Mammalian ZAP and KHNYN independently restrict CpG-enriched avian viruses

Jordan T. Becker[1,5,6], Clayton K. Mickelson[1], Lauren M. Pross[1], Autumn E. Sanders[1], Esther R. Vogt[1], Frances K. Shepherd[1], Chloe Wick[2], Alison J. Barkhymer[1], Stephanie L. Aron[1], Elizabeth J. Fay[1,2], Reuben S. Harris[3,4,5], & Ryan A. Langlois[1,5]

[1]Department of Microbiology and Immunology, University of Minnesota – Twin Cities, Minneapolis, MN, USA, 55455

[2]Department of Biochemistry, Molecular Biology, and Biophysics, University of Minnesota – Twin Cities, Minneapolis, MN, USA, 55455

[3]Department of Biochemistry and Structural Biology, University of Texas Health, San Antonio, TX, USA, 78229

[4]Howard Hughes Medical Institute, University of Texas Health, San Antonio, TX, USA, 78229

[5]Co-corresponding authors

[6]Lead contact

Address correspondence to: beck1169@umn.edu, langlois@umn.edu, or rsh@uthscsa.edu

This manuscript was deposited on the preprint server, bioRxiv, under a CC-BY license.

CLASSIFICATIONS: BIOLOGICAL SCIENCES; MICROBIOLOGY

KEYWORDS: ZAP, KHNYN, influenza A virus, Rous sarcoma virus, host-pathogen interactions

HIGHLIGHTS:

- CpG content is depleted in human and swine IAV relative to avian IAV
- Human ZAP-S and KHNYN but not chicken ZAP independently restrict CpG-rich IAV.
- Mammalian ZAP-L and KHNYN but not chicken ZAP independently restrict ROSV.
- Platypus KHNYN potently restricts retroviruses, including IAV, HIV-1, MLV, and ROSV.

Jump to:

RESULTS

DISCUSSION

METHODS

REFERENCES

FIGURE LEGENDS

ABSTRACT

Zoonotic viruses are an omnipresent threat to global health. Influenza A virus (IAV) transmits between birds, livestock, and humans. Proviral host factors involved in the cross-species interface are well known. Less is known about antiviral mechanisms that suppress IAV zoonoses. We observed CpG dinucleotide depletion in human IAV relative to avian IAV. Notably, human ZAP selectively depletes CpG-enriched viral RNAs with its cofactor KHNYN. ZAP is conserved in tetrapods but we uncovered that avian species lack KHNYN. We found that chicken ZAP does not affect IAV (PR8) or CpG enriched IAV. Human ZAP or KHNYN independently restricted CpG enriched IAV by overexpression in chicken cells or knockout in human cells. Additionally, mammalian ZAP-L and KHNYN also independently restricted an avian retrovirus (ROSV). Curiously, platypus KHNYN, the most divergent from eutherian mammals, was also capable of direct restriction of multiple diverse viruses. We suggest that mammalian KHNYN may be a *bona fide* restriction factor with cell-autonomous activity. Furthermore, we speculate that through repeated contact between avian viruses and mammalian hosts, protein changes may accompany CpG-biased mutations or reassortment to evade mammalian ZAP and KHNYN.

SIGNIFICANCE

Viruses adapt to hosts to replicate successfully. We show that two mammalian proteins, ZAP and KHNYN, restrict CpG-enriched avian viruses. Mammalian KHNYN may be a *bona fide* restriction factor with cell-autonomous activity. We also identified a platypus KHNYN with potent and broad antiviral activity highlighting a significant need to investigate antiviral mechanisms in novel and understudied species. Ongoing efforts to understand viruses with zoonotic potential will benefit from further identification of species-/class-specific restriction factors and their antiviral preferences. Furthermore, we speculate that evolving viral nucleotide composition indicates zoonotic potential and adaptation to mammals requires dinucleotide and amino acid changes.

INTRODUCTION

Influenza A virus (IAV) is a constantly circulating infectious burden in global waterfowl populations, agricultural animals, and among humans seasonally. Avian IAV occasionally transmits into domesticated mammals and humans. While these infections do not usually result in successful human-to-human spread, they are distinctly pathogenic with a reportedly high mortality rate. However, when novel avian IAV establishes in the human population there can be devastating consequences. Major pandemics have been the result of avian-to-human zoonoses or mixed-animal zoonoses involving avian-swine-human infection chains(1). Notably, multiple pandemics or large-scale epidemics occurred in 1918, 1957, 1968, and 2009 that resulted in massive infection incidence and mortality in human populations(2). Much is known about the host dependency factors to which avian IAV must adapt to infect human cells productively(3). These include differences in viral hemagglutinin binding to host sialic acid linkages(4) with species-specific airway localization(5), pH-sensitivity of viral RNA (vRNA) endosomal release into the cytoplasm(6), and species-specific ability to use the polymerase cofactor, ANP32(7). Relatively less is known about the network of restriction factors that might specifically target zoonotic IAV infections in mammals(3). Much of our knowledge of host-pathogen interactions and genetic conflicts has focused on amino acid level and related protein structure changes that predicate animal-to-human adaptation(8). Comparatively less is known about nucleic acid adaptations driven by host-pathogen conflicts during interspecies infections.

Viral nucleic acid sequence and structure are not only a determinant of protein coding changes, but rather they encompass a complex and sensitive network of features that determine viral fitness, stability, and pathogenesis(9, 10). In addition, host and pathogen genomes are not in equilibrium regarding the frequency and usage of nucleotide doublets(11, 12). For example, CpG dinucleotide frequency is depleted throughout the human genome(13). Commonly, this is a result of DNA methylation at CpG motifs near promoter sequences that regulate transcription(14). Mammalian genomes have also selected against encoding CpG dinucleotides in cytoplasmic-resident RNAs(13, 15). Others observed a depletion of CpG content among interferon stimulated genes (ISGs), notably in mammalian species(13, 16). Accordingly, many viral pathogens that infect mammalian species match the dinucleotide content of their host species and mammalian viruses are often CpG-depleted(11, 13, 17, 18). Others showed that avian IAV sequences exhibit elevated CpG content and that CpG content has depleted over time for IAV strains circulating in humans(13, 18). This suggests that host factors might exist in human hosts to select against CpG-enriched viral nucleic acids that are absent or defective in avian species.

95

96 Recently, several groups showed that the human restriction factor, zinc-finger antiviral protein
97 (ZAP also known as PARP13), specifically binds to and suppresses CpG dinucleotides in
98 cytoplasmic RNA(19, 20). Human ZAP is encoded by the *ZC3HAV1* gene, induced by interferon,
99 and generates multiple isoforms(21, 22), including the ZAP-L (a 902 amino acid isoform
100 preferentially localized to intracellular membranes) and ZAP-S (a 699 amino acid isoform with
101 diffuse cytoplasmic localization). Like many antiviral proteins, ZAP is evolving under positive
102 selection(23–25). ZAP antiviral activity is regulated by the interferon-inducible E3 ubiquitin ligase,
103 TRIM25(19, 20, 26, 27). KHNYN was identified as a ZAP interacting protein and putative cofactor
104 for ZAP antiviral activity(20). Moreover, KHNYN exhibits basal expression and is not interferon-
105 inducible(28). While KHNYN can exert antiviral effects through ZAP it is unknown if it can exert
106 additional direct antiviral functions. Together, human ZAP, KHNYN, and TRIM25 may function
107 cooperatively, selectively, and/or independently to restrict certain classes of viruses with a bias
108 towards viruses with CpG-enriched nucleotide content. While the function of these human
109 proteins has been well studied, less is known about their non-primate and non-mammalian
110 orthologues and their antiviral restriction ability.

111

112 Understanding the intrinsic and innate immune factors capable of restricting zoonotic viruses is
113 paramount in identifying zoonotic viruses with pandemic potential(29). Importantly, IAV is a
114 seasonal infectious burden in humans and common zoonosis among wild waterfowl, agricultural
115 birds, domesticated mammals, and humans. We hypothesized that birds are defective in RNA
116 recognition and restriction pathways and that presents a barrier to emergence in mammals. Here,
117 we tested whether restriction factors targeting CpG-enriched RNAs could affect replication of IAV
118 with an elevated CpG-content, mimicking avian IAV. We found that chicken and duck ZAP do not
119 restrict IAV replication regardless of CpG content. We show through genomic, phylogenetic, and
120 functional analyses that avian species are defective in ZAP function. Additionally, we demonstrate
121 that avian species lack KHNYN, which likely arose through gene duplication after birds and
122 mammals diverged. Thus, we used chicken cells as a gain-of-restriction system to test the
123 exogenous provision of host factors for IAV restriction. Indeed, human ZAP-S and KHNYN stably
124 expressed in chicken cells independently restricted CpG-enriched IAV replication. We showed
125 that combined knockout of ZAP and KHNYN in human cells increased replication of CpG-enriched
126 IAV to a greater extent than single knockout of ZAP or KHNYN alone. We also found that
127 mammalian but not avian (*Galloanserae*) ZAP as well as some mammalian KHNYN proteins
128 restrict a naturally CpG-enriched avian retrovirus, Rous sarcoma virus (ROSV). Unexpectedly,

129 we discovered a potently antiviral KHNYN homologue from platypus that suggests an ancient
130 restriction mechanism. We show that KHNYN is a *bona fide* restriction factor with cell-autonomous
131 activity. Furthermore, we speculate that through repeated contact between avian and human
132 hosts followed by seasonal circulation among humans, protein level changes for receptor
133 preference and polymerase function may be accompanied by dinucleotide changes and/or CpG-
134 disfavored reassortment that evade restriction by mammalian ZAP and KHNYN.
135
136 **RESULTS**
137 **Mammalian IAV sequences are CpG-depleted relative to avian IAV.** Viruses evolve to
138 replicate successfully in their host species. Previous work showed that human IAV exhibited a
139 depletion of CpG dinucleotide content relative to avian IAV sequences since the 1918 IAV
140 pandemic(13, 18, 30). Here, surveying influenza sequences in FluDB(31) collected between 1918
141 and 2020, we also observed a depletion of CpG dinucleotides per kilobase (CpG content) in
142 human and swine but not avian IAV sequences over time since 1918 (**Figure S1A**). By comparing
143 all human, swine, and avian IAV segment sequences *en masse*, we observed a lower CpG
144 content in human and swine IAV relative to avian IAV sequences (**Figure 1A**). We observed
145 similar results for ρ_{CG} (rho accounts for total C and G content(12, 13, 16)) but less so for other
146 dinucleotides or %GC content (**Figure S2A-D**). Notably, the largely human-specific influenza B
147 virus(32) (IBV) and influenza C virus(33) (ICV) are CpG-depleted relative to all IAV sequences,
148 speculatively consistent with long-term near-exclusive circulation in mammalian/human
149 individuals (**Figure 1A, S1, and S2**). We found that when comparing major HA and NA subtypes,
150 CpG content is depleted in subtypes seasonally circulating in humans (H1N1 and H3N2) relative
151 to avian sequences of the same subtype (**Figure S1C**). Conversely, CpG content in human
152 isolates of common avian-to-human zoonotic HA and NA subtypes (e.g. H5N1 and H7N9) is
153 similar to their avian counterparts. We observed that human IAV segments 1-3 encoding the
154 heterotrimeric viral polymerase (i.e. PB2, PB1, and PA) are more CpG-depleted relative to avian
155 IAV than segments 4-6 encoding HA, NP, and NA, virion proteins typically subject to adaptive
156 immune pressure(34) (**Figure S1B**). The natural history of human IAV includes multiple
157 introductions(35) resulting from avian zoonoses and/or reassortment of avian IAV segments with
158 human viruses in 1918, 1957, 1968, and 2009. We compared CpG content of avian IAV to human
159 IAV during the time periods surrounding the replacement of H1N1 with H2N2 in 1957, H2N2 with
160 H3N2 in 1968, reintroduction of H1N1 in 1977, and pandemic "swine flu" H1N1 in 2009 (**Figure**
161 **S1D, bottom**). We observed that CpG content is lower in human relative to avian IAV sequences
162 during the longer time periods following zoonoses in 1977 and 2009 (**Figure S1D, top**). This may

163 reflect periodic reintroduction of CpG-enriched avian sequences that deplete over longer time
164 frames.
165
166 **Avian ZAP does not restrict replication of CpG-enriched IAV.** We hypothesized that avian
167 species are defective in cytoplasmic CpG-sensing by ZAP and reasoned that this deficiency
168 allows CpG-maintenance of IAV sequences in avian host species. Human ZAP-L and chicken
169 ZAP are homologous proteins encoding four zinc fingers, an inactive poly-ADP-ribose
170 polymerase-like domain(36), and carboxy-terminal prenylation motif(22, 37) (**Figure 1B, top**).
171 Previous structural and functional analysis of human ZAP identified zinc finger 2 (ZF2, residues
172 88-110) as responsible for directly recognizing and binding to CpG dinucleotides in the context of
173 short (3-20 nucleotide) and long (full-length 9 kilobase HIV-1 genome) RNA molecules(19, 38,
174 39). Using two structural prediction algorithms, we generated models of the four zinc fingers of
175 chicken ZAP (ggaZAP, residues 1-220) by both comparative modeling based on human ZAP
176 (Robetta-CM(40), PDB: 6UEJ(38)) and *ab initio* folding by RoseTTAFold(41) (**Figure S3A-B**). We
177 also compared these structures to the model of chicken ZAP predicted by AlphaFold2(42) (**Figure**
178 **S3A-B**). As expected, we observed that both *ab initio* predicted structures (RoseTTAFold and
179 AlphaFold2) exhibited lower structural similarity (as measured by root mean squared deviation,
180 RMSD) to human ZAP relative to those generated with comparative modeling (**Figure S3B**). In
181 addition, we observed structural differences between human ZAP and models of duck ZAP as
182 well as quail ZAP (**Figure S3A-B**). Alignment of the amino acid sequences of ZAP homologues
183 revealed multiple amino acid differences between ZF2 of avian and mammalian species (**Figure**
184 **1B**). Therefore, we hypothesized that chicken ZAP is unable to restrict CpG-enriched viruses and
185 tested this hypothesis using the common IAV laboratory strain A/Puerto Rico/8/1934//H1N1 (PR8)
186 and a CpG-enriched version of PR8(43) ($PR8_{CG}$, orange dashed box; **Figure 1C**).
187
188 In chicken DF-1 cells stably expressing chicken ZAP (ggaZAP) as well as mNeonGreen-tagged
189 chicken ZAP (mNG-ggaZAP), we found no difference in virus replication over 48 hours for PR8
190 or $PR8_{CG}$ relative to vector-transduced (Vector) and mNeonGreen-transduced (mNG) cells
191 (**Figure 1D**). The chicken *ZC3HAV1* gene encodes a single ZAP isoform of 713 amino acids with
192 a carboxy-terminal prenylation motif (amino acids: $_{710}$CIVC$_{713}$) that is homologous to the long
193 isoform of human ZAP (ZAP-L). The short isoform of human ZAP (ZAP-S) is generated by
194 alternative splicing and lacks the carboxy-terminal prenylation motif and the inactive poly-ADP-
195 ribose polymerase-like domain. To test if either prenylation of chicken ZAP or if the divergent ZF2
196 regulate restriction of CpG-enriched IAV, we generated mutants of chicken ZAP lacking the critical

197 cysteine residue (C710S) for membrane-targeting prenylation, a chimera encoding the human
198 ZF2 (hZF2), or a combination (hZF2.C710S) of these two mutants all of which expressed and
199 localized as expected (**Figure 1E**). However, no form of chicken ZAP restricted PR8 or $PR8_{CG}$
200 replication (**Figure 1F**). Finally, we found that pooled CRISPR/Cas9-mediated knockout of
201 endogenous ZAP in chicken DF-1 and duck CCL-141 cells had no effect on PR8 or $PR8_{CG}$
202 replication (**Figure 1G-H** and **S4**). Similarly, we observed that knockdown of endogenous ZAP by
203 stable shRNA in chicken DF-1 cells or transient siRNA in duck CCL-141 cells had no effect on
204 PR8 or $PR8_{CG}$ replication (**Figure S3C-E**). qPCR analysis showed endogenous chicken ZAP is
205 expressed in DF-1 cells, induced by interferon, and can be suppressed by shRNA (**Figure S3C**).
206 Given that human ZAP can use chicken TRIM25 as a cofactor(44), we reasoned that chicken DF-
207 1 cells are deficient in ZAP-dependent antiviral activity targeting this CpG-enriched IAV and could
208 serve as a gain-of-restriction system to test other host factors for the ability to restrict CpG-
209 enriched viruses.
210
211 **Human ZAP-S and KHNYN independently restrict CpG-enriched IAV.** Restriction of CpG RNA
212 has been attributed to human ZAP along with its cofactors KHNYN and/or TRIM25. However, we
213 found that avian species lacked a *bona fide* KHNYN orthologue. There are no annotated avian
214 KHNYN orthologues in ensembl, NCBI, or UCSC Genome Browser and a ~400,000bp syntenic
215 region surrounding human KHNYN is missing in avian species, including the KHNYN paralogue
216 – NYNRIN (**Figure S5A**). Human KHNYN and NYNRIN are encoded in tandem on chromosome
217 14 and NYNRIN is the result of a KHNYN gene duplication and fusion to a retroviral integrase
218 domain(45) (**Figure S5B**). KHNYN and NYNRIN are paralogous to N4BP1(46), a gene conserved
219 in birds, reptiles, and fish (**Figure S5B-C**). We estimated a maximum likelihood phylogeny of
220 human KHNYN, NYNRIN, and N4BP1 orthologues nucleotide sequences from ensembl (**Figure**
221 **S5C**). We confirmed that N4BP1 is conserved in *Gnathostomats* (i.e. jawed vertebrates including
222 fish, birds, reptiles, and mammals) and is closely related to a fish-specific N4BP1 paralogue
223 (sometimes annotated as *khnyn* or *khnyn-like*, e.g. *Danio rerio khnyn*, ENSDARG00000092488).
224 However, KHNYN found in mammalian species (i.e. monotremes, marsupials, and placentals)
225 and NYNRIN found in therian mammals (i.e. marsupials and placentals) branched together and
226 are more like each other than to the N4BP1 genes (**Figure S5C**). We noted that NCBI and
227 ensembl contain reptile sequences annotated as *KHNYN* (e.g. *Alligator mississippiensis KHNYN*
228 XM_014607035.3 or *Anolis carolinensis khnyn* ENSACAG00000009105) that may represent
229 ancestral KHNYN orthologues based on nucleotide sequence (**Figure S5C**; ancient KHNYN
230 orthologues), exon organization, and syntenic neighborhood (**Figure S5D**). Thus, we reasoned

231 that avian species are defective and/or deficient in two restriction factor genes: ZAP as well as
232 KHNYN. As such, we generated chicken DF-1 cells that stably expressed human ZAP-L, ZAP-S,
233 KHNYN, or TRIM25 (schematic in **Figure 2A**). We included a mutant of KHNYN that is expected
234 to be enzymatically inactive within its endonuclease domain(20) due to substitution of three critical
235 aspartic acid residues (DDD443/524/525AAA, dKHNYN). We found that ZAP-S as well as
236 KHNYN modestly restricted PR8 relative to vector transduced cells (**Figure 2B, left**). KHNYN has
237 not been shown previously to act independently of ZAP. Surprisingly, we found that human ZAP-
238 S and KHNYN (but not ZAP-L, dKHNYN, or TRIM25) independently restricted $PR8_{CG}$ relative to
239 vector (**Figure 2B, right**).
240
241 **CpG-targeted restriction requires intact RNA-binding domains of ZAP and KHNYN.** To
242 confirm the role of ZAP-S and KHNYN in restricting $PR8_{CG}$ we generated mutants of each, stably
243 expressed them in DF-1 cells, and performed multicycle IAV replication assays. To test the CpG
244 specificity of ZAP-S in restricting $PR8_{CG}$, we generated a Y108A mutant previously shown to
245 reduce CpG specificity(38) as well as a chimeric human ZAP-S encoding the ZF2 of chicken ZAP
246 (gZF2). Only WT human ZAP-S remained capable of restricting $PR8_{CG}$, an ability that was lost by
247 the Y108A and gZF2 mutants (**Figure 2C**). To test if the prenylation motif (amino acids: CVIS)
248 and corresponding membrane localization of ZAP regulated restriction activity against $PR8_{CG}$, we
249 generated a mutant of ZAP-S that encoded a carboxy-terminal CVIS motif (+CVIS) as well as a
250 C899S mutant of ZAP-L(22). While the C899S mutant of human ZAP-L did not gain the ability to
251 restrict $PR8_{CG}$, human ZAP-S +CVIS lost the ability to restrict $PR8_{CG}$ as expected (**Figure 2D**).
252 Next, to delineate the manner of KHNYN restriction of $PR8_{CG}$ we mutated its KH domain. KH-
253 domain containing proteins encode a GXXG motif at the top of a central alpha-helix and mutations
254 in the GXXG motif reduced RNA-interactions in another KH-domain containing protein,
255 HNRNPK(47). We generated models of the KH domain of KHNYN using comparative modeling
256 based on the third KH-domain of human HNRNPK(48) (Robetta-CM; PDB: 1ZZK) as well as by
257 *ab initio* folding (RoseTTAFold) (**Figure 2E; Figure S5E**). Both models showed a central alpha-
258 helix with the $_{94}$GAQG$_{97}$ motif at its apex adjacent to a three-helix bundle (**Figure S5E**). We found
259 that only WT human KHNYN was capable of restricting $PR8_{CG}$ over 48 hours of replication in
260 chicken DF-1 cells but not the catalytically inactive dKHNYN, a AQ95/96DE mutant (AQDE), or a
261 catalytically inactive AQ95/96DE (dAQDE) (**Figure 2F**). Together, these results indicated that
262 intact human ZAP-S as well as KHNYN can independently restrict $PR8_{CG}$.
263

264 **ZAP-S and KHNYN do not alter chicken or IAV mRNA.** To understand how these mammalian
265 restriction factors functionally inhibit IAV, we examined transcriptomic changes during IAV
266 infection. We performed RNA-seq on uninfected as well as PR8 or $PR8_{CG}$ infected chicken DF-1
267 cells stably expressing human ZAP-L, ZAP-S, KHNYN, or dKHNYN to evaluate if these human
268 proteins modulated the chicken transcriptome and viral mRNA expression relative to vector
269 transduced cells. We found that the transcriptomes of uninfected cells were largely unperturbed
270 by whatever human factor was overexpressed (**Figure S6A**) suggesting antiviral activity may not
271 be caused by changing the underlying gene expression profile. Next, we hypothesized that by
272 interacting with and depleting CpG-enriched viral RNAs during multicycle IAV replication
273 experiments, ZAP-S and KHNYN might also negatively bias CpG-enriched segment accumulation
274 during replication. We evaluated relative levels of IAV RNA segments in PR8 versus $PR8_{CG}$
275 infected cells in the presence of these different human factors and found similar levels of viral
276 mRNA within segments (**Figure S6B**). Infection with either PR8 or $PR8_{CG}$ resulted in similar gene
277 expression patterns for host cell factors between the two viruses, suggesting that the mechanism
278 of restriction is not through differential induction of host antiviral responses (**Figure S6C**).
279 Together these results suggested initial mechanistic insights into how ZAP and KHNYN might
280 restrict IAV during zoonotic infections although the restriction may not directly target/deplete viral
281 mRNAs.
282
283 **Endogenous human ZAP and KHNYN restrict CpG-enriched viruses.** We next tested the
284 ability of endogenous human ZAP and/or KHNYN to restrict CpG-enriched IAV in human lung
285 cells. Therefore, we generated multiple human lung adenocarcinoma A549 clonal cell lines
286 depleted for ZAP, KHNYN, or both by CRISPR/Cas9 genome editing as well as a vector control
287 (Vector) and a negative control guide targeting *E. coli* LacZ beta-galactosidase (LacZ) clonal cell
288 lines (**Figure 3A-B**). We tested PR8 and $PR8_{CG}$ by multicycle replication assay in these cell lines.
289 We found that single knockout of only ZAP or only KHNYN failed to increase PR8 or $PR8_{CG}$
290 replication relative to LacZ or parental A549 cells (**Figure 3C**). We reasoned that this was
291 consistent with our results by overexpression in DF-1 cells such that either ZAP or KHNYN are
292 capable of independently restricting $PR8_{CG}$. Next, we tested multicycle replication of PR8 and
293 $PR8_{CG}$ in double knockout (ZAP and KHNYN depleted) cells revealing that genetic loss of both
294 ZAP and KHNYN increased replication of $PR8_{CG}$ relative to parental A549 or LacZ cell lines
295 (**Figure 3D**) supporting the hypothesis that ZAP and KHNYN can independently restrict of CpG
296 enriched IAV. We also found replication of PR8 or $PR8_{CG}$ was unaffected by knockout of

297 endogenous TRIM25 (**Figure 3E**). Altogether, these results support the hypothesis that ZAP and
298 KHNYN can act independently against IAV.
299
300 **ROSV is CpG-enriched and susceptible to mammalian CpG-dependent restriction.** Avian
301 IAV exhibits an enrichment for CpG content relative to human IAV (**Figure 1A, Figures S1-2**).
302 We next asked if other avian viruses with different levels of human or mammalian zoonotic
303 interactions might exhibit higher levels of CpG enrichment. We noticed that ROSV (a tractable
304 retrovirus model in chicken) exhibits a higher CpG content relative to mammalian retroviruses,
305 including HIV-1 or MLV. We observed that ROSV is CpG-enriched along its entire length relative
306 to mammalian retroviruses including HIV-1 (**Figure 4A**). To determine whether ROSV was subject
307 to CpG-dependent restriction, we performed retroviral assembly and single-cycle infectivity
308 assays (schematic in **Figure 4B**) by co-transfecting chicken DF-1 cells with a full-length ROSV
309 reporter virus expression plasmid and VSV-G along with increasing amounts of control, human
310 ZAP-L, human ZAP-S, KHNYN, or dKHNYN plasmid. We found that human ZAP-L inhibited
311 ROSV, but ZAP-S did not (**Figure 4C**). Using mutants of human ZAP-L and ZAP-S, we confirmed
312 that for ROSV, ZAP restriction required an intact ZF2 and prenylation motif (**Figure 4C**).
313 Specifically, the C899S prenylation mutant of ZAP-L failed to restrict ROSV while ZAP-S fused to
314 the four amino acid ZAP-L prenylation motif (+CVIS) gained the ability to restrict ROSV. Next, we
315 tested human ZAP-L, chicken ZAP, and duck ZAP as well as prenylation mutants of each for
316 restriction of ROSV using fusions to mNeongreen. We found that human ZAP-L restricted ROSV
317 but not either of the avian ZAP proteins (**Figure 4D**). Finally, we tested mammalian orthologues
318 of ZAP from dog, pig, and platypus and found that the prenylated forms of placental mammal ZAP
319 restricted ROSV, while platypus ZAP or a prenylation mutant (C653S) were less restrictive
320 (**Figure S7A**).
321
322 **Mammalian KHNYN restricts ROSV independent of ZAP.** Next, we tested human KHNYN,
323 mutant KHNYN, and mammalian KHNYN orthologues from pig, dog, and platypus for restriction
324 of ROSV. Pigs and dogs are domesticated mammals frequently in contact with agricultural poultry,
325 a common source of avian-to-mammal zoonoses(3). We selected platypus KHNYN to test
326 because it is the only mammalian species to encode KHNYN but lack the NYNRIN gene(45)
327 (**Figure S3C**). We found that pig KHNYN, but not dog KHNYN, similarly restricted ROSV as
328 human KHNYN while platypus KHNYN potently restricted ROSV (**Figure 4E, Figure S7B**). We
329 also tested these proteins for ROSV restriction tagged with mNeongreen at their amino terminus
330 relative to catalytically inactivated mutants. Curiously, mutation of the aspartic acid residues in

331 the NYN endonuclease domains inhibited the restrictive ability for pig KHNYN (DD508/508AA)
332 but not platypus KHNYN (DDD416/497/498AAA) (**Figure S7B**). We tested multiple mutants and
333 truncated forms of platypus KHNYN and found that mutation of aspartic acid residues critical for
334 endonuclease activity in other KHNYN orthologues failed to ameliorate platypus KHNYN activity
335 while truncation mutants which deleted the endonuclease domain lost restriction activity (**Figure**
336 **S7C**). We generated structural models of platypus KHNYN and noted a substantial structural
337 difference in the loop above the apex of the putative XG motif in platypus KHNYN encoding a
338 notably shorter loop relative to human, pig, and dog KHNYN (**Figure S5E**). Curiously, we found
339 that platypus KHNYN also restricted HIV-1, a CpG-enriched HIV-1, and MLV, independent of
340 glycoprotein provided for pseudo-typing (**Figure S7D**). Finally, we tested if platypus KHNYN could
341 restrict IAV replication in human A549 cell lines expressing YFP-tagged platypus KHNYN.
342 Relative to YFP-alone expressing cells, platypus KHNYN modestly inhibited single cycle PR8,
343 pandemic Cal04, and avian-origin OH175 (**Figure 4F**). Together, these results support the
344 hypothesis that mammalian ZAP and KHNYN can restrict CpG-enriched avian viruses, including
345 IAV and ROSV. Furthermore, we discovered an orthologue of KHNYN in platypus that potently
346 restricts multiple viruses from distinct virus families independent of CpG content.
347

348 **DISCUSSION**

349 KHNYN was identified as a ZAP-interacting cofactor that required ZAP to exert a combined
350 antiviral effect against a CpG-enriched HIV-1(20). Human ZAP functions independent of KHNYN
351 against multiple viruses including HIV-1(19, 37, 49), Ebolavirus(50, 51), Zikavirus(52), and
352 Sindbis virus(27). Here, we provide the first evidence that KHNYN may be capable of exerting an
353 antiviral effect independent of ZAP. We note that, by definition, restriction factors act in a cell-
354 autonomous fashion(53, 54). As such, KHNYN may be a novel *bona fide* restriction factor with
355 unknown RNA preferences Furthermore, multiple groups showed different, sometimes conflicting,
356 results regarding ZAP restriction of IAV(55–58). Here, we observed that human IAV exhibits a
357 marked decrease in CpG content relative to avian IAV (**Figure 1** and **Figure S1-2**). We showed
358 that chicken ZAP as well as duck ZAP do not restrict IAV (**Figure 1** and **Figure S3-4**) or ROSV
359 (**Figure 4** and **Figure S7**). We reasoned that chicken cells (e.g. DF-1) could be used as a gain-
360 of-restriction system to test restriction factors for their ability to inhibit IAV and other CpG-enriched
361 viruses. In DF-1 cells stably expressing human ZAP-L, ZAP-S, KHNYN, a catalytically inactive
362 KHNYN, or TRIM25, we found that ZAP-S and KHNYN modestly restricted PR8 and significantly
363 restricted $PR8_{CG}$ (**Figure 2**). We also confirmed that this restriction required an intact CpG-
364 recognizing ZF2 of ZAP or a catalytically active KHNYN. We found that human ZAP and KHNYN

365 do not appreciably manipulate the host transcriptome in chicken cells but negatively affect IAV
366 RNA levels (**Figure S6**). We supported the hypothesis that each of these restriction mechanisms
367 can act independently in human cells wherein only combined knockout of both ZAP and KHNYN
368 in human A549 cells increased replication of $PR8_{CG}$ (**Figure 3**). We also found that another avian
369 virus is subject to ZAP and KHNYN restriction, ROSV (**Figure 4** and **Figure S7**). Finally, we
370 discovered a potent antiviral KHNYN homologue in platypus capable of restricting multiple
371 divergent retroviruses as well as IAV (**Figure 4** and **Figure S7**). Altogether, we hypothesize that
372 these mammalian restriction factors act against CpG-enriched viruses including avian viruses and
373 select for CpG-depleted nucleotide mutations.
374
375 Multiple groups have analyzed the dinucleotide content of influenza sequences, other viruses, as
376 well as host transcripts(11, 13, 16, 18). Here, we analyzed influenza virus segment sequences
377 from different host species collected over the past century and found the CpG content is depleted
378 over time in mammalian species and that this occurs in polymerase segments more so than HA
379 or NP segments (**Figure 1**). We speculate that over a decades-to-centuries long timeframe, CpG-
380 depletion is a relatively slow adaptative process compared to single amino acid substitutions in
381 PB2, PA, NP, NA, or HA traditionally associated with host species adaptation(59–64). We
382 observed that H1N1 and H3N2 human IAV sequences exhibited a lower CpG content than
383 subtype-matched avian sequences. However, H5N1 and H7N9 sequences were essentially
384 identical in CpG content between human and avian samples likely reflecting the limited amount
385 of time and replication within human patients. Yet, it is unclear how CpG content regulates viral
386 RNA intrinsic stability or viral replication. Recent work showed that human ZAP targets viral RNA
387 molecules encoding a minimum threshold number of CpG dinucleotides within a length of RNA
388 that is neither too dense nor too sparse(65). This likely reflects a requirement for some minimum
389 number of ZAP molecules to bind RNA and if CpG dinucleotides are too close these sites occlude
390 each other while if too sparse or too distributed this dilutes a necessary yet undefined minimum
391 of ZAP clustering. Here, we showed that human ZAP-S can target an artificially CpG-enriched
392 PR8 IAV and to a lesser extent PR8 IAV. Future work should examine how CpG number, density
393 (number per length), segment-specificity, and influenza RNA class (viral messenger,
394 complementary, or genomic RNA) are targeted by ZAP.
395
396 Human ZAP-S and ZAP-L can restrict different viruses and human transcripts(66), including those
397 associated with the interferon response(16), based on their differential cellular localization as
398 determined by the carboxy-terminal prenylation motif(22, 37, 58). KHNYN was identified by

399 protein-protein interaction with ZAP in a yeast two-hybrid screen and functionally shown to be a
400 required cofactor for ZAP restriction of an artificially CpG-enriched HIV-1 construct(20).
401 Subsequent work suggested that ZAP-based restriction of different CpG-enriched HIV-1
402 constructs were not uniformly KHNYN-dependent(49). Curiously, one group showed that ZAP-L
403 requires the C-terminal prenylation motif to restrict CpG-enriched HIV-1(37) while another found
404 that a C-terminally HA-tagged ZAP-L (that effectively eliminates prenylation) restricted different
405 CpG-enriched HIV-1 mutants(19, 44). It is unclear whether the differences in CpG number and
406 density or other technical differences might explain these differences. Previous work
407 demonstrated that human ZAP-L inhibited IAV (including PR8) by binding viral PB2 and PA
408 proteins but is counteracted by PB1(56). However, others found that human ZAP-S inhibits IAV
409 (including PR8) by reducing viral mRNA levels but is counteracted by NS1(57). These two results
410 predated identification of human ZAP preference for CpG dinucleotides. Finally, recent work
411 showed that ZAP-S inhibits PR8 with a CpG-enriched segment 1 but not PR8 with a CpG-enriched
412 segment 5(58). Here, we show that human ZAP-S and KHNYN independently restrict IAV with a
413 CpG-enriched segment 5 ($PR8_{CG}$; **Figure 2-3**) while ZAP-L and KHNYN independently restrict
414 ROSV (**Figures 4**). Curiously, the original discovery of ZAP-dependent restriction of MLV
415 identified and used a form of ZAP lacking the carboxy-terminal prenylation motif(67). KHNYN has
416 not been previously shown to independently restrict any viruses. Altogether, we suggest that
417 antiviral activities of ZAP and KHNYN might be context dependent. For example, a subset of
418 viruses might be restricted by the membrane-associated ZAP-L and others restricted by the
419 cytosolic ZAP-S because the antiviral activity of many restriction factors is localization-dependent.
420 In addition, ZAP might restrict some sequences, KHNYN might restrict other sequences, and they
421 may act together against yet a third group of sequences. In fact, the results from Sharp, et al.
422 suggest mutants of a specific virus strain may exhibit differential susceptibility to ZAP and/or
423 KHNYN(58). In the case of segmented viruses, including IAV, this suggests an impact on genetic
424 drift (mutation) as well as genetic shift (reassortment) over time.
425
426 In contrast to ZAP, relatively less is known about the function, preferences, and regulation of
427 KHNYN. KHNYN is one member of family of paralogous genes encoding a putative RNA-binding
428 domain fused to a NYN-endonuclease domain(46). Our phylogenetic analysis (**Figure S5**)
429 support the hypothesis that KHNYN may be derived from N4BP1 occurring prior to the radiation
430 of mammalian species and NYNRIN is the result of a gene duplication and fusion/integration of
431 KHNYN to Metaviral *Pol* gene occurring between the emergence of monotreme and therian
432 mammals(45). We showed here that mutation of key aspartic acid residues in the NYN domain of

433 KHNYN reduce its restriction ability, similar to others(20). We also show that mutations in the KH-
434 domain of KHNYN reduces its restriction ability(47). Recent structural and biochemical studies
435 showed the extended di-KH domain of KHNYN does not bind RNA but the full-length KHNYN
436 protein binds RNA without CpG preferences and the NYN domain exhibits single-stranded
437 endonuclease specificity(68). Furthermore, how these preferences and activity change over
438 evolutionary time in unclear. Finally, while we identified serendipitously the potent platypus
439 KHNYN restriction factor, its mechanism of action is unclear. We speculate that this potent
440 ancestral KHNYN required attenuation for the genesis of retroelement-derived NYNRIN and
441 retroviral-associated placentation acquired during the rapid evolutionary leap from monotremes
442 to therian mammals between 180-160 million years ago(45, 69).
443
444 **LIMITATIONS**
445 We appreciate limitations exist in this study attempting to synthesize and understand the multiple
446 (and sometimes conflicting) reports regarding CpG-depletion in cytoplasmic-resident RNAs and
447 viral adaptation driven by ZAP, KHNYN, and related antiviral genes. Throughout this study we
448 primarily use a common laboratory strain of IAV and a previously published CpG-enriched
449 segment 5 mutant strain (PR8 and $PR8_{CG}$). Future work will benefit from examining ZAP, KHNYN,
450 and TRIM25 susceptibility of many influenza strains derived from wild birds, agricultural poultry,
451 and domestic mammals as well as emerging mammalian H5N1 IAV isolates. This work largely
452 uses chicken cell lines and relies on overexpression of many host restriction factors by
453 transfection or stable transduction. While aspects of our work were substantiated by knockdown
454 and/or knockout of endogenous host cell factors, future work will benefit from using primary cells.
455 Furthermore, future work will also benefit from cell systems derived from avian reservoir species
456 including shorebirds that currently do not exist as tractable laboratory models. Finally, we
457 acknowledge two aspects of KHNYN antiviral activity shown here are limited. Specifically, the
458 mechanisms of KHNYN antiviral activity should be fully elucidated in the future including potential
459 co-factors. In addition, the potent antiviral platypus KHNYN orthologue is derived from a predicted
460 mRNA and protein product and investigation of its mechanism of action will be vastly informative.
461 However, the sum of the results shown here should spur multiple important research directions
462 and warrant intense interest in multiple fields including host-pathogen interactions, virus evolution,
463 mammalian evolution, and embryonic development.

464 **METHODS**

465

466 **Cell lines.** Mammalian cells, chicken DF-1 cells, and duck CCL-141 cells were grown in DMEM
467 supplemented with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin. Chicken DT40
468 cells were grown in DMEM supplemented with 10% FBS, 10% tryptose phosphate buffer, 1%
469 penicillin/streptomycin, and 50μM beta-mercaptomethanol. Cells were maintained at 37°C, 5%
470 CO_2, and 50% humidity. HEK293T (male human epithelial kidney, ATCC, CRL-3216), A549 (male
471 human lung carcinoma, ATCC, CCL-185), PK-15 (male porcine kidney, ATCC, CCL-33), MDCK
472 (female canine kidney, ATCC, CCL-34), DF-1 (unspecified sex chicken embryonic fibroblast,
473 ATCC, CRL-12203), DT40 (unspecified sex chicken lymphoblast, ATCC, CRL-2111), and CCL-
474 141 (unspecified sex duck embryonic fibroblast, ATCC) cells were obtained from ATCC and used
475 without extensive passaging. Sup-T11 (male human T lymphoblast clone derived from Sup-T1,
476 ATCC, CRL-1942) cells have been described previously(70).

477

478 **Plasmids.** A CpG-enriched IAV segment 5 (encoding nucleoprotein, NP) derived from Gaunt, et
479 al.(43), ordered from IDT as a gBlock, and cloned into a pDZ bidirectional influenza vRNA/mRNA
480 expression plasmid(71) via In-Fusion (Takara Bio). Expression plasmids for human ZAP-L, ZAP-
481 S, KHNYN, and TRIM25 as well as pig ZAP, chicken ZAP, and pig KHNYN were generated using
482 conventional molecular biology techniques. Specifically, RNA was extracted from human A549
483 cells, chicken DF-1, or pig PK15 cells using Qiagen RNeasy Mini Kit, cDNA synthesized using
484 QIAGEN Quantitect cDNA Synthesis Kit, and genes of interest were amplified by PCR from first-
485 strand cDNA, restriction enzyme digested, and ligated into bespoke MLV-based retroviral vectors
486 also harboring blasticidin genes downstream of an IRES for bicistronic expression(72–74) or
487 bespoke pcDNA3-based expression vectors. Dog and platypus KHNYN as well as dog and
488 platypus ZAP were ordered from IDT as gBlocks. Duck ZAP was synthesized by Twist
489 Biosciences followed by subcloning into additional plasmid backbones. Mutants, chimeras, and
490 fluorescently tagged versions of all plasmids were generated by site-directed mutagenesis and/or
491 overlapping PCR and subcloned into expression vectors using Phusion DNA polymerase and T4
492 DNA ligase (NEB). RCAS-GFP plasmid was a gift of Connie Cepko (Addgene #13878). ROSV-
493 mCherry was generated by replacing EGFP in RCAS-GFP with mCherry followed by site-directed
494 mutagenesis to inactivate the Env glycoprotein open reading reframe. HIV-mCherry has been
495 described previously(75). HIV_{CG}-mCh was cloned by inserting a gBlock (IDT) encoding env86-
496 561CpG(20) (from NCBI: MN685350.1) into an HIV-1 NL4-3 Env- Vpr- Nef- mCherry+ (HIV-mCh)
497 reporter virus expression plasmid. MLV-Gag-mCh was a kind gift of Nathan Sherer(76). Codon-

498 optimized HIV-1 NL4-3 Env (SynEnv) was synthesized as a gBlock by IDT prior to subcloning into
499 pcDNA3.2. pCAG-RABV-G (Addgene #36398) expressing Rabies virus glycoprotein (RABV-G)
500 was a kind gift of Connie Cepko(77). CRISPR/Cas9 editing of human A549 cells was performed
501 by lentiviral transduction using pLentiCRISPR1000(78). Attempts at genetic editing in avian cells
502 with pLentiCRISPR1000 were unsuccessful. CRISPR/Cas9 editing of avian cells was performed
503 by sequential retroviral transduction of Cas9-GFP expression vector followed by single guide RNA
504 (sgRNA) expression vector – both bespoke plasmids generated for this study. Briefly, guide RNAs
505 were designed from the literature, Synthego design tool, or RGEN Cas-Designer(79),
506 complementary oligos with Esp3I restriction enzyme compatible ends were annealed, and
507 subsequently ligated into pLentiCRISPR1000 or pMLV.miRFP670.iU6sgRNA via Golden-Gate
508 ligation with T4 DNA ligase. IAV bidirectional rescue plasmids (pDZ) have been described
509 previously(80). Lentiviral constructs for shRNA knockdown of chicken ZAP were cloned by
510 annealing complementary oligos with AgeI and EcoRI restriction enzyme compatible ends and
511 subsequently ligated into pLKO.2 via Golden-Gate ligation with T4 DNA ligase. shRNAs were
512 designed targeting chicken ZC3HAV1 using the Broad Institute's Genetic Perturbation Platform.
513 DNA constructs were sequence confirmed by whole plasmid sequencing performed by
514 Plasmidsaurus using Oxford Nanopore Technology or Sanger sequencing performed by Eurofins
515 Scientific or GeneWiz from Azenta. Plasmid, shRNA, siRNA, and sgRNA sequences are provided
516 in **Supplemental File 1**.
517
518 **Retroviral transductions and stable cell generation.** Simple retroviruses for transduction were
519 generated by transfecting HEK293T cells with 1µg vector plasmid(73, 74), 1µg pMD-gagpol(81),
520 and 200ng pMD-VSVG(82) using Transit-LT1 (Mirus Bio), changing media after 24 hours, and
521 harvesting culture supernatant at 48 hours post-transfection followed by 0.45µm syringe filtration.
522 Lentiviruses for transduction were generated by transfecting HEK293T with 1.25µg
523 pLentiCRISPR1000 or pLKO.2_shRNA plasmid, 750ng psPAX2 (Addgene plasmid #12260;
524 encoding Gag/GagPol, Rev, Tat), and 200ng pMD-VSVG, changing media after 24 hours, and
525 harvesting culture supernatant at 48 hours post-transfection followed by 0.45µm syringe filtration.
526 Transducing virus preps were aliquoted and stored at -20°C. Stable cells were generated as
527 previously described(75, 83). Briefly, approximately 2500 target cells were seeded into a 96-well
528 flat bottom plate, allowed to adhere overnight, and 20-200µL of transducing viral supernatant with
529 up to 10µg/mL polybrene added to each well. After 2-3 days in culture, transduced cells are
530 washed with PBS, and fresh media containing antibiotic (2µg/mL puromycin, 2µg/mL blasticidin,
531 or 200µg/mL hygromycin; GoldBio, Inc) was added. Cells were grown and expanded in the

532 ongoing presence of antibiotic. In the case of CRISPR/Cas9 lentiviral transduction, A549 cells
533 were single cell cloned by serial dilution (re-plating from pooled transducing culture into 96-well
534 flat bottom plates at 25, 5, 1, and 0.2 cells per well or by gradient dilution down and across 96-
535 well plates). After 2-4 days, single cell-derived colonies were marked and followed regularly,
536 expanded, and then screened initially by western blot or PCR. Multiple independent clones were
537 screened for each knockout. Genetic knockout in A549 cells was confirmed by western blot
538 following genomic DNA extraction, amplification of Cas9 target site, blunt cloning into pJET1.2
539 (Thermo Scientific), and Sanger sequencing multiple transformed clones. Double knockout cells
540 were transduced simultaneously (A549.ZK1 and A549.ZK2) with ZAP-targeting (blasticidin-
541 resistant) and KHNYN-targeting (puromycin-resistant) CRISPR/Cas9 lentiviruses, selected with
542 both antibiotics after 48 hours, single cell cloned, screened by western blot or PCR, and confirmed
543 by Sanger sequencing. Avian cells were transduced with simple retrovirus Cas9-GFP (puromycin-
544 resistant). Avian cell knockouts were maintained as pools stably expressing Cas9 and sgRNAs
545 (blasticidin-resistant). Genetic knockout in avian cells was determined by genomic DNA
546 extraction, PCR amplification of Cas9 target site, Sanger sequencing of pooled amplicons, and
547 computational deconvolution of overlapping chromatograms by TIDE analysis(84).
548
549 **Influenza viruses.** Wild-type Influenza A/Puerto Rico/8/1934(H1N1) (PR8) was rescued by
550 plasmid-based transfection into HEK293T cells using Lipofectamine 3000 (ThermoFisher
551 Scientific) and amplified in 10-day old embryonated chicken eggs. $PR8_{CG}$ virus was made by
552 plasmid-based transfection with the WT segment 5 plasmid replaced with CpG-enriched segment
553 5 and amplified in embryonated chicken eggs. Single-cycle PR8 and Cal04
554 (A/California/04/2009(H1N1)) lack virus-encoded HA and instead express mCherry from the HA
555 segment were rescued by plasmid-based transfection into HEK293T cells and amplified in MDCK
556 cells stably expressing PR8 HA(85). Chimeric OH175 Influenza A/Green-Winged
557 Teal/Ohio/175/1986(H2N1) viruses(61) were rescued by plasmid-based transfection into
558 HEK293T cells using six internal segment plasmids from OH175, the HA, and NA segment
559 plasmids from PR8 for biosafety reasons and subsequently amplified in MDCK cells. PR8, $PR8_{CG}$,
560 and OH175 transfections were performed in combination with pCAGGS expression plasmids as
561 needed. OH175 plasmids (also known as S009) were a kind gift of Andrew Mehle. Viral titers
562 were determined by plaque assay in MDCK cells as previously described(86).
563
564 **Influenza multicycle replication assays.** For multicycle IAV infections, 500,000 A549 or DF-1
565 cells were plated in 6-well plates (or 100,000 CCL-141 in 12-well) and allowed to grow overnight

566 for up to 24 hours. Cells were infected at an MOI of 0.05 infecting a confluent culture of cells, as
567 previously described(86). For each cell condition (e.g. DF-1 Vector, ZAP, KHNYN) and for each
568 virus (e.g. PR8 or $PR8_{CG}$) there were three independently infected well replicates. Briefly, cells
569 were washed with PBS prior to addition of 1mL infection media (1x PBS, 2.5% bovine serum
570 albumin, 1% calcium/magnesium). Viral stocks were diluted such that 40µL of inoculum could be
571 added to each well yielding an MOI of 0.05 plaque forming units per cell (pfu/cell). After addition
572 of viral inoculum, cells were incubated at 37°C for 1 hour. Next, cells were washed with PBS prior
573 to addition of 2.5mL viral growth medium (1x DMEM, 2.5% HEPES buffer, 2.5% bovine serum
574 fraction V [7.5%], 1% penicillin/streptomycin) supplemented with TPCK-resistant trypsin (1µg/mL
575 for A549 as well as 0.625 µg/mL for DF-1 and CCL-141). Culture supernatants (at least two
576 aliquots of 200µL each) were collected immediately after addition to wells (0 hours post-infection)
577 and at 24- and 48-hours post-infection (hpi) and optionally at 12-hours post-infection.
578 Supernatants were stored at -80°C until viral titer was determined by plaque assay in MDCK cells
579 as previously described(86). Multicycle replication assays were performed in triplicate with key
580 results being replicated in independent assays (e.g. ZAP-S relative to Vector in Figure 2B, 2C,
581 and 2D). Plaque assays were performed and quantified single-blinded. For transient knockdown
582 by siRNA, 100,000 CCL-141 cells were plate in 12-well plates and transfected with mock, 5nM
583 scramble control, and 1nM or 10nM pools of 10 duplex siRNAs targeting *Anas platyrhynchus*
584 ZC3HAV1 (aplZAP) using RNAiMAX (ThermoFisher). siRNAs were designed using IDT dsiRNA
585 design tool and sequences are provided in **Supplemental File 1**.
586
587 **Single-cycle retrovirus assembly and infectivity assay.** For ROSV, 500,000 DF-1 cells were
588 plated into 6-well plates in 2mL media, cultured overnight, and transfected with 1.6µg RCAS-
589 mCherry, 200ng pMD-VSVG, 100/200/400ng gene of interest plasmid, and topped up to 2.2µg
590 total plasmid DNA mass with pBlueScript using TransIT-LT1 (Mirus Bio). At 24 hours post-
591 transfection, we exchanged media. At 48 hours post-transfection, virus-containing supernatants
592 were harvested and filtered using 3mL syringes and 0.45µm syringe PVDF filters. Transfections
593 were performed as dose-titrations of gene of interest plasmid amounts rather than triplicates of a
594 single plasmid amount. Culture supernatants were used to infect chicken DT40 suspension cells
595 in triplicate and the remaining volume was stored at -80°C. We performed flow cytometry to detect
596 infected (i.e. mCherry-positive) DT40 cells at 48 hours post-infection using a Becton Dickinson
597 FACS Canto II with a high-throughput plate-based sample aspirator. For MLV, HEK293T cells
598 were plated and transfected with 500ng pMD-gagpol, 1100ng pMLV-Gag-mCherry, 200ng pMD-
599 VSVG, 100/200/400ng gene of interest plasmid, and topped up to 2.2µg total plasmid DNA mass

600 with pBlueScript. For HIV-1, HEK293T cells were plated and transfected with 1600ng pHIV-
601 mCherry, 200ng glycoprotein-expression plasmid (pMD2-VSVG, pCAG-RABVG, or
602 pcDNA.HIV.SynEnv), 100/200/400ng gene of interest plasmid, and topped up to 2.2µg total
603 plasmid DNA mass with pBlueScript. MLV single-cycle infectivity was detected by infecting
604 HEK239T cells and HIV-1 infectivity was detected using SupT11 cells (for VSV-G and HIV Env
605 pseudo-typed virus) or HEK293T cells (for RABV-G pseudo-typed virus).
606
607 **Western blotting and antibodies.** A549 cells were treated with 1000U/mL of universal
608 interferon-alpha (PBL Assay Sciences, #11200-2) for 24 hours prior to pelleting cells and lysing
609 in 2.5X reducing sample buffer (RSB; 125 mM Tris HCl, 20% glycerol, 7.5% SDS, 5% β-
610 mercaptoethanol, 250 mM DTT, 0.05% Orange G, pH 6.8). Western blots were performed
611 conventionally using commercial precast polyacrylamide gels (26-well Bio-Rad, 4-20% Criterion
612 TGX, #5671095) for electrophoresis (60 volts for 60 minutes followed by 120 volts for 60 minutes)
613 and transferred to 0.2µm nitrocellulose membrane using Invitrogen Power Blotter XL
614 (ThermoFisher Scientific; ≤25 volts for 12 minutes). Membranes were washed in PBS and PBST
615 (PBS with 0.1% Tween-20) followed by blocking (2% dry milk in PBST). With constant rocking at
616 all steps, membranes were incubated with primary antibodies overnight at 4°C, followed by three
617 washes for five minutes each with PBST, incubation with secondary antibodies in 2% milk in PBST
618 for at least one hour at room temperature, followed by three washes for five minutes each in
619 PBST, and one final wash in PBS. Primary antibodies used include rabbit anti-ZAP (Abcam,
620 ab154680, 1:5,000 dilution), rabbit anti-KHNYN (Invitrogen, PA5-31660, 1:500), rabbit anti-
621 TRIM25 (Abcam, ab167154, 1:10,000), and rabbit anti-H3 (Abcam, ab1791, 1:10,000).
622 Secondary antibodies used include goat anti-mouse IgG IRDye 680LT and goat anti-rabbit IgG
623 IRDye 800CW (both used at 1:10,000 dilution in 2% milk in PBST). Membranes were imaged
624 using a LiCor Odyssey Fc infrared fluorescent imager for 10 minutes per channel.
625
626 **Fluorescence microscopy.** Fixed-cell images were obtained on a BioTek Cytation 5 (Agilent)
627 using a 4X (NA=0.13, #1220519) or 20X (NA=0.40, #1220517) Plan Fluorite objective and LED-
628 filter cube sets for DAPI (excitation = 377±50nm; emission = 447±60nm; #1225007 and
629 #1225100), GFP for mNeonGreen (excitation = 469±35; emission = 525±39; #1225101 and
630 #1225001), and TRITC for mCherry and AlexaFluor568 (excitation = 556±29; emission = 600±37;
631 #1225125 and #1225012). We quantified single-cycle IAV infections by imaging and used the
632 BioTek Gen5 software to calculate percent infectivity using scPR8 (MOI=5) and scCal04 (MOI=5)
633 encoded mCherry as a readout for infectivity (# TRITC+ cells / # DAPI-stained nuclei). To quantify

634 OH175 non-replicating infections, we performed immunofluorescence on infected cells (MOI=0.5
635 without TPCK trypsin to block virus transmission) fixed with 4% paraformaldehyde at 24 hours
636 post-infection. Briefly, we subjected cells to permeabilization (PBS with 0.25% Triton X-100) for
637 10 minutes, blocking (PBS with 3% bovine serum albumin), primary antibody staining with mouse
638 anti-NP (BEI, NR-43899, 1:500) in blocking buffer (3% bovine serum albumin) for one hour,
639 secondary antibody staining with goat anti-rabbit AlexaFlour-568 (Invitrogen, #A-11031, 1:500),
640 and nuclear counterstaining with DAPI (Sigma, D9542, 100ng/mL).
641
642 **Structural models of ZAP and KHNYN.** For structural models of ZAP, amino acid sequences
643 for the four zinc finger domains of different species and mutants of ZAP were used as input for
644 RobettaCM using the human ZAP crystal structure (PDB: 6UEJ) for comparative modeling. We
645 noted that in most cases, *ab initio* folding by AlphaFold or RoseTTAFold may not capture metal
646 ion-coordinated finger structures. Sequences used are provided in **Supplemental File 1**. Models
647 generated and 6UEJ are shown in **Supplemental File 4**. For structural models of KHNYN, amino
648 acid sequences for the KH domain of KHNYN were used as input for RoseTTAFold using default
649 settings as well as RobettaCM using HNRNPK KH-domain (PDB: 1ZZK) for comparative
650 modeling. Sequences used are shown in **Supplemental File 1**. Both RoseTTAFold and
651 RobettaCM are publicly available at robetta.bakerlab.org. All models generated, 1ZZK, and 6Q3V
652 (N4BP1 KH-domain) are provided in **Supplemental File 5 and 6**. Models were visualized in
653 Chimera (v1.13.1, build 41965, UCSC) and aligned using MatchMaker with a Needleman-Wunsch
654 alignment algorithm and BLOSUM-62 matrix.
655
656 **Phylogenetic analyses.** Coding sequences (CDS) of human KHNYN, NYNRIN, N4BP1, and
657 their orthologues were downloaded from ensembl. CDS of ZAP homologues were downloaded
658 from ensemble. CDS were aligned using ClustalOmega in Seaview (V5.0.4). Phylogenetic trees
659 were generated in Seaview using PhyML with 100 bootstraps using CDS. Synteny schematics
660 were generated manually in Adobe Illustrator using data from UCSC ENCODE Genome Browser,
661 NCBI, and ensembl. Alignments and phylogenetic trees are available in **Supplementary Data 7**.
662
663 **Virus bioinformatic analysis.** Full-length influenza virus sequences were downloaded from the
664 Influenza Research Database for each animal host species (human A, avian A, swine A, human
665 B, and human C) including the nucleotide sequence, host, collection date, and subtype.
666 Dinucleotide values were counted by string count. Data were plotted using ggplot2. Data and R

667 code are available in **Supplemental File 3**. Sliding window CpG data were analyzed using an in-
668 house custom script (cpg_counts.R) available in **Supplemental File 3**.
669
670 **RNA-seq**. RNA was extracted from mock and infected samples at 24 hours post-infection
671 (MOI=0.05) using Qiagen RNeasy Kit. Samples were submitted for polyA-selected RNA (mRNA)
672 sequencing using an Illumina NovaSeq 6000 with 150 bp paired end reads at the University of
673 Illinois Urbana-Champaign Roy J. Carver Biotechnology Center. Reads were trimmed of adapter
674 sequences and low-quality reads eliminated using Trimmomatic (V0.32). Trimmed reads were
675 mapped to a hybrid reference genome (Gallus gallus GRCg7b, PR8, and $PR8_{CG}$) using STAR
676 (V2.7.11b) and quantified per feature using HTSeq (V2.0). Differential expression analysis was
677 performed using DESeq2 (v1.34.0). Data were processed and analyzed in R-Studio version
678 1.3.1073 using R language version 3.6.3. RNA-seq data are available at the NCBI Sequence
679 Read Archive (SRA) under accession code XXX.
680
681 **Quantitative reverse-transcription PCR.** To quantify endogenous chicken *ZC3HAV1* and
682 *TUBA4A* mRNA transcripts in shRNA-expressing stable cells following chicken interferon-alpha
683 treatment (1000U/mL; Bio-Rad) for 24 hours, 500,000 cells were plated in 6-well plates and
684 allowed to adhere overnight in culture media followed by addition of chicken interferon or vehicle.
685 After 24 hours, cells were harvested, RNA extracted using Qiagen RNeasy kit, and cDNA
686 synthesized used Qiagen Quantitect Reverse Transcription Kit using random hexamers and poly-
687 dT primers with genomic DNA wipeout. Quantitative PCR was performed in technical triplicate on
688 a LightCycler 480 II (Roche Life Science) using Ssofast Eva Green Supermix (Bio-Rad).
689
690 **Statistical analyses.** Plaque assay results from IAV infections were plotted in GraphPad Prism
691 10 and one-way ANOVA was used at 48 hours post-infection or two-way ANOVA for multiple time
692 points to determine statistical significance. Single-cycle ROSV and IAV infectivity results were
693 plotted in GraphPad Prism 10. Unpaired t tests were used to calculate statistical significance for
694 IAV single-cycle infectivity assays. Statistical analyses are available in **Supplemental File 2**.
695

696 **ACKNOWLEDGEMENTS**
697 We wish to thank members of both Langlois and Harris labs for helpful conversations. We wish
698 to acknowledge and thank Andrew Mehle for the A/green-wing teal/Ohio/1751986(H2N1)
699 OH175/S009 plasmids. This work was supported by NIAID grants R56-AI150965, R01-AI148669,
700 and NIH Centers for Excellence in Influenza Research and Response NHH75N93021C00017
701 awarded to RAL. Salary support for JTB was provided by NIAID F32-AI147813. RSH is an
702 Investigator of the Howard Hughes Medical Institute and the Ewing Halsell President's Council
703 Distinguished Chair. We thank the University of Illinois Roy J. Carver Biotechnology Center for
704 RNA and amplicon sequencing. We report no conflicts of interest.
705
706 **AUTHOR CONTRIBUTIONS**
707 JTB and RAL conceptualized the study. JTB performed all experiments, cloned plasmid
708 constructs, performed phylogenetic and viral CpG content analysis, curated the data, performed
709 formal data analysis, generated figures, and wrote the first complete draft of this manuscript. CKM
710 quantified IAV multicycle replication assays by plaque assay with assistance from AS, LM, EV,
711 AJB, and JTB. FKS developed RNA-seq analysis pipeline and assisted with data analysis and
712 interpretation. CW performed qPCR assay. AJB cloned pig ZAP and KHNYN from *Sus scrofa* PK-
713 15 cells. SLA cloned CpG-enriched IAV PR8 NP Segment 5 gBlock into bidirectional pDZ plasmid.
714 CKM generated single-cycle human IAV viruses in MDCK cells and titered viruses. CKM
715 generated OH175 IAV viruses and titered viruses. SLA and EJF helped generate $PR8_{CG}$ virus
716 stocks in embryonated eggs, sequence confirmed viral stocks, and titered viruses. RSH provided
717 funding, supervision, and resources. RAL provided funding, resources, refined the manuscript,
718 and informed data analysis. All authors contributed to revision of this manuscript.
719
720 **DECLARATION OF INTERESTS**
721 The authors have no competing interests to declare.
722

723 **REFERENCES**

724 1. J. K. Taubenberger, D. M. Morens, Influenza: the once and future pandemic. *Public Health*
725 *Rep. Wash. DC 1974* **125 Suppl 3**, 16–26 (2010).
726 2. J. K. Taubenberger, D. M. Morens, 1918 Influenza: the mother of all pandemics. *Emerg.*
727 *Infect. Dis.* **12**, 15–22 (2006).
728 3. J. S. Long, B. Mistry, S. M. Haslam, W. S. Barclay, Host and viral determinants of influenza
729 A virus species specificity. *Nat. Rev. Microbiol.* **17**, 67–81 (2019).
730 4. Y. Shi, Y. Wu, W. Zhang, J. Qi, G. F. Gao, Enabling the "host jump": structural determinants
731 of receptor-binding specificity in influenza A viruses. *Nat. Rev. Microbiol.* **12**, 822–831
732 (2014).
733 5. T. Walther, *et al.*, Glycomic analysis of human respiratory tract tissues and correlation with
734 influenza virus infection. *PLoS Pathog.* **9**, e1003223 (2013).
735 6. S. E. Galloway, M. L. Reed, C. J. Russell, D. A. Steinhauer, Influenza HA subtypes
736 demonstrate divergent phenotypes for cleavage activation and pH of fusion: implications
737 for host range and adaptation. *PLoS Pathog.* **9**, e1003151 (2013).
738 7. J. S. Long, *et al.*, Species difference in ANP32A underlies influenza A virus polymerase host
739 restriction. *Nature* **529**, 101–104 (2016).
740 8. L. Van Valen, A new evolutionary law. *Evol. Theory* **1**, 1–30 (1973).
741 9. R. P. Smyth, M. Negroni, A. M. Lever, J. Mak, J. C. Kenyon, RNA Structure-A Neglected
742 Puppet Master for the Evolution of Virus and Host Immunity. *Front. Immunol.* **9**, 2097
743 (2018).
744 10. Y. Bermudez, D. Hatfield, M. Muller, A Balancing Act: The Viral-Host Battle over RNA
745 Binding Proteins. *Viruses* **16**, 474 (2024).
746 11. S. Karlin, W. Doerfler, L. R. Cardon, Why is CpG suppressed in the genomes of virtually all
747 small eukaryotic viruses but not in those of large eukaryotic viruses? *J. Virol.* **68**, 2889–
748 2897 (1994).
749 12. S. Karlin, I. Ladunga, B. E. Blaisdell, Heterogeneity of genomes: measures and values.
750 *Proc. Natl. Acad. Sci. U. S. A.* **91**, 12837–12841 (1994).
751 13. B. D. Greenbaum, A. J. Levine, G. Bhanot, R. Rabadan, Patterns of evolution and host gene
752 mimicry in influenza and other RNA viruses. *PLoS Pathog.* **4**, e1000079 (2008).
753 14. D. N. Cooper, S. Gerber-Huber, DNA methylation and CpG suppression. *Cell Differ.* **17**,
754 199–205 (1985).
755 15. P. Simmonds, W. Xia, J. K. Baillie, K. McKinnon, Modelling mutational and selection
756 pressures on dinucleotides in eukaryotic phyla--selection against CpG and UpA in
757 cytoplasmically expressed RNA and in RNA viruses. *BMC Genomics* **14**, 610 (2013).
758 16. A. E. Shaw, *et al.*, The antiviral state has shaped the CpG composition of the vertebrate
759 interferome to avoid self-targeting. *PLoS Biol.* **19**, e3001352 (2021).
760 17. N. J. Atkinson, J. Witteveldt, D. J. Evans, P. Simmonds, The influence of CpG and UpA
761 dinucleotide frequencies on RNA virus replication and characterization of the innate
762 cellular pathways underlying virus attenuation and enhanced replication. *Nucleic Acids*
763 *Res.* **42**, 4527–4545 (2014).
764 18. H. Gu, R. L. Y. Fan, D. Wang, L. L. M. Poon, Dinucleotide evolutionary dynamics in
765 influenza A virus. *Virus Evol.* **5**, vez038 (2019).
766 19. M. A. Takata, *et al.*, CG dinucleotide suppression enables antiviral defence targeting non-
767 self RNA. *Nature* (2017). https://doi.org/10.1038/nature24039.
768 20. M. Ficarelli, *et al.*, KHNYN is essential for the zinc finger antiviral protein (ZAP) to restrict
769 HIV-1 containing clustered CpG dinucleotides. *eLife* **8**, e46767 (2019).
770 21. M. M. H. Li, *et al.*, Characterization of Novel Splice Variants of Zinc Finger Antiviral Protein
771 (ZAP). *J. Virol.* **93**, e00715-19 (2019).

22. J. Schwerk, *et al.*, RNA-binding protein isoforms ZAP-S and ZAP-L have distinct antiviral and immune resolution functions. *Nat. Immunol.* **20**, 1610–1620 (2019).
23. J. A. Kerns, M. Emerman, H. S. Malik, Positive selection and increased antiviral activity associated with the PARP-containing isoform of human zinc-finger antiviral protein. *PLoS Genet.* **4**, e21 (2008).
24. M. D. Daugherty, J. M. Young, J. A. Kerns, H. S. Malik, Rapid evolution of PARP genes suggests a broad role for ADP-ribosylation in host-virus conflicts. *PLoS Genet.* **10**, e1004403 (2014).
25. S. Huang, *et al.*, Positive selection analyses identify a single WWE domain residue that shapes ZAP into a more potent restriction factor against alphaviruses. *PLoS Pathog.* **20**, e1011836 (2024).
26. X. Zheng, *et al.*, TRIM25 Is Required for the Antiviral Activity of Zinc Finger Antiviral Protein. *J. Virol.* **91**, e00088-17 (2017).
27. M. M. H. Li, *et al.*, TRIM25 Enhances the Antiviral Action of Zinc-Finger Antiviral Protein (ZAP). *PLoS Pathog.* **13**, e1006145 (2017).
28. A. E. Shaw, *et al.*, Fundamental properties of the mammalian innate immune system revealed by multispecies comparison of type I interferon responses. *PLoS Biol.* **15**, e2004086 (2017).
29. C. J. Carlson, *et al.*, The future of zoonotic risk prediction. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **376**, 20200358 (2021).
30. R. Rabadan, A. J. Levine, H. Robins, Comparison of avian and human influenza A viruses reveals a mutational bias on the viral genomes. *J. Virol.* **80**, 11887–11891 (2006).
31. Y. Zhang, *et al.*, Influenza Research Database: An integrated bioinformatics resource for influenza virus research. *Nucleic Acids Res.* **45**, D466–D474 (2017).
32. M. J. Pekarek, E. A. Weaver, Existing Evidence for Influenza B Virus Adaptations to Drive Replication in Humans as the Primary Host. *Viruses* **15**, 2032 (2023).
33. B. K. Sederdahl, J. V. Williams, Epidemiology and Clinical Characteristics of Influenza C Virus. *Viruses* **12**, 89 (2020).
34. K. E. Kistler, T. Bedford, An atlas of continuous adaptive evolution in endemic human viruses. *Cell Host Microbe* **31**, 1898-1909.e3 (2023).
35. A. C. Lowen, Constraints, Drivers, and Implications of Influenza A Virus Reassortment. *Annu. Rev. Virol.* **4**, 105–121 (2017).
36. T. Karlberg, *et al.*, Structural basis for lack of ADP-ribosyltransferase activity in poly(ADP-ribose) polymerase-13/zinc finger antiviral protein. *J. Biol. Chem.* **290**, 7336–7344 (2015).
37. D. Kmiec, M. J. Lista, M. Ficarelli, C. M. Swanson, S. J. D. Neil, S-farnesylation is essential for antiviral activity of the long ZAP isoform against RNA viruses with diverse replication strategies. *PLoS Pathog.* **17**, e1009726 (2021).
38. J. L. Meagher, *et al.*, Structure of the zinc-finger antiviral protein in complex with RNA reveals a mechanism for selective targeting of CG-rich viral sequences. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 24303–24309 (2019).
39. X. Luo, *et al.*, Molecular Mechanism of RNA Recognition by Zinc-Finger Antiviral Protein. *Cell Rep.* **30**, 46-52.e4 (2020).
40. Y. Song, *et al.*, High-resolution comparative modeling with RosettaCM. *Struct. Lond. Engl. 1993* **21**, 1735–1742 (2013).
41. M. Baek, *et al.*, Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871–876 (2021).
42. M. Varadi, *et al.*, AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **50**, D439–D444 (2022).
43. E. Gaunt, *et al.*, Elevation of CpG frequencies in influenza A genome attenuates pathogenicity but enhances host response to infection. *eLife* **5**, e12735 (2016).

823 44. D. Gonçalves-Carneiro, M. A. Takata, H. Ong, A. Shilton, P. D. Bieniasz, Origin and
824 evolution of the zinc finger antiviral protein. *PLoS Pathog.* **17**, e1009545 (2021).
825 45. A. Plianchaisuk, *et al.*, Origination of LTR Retroelement-Derived NYNRIN Coincides with
826 Therian Placental Emergence. *Mol. Biol. Evol.* **39**, msac176 (2022).
827 46. V. Anantharaman, L. Aravind, The NYN domains: novel predicted RNAses with a PIN
828 domain-like fold. *RNA Biol.* **3**, 18–27 (2006).
829 47. Z. Yin, *et al.*, RNA-binding motifs of hnRNP K are critical for induction of antibody
830 diversification by activation-induced cytidine deaminase. *Proc. Natl. Acad. Sci. U. S. A.*
831 **117**, 11624–11635 (2020).
832 48. P. H. Backe, A. C. Messias, R. B. G. Ravelli, M. Sattler, S. Cusack, X-ray crystallographic
833 and NMR studies of the third KH domain of hnRNP K in complex with single-stranded
834 nucleic acids. *Struct. Lond. Engl. 1993* **13**, 1055–1067 (2005).
835 49. M. Ficarelli, *et al.*, CpG Dinucleotides Inhibit HIV-1 Replication through Zinc Finger Antiviral
836 Protein (ZAP)-Dependent and -Independent Mechanisms. *J. Virol.* **94**, e01337-19 (2020).
837 50. R. P. Galão, *et al.*, TRIM25 and ZAP target the Ebola virus ribonucleoprotein complex to
838 mediate interferon-induced restriction. *PLoS Pathog.* **18**, e1010530 (2022).
839 51. S. Müller, *et al.*, Inhibition of filovirus replication by the zinc finger antiviral protein. *J. Virol.*
840 **81**, 2391–2400 (2007).
841 52. J. J. Fros, *et al.*, The dinucleotide composition of the Zika virus genome is shaped by
842 conflicting evolutionary pressures in mammalian hosts and mosquito vectors. *PLoS Biol.*
843 **19**, e3001201 (2021).
844 53. R. S. Harris, J. F. Hultquist, D. T. Evans, The restriction factors of human immunodeficiency
845 virus. *J. Biol. Chem.* **287**, 40875–40883 (2012).
846 54. M. H. Malim, P. D. Bieniasz, HIV Restriction Factors and Mechanisms of Evasion. *Cold*
847 *Spring Harb. Perspect. Med.* **2**, a006940 (2012).
848 55. V. Odon, S. R. Fiddaman, A. L. Smith, P. Simmonds, Comparison of CpG- and UpA-
849 mediated restriction of RNA virus replication in mammalian and avian cells and
850 investigation of potential ZAP-mediated shaping of host transcriptome compositions. *RNA*
851 *N. Y. N* **28**, 1089–1109 (2022).
852 56. C.-H. Liu, L. Zhou, G. Chen, R. M. Krug, Battle between influenza A virus and a newly
853 identified antiviral activity of the PARP-containing ZAPL protein. *Proc. Natl. Acad. Sci. U.*
854 *S. A.* **112**, 14048–14053 (2015).
855 57. Q. Tang, X. Wang, G. Gao, The Short Form of the Zinc Finger Antiviral Protein Inhibits
856 Influenza A Virus Protein Expression and Is Antagonized by the Virus-Encoded NS1. *J.*
857 *Virol.* **91**, e01909-16 (2017).
858 58. C. P. Sharp, *et al.*, CpG dinucleotide enrichment in the influenza A virus genome as a live
859 attenuated vaccine development strategy. *PLoS Pathog.* **19**, e1011357 (2023).
860 59. O. T. Gorman, W. J. Bean, Y. Kawaoka, R. G. Webster, Evolution of the nucleoprotein gene
861 of influenza A virus. *J. Virol.* **64**, 1487–1497 (1990).
862 60. M. Gammelin, *et al.*, Phylogenetic analysis of nucleoproteins suggests that human influenza
863 A viruses emerged from a 19th-century avian ancestor. *Mol. Biol. Evol.* **7**, 194–200 (1990).
864 61. A. Mehle, J. A. Doudna, Adaptive strategies of the influenza virus polymerase for replication
865 in humans. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 21312–21316 (2009).
866 62. I. Kosik, J. W. Yewdell, Influenza Hemagglutinin and Neuraminidase: Yin¯Yang Proteins
867 Coevolving to Thwart Immunity. *Viruses* **11**, 346 (2019).
868 63. Y. Chen, *et al.*, Rare variant MX1 alleles increase human susceptibility to zoonotic H7N9
869 influenza virus. *Science* **373**, 918–922 (2021).
870 64. A. Mehle, V. G. Dugan, J. K. Taubenberger, J. A. Doudna, Reassortment and mutation of
871 the avian influenza virus polymerase PA subunit overcome species barriers. *J. Virol.* **86**,
872 1750–1757 (2012).

873 65. D. Gonçalves-Carneiro, *et al.*, Rational attenuation of RNA viruses with zinc finger antiviral
874 protein. *Nat. Microbiol.* **7**, 1558–1567 (2022).
875 66. K. Q. de Andrade, C. C. Cirne-Santos, Antiviral Activity of Zinc Finger Antiviral Protein (ZAP)
876 in Different Virus Families. *Pathog. Basel Switz.* **12**, 1461 (2023).
877 67. G. Gao, X. Guo, S. P. Goff, Inhibition of retroviral RNA production by ZAP, a CCCH-type
878 zinc finger protein. *Science* **297**, 1703–1706 (2002).
879 68. Z. C. Yeoh, *et al.*, A minimal complex of KHNYN and zinc-finger antiviral protein binds and
880 degrades single-stranded RNA. *Proc. Natl. Acad. Sci. U. S. A.* **121**, e2415048121 (2024).
881 69. W. E. Johnson, Origins and evolutionary consequences of ancient endogenous retroviruses.
882 *Nat. Rev. Microbiol.* **17**, 355–370 (2019).
883 70. E. W. Refsland, *et al.*, Quantitative profiling of the full APOBEC3 mRNA repertoire in
884 lymphocytes and tissues: implications for HIV-1 restriction. *Nucleic Acids Res.* **38**, 4274–
885 4284 (2010).
886 71. M. Quinlivan, *et al.*, Attenuation of equine influenza viruses through truncations of the NS1
887 protein. *J. Virol.* **79**, 8431–8439 (2005).
888 72. S. Gallois-Montbrun, *et al.*, Antiviral protein APOBEC3G localizes to ribonucleoprotein
889 complexes found in P bodies and stress granules. *J. Virol.* **81**, 2165–2178 (2007).
890 73. E. L. Evans, J. T. Becker, S. L. Fricke, K. Patel, N. M. Sherer, HIV-1 Vif's Capacity To
891 Manipulate the Cell Cycle Is Species Specific. *J. Virol.* **92** (2018).
892 74. J. T. Becker, A. A. Auerbach, R. S. Harris, APEX3 – An Optimized Tool for Rapid and
893 Unbiased Proximity Labeling. *J. Mol. Biol.* **435**, 168145 (2023).
894 75. J. T. Becker, N. M. Sherer, Subcellular Localization of HIV-1 gag-pol mRNAs Regulates
895 Sites of Virion Assembly. *J. Virol.* **91** (2017).
896 76. F. Li, X. Sewald, J. Jin, N. M. Sherer, W. Mothes, Murine leukemia virus Gag localizes to the
897 uropod of migrating primary lymphocytes. *J. Virol.* **88**, 10541–10555 (2014).
898 77. K. T. Beier, *et al.*, Anterograde or retrograde transsynaptic labeling of CNS neurons with
899 vesicular stomatitis virus vectors. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 15414–15419 (2011).
900 78. M. A. Carpenter, E. K. Law, A. Serebrenik, W. L. Brown, R. S. Harris, A lentivirus-based
901 system for Cas9/gRNA expression and subsequent removal by Cre-mediated
902 recombination. *Methods San Diego Calif* **156**, 79–84 (2019).
903 79. J. Park, S. Bae, J.-S. Kim, Cas-Designer: a web-based tool for choice of CRISPR-Cas9
904 target sites. *Bioinforma. Oxf. Engl.* **31**, 4014–4016 (2015).
905 80. N. E. Sanjana, O. Shalem, F. Zhang, Improved vectors and genome-wide libraries for
906 CRISPR screening. *Nat. Methods* **11**, 783–784 (2014).
907 81. D. S. Ory, B. A. Neugeboren, R. C. Mulligan, A stable human-derived packaging cell line for
908 production of high titer retrovirus/vesicular stomatitis virus G pseudotypes. *Proc. Natl.*
909 *Acad. Sci. U. S. A.* **93**, 11400–11406 (1996).
910 82. L. Naldini, *et al.*, In vivo gene delivery and stable transduction of nondividing cells by a
911 lentiviral vector. *Science* **272**, 263–267 (1996).
912 83. J. Wang, *et al.*, The Role of RNA in HIV-1 Vif-Mediated Degradation of APOBEC3H. *J. Mol.*
913 *Biol.* **431**, 5019–5031 (2019).
914 84. E. K. Brinkman, T. Chen, M. Amendola, B. van Steensel, Easy quantitative assessment of
915 genome editing by sequence trace decomposition. *Nucleic Acids Res.* **42**, e168 (2014).
916 85. L. E. Sjaastad, *et al.*, Distinct antiviral signatures revealed by the magnitude and round of
917 influenza virus replication in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 9610–9615 (2018).
918 86. B. M. Waring, *et al.*, MicroRNA-Based Attenuation of Influenza Virus across Susceptible
919 Hosts. *J. Virol.* **92** (2018).
920
921

922 **FIGURE LEGENDS**

923

924 **Figure 1 Avian ZAP does not restrict replication of CpG-enriched IAV.**

925 **(A)** IAV sequences exhibit depletion of CpG content (CpG per kb = # CpG / segment length in
926 kilobases [kb]). Box and whisker plots showing CpG content of all IAV sequences from avian
927 (orange), human (blue), and swine (green) hosts as well as IBV (pink) and ICV (purple). Thick
928 black bars represent median CpG content, boxes define the 25th to 75th percentile, and whiskers
929 represent range excluding outliers. Horizontal dashed gray lines highlight median avian IAV CpG
930 content as a reference.

931 **(B)** Schematic of human ZAP-L, human ZAP-S and chicken ZAP (ggaZAP) protein domain
932 architecture including four CCCH zinc fingers and inactive poly-ADP ribose polymerase (iPARP).
933 Amino acid alignment of ZAP zinc finger 2 (ZF2; human residues 88-110) from multiple species.
934 Conserved CCCH residues are highlighted with green vertical bars.

935 **(C)** CpG content of PR8 (shades of gray by segment length), wild-type NP (blue), and CpG-
936 enriched NP (orange) calculated as # CpG per 100 nucleotides (nt) over a sliding window (top).
937 Schematic of PR8 and PR8$_{CG}$ viruses and displaying CpG content for each virus and total CpG
938 number encoded by each NP segment (bottom).

939 **(D)** Titers of PR8 and PR8$_{CG}$ at 48 hours post infection (hpi) in chicken DF-1 cells stably
940 expressing mNeonGreen (mNG, gray), chicken ZAP (ggaZAP, yellow), or mNG-ggaZAP (orange)
941 as well as control transduced cells (Vector, black). Bar represents mean plaque forming units per
942 mL (PFU/mL) with error bars representing standard deviation of the mean from three biological
943 replicates. Multiplicity of infection (MOI) = 0.05 (pfu per cell).

944 **(E)** Fluorescence images of DF-1 cells expressing mNG as well as fusion proteins of mNG-
945 ggaZAP and mutants indicated (green, top). Nuclei counterstained with DAPI (blue) and shown
946 in merge image (bottom). Scale bars represent 10 microns.

947 **(F)** Titers of PR8 and PR8$_{CG}$ over 48 hpi in chicken DF-1 cells stably expressing mNG-ggaZAP
948 mutants shown in (E). MOI=0.05.

949 **(G)** Titers of PR8 and PR8$_{CG}$ at 48 hpi in parental untransduced DF-1 cells (DF-1, black), pooled
950 CRISPR/Cas9 knockout chicken DF-1 cells expressing a control sgRNA targeting E. coli LacZ
951 (LacZ, gray), or sgRNAs targeting endogenous chicken ZC3HAV1 (ggaZAP1-3 KO, shades of
952 orange). MOI=0.05.

953 **(H)** Titers of PR8 and PR8$_{CG}$ at 48 hpi in parental untransduced CCL-141 cells (CCL-141, black),
954 pooled CRISPR/Cas9 knockout duck CCL-141 cells expressing a control sgRNA targeting E. coli

955 LacZ (LacZ, gray), or sgRNAs targeting endogenous duck ZC3HAV1 (aplZAP1-3 KO, shades of
956 yellow). MOI=0.05.
957
958 **Figure 2 Human ZAP-S and KHNYN independently restrict CpG-enriched IAV.**
959 **(A)** Schematic showing lack of CpG-targeting activity in chicken cells (left) and restriction of
960 CpG-enriched cytoplasmic RNAs by human ZAP, KHNYN, or TRIM25 transduced into chicken
961 cells (right).
962 **(B)** Titers of PR8 and $PR8_{CG}$ over 48 hpi in chicken DF-1 cells stably expressing human ZAP-L
963 (green), ZAP-S (lime green), KHNYN (blue), dKHNYN (gray), or TRIM25 (pink) relative to Vector
964 (black). MOI=0.05. *** indicates $p \leq 0.0001$ at 48hpi by two-way ANOVA.
965 **(C)** Titers of PR8 and $PR8_{CG}$ over 48 hpi in chicken DF-1 cells stably expressing RNA-binding
966 mutants of ZAP-S (lime green) including Y108A (gray with solid line) and chimeric encoding
967 chicken ZF2 (gZF2; gray with black outline and dashed line) relative to vector (black). MOI=0.05.
968 ** indicates p=0.0093 at 48hpi by two-way ANOVA.
969 **(D)** Titers of PR8 and $PR8_{CG}$ over 48 hpi in chicken DF-1 cells stably expressing C-terminal
970 mutants of human ZAP-L (green) and ZAP-S (lime green) including ZAP-L C899S (green with
971 gray outline and dashed line) and ZAP-S +CVIS motif (CVIS; lime green with gray outline and
972 dashed line). MOI=0.05. * indicates p=0.0108 at 48hpi by two-way ANOVA.
973 **(E)** Schematic of human KHNYN protein showing KH-like domain (tan), NYN endonuclease
974 domain (pink), and CUBAN domain (gray) as well as the GXXG motif with amino acid locations
975 indicated (top). Structure and model of human KH-domains including KHNYN (residues 58-141,
976 blue, RobettaCM model), HNRNPK (green, PDB: 1ZZK), and N4BP1 (pink, PDB: 6Q3V) from two
977 views (bottom). Putative GXXG motif is indicated at top of the central alpha-helix of the three
978 structures.
979 **(F)** Titers of PR8 and $PR8_{CG}$ over 48 hpi in chicken DF-1 cells stably expressing mutants of
980 KHNYN including catalytically inactive dKHNYN (gray), $_{95}AQ/DE_{96}$ mutant (AQDE; blue dot with
981 black outline and light blue dashed line), and combination mutant of dKHNYN with AQDE
982 (dAQDE; gray dot with black outline and gray dashed line). MOI=0.05. * indicates p=0.0158 at
983 48hpi by two-way ANOVA.
984
985 **Figure 3 Endogenous human ZAP and KHNYN independently restrict CpG-enriched IAV.**
986 **(A)** Schematic depicting exon organization of human ZAP (green), KHNYN (blue), and TRIM25
987 (pink) genes and sites targeted by CRISPR single guide RNAs (sgRNAs; scissors). Colored boxes
988 indicate coding exons and white boxes indicate noncoding portions of exons.

989 **(B)** Western blots of CRISPR/Cas9 engineered knockout (KO) human A549 cell lines including
990 parental A549 cells, control clone transduced with pLentiCRISPR1000 (Vector), control clone
991 targeting E. coli LacZ (LacZ), two ZAP KO clones (ZAP1 and ZAP2), two KHNYN KO clones (K1
992 and K2), two ZAP and KHNYN dual KO clones (ZK1 and ZK2), and two TRIM25 KO clones (T1
993 and T2). Cells were treated with vehicle (PBS) or 1000 units per mL (U/mL) universal interferon-
994 alpha (IFNα) for 24 hours. Histone H3.3 detected as loading control.
995 **(C)** Titers of PR8 and $PR8_{CG}$ at 48 hpi in human A549 KO cells including parental (black), Vector
996 and LacZ (gray), ZAP KO (green), and KHNYN KO (blue). MOI=0.05.
997 **(D)** Titers of PR8 and $PR8_{CG}$ at 48 hpi in human A549 KO cells including parental (black), Vector
998 (gray), and ZAP+KHNYN KO (green and blue striped). MOI=0.05. *** indicates $p \leq 0.0001$ and **
999 indicates p=0.0045 by one-way ANOVA.
1000 **(E)** Titers of PR8 and $PR8_{CG}$ at 48 hpi in human A549 KO cells including parental (black), Vector
1001 (gray), and TRIM25 KO (pink). MOI=0.05.
1002
1003 **Figure 4 Prenylated human ZAP and mammalian KHNYN independently restrict naturally**
1004 **CpG-enriched ROSV.**
1005 **(A)** CpG content of retroviruses (ROSV in black, ROSV-mCherry reporter virus in pink dashed
1006 line, MLV in orange, and HIV-1 in blue) calculated as # CpG per 100 nucleotides over a sliding
1007 window (graph). Schematic of retrovirus open reading frame (ORF) organization (top).
1008 **(B)** Schematic of retroviral assembly and single-cycle infectivity assay including plasmid
1009 transfection into cells, cell lysate and supernatant harvest at 48 hours post-transfection (hpt), and
1010 storage as well as DT40 suspension cell infection and flow cytometry as a measure of infectivity.
1011 **(C)** Retroviral assembly and single-cycle infectivity results of human ZAP-L, ZAP-S, and mutants
1012 effects on ROSV showing relative percent ROSV infected cells (% mCherry positive; % mCh+)
1013 as determined by flow cytometry. DF-1 cells transfected with increasing plasmid amounts of
1014 indicated host proteins (100, 200, or 400ng), 1500ng R, and 200ng VSV-G. Infectivity is relative
1015 to backbone pcDNA3.2 transfections (Mock).
1016 **(D)** Retroviral assembly and single-cycle infectivity results of human ZAP-L, chicken (gga) ZAP,
1017 duck (apl) ZAP, and prenylation mutant on ROSV infectivity.
1018 **(E)** Retroviral assembly and single-cycle infectivity results of human KHNYN, mutants, and
1019 mammalian KHNYN homologues (from pig, dog, and platypus) on ROSV infectivity.
1020 **(F)** Single-cycle infectivity of YFP alone (black) and YFP-tagged platypus KHNYN (blue) effects
1021 on single-cycle IAV including laboratory PR8, pandemic Cal04, and avian-origin OH175. ***
1022 indicates p=0.000832, * indicates p=0.025875, and ** indicates p=0.002263 by unpaired t test.

1023 **SUPPLEMENTAL FIGURE LEGENDS**

1024

1025 **Figure S1 Mammalian IAV sequences are CpG-depleted over time relative to avian IAV.**

1026 **(A)** IAV sequences from human (blue) and swine (green) hosts exhibit depletion of CpG content
1027 (CpG per kb = # CpG / segment length in kilobases (kb)) relative to avian (orange) IAV over time
1028 since 1918. IBV (pink) and ICV (purple) exhibit relatively stable but dramatically lower CpG
1029 content than all IAV segments. Thick lines represent linear trend of data surrounded by 95%
1030 confidence interval in lighter shading.

1031 **(B)** CpG content of IAV sequences from avian, human, and swine segregated by virus segment.
1032 Segments 1-3 include PB2, PB1, and PA encoding the viral Polymerase. Segments 4-6 include
1033 HA, NP, and NA encoding virion proteins often subject to adaptive immune pressure. Segments
1034 7-8 include M and NS.

1035 **(C)** CpG content of IAV sequences from avian, human, and swine hosts segregated by common
1036 HA and NA subtypes.

1037 **(D)** CpG content of IAV sequences segregated by time periods indicated on the x-axis (top).
1038 Schematic of IAV natural history showing avian H1N1, H2N1, and H3N2 in orange shaded lines,
1039 human circulating subtypes in shaded blue lines, 2009 H1N1 pandemic swine flu in green lines,
1040 (middle) and source of reassortant segments in matched colors (bottom).

1041 Thick black bars represent median CpG content, boxes define the 25th to 75th percentile, and
1042 whiskers represent range excluding outliers. Horizontal dashed gray lines highlight median avian
1043 IAV CpG content as a reference.

1044

1045 **Figure S2 Dinucleotide content of influenza viruses.**

1046 **(A)** CpG content as the normalized rho (ρ = #CpG / length divided by the product of #C / length
1047 and #G per length) of all IAV sequences from avian, human, and swine hosts as well as IBV and
1048 ICV.

1049 **(B)** Percent GC content (%GC = (#G + #C) / total length) all IAV sequences from avian, human,
1050 and swine hosts as well as IBV and ICV.

1051 **(C)** GpC content of all IAV sequences from avian, human, and swine hosts as well as IBV and
1052 ICV.

1053 **(D)** Remaining dinucleotide content of all IAV sequences from avian, human, and swine hosts as
1054 well as IBV and ICV.

1055

1056 **Figure S3 Structural comparison of avian ZAP proteins and data for avian ZAP**
1057 **knockdowns.**
1058 **(A)** Structure and model of human (blue; PDB: 6UEJ), chicken (RobettaCM, orange), chicken
1059 (RoseTTAFold, medium orange), and chicken (AlphaFold2, light orange) ZAP zinc fingers (1-4)
1060 with a CpGpU RNA oligonucleotide (gray; PDB: 6UEJ).
1061 Structure and model of human (blue; PDB: 6UEJ), duck (RobettaCM, green), duck
1062 (RoseTTAFold, medium green), and duck (AlphaFold2, light green) ZAP zinc fingers (1-4) with a
1063 CpGpU RNA oligonucleotide (gray; PDB: 6UEJ).
1064 Structure and model of human (blue; PDB: 6UEJ), quail (RobettaCM, pink), quail (RoseTTAFold,
1065 medium pink), and quail (AlphaFold2, light pink) ZAP zinc fingers (1-4) with a CpGpU RNA
1066 oligonucleotide (gray; PDB: 6UEJ).
1067 Structure and model of human (blue; PDB: 6UEJ), chicken (RobettaCM, orange), duck
1068 (RobettaCM, green), and chicken (AlphaFold2, light pink) ZAP zinc fingers (1-4) with a CpGpU
1069 RNA oligonucleotide (gray; PDB: 6UEJ).
1070 **(B)** Root mean squared deviation (RMSD) of structures and models in (A-D).
1071 **(C)** qPCR results detecting endogenous gga *ZC3HAV1* mRNA in chicken DF-1 cells transduced
1072 with lentiviruses expressing shRNAs indicated (shades of gray) following treatment with 1000
1073 U/mL of chicken interferon-alpha (IFN) relative to mock treated cells normalized to gga *TUBA1A*.
1074 **(D)** Titers of PR8 and $PR8_{CG}$ over 48 hpi in chicken DF-1 cells stably expressing shRNAs targeting
1075 endogenous ggaZAP (shggaZAP1-3, shades of gray), a control scrambled sequence siRNA
1076 (shScramble, dark gray), and parental untransduced DF-1 cells (black). MOI=0.05.
1077 **(E)** Titers of PR8 and $PR8_{CG}$ at 48 hpi in duck CCL-141 cells transiently transfected with 1nM or
1078 10nM siRNAs targeting endogenous aplZAP (siZAP, shades of yellow), a control scrambled
1079 sequence siRNA (siNeg, 5nM), and parental untransfected CCL-141 cells. MOI=0.05.
1080
1081 **Figure S4 TIDE analysis of pooled CRISPR/Cas9 chicken DF-1 and duck CCL-141**
1082 **ZC3HAV1/ZAP knockouts.**
1083 **(A-B)** Schematic depicting exon organization of chicken ZC3HAV1/ZAP (orange) and duck
1084 ZC3HAV1/ZAP (yellow) as well as sites targeted by CRISPR single guide RNAs (sgRNAs;
1085 scissors). Colored boxes indicate coding exons and white boxes indicate noncoding portions of
1086 exons. Genomic sequence targeted by sgRNAs (underlined) and protospacer adjacent motif
1087 (PAM, lowercase) and amino acids encoding by that region (bold residues overlap sgRNA target).
1088 Genomic distances are relative and total locus size indicated.

1089 **(C)** Total knockout efficiency as determined by tracking of indels by decomposition (TIDE)
1090 analysis (% Knockout) in each pool relative to LacZ control cells.
1091 **(D-F)** TIDE analysis showing percent of sequences in pool exhibiting indicated indel (-10 to +10
1092 bp from expected cut site) in chicken ZC3HAV1-targeted KO DF-1 cells.
1093 **(G-I)** TIDE analysis showing percent of sequences in pool exhibiting indicated indel (-10 to +10
1094 bp from expected cut site) in duck ZC3HAV1-targeted KO CCL-141 cells.
1095 Red bars indicate disruptive indels (CRISPR-mediated insertions or deletions resulting in out-of-
1096 frame mutations). Gray bars indicate un-edited sequences. Black bars indicate in-frame indels
1097 (CRISPR-mediated insertions or deletions resulting in amino acid deletions).
1098
1099 **Figure S5 Genetic and genomic comparisons of KHNYN genes**
1100 **(A)** Synteny diagrams of KHNYN genomic organization in representative species. KHNYN gene
1101 in blue, NYNRIN in magenta, neighboring human/pig/koala genes in black, neighboring platypus
1102 genes in red, and human/chicken genes syntenic to red genes in gray. Distances are relative and
1103 approximate size of locus shown on the right in kilobasepairs (Kbp).
1104 **(B)** Schematic of exon organization (left) and protein domain architecture (right) of human N4BP1
1105 (pink), KHNYN (blue), and NYNRIN (magenta). KH domain represents the extended HNRNPK-
1106 like, GXXG-like motif residues typical of KH-domains are indicated, NYN represents the NYN/PIN
1107 endonuclease domain, and CU represents cullin-binding domain associated with NEDD8
1108 (KHNYN and NYNRIN) or cousin of CUBAN domain (N4BP1).
1109 **(C)** Maximum likelihood phylogeny of all ensembl-annotated orthologues of human KHNYN,
1110 NYNRIN, and N4BP1.
1111 **(D)** Synteny diagrams and schematics of exon organization for alligator, anole, and zebrafish
1112 *"khnyn-like"* genes.
1113 **(E)** Structural predictions in three views of human (dark blue), pig (medium blue), dog (cyan), and
1114 platypus (lime green) KHNYN (residues 1-213). Putative GXXG motif highlighted by dashed black
1115 circle. Platypus-specific short loop highlighted by dashed black square.
1116
1117 **Figure S6 Human transgenes do not affect IAV mRNA or chicken transcriptome.**
1118 **(A)** Principal component analysis (PCA) of RNA-seq data.
1119 **(B)** Heatmap showing differentially expressed IAV mRNAs.
1120 **(C)** Heatmap showing most differentially expressed chicken mRNAs.
1121
1122 **Figure S7 Platypus KHNYN is a potent and broadly acting antiviral protein.**

1123 **(A)** Retroviral assembly and single-cycle infectivity results of mammalian ZAP-L and ZAP-S
1124 homologues including pig, dog, and platypus on ROSV infectivity. Platypus encodes a well
1125 annotated ZAP-L homologue but not a ZAP-S, a C653S prenylation mutant was included.
1126 **(B)** Retroviral assembly and single-cycle infectivity results of mammalian KHNYN homologues
1127 including pig, dog, and platypus as well as catalytically inactivated (*"d"*) mutants of each on ROSV
1128 infectivity.
1129 **(C)** Retroviral assembly and single-cycle infectivity results of platypus KHNYN including truncation
1130 mutants on ROSV infectivity (top). Schematic of protein domain architecture of platypus KHNYN
1131 with arrows indicating position of truncating stop mutation (bottom).
1132 **(D)** Retroviral assembly and single-cycle infectivity results of platypus KHNYN including truncation
1133 mutants on ROSV infectivity (top). Indication (bottom) of virus tested (HIV-mCh, HIV_{CG}-mCh, or
1134 MLV-mCh) and glycoprotein used for pseudotyping (HIV-1 Env, VSV-G, or RABV-G).
1135
1136 **Supplemental File 1. Key reagents.** Spreadsheet
1137 **Supplemental File 2. Statistical analyses.** Spreadsheet
1138 **Supplemental File 3. Influenza sequences, metadata, and R code.** Spreadsheet, text files
1139 **Supplemental File 4. Human and avian ZAP structural models.** Python files open in UCSC
1140 Chimera
1141 **Supplemental File 5. Human KHNYN, N4BP1, and HNRNPK structural models.** Python files
1142 open in UCSC Chimera
1143 **Supplemental File 6. Mammalian KHNYN structural models.** Python files open in UCSC
1144 Chimera
1145 **Supplemental File 7. N4BP1, KHNYN, and NYNRIN phylogenetic data.** Clustal aln
1146 alignment, pdf of phylogenetic trees
1147

Figure 1

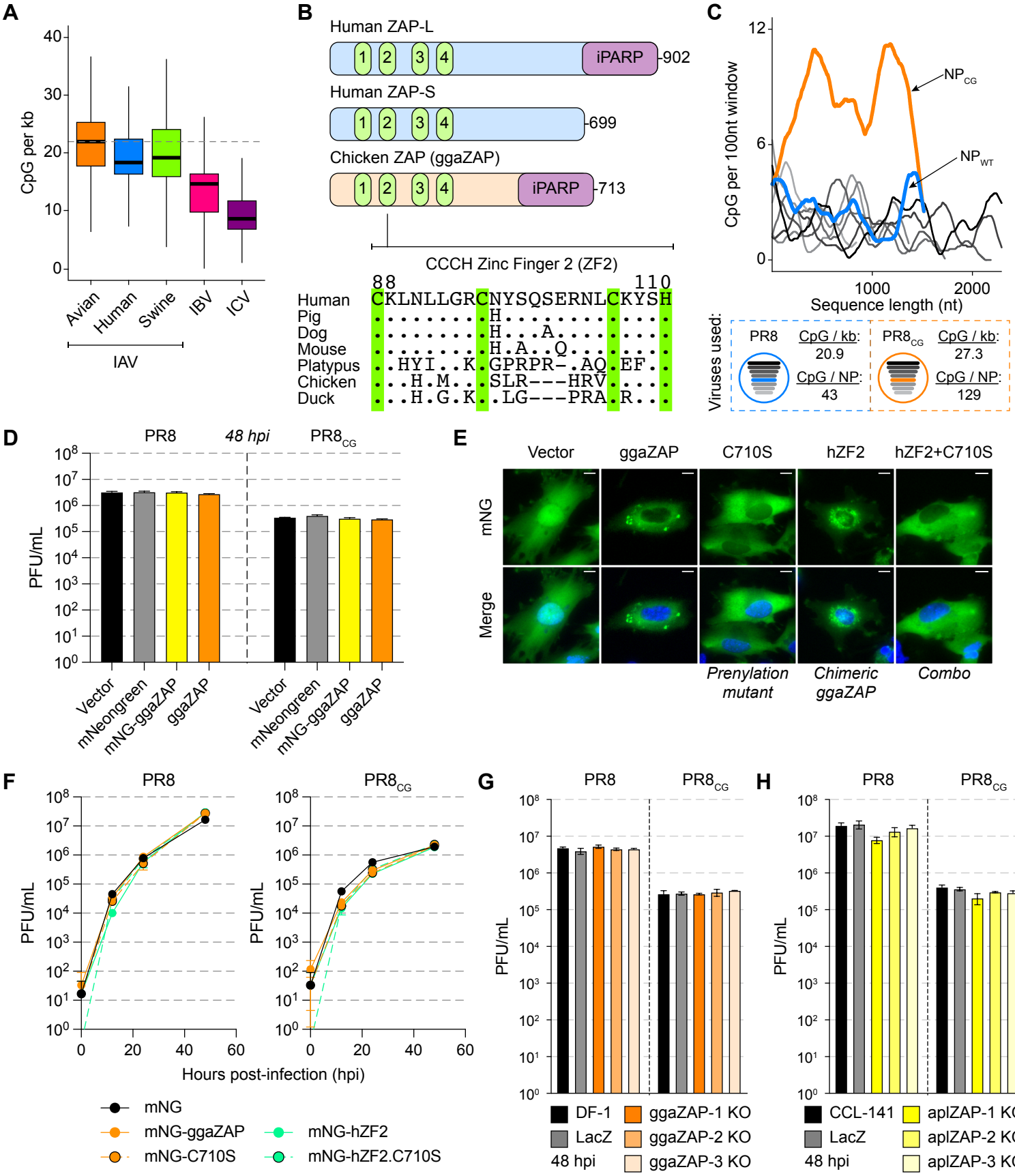

A
CpG per kb
Avian
Human
Swine
IBV
ICV
IAV
B
Human ZAP-L
iPARP
902
Human ZAP-S
699
Chicken ZAP (ggaZAP)
713
CCCH Zinc Finger 2 (ZF2)
88
110
Human CKLNLLGRCNYSQSERNLCKYSH
Pig
Dog
Mouse
Platypus
Chicken
Duck
C
CpG per 100nt window
NP_CG
NP_WT
Sequence length (nt)
Viruses used:
PR8
CpG / kb: 20.9
CpG / NP: 43
PR8_CG
CpG / kb: 27.3
CpG / NP: 129
D
PR8
48 hpi
PR8_CG
PFU/mL
Vector
mNeongreen
mNG-ggaZAP
ggaZAP
E
Vector
ggaZAP
C710S
hZF2
hZF2+C710S
mNG
Merge
Prenylation mutant
Chimeric ggaZAP
Combo
F
PR8
PR8_CG
PFU/mL
Hours post-infection (hpi)
mNG
mNG-ggaZAP
mNG-C710S
mNG-hZF2
mNG-hZF2.C710S
G
PR8
PR8_CG
PFU/mL
DF-1
LacZ
48 hpi
ggaZAP-1 KO
ggaZAP-2 KO
ggaZAP-3 KO
H
PR8
PR8_CG
PFU/mL
CCL-141
LacZ
48 hpi
aplZAP-1 KO
aplZAP-2 KO
aplZAP-3 KO

Figure 2

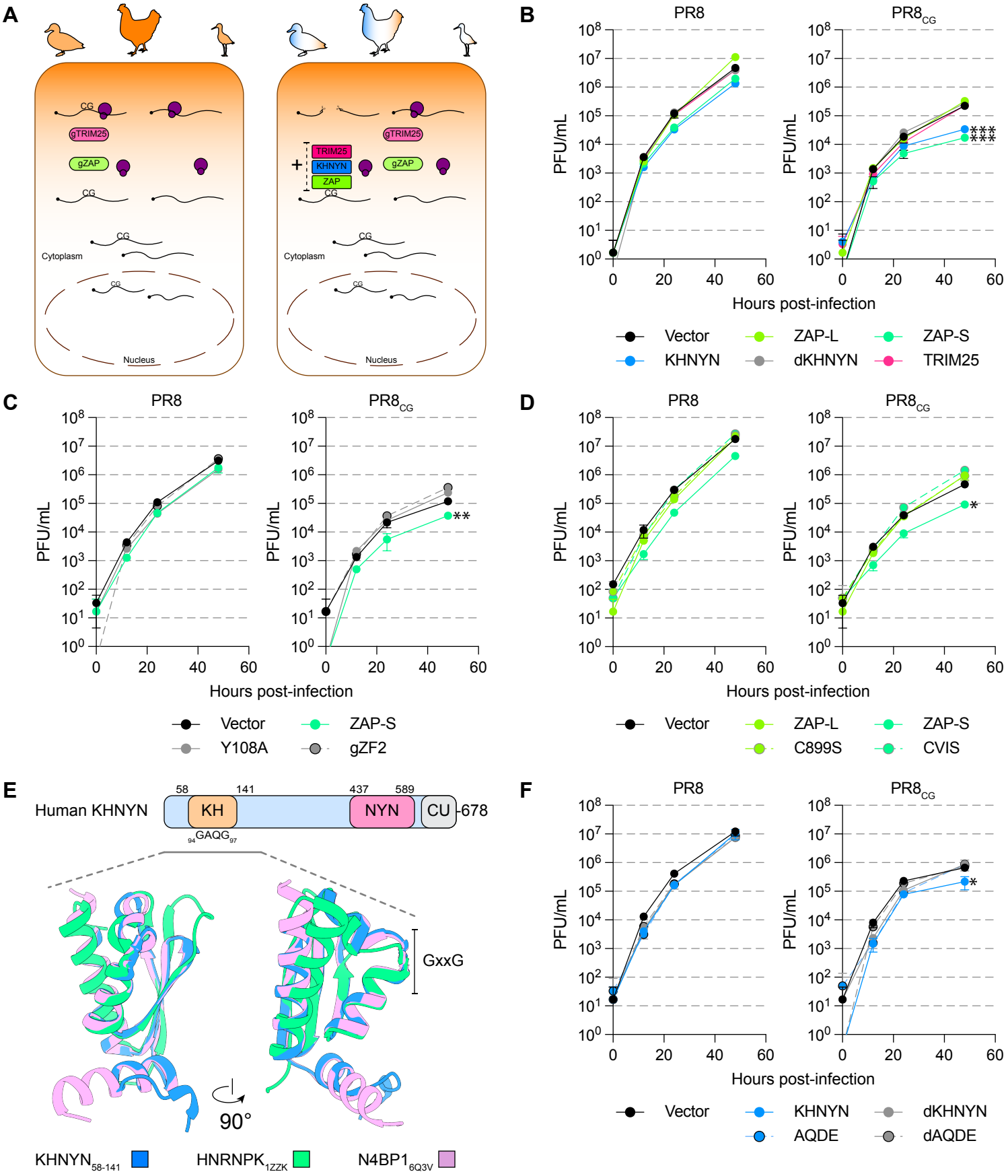

A
CG
gTRIM25
gZAP
Cytoplasm
Nucleus
TRIM25
KHNYN
ZAP
B
PR8
PR8CG
PFU/mL
Hours post-infection
Vector
ZAP-L
ZAP-S
KHNYN
dKHNYN
TRIM25
C
Y108A
gZF2
D
C899S
CVIS
E
Human KHNYN
58
141
437
589
678
KH
NYN
CU
94GAQG97
GxxG
90°
KHNYN58-141
HNRNPK1ZZK
N4BP16Q3V
F
AQDE
dAQDE

Figure 3

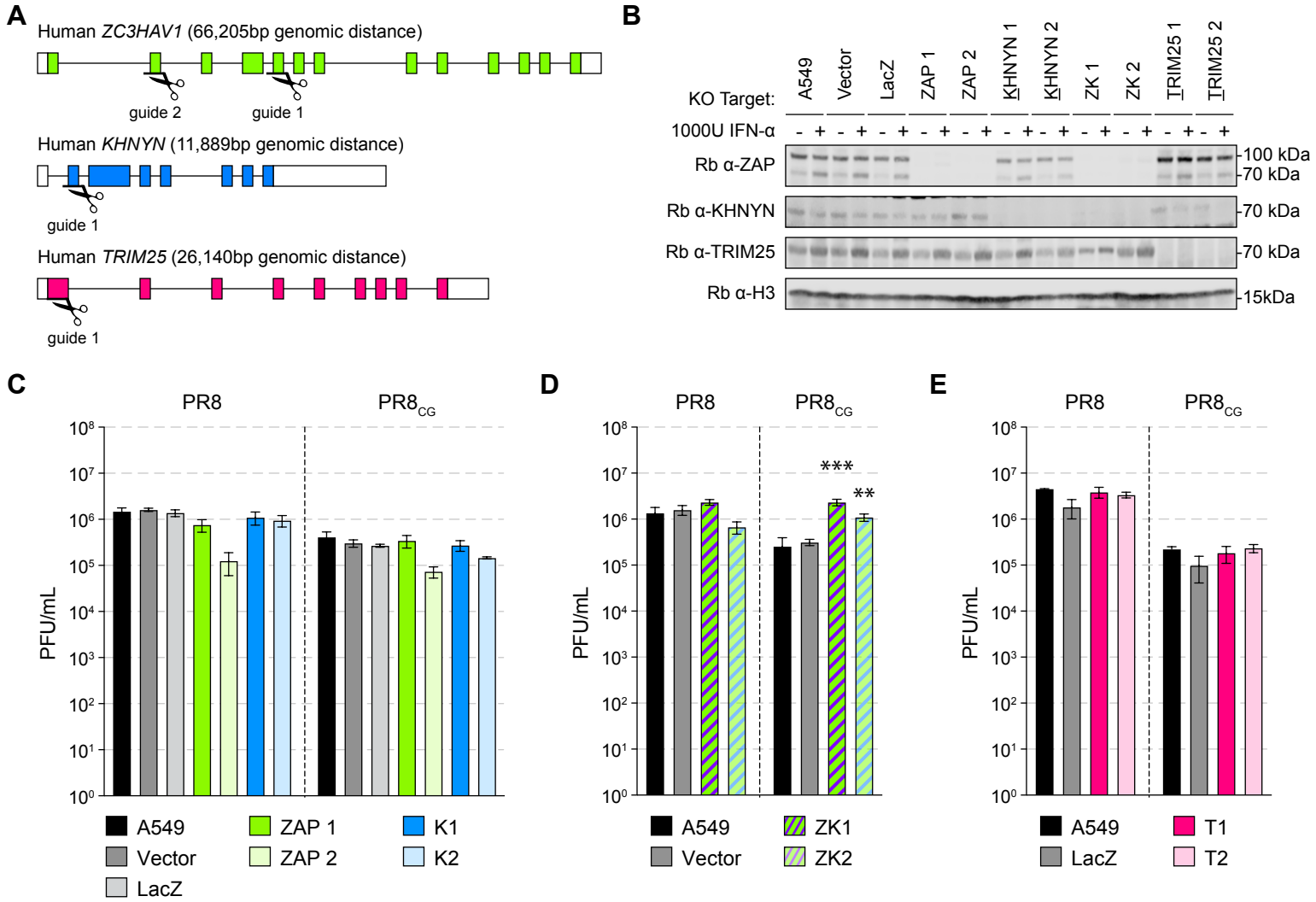

A
Human ZC3HAV1 (66,205bp genomic distance)
guide 2
guide 1
Human KHNYN (11,889bp genomic distance)
guide 1
Human TRIM25 (26,140bp genomic distance)
guide 1
B
KO Target:
A549
Vector
LacZ
ZAP 1
ZAP 2
KHNYN 1
KHNYN 2
ZK 1
ZK 2
TRIM25 1
TRIM25 2
1000U IFN-α
Rb α-ZAP
Rb α-KHNYN
Rb α-TRIM25
Rb α-H3
100 kDa
70 kDa
70 kDa
70 kDa
15kDa
C
PR8
PR8CG
PFU/mL
A549
Vector
LacZ
ZAP 1
ZAP 2
K1
K2
D
PR8
PR8CG
***
**
PFU/mL
A549
Vector
ZK1
ZK2
E
PR8
PR8CG
PFU/mL
A549
LacZ
T1
T2

Figure 4

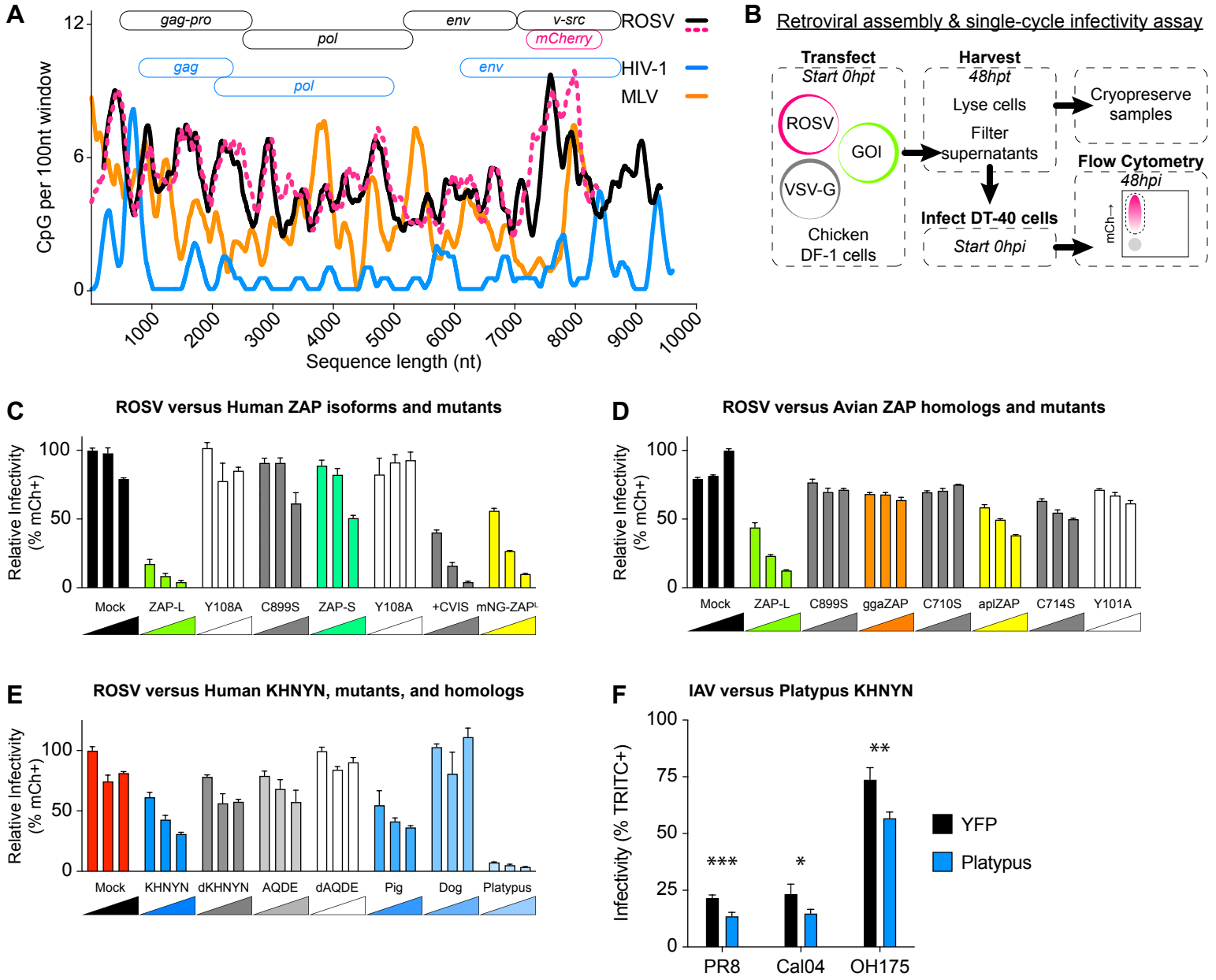

A
gag-pro
pol
env
v-src
mCherry
ROSV
gag
pol
env
HIV-1
MLV
CpG per 100nt window
Sequence length (nt)
B
Retroviral assembly & single-cycle infectivity assay
Transfect
Start 0hpt
ROSV
GOI
VSV-G
Chicken DF-1 cells
Harvest
48hpt
Lyse cells
Filter supernatants
Cryopreserve samples
Infect DT-40 cells
Start 0hpi
Flow Cytometry
48hpi
mCh→
C
ROSV versus Human ZAP isoforms and mutants
Relative Infectivity (% mCh+)
Mock
ZAP-L
Y108A
C899S
ZAP-S
Y108A
+CVIS
mNG-ZAPL
D
ROSV versus Avian ZAP homologs and mutants
Relative Infectivity (% mCh+)
Mock
ZAP-L
C899S
ggaZAP
C710S
aplZAP
C714S
Y101A
E
ROSV versus Human KHNYN, mutants, and homologs
Relative Infectivity (% mCh+)
Mock
KHNYN
dKHNYN
AQDE
dAQDE
Pig
Dog
Platypus
F
IAV versus Platypus KHNYN
Infectivity (% TRITC+)
YFP
Platypus
PR8
Cal04
OH175

Figure S1

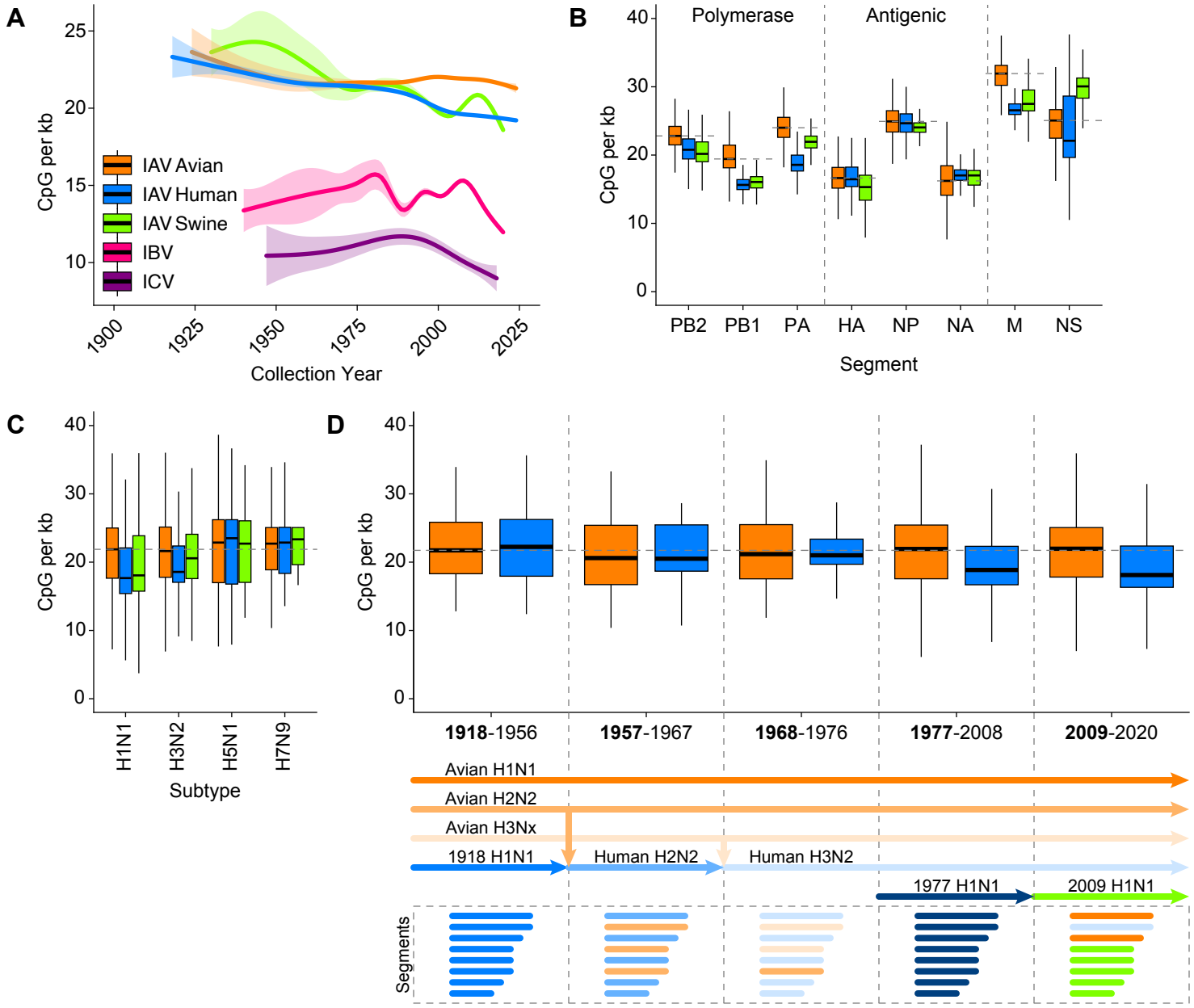

A
CpG per kb
Collection Year
IAV Avian
IAV Human
IAV Swine
IBV
ICV
B
Polymerase
Antigenic
CpG per kb
PB2
PB1
PA
HA
NP
NA
M
NS
Segment
C
CpG per kb
H1N1
H3N2
H5N1
H7N9
Subtype
D
CpG per kb
1918-1956
1957-1967
1968-1976
1977-2008
2009-2020
Avian H1N1
Avian H2N2
Avian H3Nx
1918 H1N1
Human H2N2
Human H3N2
1977 H1N1
2009 H1N1
Segments

Figure S2

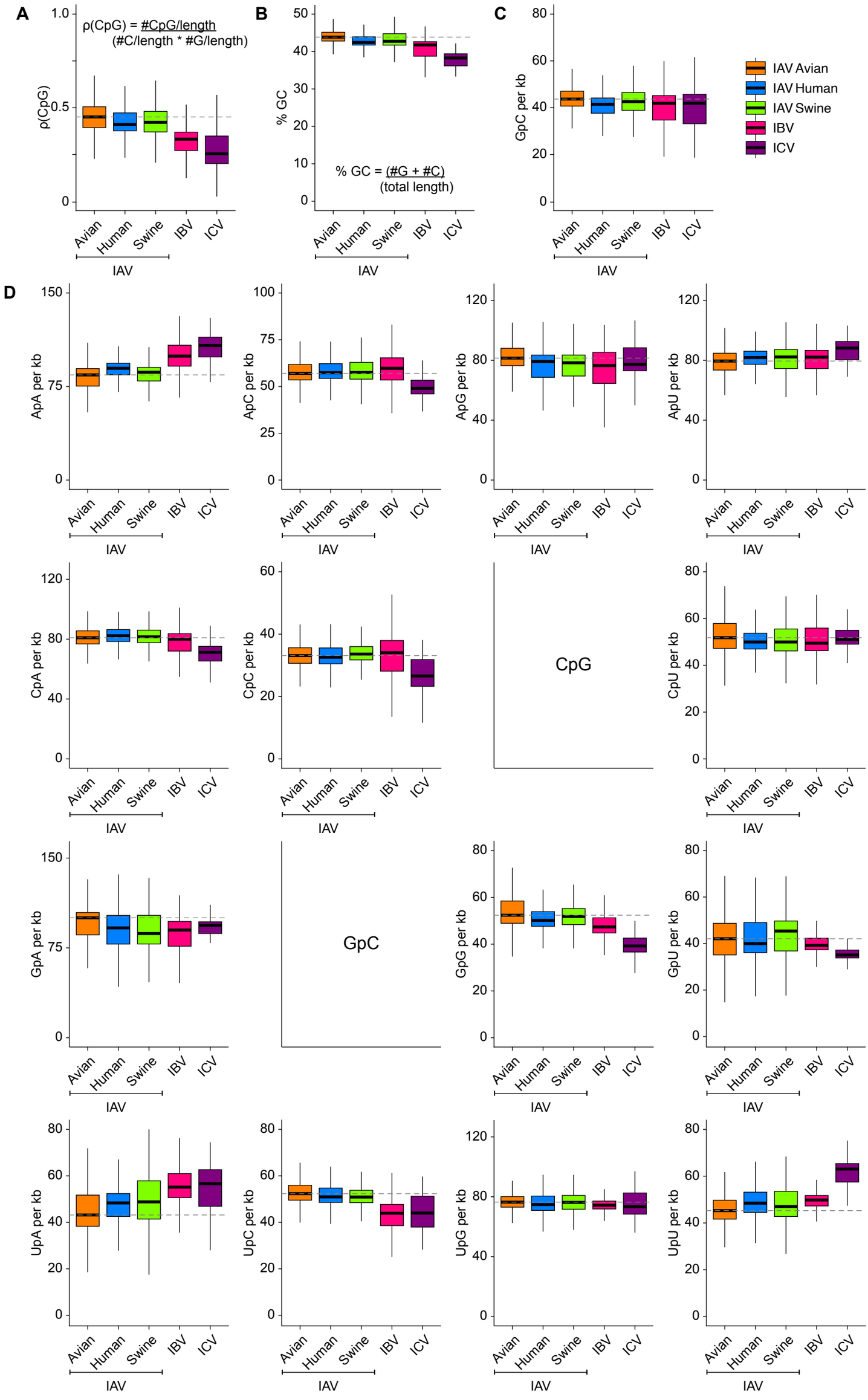

A
ρ(CpG) = #CpG/length
(#C/length * #G/length)
ρ(CpG)
B
% GC
% GC = (#G + #C)
(total length)
C
GpC per kb
IAV Avian
IAV Human
IAV Swine
IBV
ICV
D
ApA per kb
ApC per kb
ApG per kb
ApU per kb
CpA per kb
CpC per kb
CpG
CpU per kb
GpA per kb
GpC
GpG per kb
GpU per kb
UpA per kb
UpC per kb
UpG per kb
UpU per kb
Avian
Human
Swine
IBV
ICV
IAV

Figure S3

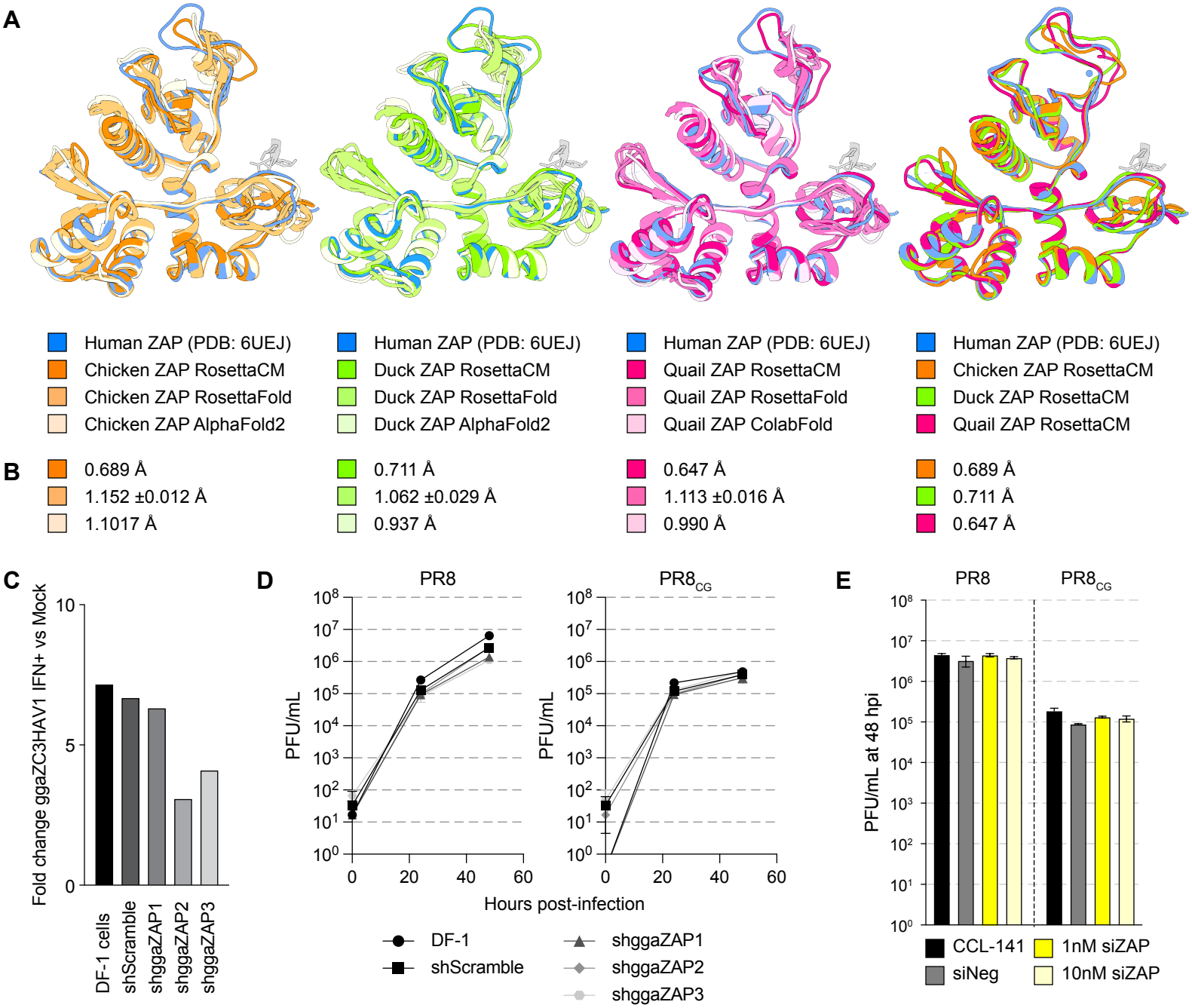

A
Human ZAP (PDB: 6UEJ)
Chicken ZAP RosettaCM
Chicken ZAP RosettaFold
Chicken ZAP AlphaFold2
Human ZAP (PDB: 6UEJ)
Duck ZAP RosettaCM
Duck ZAP RosettaFold
Duck ZAP AlphaFold2
Human ZAP (PDB: 6UEJ)
Quail ZAP RosettaCM
Quail ZAP RosettaFold
Quail ZAP ColabFold
Human ZAP (PDB: 6UEJ)
Chicken ZAP RosettaCM
Duck ZAP RosettaCM
Quail ZAP RosettaCM
B
0.689 Å
1.152 ±0.012 Å
1.1017 Å
0.711 Å
1.062 ±0.029 Å
0.937 Å
0.647 Å
1.113 ±0.016 Å
0.990 Å
0.689 Å
0.711 Å
0.647 Å
C
Fold change ggaZC3HAV1 IFN+ vs Mock
DF-1 cells
shScramble
shggaZAP1
shggaZAP2
shggaZAP3
D
PR8
PR8CG
PFU/mL
Hours post-infection
DF-1
shScramble
shggaZAP1
shggaZAP2
shggaZAP3
E
PR8
PR8CG
PFU/mL at 48 hpi
CCL-141
siNeg
1nM siZAP
10nM siZAP

Figure S4

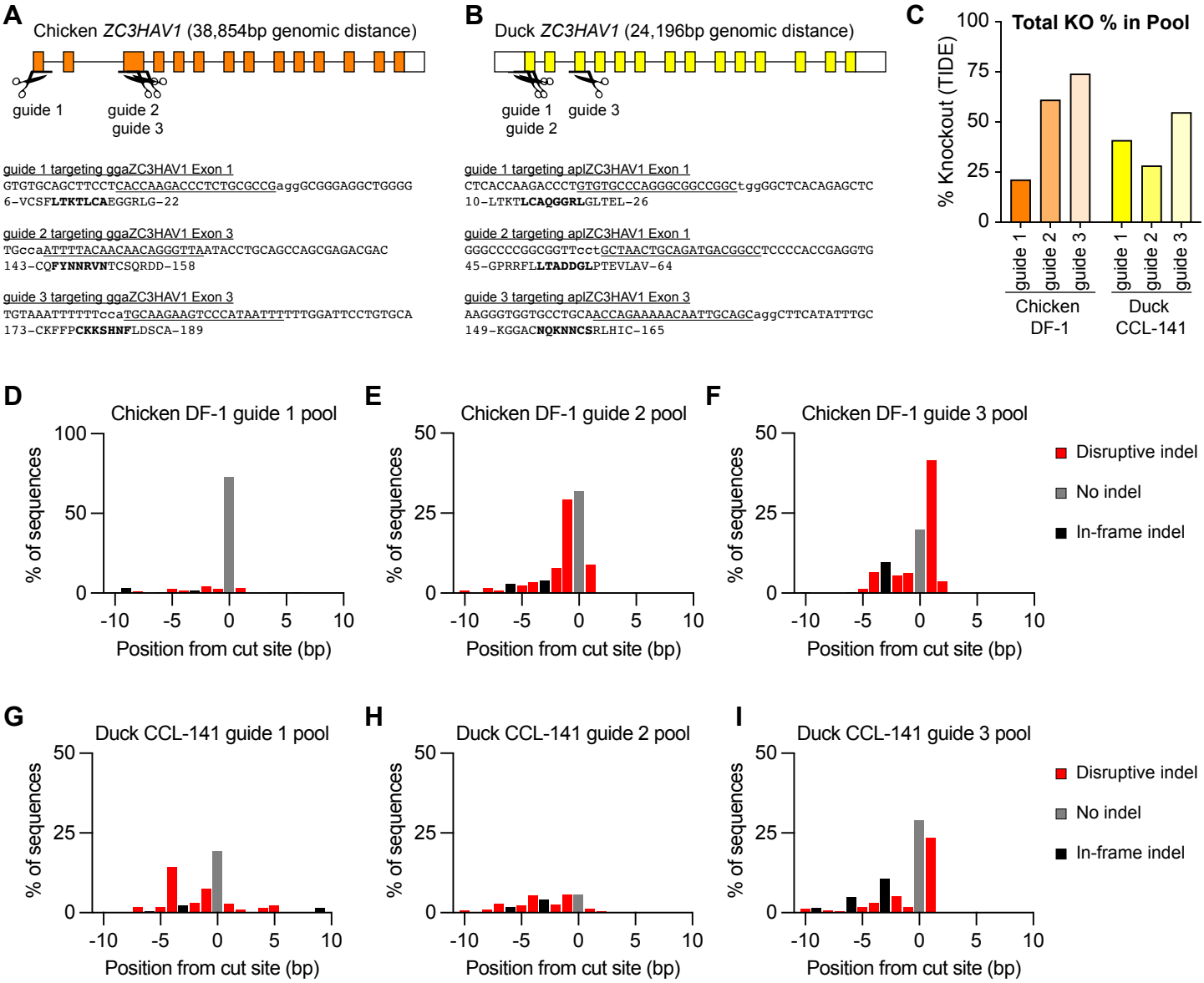

A
Chicken ZC3HAV1 (38,854bp genomic distance)
guide 1
guide 2
guide 3
guide 1 targeting ggaZC3HAV1 Exon 1
GTGTGCAGCTTCCTCACCAAGACCCTCTGCGCCGaggGCGGGAGGCTGGGG
6-VCSFLTKTLCAEGGRLG-22
guide 2 targeting ggaZC3HAV1 Exon 3
TGccaATTTTACAACAACAGGGTTAATACCTGCAGCCAGCGAGACGAC
143-CQFYNNRVNTCSQRDD-158
guide 3 targeting ggaZC3HAV1 Exon 3
TGTAAATTTTTTccaTGCAAGAAGTCCCATAATTTTTTGGATTCCTGTGCA
173-CKFFPCKKSHNFLDSCA-189
B
Duck ZC3HAV1 (24,196bp genomic distance)
guide 1
guide 2
guide 3
guide 1 targeting aplZC3HAV1 Exon 1
CTCACCAAGACCCTGTGTGCCCAGGGCGGCCGGCtggGGCTCACAGAGCTC
10-LTKTLCAQGGRLGLTEL-26
guide 2 targeting aplZC3HAV1 Exon 1
GGGCCCCGGCGGTTcctGCTAACTGCAGATGACGGCCTCCCCACCGAGGTG
45-GPRRFLLTADDGLPTEVLAV-64
guide 3 targeting aplZC3HAV1 Exon 3
AAGGGTGGTGCCTGCAACCAGAAAAACAATTGCAGCaggCTTCATATTTGC
149-KGGACNQKNNCSRLHIC-165
C
Total KO % in Pool
% Knockout (TIDE)
100
75
50
25
0
guide 1
guide 2
guide 3
guide 1
guide 2
guide 3
Chicken DF-1
Duck CCL-141
D
Chicken DF-1 guide 1 pool
E
Chicken DF-1 guide 2 pool
F
Chicken DF-1 guide 3 pool
% of sequences
Position from cut site (bp)
Disruptive indel
No indel
In-frame indel
G
Duck CCL-141 guide 1 pool
H
Duck CCL-141 guide 2 pool
I
Duck CCL-141 guide 3 pool

Figure S5

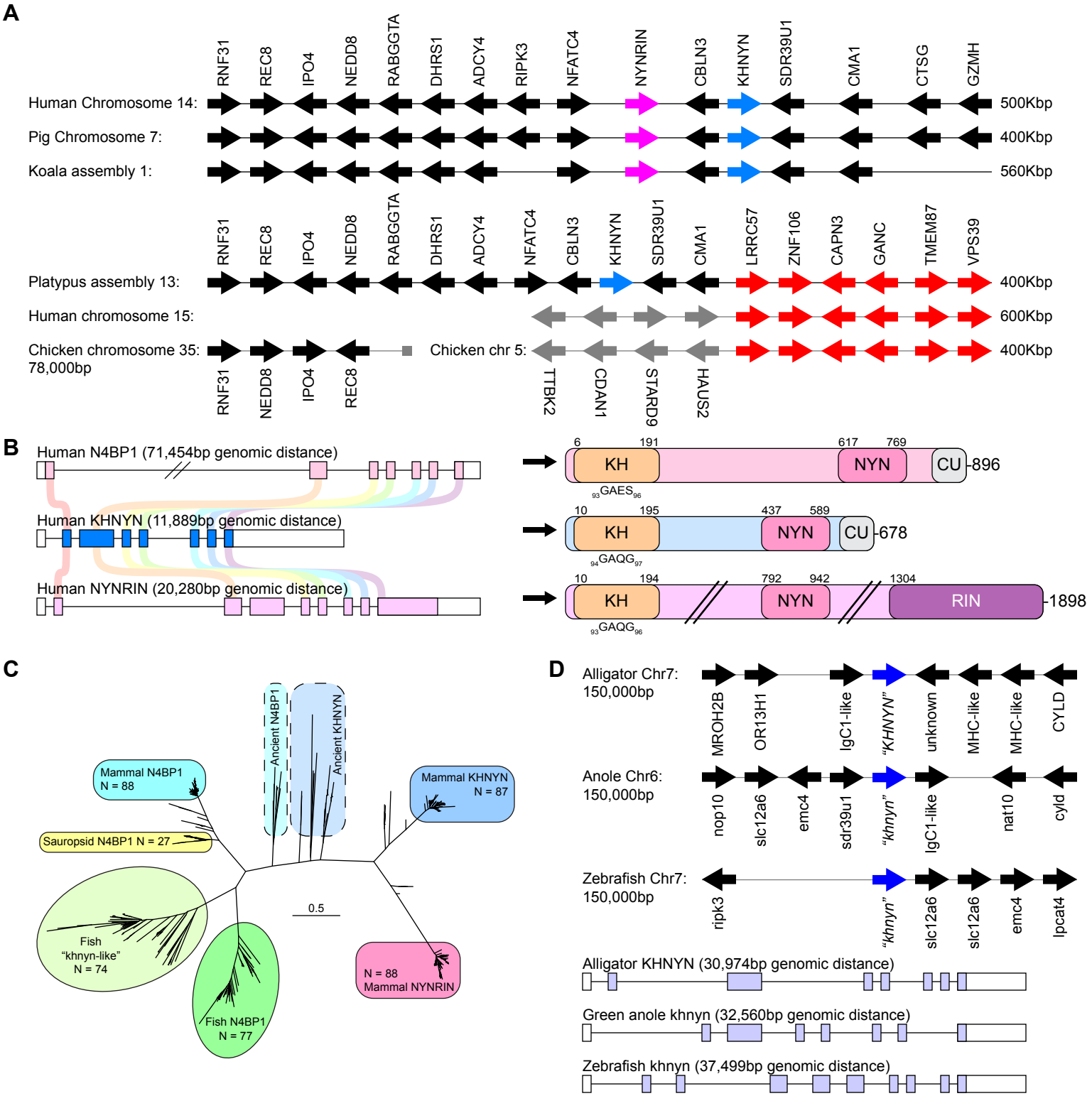

A
RNF31
REC8
IPO4
NEDD8
RABGGTA
DHRS1
ADCY4
RIPK3
NFATC4
NYNRIN
CBLN3
KHNYN
SDR39U1
CMA1
CTSG
GZMH
Human Chromosome 14:
500Kbp
Pig Chromosome 7:
400Kbp
Koala assembly 1:
560Kbp
RNF31
REC8
IPO4
NEDD8
RABGGTA
DHRS1
ADCY4
NFATC4
CBLN3
KHNYN
SDR39U1
CMA1
LRRC57
ZNF106
CAPN3
GANC
TMEM87
VPS39
Platypus assembly 13:
400Kbp
Human chromosome 15:
600Kbp
Chicken chromosome 35:
78,000bp
Chicken chr 5:
400Kbp
RNF31
NEDD8
IPO4
REC8
TTBK2
CDAN1
STARD9
HAUS2
B
Human N4BP1 (71,454bp genomic distance)
Human KHNYN (11,889bp genomic distance)
Human NYNRIN (20,280bp genomic distance)
6
191
617
769
KH
NYN
CU
896
93GAES96
10
195
437
589
KH
NYN
CU
678
94GAQG97
10
194
792
942
1304
KH
NYN
RIN
1898
93GAQG96
C
Ancient N4BP1
Ancient KHNYN
Mammal N4BP1
N = 88
Mammal KHNYN
N = 87
Sauropsid N4BP1 N = 27
0.5
Fish
"khnyn-like"
N = 74
Fish N4BP1
N = 77
N = 88
Mammal NYNRIN
D
Alligator Chr7:
150,000bp
MROH2B
OR13H1
IgC1-like
"KHNYN"
unknown
MHC-like
MHC-like
CYLD
Anole Chr6:
150,000bp
nop10
slc12a6
emc4
sdr39u1
"khnyn"
IgC1-like
nat10
cyld
Zebrafish Chr7:
150,000bp
ripk3
"khnyn"
slc12a6
slc12a6
emc4
lpcat4
Alligator KHNYN (30,974bp genomic distance)
Green anole khnyn (32,560bp genomic distance)
Zebrafish khnyn (37,499bp genomic distance)
E
30°
15°
Human KHNYN
Pig KHNYN
Dog KHNYN
Platypus KHNYN
GxxG motif (human 94-GAQG-97)
Platypus short loop
(145-GGGALVGD-152)

Figure S6

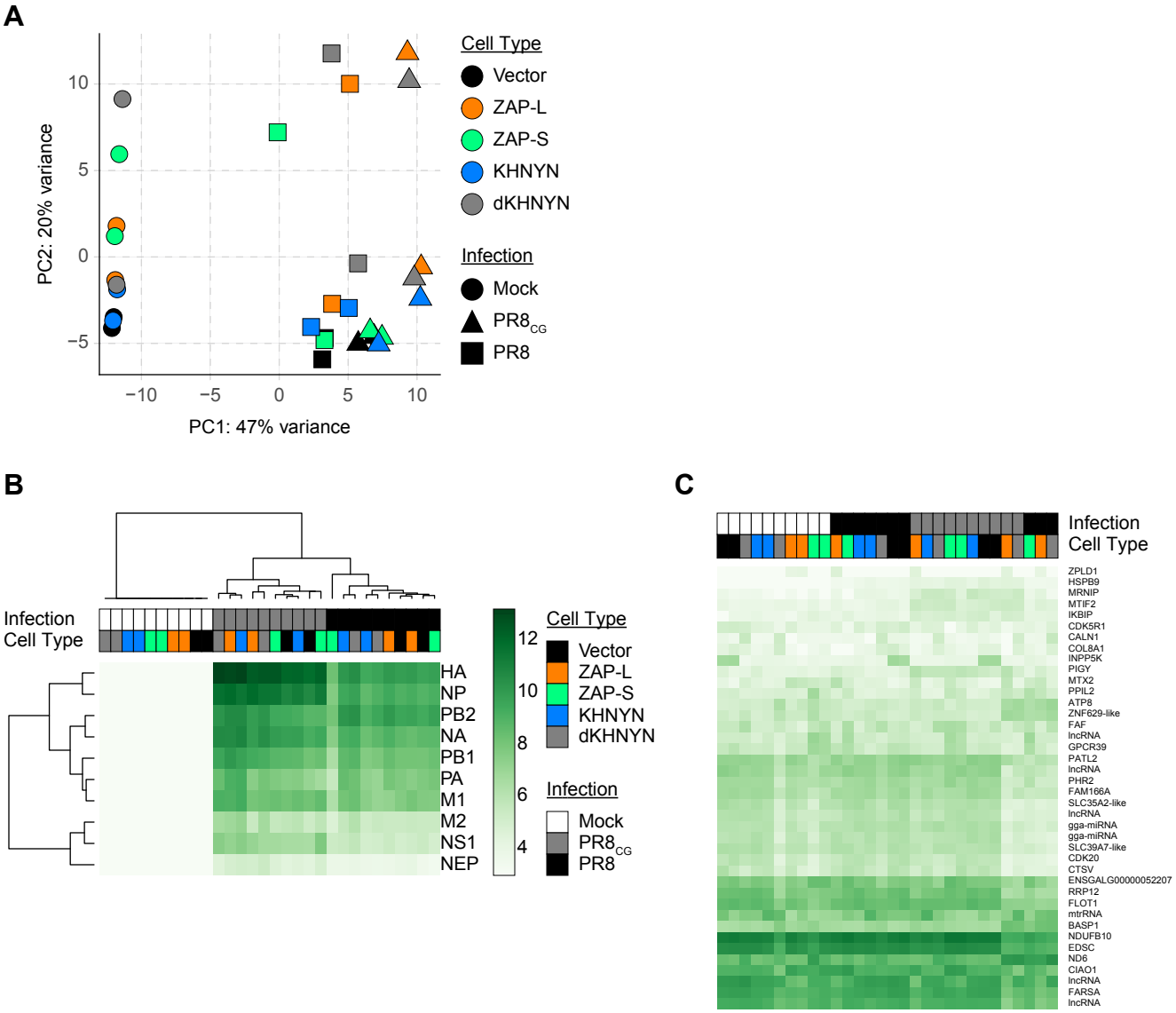
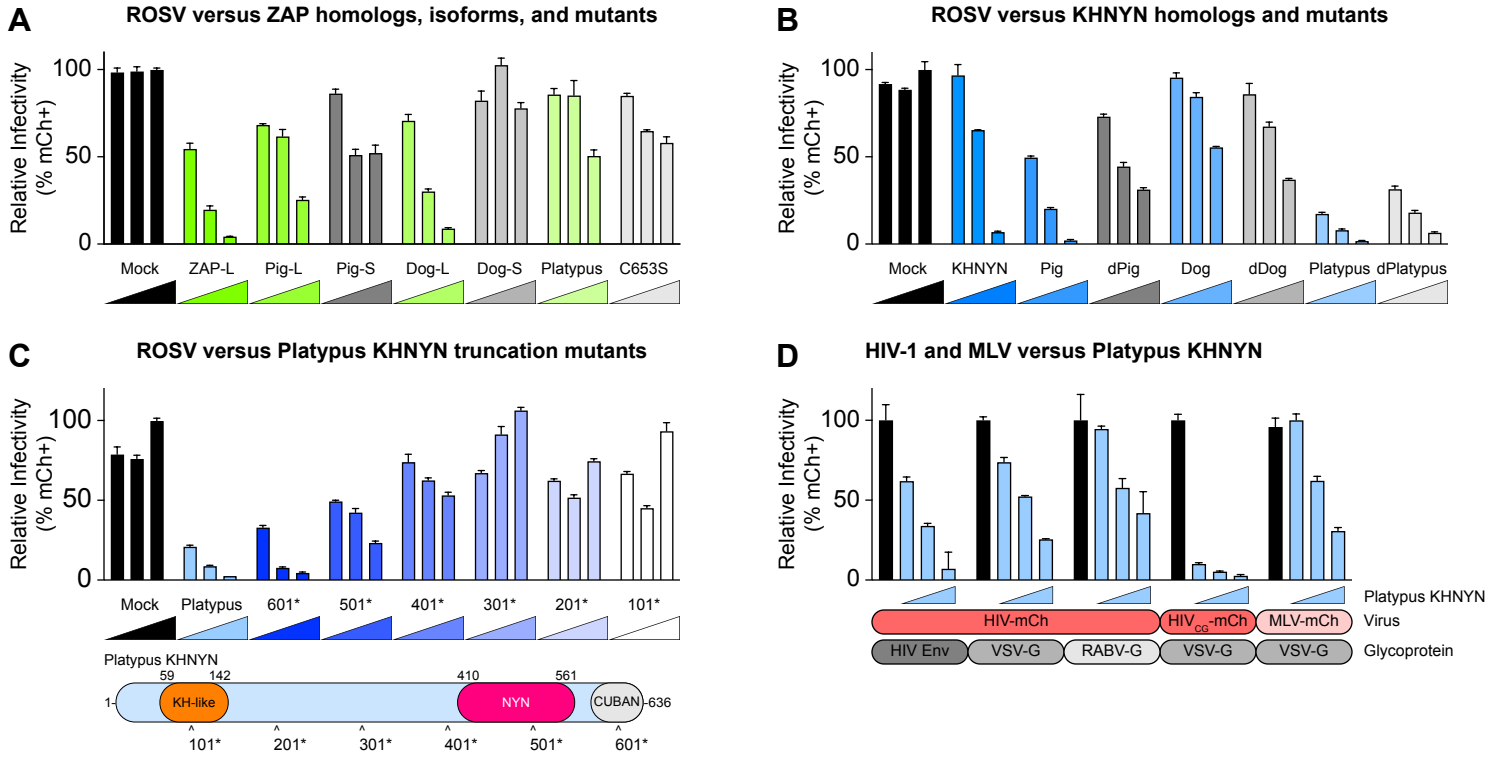

A
PC2: 20% variance
PC1: 47% variance
10
5
0
−5
−10
−5
0
5
10
Cell Type
Vector
ZAP-L
ZAP-S
KHNYN
dKHNYN
Infection
Mock
PR8CG
PR8
B
Infection
Cell Type
HA
NP
PB2
NA
PB1
PA
M1
M2
NS1
NEP
12
10
8
6
4
Cell Type
Vector
ZAP-L
ZAP-S
KHNYN
dKHNYN
Infection
Mock
PR8CG
PR8
C
Infection
Cell Type
ZPLD1
HSPB9
MRNIP
MTIF2
IKBIP
CDK5R1
CALN1
COL8A1
INPP5K
PIGY
MTX2
PPIL2
ATP8
ZNF629-like
FAF
lncRNA
GPCR39
PATL2
lncRNA
PHR2
FAM166A
SLC35A2-like
lncRNA
gga-miRNA
gga-miRNA
SLC39A7-like
CDK20
CTSV
ENSGALG00000052207
RRP12
FLOT1
mtrRNA
BASP1
NDUFB10
EDSC
ND6
CIAO1
lncRNA
FARSA
lncRNA

Figure S7


A
ROSV versus ZAP homologs, isoforms, and mutants
Relative Infectivity (% mCh+)
100
50
0
Mock
ZAP-L
Pig-L
Pig-S
Dog-L
Dog-S
Platypus
C653S
B
ROSV versus KHNYN homologs and mutants
Relative Infectivity (% mCh+)
100
50
0
Mock
KHNYN
Pig
dPig
Dog
dDog
Platypus
dPlatypus
C
ROSV versus Platypus KHNYN truncation mutants
Relative Infectivity (% mCh+)
100
50
0
Mock
Platypus
601*
501*
401*
301*
201*
101*
Platypus KHNYN
59
142
410
561
1-
KH-like
NYN
CUBAN
-636
101*
201*
301*
401*
501*
601*
D
HIV-1 and MLV versus Platypus KHNYN
Relative Infectivity (% mCh+)
100
50
0
Platypus KHNYN
HIV-mCh
HIV$_{CG}$-mCh
MLV-mCh
Virus
HIV Env
VSV-G
RABV-G
VSV-G
VSV-G
Glycoprotein