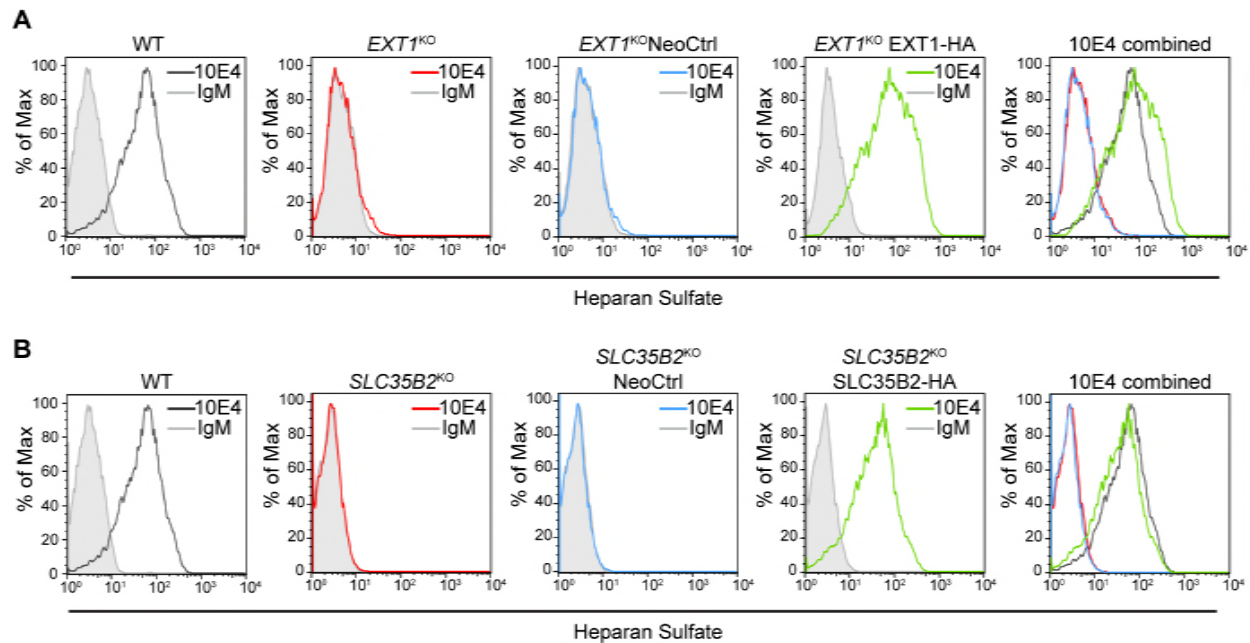858   from a single representative experiments.

859

860

861

Fig 5. Proposed model of HERV-K entry. We propose a 3-step model for HERV-K entry: 1. HERV-K binds heparan sulfate on the cell surface to attach to the cell. 2. The virus is taken up by dynamin-dependent, clathrin-independent endocytosis. 3. Exposure to low pH following endosomal acidification triggers Env to fuse the viral and cellular membranes, releasing the viral core into the cytoplasm.
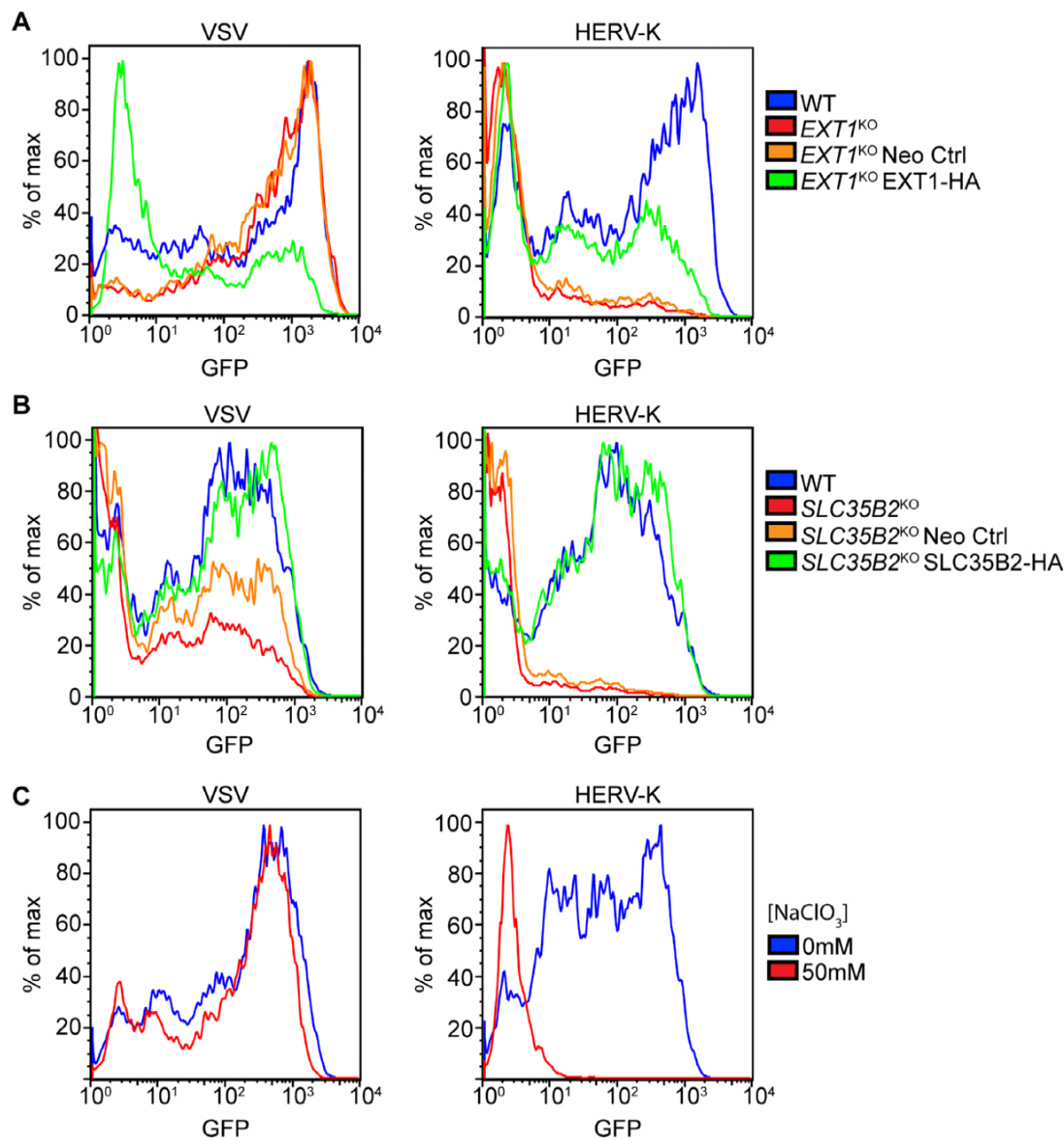
867

S1 Fig. Cartoon schematic of the glycosaminoglycan (GAG) synthesis pathway. GAGs are added to a core protein (in grey). There is a core linkage of 4 sugars. The pathway then splits into the heparan sulfate/heparin pathway and the chondroitin sulfate/dermatan sulfate pathway. The enzymes that catalyze the sugar addition are written above/below the sugars. Sulfation is catalyzed by enzymes NDST1, HS2ST1, HS3ST1, and HS6ST1-3. Each enzyme adds a sulfate to a different position on the sugar. The sulfate donor, PAPS is transported into the Golgi by SLC35B2. Genes highlighted in red were identified as hits in the haploid screen. The significance score for each hit, rounded to the nearest integer, is indicated in parentheses.

876
877 **S2 Fig. Histograms of GFP expression and infectivity in cells from Fig 1. (A)** Representative

878 histograms from experiments in Fig 1D. Top: VSV infected cells. Bottom: VSV-HERVK infected

879 cells. Histograms are from a single representative experiment. **(B)** Infectivity of VSV and VSV-

880 HERVK in gene edited cells. Data are from the same experiment as Fig 1D and are normalized to

881 infectivity in WT cells. **(C)** MFI of cells in (B). MFI of all cells for each condition are normalized to

882 that of WT cells.

45

883

S3 Fig. Heparan sulfate expression of *EXT1*KO and *SLC35B2*KO HAP1 cells. The indicated cell lines were stained with 10E4, a heparan sulfate-specific antibody, or mouse IgM isotype control antibody, and analyzed by flow cytometry. **(A)** WT HAP1, *EXT1KO*, *EXT1KO*Neo[r] Control, and EXT1KO+EXT1-HA cells. **(B)** WT HAP1, *SLC35B2KO*, *SLC35B2KO*Neo[r] Control, and *SLC35B2KO*+SLC35B2-HA cells. Representative histograms are shown from a single experiment. The WT histograms in (A) and (B) are the same sample and therefore identical. Histograms are from a single representative experiment.

891

S4 Fig. Histograms of GFP expression in cells from Fig 2. Representative histograms are shown from experiments from experiments in Fig 2A, 2B, and 2C. **(A)** $EXT1^{KO}$ cells. **(B)** $SLC35B2^{KO}$ cells. **(C)** Sodium chlorate treated cells. Representative histograms are shown from a single experiment.
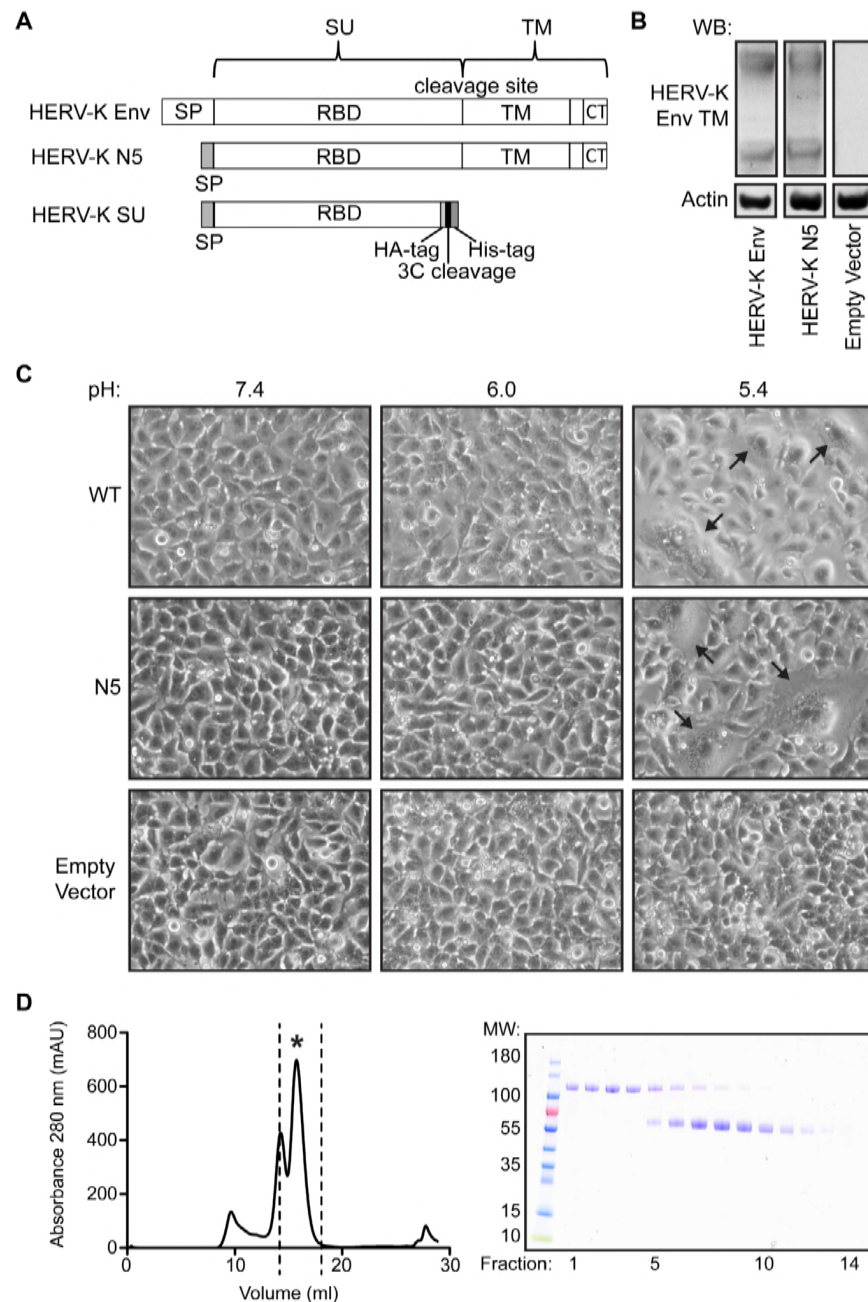
A

```
          N-Terminal Truncation          SU C-Terminal Truncation
AA:      86                  106     418                                           461
         |                   |       |                                             |
WT:      MVVSLPMPAGAAAANYTYWAY...     |                                             |
N1:                       WAY...      |                                             |
N2:                      YWAY...      |                                             |
N3:                     YTYWAY...     |                                             |
N4:                    NYTYWAY...     |                                             |
N5:                  AAAANYTYWAY...   |                                             |
N6:            LPMPAGAAAANYTYWAY...   |                                             |
N7:      MVVSLPMPAGAAAANYTYWAY...     |                                             |
N5C0:                AAAANYTYWAY..... FNWQHRILLVRAREGVWIPVSMDRPWEASPSIHILTEVLKGVLN
N4C0:                  NYTYWAY..... FNWQHRILLVRAREGVWIPVSMDRPWEASPSIHILTEVLKGVLN
N5C1:                AAAANYTYWAY..... FNWQHRILLVRAREGVWIPVSMDRPWE
N5C2:                AAAANYTYWAY..... FNWQHRILLVRAREGVWIPVS
N5C3:                AAAANYTYWAY..... FNWQHRILLVRAREG
N5C4:                AAAANYTYWAY..... FNW
N3C4:                   YTYWAY..... FNW
```

B

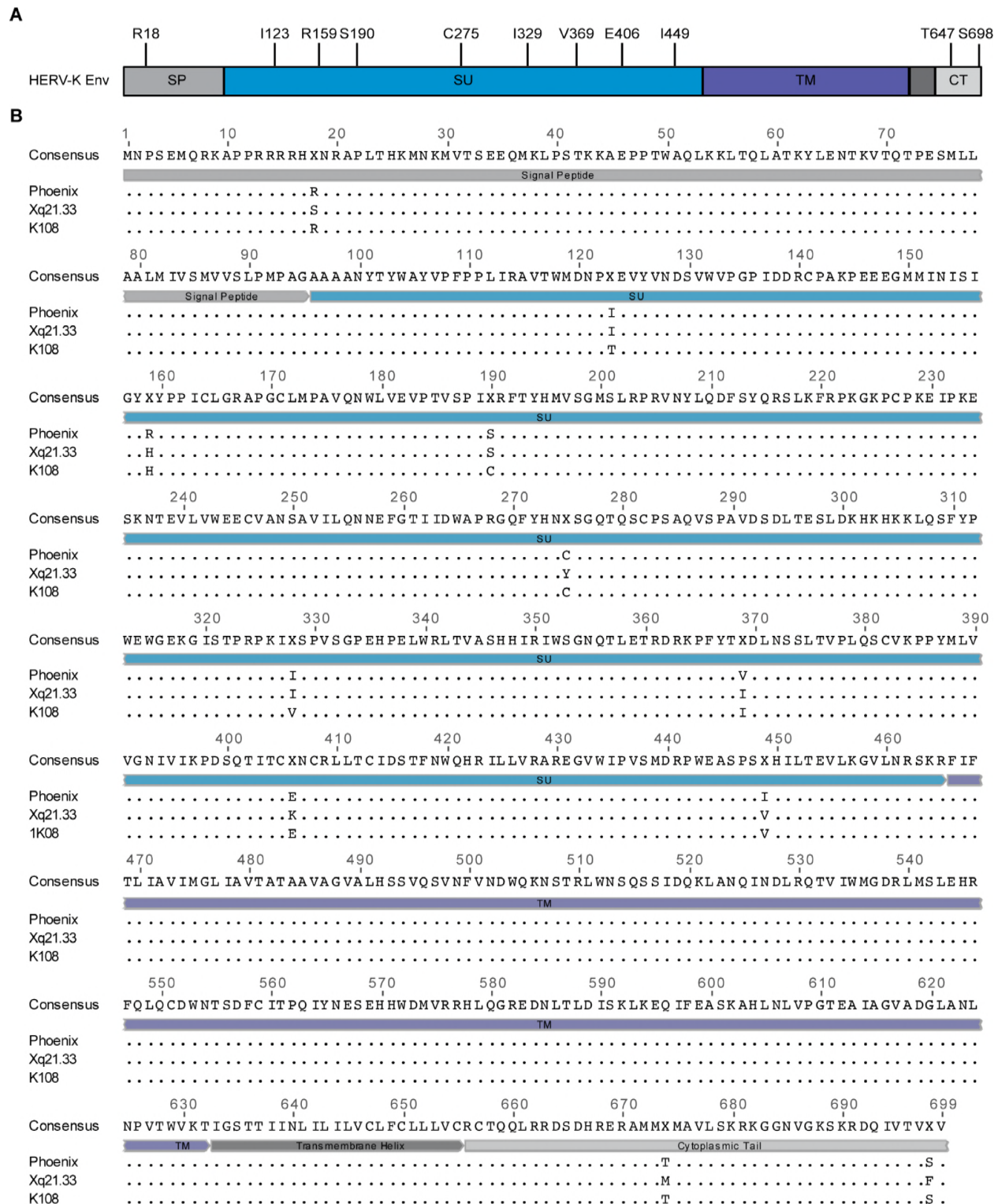| Construct	Expressed	Proccessed	Fusogenic	pH of fusion	Aggregate/ insoluble
WT	+++	+++	+++	5.4	
N1	+++	-	-		
N2	+++	-	-		
N3	+++	+	+++	5.4	
N4	+++	+++	+++	5.4	
N5	+++	+++	+++	5.4	
N6	+++	+++	++++	6	
N7	+++	+++	++++	6	
N4C0	+				+++
N5C0	+				+++
N5C1	-				
N5C2	N.D				
N5C3	+++				-
N5C4	++				+
N3C4	-				

897

898 **S5 Fig. Generation and characterization of soluble HERV-K SU. (A)** Schematic of the sequences

899 of the various N- and C-terminal truncations tested. The tissue plasminogen activator signal

900 peptide was introduced at the N-terminus of all of the truncations. N1-N7 were made in

901 otherwise full-length sequences. C-terminal truncations were further modified with an HA-tag,

902 a 3C protease cleavage site, and a tandem His_{8X}-His_{6X} tag. Amino acid residue numbers are

903 indicated above the sequences, with 1 being the initiating methionine. **(B)** Characteristics of

904   HERV-K Env truncations. N1-N7 were expressed in BSRT7 cells and tested by Western blot for

905   expression and proteolytic processing, and by cell-cell fusion assay for fusogenicity and pH

906   dependence. C-terminal truncations were expressed in 293T cells. Protein in supernatant was

907   isolated over cobalt resin and tested for expression, solubility and oligomerization state. N.D.:

908   Not determined. Empty boxes: Assay not applicable to given construct. +: 1-30% of WT levels.

909   ++: 31-60% of WT levels. +++ 61-100% of WT levels. ++++: 101-130% of WT levels. For C-

910   terminal truncations, values are compared to N5C3. pH of fusion: Highest pH at which cell-cell

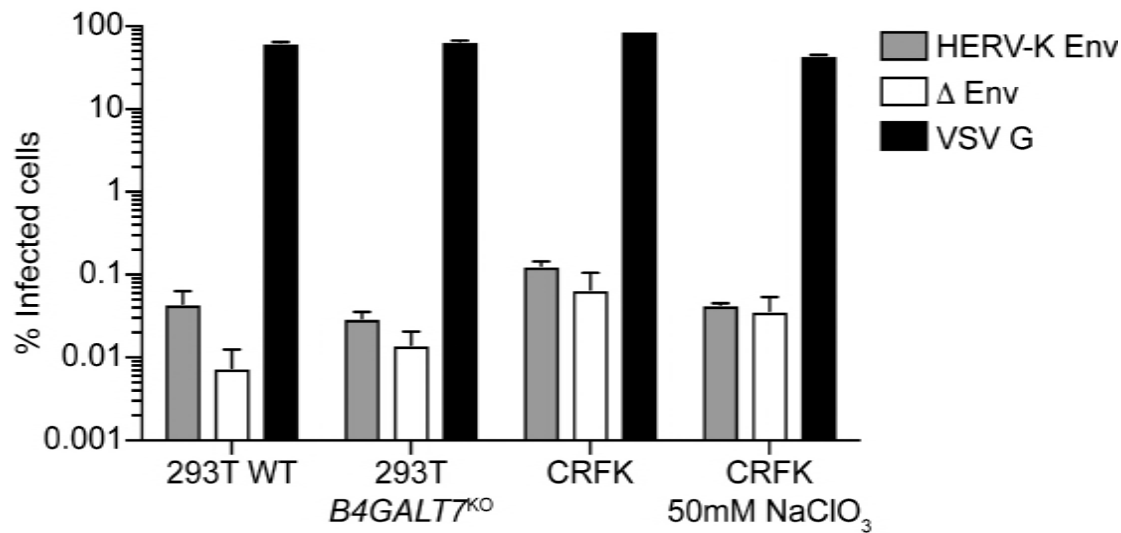911   fusion was observed.

912

S6 Fig. Validation of HERV-K SU. (A) Schematic of HERV-K Env, HERV-K N-terminal truncation

N5, and HERV-K SU used in this study. **(B)** HERV-K Env and HERV-K N5 were transfected into

BSRT7 cells and cell lysates were subjected to Western blot against HERV-K Env TM subunit and

actin, to assess expression and proteolytic processing. For HERV-K Env blot: top band,

uncleaved Env; bottom band, TM subunit. **(C)** BSRT7 cells were transfected with HERV-K Env,

918   HERV-K N5, and empty vector. Cells were exposed to the indicated pH and assessed for the

919   presence of multinucleated syncytia (indicated by arrows) **(D)** FPLC trace of HERV-K SU from gel

920   filtration chromatography. The major peak (at approximately 15 ml, indicated with an asterisk)

921   corresponds to monomeric SU. The peak at 13 ml corresponds to dimeric SU, and the peak at 9

922   ml is an aggregate of SU. Fractions from the FPLC (indicated with dashed lines) were run on a

923   non-reducing SDS-PAGE and coomassie stained. Fraction 1 corresponds to 14.16 ml and fraction

924   14 corresponds to 18.06 ml. The top band at approximately 120 kDa represents the dimeric

925   species and the lower band at approximately 60kDa represents monomer. Only fractions

926   containing only monomer were used for pull-down experiments.

927

S7 Fig. Alignment of Env sequences from Phoenix, Xq21.33, and HERV-K 108. (A) Schematic of

HERV-K Env. Positions of amino acids with differences between Phoenix and either Xq21.33 or

52

930   K108 are shown with the amino acid identity in Phoenix indicated. SP: signal peptide. SU:

931   surface subunit. TM: transmembrane subunit. CT: cytoplasmic tail. **(B)** Alignment of Phoenix,

932   Xq21.33, and K108 Envs.

933

934

935

S8 Fig. Relative infectivity of lentiviral pseudotypes. Lentivirus was produced as described

above and particle concentration determined by p24 ELISA. The indicated cell lines were

inoculated with equal particle numbers, based on p24 levels. % infected cells was determined

by flow cytometry to determine the % GFP positive cells. Pseudotypes bearing HERV-K Env have

an approximately 4-log defect in relative infectivity compared to those bearing VSV G, and have

relative infectivities close to that of "bald" (Δ Env) particles.