CpG dinucleotide enrichment in the influenza A virus genome as a live attenuated vaccine development strategy

Colin P Sharp[1], Beth H Thompson[1], Tessa J Nash[1], Ola Diebold[1], Rute M Pinto[1], Luke Thorley[1], Yao-Tang Lin[1], Samantha Sives[1], Helen Wise[2], Sara Clohisey Hendry[1], Finn Grey[1], Lonneke Vervelde[1], Peter Simmonds[3], Paul Digard[1], Eleanor R Gaunt[1*]

*Corresponding author

[1] The Roslin Institute, The University of Edinburgh, Easter Bush Campus, Midlothian, EH25 9RG

[2] Royal Infirmary of Edinburgh, NHS Lothian, Edinburgh, EH16 4SA

[3] Nuffield Department of Medicine, University of Oxford, South Parks Road, Oxford, OX1 3SY

ABSTRACT

Synonymous recoding of RNA virus genomes is a promising approach for generating attenuated viruses to use as vaccines. Problematically, recoding typically hinders virus growth, but this may be rectified using CpG dinucleotide enrichment. CpGs are recognised by cellular zinc-finger antiviral protein (ZAP), and so in principle, removing ZAP sensing from a virus propagation system will reverse attenuation of a CpG-enriched virus, enabling high titre yield of a vaccine virus. We tested this using a vaccine strain of influenza A virus (IAV) engineered for increased CpG content in genome segment 1. Virus attenuation was mediated by the short isoform of ZAP, correlated with the number of CpGs added, and was enacted via turnover of viral transcripts. The CpG-enriched virus was strongly attenuated in mice, yet conveyed protection from a potentially lethal challenge dose of wildtype virus. Importantly for vaccine development, CpG-enriched viruses were genetically stable during serial passage. Unexpectedly, in both MDCK cells and embryonated hens' eggs that are used to propagate live attenuated influenza vaccines, the ZAP-sensitive virus was fully replication competent. Thus, ZAP-sensitive CpG enriched viruses that are defective in human systems can yield high titre in vaccine propagation systems, providing a realistic, economically viable platform to augment existing live attenuated vaccines.

AUTHOR SUMMARY

CpG dinucleotides are under-represented in vertebrate genomes, wherein cytosines in the CpG conformation are methylated to regulate transcription. Methylated cytosines are prone to deamination, resulting in TpG dinucleotides replacing CpGs. The resultant CpG suppression has provided a route by which vertebrate cells can recognise RNA from invading pathogens, using cellular Zinc-finger Antiviral Protein (ZAP) as a CpG sensor. Vertebrate-infecting RNA viruses also genomically suppress CpGs, and it is believed that this is an evolved trait to evade detection by ZAP. Here, we engineered an influenza A virus (IAV) with elevated CpG content and characterised how this impacts viral replication. CpG addition resulted in viral attenuation, mediated by ZAP activity. CpG suppression is conserved in dog and chicken genomes (relevant for live attenuated IAV vaccine propagation), and it is logical to predict that ZAP-mediated CpG sensing would also be conserved in these species. However, when we propagated ZAP-sensitive IAV in cognate culture systems, we saw no replication defect. This

39 unexpected result raises questions about why viruses infecting these species suppress CpG in their
40 genomes, and importantly delivers a new, tractable approach to augment rational live attenuated IAV
41 vaccine design.

42 **INTRODUCTION**

43 Single stranded RNA (ssRNA) viruses typically under-represent CpG dinucleotides in their genomes
44 (1-3). CpG dinucleotides in viral transcripts may be bound by zinc-finger antiviral protein (ZAP) (4),
45 which causes an antiviral state in the infected cell. This provides a basis for observations from several
46 studies showing that engineering viral genomes enriched for CpG yields replication-defective viruses
47 (reviewed in (5)).

48 Two major isoforms of ZAP, ZAPS (short) and ZAPL (long) are produced by alternative splicing (6), with
49 a common N-terminal region that contains four zinc-finger motifs with RNA-binding function (7-9). While
50 CpG-binding activity was originally found to be specific to ZAPS (4), CpG-specific antiviral activity has
51 since also been demonstrated for ZAPL (10). Therefore, a consensus mechanism for CpG sensing by
52 ZAP remains unresolved. ZAPS is an interferon-stimulated gene (ISG) (9,11,12), whereas ZAPL is
53 constitutively expressed (6). For its antiviral function, ZAPS must interact with its binding partner, the
54 RNA-binding E3 ubiquitin ligase-like tripartite motif containing protein 25 (TRIM25) (13,14). The current
55 understanding is that ZAPS binds CpG motifs on viral RNA, and upon subsequent binding by TRIM25
56 and catalytic activation, a downstream cascade leads to the induction of type I and type II interferons
57 (IFN) (11). ZAPS may also directly bind RNA leading to exonuclease-mediated degradation (15), and
58 so the antiviral activity of ZAPS appears multi-faceted (reviewed in (5)). A newly discovered ZAPS
59 binding partner, KHNYN, has so far only been found to be required for ZAP-mediated antiviral activity
60 during lentivirus infections (16,17). While ZAPS has so far been more extensively studied, interactions
61 of ZAPL with TRIM25 and KHNYN have been shown to be more efficient than those of ZAPS (10).
62 CpG-specific binding by ZAP is likely an evolutionarily conserved trait across tetrapods that is not
63 retained in avian species (18); nevertheless, chickens express ZAPL but not the shorter isoform (19).

64 Due to the replication defects imparted by CpG introduction into viral genomes, CpG enrichment has
65 widely been proposed as a potential strategy for the development of live attenuated vaccines. However,
66 if viral growth is impaired, then the large-scale virus propagation required for vaccine manufacturing
67 (20) is not viable. The identification of ZAP as a CpG sensor (4) delivered the possibility that CpG-
68 enriched viruses could be grown in ZAP knockout (KO) systems, thus restoring viral titre in an artificial
69 system, with the replication defect maintained in a vaccine recipient (5). This principle has been

70 validated using enterovirus-71 as an exemplar (21), although there is currently no licensed vaccine
71 targeting this virus.

72 Live attenuated vaccines for influenza A viruses (IAV) are used in children in the USA and Europe
73 (22,23), with up to 14 million doses delivered annually (24). IAV vaccines represent an ideal target for
74 augmentation by incorporating additional CpGs, as this approach has the potential to overcome some
75 of the significant challenges arising during vaccine manufacture. IAV live vaccines are propagated in
76 either embryonated hens' eggs or Madin-Darby canine kidney (MDCK) cells, and both IAV vaccine
77 propagation systems can be subjected to genome editing (25,26) to improve virus yield. For egg based
78 vaccines, unpredictable supply and ethical challenges mean that there is a consensus about the need
79 to move towards cell culture generated vaccines (27,28). However, MDCK based vaccines are more
80 expensive and large-scale manufacture will require establishment of new production pipelines (29,30);
81 therefore, this transition is not straightforward and egg-based vaccines will be commercialised for the
82 foreseeable future.

83 Recoding of the IAV genome (i.e., synonymous alterations of the genetic sequence) is achieved using
84 the facile reverse genetics system (31,32), in which the 8 segments of negative sense RNA genome
85 are encoded on plasmids that upon co-transfection into permissive cells yield infectious virus. The IAV
86 genome has been thoroughly characterised, with well-defined packaging signals (33-35) and gene
87 annotation (including accessory proteins, reviewed in (36)), providing sufficient clarity about which
88 genome regions may be suitable for recoding and those that must remain unaltered.

89 To determine if CpG enrichment can enhance live attenuated vaccines, we used an IAV vaccine strain
90 to test the effects of adding CpGs to segment 1, which encodes the polymerase protein PB2 and the
91 accessory protein PB2-S1. This CpG enrichment resulted in a virus that was attenuated in human cell
92 systems, with wildtype fitness restored by ZAP KO. The CpG-high virus was genetically stable in ZAP-
93 positive cells, and showed no growth defects in embryonated hens' eggs or MDCK cells. Therefore, we
94 have provided a pathway to develop CpG-enriched live attenuated IAV that can be propagated to high
95 yields using standard approaches.

96

97 **RESULTS**

98 **Addition of CpG dinucleotides to the IAV genome causes ZAPS-dependent replication defects.**

99 As ZAP recognises CpG-enriched transcripts (4) and the IAV genome naturally suppresses CpG, we
100 investigated whether adding CpGs to the IAV genome resulted in viral replication deficiencies and
101 whether attenuation could be offset in ZAP -/- cells. We engineered a CpG-high ('CpGH') PR8 (A/Puerto
102 Rico/8/1934 H1N1) virus through synonymous addition of 126 CpGs into segment 1, including
103 compensatory mutations to maintain individual base frequencies across the recoded regions.
104 Additionally, the same viral genomic region was synonymously recoded by codon rearrangement to
105 create a control ('CDLR') virus. This allows the detection of any attenuation caused by disruption of e.g.
106 RNA structures or alternative open reading frames. The panel of recoded viruses is summarised (**Table**
107 **1**).

108 The CpGH IAV displayed a significant ~1.5-$\log_{10}$ defect in production of infectious virus progeny in A549
109 cells, whereas the CDLR virus was fully replication competent (**Fig. 1A**). In paired ZAP -/- cells (a kind
110 gift from the laboratory of Prof. Sam Wilson, MRC-University of Glasgow Centre for Virus Research,
111 UK; validated in **Fig. S1A**) (37), the CpGH virus replicated as well as wild type (WT) PR8 (**Fig. 1A**),
112 suggesting that the replication defect in the CpGH virus is ZAP dependent.

113 ZAP binding partner TRIM25 is required for ZAP antiviral activity (13,14) and so replication of WT and
114 engineered viruses was measured in TRIM25-/- cells (a kind gift from the laboratory of Prof. Gracjan
115 Michlewski, International Institute of Molecular and Cell Biology, Warsaw ,Poland; validated in **Fig. S1B**)
116 (38). In WT HEK293 cells, CpGH IAV was again significantly replication-impaired. As with ZAP KO,
117 TRIM25 KO also resulted in rescued replication of CpGH virus (**Fig. 1B**).

118 To determine whether the observed phenotypes could be recapitulated in primary cells, siRNA
119 knockdown of ZAP or TRIM25 was applied in Normal Human Dermal Fibroblasts (NHDF; primary
120 human foreskin fibroblasts) as previously described (39) (**Fig. 1C**). While ZAP knockdown recapitulated
121 rescue of CpGH virus fitness, TRIM25 knockdown did not (**Fig. 1D**). TRIM25 knockdown was
122 incomplete (Fig. 1C), and so this may indicate that sub-physiological levels of TRIM25 are sufficient for
123 its role in CpG recognition.

124 KHNYN is a recently identified ZAP-binding partner (16) and so its impact on replication of the virus
125 panel was also tested (cells were a kind gift from the laboratory of Dr Chad Swanson, King's College
126 London, UK; **Fig. S1C**). As in Fig. 1A, a significant defect in the replication of CpGH virus was seen in

127 WT A549 cells. In paired KHNYN-/- cells, CpGH virus titres remained significantly lower than those of
128 WT/CDLR PR8. However, the fold difference in titre between WT PR8 and CpGH was significantly
129 reduced, indicating a partial rescue of CpGH virus replication in the absence of KHNYN and suggesting
130 that KHNYN is required for optimal CpG sensing (**Fig. 1E**).

131 To determine whether we could reduce the number of CpGs in the CpGH construct and maintain the
132 same level of replication inhibition, the recoded region of segment 1 was split so that either only the 5'
133 two thirds of recoding were present (as viewed in positive polarity; '5'-CpGH'), or the 3' third ('3'-CpGH';
134 **Fig. S1D-E**). As an alternative, CpGs were introduced across the entire recoded region at a lower
135 density than CpGH but at the same frequency as 3'-CpGH ('CpG-M'), to test whether CpG frequency
136 or density were of primary importance. This virus panel was used to infect A549 cells. As previously,
137 the CpGH virus was significantly attenuated. However, reduction in CpG frequency or density
138 diminished this attenuation, with a dose-dependent effect observed (**Fig. 1F**). No attenuation was
139 apparent for either 3'-CpGH or CpG-M, which had the same number of CpGs introduced at high and
140 low density respectively.

141 To determine whether ZAP sensitivity of the CpGH virus was mediated by a specific ZAP isoform,
142 expression of ZAPL or ZAPS was reconstituted in A549 ZAP -/- cells using pSCRPSY lentiviral vectors
143 (**Fig. 1G**, **S2**). In ZAP-/- cells transduced with empty pSCRPSY no differences in replication were
144 observed between WT PR8, CDLR and CpGH viruses, indicating that the lentivirus vector did not impair
145 CpGH virus replication (**Fig. 1H**). When ZAPL expression was reconstituted, all viruses showed a
146 modest increase in replication, with CpGH virus displaying similar titres to those of WT PR8. However,
147 in cells overexpressing ZAPS, all viruses in the panel showed a replication defect. Nevertheless, CpGH
148 virus titres were significantly lower than WT PR8, indicating that ZAPS is the CpG sensor for CpGH
149 IAV.

150 **CpG enrichment does not impair IAV virion assembly.**

151 The mutations introduced into the PR8 genome were designed to avoid classically described packaging
152 signals present at the termini of viral genome segments (33,40). However, recent work has highlighted
153 the potential for the middle regions of different segments to interact during virion assembly (35).
154 Therefore, we questioned if the mutations made in segment 1 disrupted viral genome packaging. Firstly,
155 genome copy:PFU ratios were determined for the panel of virus stocks. RNA copy numbers for both

156 segment 1 and segment 5 were analysed. A previously characterised packaging-defective mutant,
157 '4c6c' (33,41) was included as a control. Relative to WT PR8, no significant impact on RNA:PFU ratios
158 was observed for either the CDLR or CpGH viruses (**Fig. 2A**).

159 Secondly, electrophoretic separation of RNA extracted from purified virus particles was performed, to
160 visually determine whether synonymous recoding altered the amount of RNA packaged for each
161 segment. For each virus, equal numbers of infectious particles were pelleted through sucrose, then
162 RNA was extracted from viral pellets and separated by urea-PAGE. As expected, the PR8 WT virus
163 yielded 7 bands corresponding with 8 segments of RNA (segments 1 and 2 co-migrate as they are the
164 same size; **Fig. 2B**). The CDLR control band densities were visually similar to those of the WT virus.
165 For CpGH virus, while it did not appear that there were any differences in band densities (in keeping
166 with the genome copy:PFU ratios), an apparent retardation in the electrophoretic mobility of CpG-
167 enriched segment 1 was observed.

168 As the CpGH construct was compositionally corrected so that all bases are equally represented, the
169 electric charges should be identical for all the segment 1 RNAs. RNAs were run on denaturing gels,
170 and so RNA structure should also not influence mobility. To determine whether a post-transcriptional
171 modification might explain the diminished electrophoretic mobility of the CpG-enriched RNA, segment
172 1 transcripts were synthesised by *in vitro* transcription, so that no cellular-derived post-transcriptional
173 modifications could occur. When equimolar CDLR and CpGH constructs were mixed together and
174 subjected to electrophoretic separation, their different migration phenotypes was maintained (**Fig. 2C**).
175 Most likely, CpG introduction has altered secondary structures in the CpG-enriched RNA transcripts
176 that survived denaturation and retarded electrophoretic migration. Nevertheless, we saw no evidence
177 that increasing CpG frequencies deleteriously affected vRNA packaging.

178 **CpG enrichment does not impact methylation status of IAV vRNAs.**

179 Methylation of cytosines in CpG dinucleotides of DNA regulates transcription (42), and cytosine
180 methylation of RNAs has been described (43,44). We therefore hypothesised that adding CpG sites to
181 RNA may increase the frequency of cytosine-associated methylation. DNA-encoded cytosines in CpG
182 conformations are prone to spontaneous deamination, creating a cytosine to thymine nucleotide
183 substitution (45), and if the same were true for CpGs in RNA, CpG enrichment of viral RNA genomes
184 could result in hypermutation. Viral genomic RNA was purified from virions propagated in embryonated

185 hens' eggs, and bisulphite conversion reactions, in which cytosines in purified RNA are converted to
186 uracil unless protected from conversion by a methyl group, were performed. Converted RNAs were
187 then sequenced and compared to non-converted sequences to identify potential methylation sites. A
188 strongly methylated site at position 3256 of chicken 28S ribosomal RNA (GenBank accession #
189 XR_003078040.1) served as a positive control. Methylation frequencies of 97.0 - 98.8% were detected
190 at this site across all samples, indicating that the protocol worked efficiently (**Fig. 3A**).

191 Deep sequencing reads of the bisulphite converted RNAs were mapped to the *in silico* converted
192 reference sequences, which yielded maximal methylation signals of ~13% across all nucleotide
193 positions (**Fig. 3B-D;** PR8, CDLR and CpGH respectively). However, in WT PR8, three obvious peaks
194 in methylation signal were identified, each occurring over regions of ~50 bases, that were evident in up
195 to approximately 10% of sequences (**Fig. 3B**). The first of these clusters was identified in a region that
196 was not mutated in CDLR or CpGH viruses and the signal was maintained across all three viruses,
197 indicating a biologically reproducible phenomenon. The second two clusters at nucleotide positions
198 ~800 and ~1400 were not retained in the mutant viruses that had been recoded in these regions (**Fig.**
199 **3C&D**), indicating that mutation at these sites removed the methylation signals. Unlike the single base
200 methylation in the chicken rRNA, these sites of apparent viral RNA methylation were centred around
201 ~50 nucleotide regions with diminishing signal flanking the peak. This suggests that these clustered
202 methylation signals may be false positives, possibly generated by RNA secondary structures that did
203 not allow RNA unfolding, thus protecting cytosines from base conversion. To test whether the
204 methylation signal was retained in RNAs synthesised in a cell-free system where post-transcriptional
205 modifications cannot occur, *in vitro* transcribed segment 1 RNAs for WT PR8, CDLR and CpGH were
206 generated. These RNAs, along with paired vRNAs, were then subjected to bisulphite conversion, PCR
207 amplification and sequencing across the target region. Methylation signals were retained in amplicons
208 generated from the *in vitro* transcribed RNAs (**Fig. 3E**), confirming that the methylation clusters were
209 artefacts. The methylation signal was higher from this targeted sequencing, most likely because PCR
210 amplification biased the sequencing output. Through whole genome sequencing of bisulphite converted
211 RNA, we found no evidence of high efficiency methylation at any site in the PR8 genome (**Fig. S3**), and
212 CpG enrichment did not introduce methylation sites.

213 **ZAP-mediated inhibition of CpG enriched viruses is independent of the type I IFN response.**

214 ZAPS is a known ISG, and so we tested whether the ZAPS-sensitive CpG-enriched IAV induced IFN
215 (46,47). Supernatants from infected A549 cells were used in a HEK-Blue assay, which gives a
216 colorimetric signal proportional to the levels of IFN-α/β present. As anticipated, by expressing a
217 functional NS1 (48), WT PR8 did not induce detectable type I IFN (**Fig. 4A**). In contrast, a control virus
218 mutated to impair function (R38K41A; (49)) strongly induced IFN signalling. The CDLR control and
219 CpGH viruses also did not induce type I IFN above background signal. This suggests that the
220 attenuation of CpGH virus by ZAP is independent of any type I IFN-stimulating activity by ZAPS (11).

221 It was possible that IAV NS1 suppressed ZAPS activity sufficiently to block ZAP-mediated IFN induction
222 by CpGH IAV. To test this, we introduced mutations into NS1 that have been described to specifically
223 inhibit its anti-ZAP or anti-TRIM25 activity ((50,51); **Table 2**). Introducing these NS1 mutations into the
224 CpGH virus should alleviate NS1-dependent suppression of the ZAP/TRIM25 pathway, and so if NS1
225 does inhibit CpG sensing by ZAP, the defect in replication of the CpGH virus would be exacerbated by
226 these mutations. However, this was not the case. In low MOI infections in A549 cells, viruses with
227 mutations in NS1 at positions 95/99 (purportedly unblocking anti-TRIM25 activity (50)) all had similar
228 titres to those of viruses without NS1 mutations (**Fig. 4B**). For all viruses with mutations in NS1 at
229 positions 96/97 (reportedly alleviating anti-ZAP activity), titres were down by 2-3 $\log_{10}$, indicating a
230 general replication defect consistent with reports indicating these mutations have a deleterious effect
231 on NS1 stability (50,52). Viruses with mutations in NS1 at positions 107-110 (targeting anti-ZAP
232 functionality) had similar titres to those of viruses without NS1 mutations. The replication defect of CpGH
233 IAV relative to WT PR8 was consistent regardless of the NS1 mutations introduced, indicating a lack of
234 specificity of these NS1 sites for ZAP/ TRIM25 regulation.

235 Across the virus panel, CpGH viral protein accumulation (PB2 and NP) was lower than for counterpart
236 WT viruses, reflecting the defective replication of this virus (**Fig. 4C**). WT PR8, CDLR and CpGH viruses
237 all induced ZAPS expression to similar levels (normalised to tubulin).TRIM25 was also consistently
238 expressed across samples.

239 Type I IFN induction by the segment 1-segment 8 mutant virus panel was examined (**Fig. 4D**). For
240 viruses with mutations introduced into NS1 amino acid residues 95 and 99, no significant type I IFN
241 activity was detected in any of the segment 1-segment 8 virus panel. Conversely, when NS1 residues
242 96 and 97 were mutated, all viruses induced IFN to similar levels. NS1 mutations at positions 107-110

243 conveyed detectable IFN signal for the CpGH virus and also for the CDLR virus (though, curiously, not
244 for PR8 WT virus). Therefore, while CpGH IAV is ZAP sensitive (Fig. 1), there was no correlation of
245 ZAP anti-CpG activity with the type I IFN pathway in this system.

246 **CpG enrichment leads to IAV transcript turnover.**

247 ZAP has also been linked with transcript turnover (53), and so we next examined the impact of ZAP on
248 transcription and translation of viral proteins, and on transcript turnover. We first assessed whether viral
249 transcript and/or protein levels were affected by CpG enrichment during virus replication. Time-course
250 experiments were performed at high MOI with samples collected at 1, 2, 4, 6, 8 and 12 hpi for
251 quantification of total viral RNAs and proteins. Compared to WT PR8 and paired CDLR virus, CpGH
252 virus showed lower transcript and protein expression levels for both PB2 and NP (**Fig. 5A-C**). This result
253 was confirmed by minigenome assay, which reconstitutes the viral polymerase in cells in the absence
254 of full infection (**Fig. S4**) (54). As PB2 is a component of the polymerase complex, this did not distinguish
255 between defective transcript and protein production.

256 To test whether segment 1 transcription and/or translation efficiencies were impaired in the absence of
257 cellular machinery (e.g. ZAP), the two processes were decoupled and separately assayed in cell-free
258 conditions. For *in vitro* transcription assays, equimolar amounts of segment 1 PR8, CDLR, and CpGH
259 amplicons were used as input, and the amount of RNA produced was measured up to 90 minutes later.
260 In the absence of cellular factors (and using the RNA-dependent RNA polymerase of bacteriophage
261 rather than that of IAV), no differences in transcript production were seen across the panel (**Fig. 5D**).
262 Next, translational efficiencies were tested by titrating equimolar amounts of RNA in *in vitro* translation
263 assays. No differences in protein production were apparent across the panel (**Fig. 5E**). Thus,
264 synonymous recoding imparted no inherent defect on transcription or translation in the absence of
265 nucleic acid sensing pathways.

266 The qPCR data suggested a borderline reduction in CpGH virus RNA levels compared with those of
267 the WT virus (Fig. 5A-B), but may be confounded by the presence of partially degraded transcripts and
268 by detection of RNAs of both polarities. To identify any reduction in full-length CpGH transcripts and
269 differentiate between positive and negative polarity RNA species, northern blotting was performed.
270 A549 and paired ZAP-/- cells were infected with PR8 WT, CDLR or CpGH viruses at MOI 10 for 8 hours,
271 and RNA was isolated from cells. Positive and negative sense specific segment 1 and segment 5 probes

272 were used for detection. Both positive and negative sense transcript levels were strongly reduced in
273 WT A549 cells for the CpGH virus, whereas transcript levels were equivalent with those from the WT
274 and CDLR viruses in ZAP-/ cells (**Fig. 5F**). The most striking reduction was seen for segment 1 positive
275 sense transcripts, which were diminished below the limit of detection. This profound reduction may have
276 been somewhat masked in the qPCR data due to the presence of vRNAs and partially degraded
277 segment 1 transcripts. In contrast, segment 5 positive sense transcripts – although of reduced intensity
278 – were readily detectable. Taken together, this is indicative of a CpG-specific reduction in transcript
279 abundance, supporting the conclusion that ZAP is targeting CpGH transcripts for degradation.

280 **CpG enrichment in IAV segment 1 causes viral attenuation in mice and attenuated CpGH IAV**
281 **delivers protection from lethal dose challenge infection with WT virus.**

282 To test whether CpGH IAV is attenuated *in vivo*, groups of six, 6-week old female BALB/c mice were
283 infected with 200 PFU (high dose) of the virus panel. All virus-infected animals lost weight after day 2
284 post infection, but mice infected with CpGH lost significantly less weight than mice infected with WT
285 PR8 (**Fig. 6A**). CpGH viral loads in the lung were significantly lower than those of WT PR8 and CDLR
286 viruses (**Fig. 6B**). Therefore, CpG-driven attenuation *in vitro* is reproduced in an *in vivo* system.

287 Next, we sought to determine whether the CpGH virus could cause a mild or clinically inapparent
288 infection while offering subsequent protection from severe disease outcome, similar to a live attenuated
289 vaccine. Mice were inoculated with low doses of virus, then subsequently challenged with a lethal dose
290 of WT virus to establish the protective effect. For low dose infections, 20 PFU of either WT PR8 (n = 6),
291 CDLR (n = 10), or CpGH virus (n=10), or were mock infected (n = 10). Clinical signs were monitored
292 for 10 days, by which time all mice were fully recovered. Mice infected with PR8 WT and CDLR
293 displayed moderate clinical signs and weight loss, whereas symptoms in mice infected with CpGH virus
294 were mild or entirely absent (**Fig. 6C**). After 20 days, sera were collected from tail bleeds for antibody
295 semi-quantitation by ELISA. 6/6 PR8, 9/10 CDLR, 9/10 CpGH and 0/10 mock infected mice developed
296 IAV-specific antibody responses; the levels of response to each virus varied but were not statistically
297 different between groups (**Fig. 6D**). After 21 days, the 6 PR8 mice, and 5/10 mice from the other
298 infection groups were culled and sera were harvested for western blotting to confirm ELISA results (**Fig.**
299 **S.5**). Remaining mice were challenged with 200 PFU WT PR8 virus. After 5 days, this exposure was
300 clinically inapparent in mice inoculated with the CDLR and CpGH viruses, but mice that were initially

301 mock-inoculated developed severe clinical signs and significant weight loss (**Fig. 6E**). These mice had
302 high virus titres in the lung, whereas no virus was recovered from mice originally inoculated with CDLR
303 or CpGH virus (**Fig. 6F**). Mean antibody titres from these mice were 5-fold higher than levels at 20 days
304 after exposure to the first virus dose, though again considerable variation was apparent (**Fig. 6G**). Thus,
305 the CpGH virus displayed the vaccine-desirable traits of causing mild or clinically inapparent infection,
306 while offering full protection from challenge with a lethal virus dose.

307 **ZAP-dependent CpG sensitivity is absent from in vaccine propagation systems**

308 Live attenuated IAV vaccines are traditionally propagated in eggs, but there is an increasing impetus to
309 move towards cell culture based propagation systems for ethical, medical and reliability reasons
310 (30,55). The impact of CpG enrichment on viral titres was therefore examined in both systems. Virus
311 rescues were performed in embryonated hens' eggs, and titres obtained for the CpGH virus were
312 reproducibly equivalent to those for WT PR8 (**Fig. 7A**). Viruses were plaqued on MDCK cells; although
313 the CpGH virus yielded slightly smaller plaques in these cells compared with PR8 WT (**Fig. S6**), titres
314 from virus passage at low MOI in MDCK cells were equivalent across the virus panel (**Fig. 7A**).

315 Possibly, a replication defect of the CpGH virus was not detected in eggs as titres had become
316 saturated. To test this, eggs were inoculated at embryonic day 6 (ED6) and day 10 (ED10) with 100
317 PFU of PR8, CDLR or CpGH virus and harvested at 24 hours. Viruses replicated to higher levels in
318 eggs at ED10, but no differences in titres between WT PR8 and CpGH virus was seen at either
319 developmental time point (**Fig. 8B**). Next, eggs at ED10 were inoculated with ten-fold serially diluted
320 virus doses and embryonic survival was monitored for 4 days. Again ,no differences were apparent
321 between the WT PR8, CDLR and CpGH viruses, with all embryos dying at similar rates (**Fig. S7**).
322 Nonlinear regression curves were fit for lethality across the dose range used (**Fig. 8C**); r2 values of
323 0.92-0.99 were obtained. LD50 values were within a five-fold range (PR8 = 0.29PFU, CDLR = 1.27PFU
324 and CpGH = 0.55PFU), thereby occurring within the limits of a single dilution factor for this experiment
325 and so confirming no differences in embryonic lethality across the virus panel.

326 To determine whether the lack of fitness penalty for the CpGH virus was recapitulated in other chicken
327 cells, chicken lung epithelial cells (CLEC213s) (56) were infected with the virus panel. Again virus titres
328 were consistent across the panel (**Fig. 7D**). As IAV is tropic for the gastrointestinal tract in birds (57),
329 replication of the virus panel was also tested in chicken enteroids (58). Again, CpGH virus replicated to

330 similar titres as those of the WT virus (**Fig. 7E**) and although viral protein production was variable, there
331 was no evidence of a defect for the CpGH virus (**Fig. 7F**). There was also no evidence of differential
332 infection profiles either by confocal imaging (**Fig. S8**) or by qPCR array to analyse expression of 88
333 innate immune genes in chickens (**Fig. S9**).

334 Thus, CpGH IAV replicated equivalently to WT PR8 in the MDCK cells and in the embryonated hens'
335 eggs used for vaccine virus propagation.

336 **Large scale recoding offers a genetically stable method of virus attenuation.**

337 Synonymous recoding of viral genomes as a method of attenuation for vaccine development is reliant
338 on the modifications being genetically stable. To investigate the stability of CpG enrichment in the IAV
339 genome, the virus panel was serially passaged ten times at low MOI in A549 cells. Passages were
340 performed four times, with input virus for two experiments being derived from egg-grown stocks and
341 two from MDCK-grown stocks. The ~1 $\log_{10}$ lower titre of CpGH virus compared with the WT or CDLR
342 viruses was maintained over the course of the passage experiments, for both egg-derived (**Fig. 8A**)
343 and MDCK-derived (**Fig. 8B**) virus stocks.

344 After the tenth passage, viruses were subjected to full-genome deep amplicon sequencing to analyse
345 nucleotide changes occurring in their genomes. Sequencing coverage across the twelve viruses (4
346 replicates each of WT PR8, CDLR, and CpGH) ranged from 127-67,000 (**Fig. S10**).

347 To establish background signal from PCR and sequencing error, amplifications and sequencing were
348 performed on the PR8 plasmids used for virus rescue. In this control sample (**Fig. S10M**), the highest
349 mismatch versus database reference was a single position at 2.3% (data not shown), and so this was
350 used as the threshold above which mutations were considered real. The sum totals of mutations over
351 this threshold arising in all egg- and MDCK-derived virus stocks after ten passages were calculated to
352 include proportional representation for substitutions that had not reached fixation (for example, if a
353 mutation was present in 10% of sequences, this was counted as 0.1 mutations). Consistent numbers
354 of mutations were observed for CpGH virus compared with WT PR8 and control viruses, and so more
355 mutations were detected per infectious virion for CpG enriched viruses (**Fig. 8C**). However, the overall
356 number of nucleotide changes identified in recoded viruses did not exceed those for WT PR8, indicating
357 a genetic stability equivalent to WT virus. No difference in the mutation rates were observed between
358 the egg-derived and MDCK-derived viruses. The numbers of mutations arising were broken down by

359 segment (**Fig. 8D**); the highest density of mutations occurred in segments 4 and 6 (encoding the viral
360 surface proteins) and segment 8 (encoding NS1 and NEP), as previously described (59).

361 The recoded sites of the passaged CpGH virus were examined for evidence of mutation and reversion
362 (**Fig. 8E**). No nucleotide changes at altered sites became fixed in any of the four replicates. Sequence
363 changes across the genomes of all serially passaged viruses are presented in Supplementary **Fig. S11**.
364 Overall, the data indicated that CpG enrichment is a genetically stable recoding strategy for IAV.

365

366 **DISCUSSION**

367 Synonymous recoding of viral genomes has been proposed as a potential strategy for live attenuated
368 vaccine development (21). However, such recoding inevitably impairs virus replication and so reduces
369 vaccine yield. We have demonstrated that CpG enrichment is a viable vaccine enhancement strategy
370 for IAV as we have made a replication-defective CpG-enriched vaccine strain IAV whose propagation
371 is unimpaired in systems used for vaccine virus synthesis (MDCK cells and embryonated hens' eggs).
372 This virus causes sub-clinical infection in mice that subsequently offers complete protection from
373 challenge infection with a lethal dose of virus. CpG introduction is genetically stable. Altogether, the
374 data presented herein support this approach for live attenuated vaccine augmentation.

375 We show a fundamental difference between human and chicken systems with respect to viral CpG
376 sensing. While a homologue of ZAPL has been identified in chickens (19), no ZAPS homologue has
377 been identified. That chicken ZAP may not have the same CpG binding specificity as human ZAP has
378 previously been reported (18). However, this is the first demonstration that a ZAP-sensitive CpG-
379 enriched virus can be propagated equivalently to a WT virus in chicken systems. Possibly, chicken ZAP
380 has lost CpG-specificity (18). Although we showed in human-derived cells that ZAP is acting via
381 transcriptional targeting and not via type I interferon signalling, it is possible that species-specific type I
382 IFN signalling cascades may also explain why chicken systems cannot target the CpGH virus. For
383 example, chickens do not have RIG-I (60) and if CpG-specific responses signal via this pathway in
384 chicken cells, chickens may have lost the ability to sense aberrant CpGs. Biochemical studies
385 investigating RNA protein interactions of ZAP and its cofactors in different species are needed in order
386 to understand this further. Alternatively, PR8 may be highly effective at counteracting CpG sensing in
387 chicken systems but less so in mammalian cells and in mice.

388 Understanding the mechanism by which CpG enrichment causes viral attenuation is paramount, as with
389 large scale synonymous recoding it is possible that a mutant virus is only a few substitutions away from
390 reversion to WT virulence. For this reason, multiple methods of viral attenuation integrated into a single
391 platform should be considered; for example, the use of CpG enrichment to improve currently used live
392 attenuated IAV vaccines that are replication-incompetent at 37°C (61-63).

393 CpG sensing by ZAP was shown to require introduction of only 15 CpGs for strong attenuation in a HIV-
394 1 replicon (21). In our IAV system, addition of 46 CpGs did not impart any attenuation, and adding 80
395 CpGs had a modest effect. ZAP was originally identified as a retrovirus restriction factor (9), and so
396 possibly it more readily detects HIV-1 transcripts. Alternatively, a gene removed from HIV-1
397 pseudoviruses may counteract ZAP, IAV may more effectively neutralise ZAP than HIV-1, or ZAP
398 sensing may be more active in the MT4 cells used for HIV-1 infections than the cell lines we used for
399 IAV infections.

400 Interestingly, the anti-CpG effects of ZAP were mediated by transcript degradation and not by type I
401 interferon signalling in our system. This lack of immune cascade contrasts with the antibody responses
402 in mice to the CpGH virus that were equivalent to those elicited by WT virus. Further mechanistic
403 understanding of how CpG-mediated attenuation occurs *in vitro* and *in vivo* is needed.

404 It has previously been reported that the IAV NS1 protein can specifically inhibit ZAPS or its binding
405 partner TRIM25 (50,51) and that this can be ablated by mutation (Table 2). However, we were unable
406 to exacerbate CpG-mediated attenuation by mutations in NS1 proposed to disrupt ZAP/TRIM25
407 targeting. Possibly, our assays were not sensitive enough to detect accentuation of a defect, or NS1
408 may target ZAP in this system. Nonetheless, the NS1 96/97 pair of mutations have previously been
409 reported to lack specificity, likely due to NS1 misfolding (50,52).

410 CpG enrichment is increasingly being proposed as a strategy for live attenuated vaccine development
411 (21). We have shown that in a vaccine strain of IAV, CpG enrichment does not impair viral growth in
412 embryonated hens' eggs or cultured MDCK cells – the main systems used for propagation of live
413 attenuated IAV vaccines. This CpGH virus has desirable vaccine properties in mice, as it is clinically
414 inapparent and highly protective from virus challenge. More widely, any live attenuated vaccines
415 generated using a reverse genetics system and propagated in mammalian cells can be CpG-enriched

416 and grown in ZAP-/- cells; therefore, large scale recoding is extrapolatable to other vaccine virus
417 systems.

418

419 **MATERIALS AND METHODS**

420 **Cells.** Madin-Darby canine kidney (MDCK), human adenocarcinoma A549, human embryonic kidney
421 (HEK)293T (all Sigma, Glasgow, UK), Normal Human Dermal Fibroblasts (NHDF; Thermo Fisher) and
422 HEK293 cells and were grown in Dulbecco's Modified Eagle Medium (DMEM) (Sigma) supplemented
423 with 10% v/v foetal calf serum (FCS) (Thermo Fisher Scientific, Waltham, MA) and 1% v/v penicillin/
424 streptomycin (Thermo Fisher Scientific). CLEC213 cells were grown in 1:1 Ham's F21 nutrient broth/
425 DMEM supplemented with 5% v/v FCS and 1% v/v penicillin/ streptomycin. Cells were maintained by
426 twice-weekly passage. A549 ZAP -/- cells were a kind gift from the laboratory of Professor Sam Wilson
427 (37), HEK293 TRIM25 -/- cells were a kind gift from the laboratory of Professor Gracjan Michlewski
428 (38), and A549 KHNYN -/- cells were a kind gift from the laboratory of Dr Chad Swanson (16). CLEC213
429 cells were a kind gift from Dr Sascha Trapp and generated by the laboratory of Dr Pascale Quéré (56).
430 Cells were checked for mycoplasma contamination regularly using a MycoAlert mycoplasma detection
431 kit (Lonza, UK).

432 **Plasmids and plasmid mutagenesis.** IAV reverse genetic plasmids corresponding with each of the 8
433 segments of IAV genome were a derivative of the UK National Institute of Biological Standards & Control
434 A/Puerto Rico/8/1934 (PR8) vaccine strain (64). For synonymous introduction of additional CpGs into
435 segment 1 of the viral genome with compensatory mutations to maintain base frequencies ('CpGH') or
436 to generate a permuted control construct synonymously recoded over the same genome region but with
437 nucleotide, dinucleotide and codon frequencies maintained ('seg1-CDLR'), constructs were designed
438 using SSE software Simmonics v1.2 'mutate sequences' function (65). The region of segment 1
439 targeted for mutation was selected to avoid previously characterised genome packaging signals and
440 splice donor and acceptor sites for PB2-S1 (66). Full segment sequences (**Table S1**) are summarised.
441 For recoding of a 1.42 kb region of segment 1, synthetic plasmid constructs were ordered from GeneArt
442 (Thermo Fisher) flanked by BsmBI sites for cloning into pDUAL plasmids used for virus rescue, as
443 previously described (64,67).

444 Mutagenesis of segment 8 was also performed to disrupt interactions of NS1 with CpG sensing
445 pathways or various components of innate immune signalling; these are summarised (**Table 2**).
446 Mutations were made on pDUAL plasmids encoding the PR8 gene (64) using a QuikChange II site
447 directed mutagenesis kit (Agilent) according to manufacturer's instructions. Primers used are tabulated
448 (**Table S2A**).

449 **Table 1. Summary of the genetic properties of CpG modified influenza A viruses that were recoded in**
450 **segment 1.**

Virus backbone	Mutation	No. CpGs	Δ CpG	No. UpAs	Δ UpA	CAI
A/Puerto Rico/8/1934	Wildtype (segment 1)	51	-	108	-	0.72
A/Puerto Rico/8/1934	CDLR	52	+1	108	-	0.72
A/Puerto Rico/8/1934	CpGH	177	+126 (347%)	103	-5	0.64
A/Puerto Rico/8/1934	5'-CpGH	131	+80 (257%)	107	-1	
A/Puerto Rico/8/1934	3'-CpGH	97	+46 (187%)	104	-4	
A/Puerto Rico/8/1934	CpGH_low_density	97	+46 (187%)	103	-5	

451 CAI, codon adaptation index. CAI scores have been calculated relative to the human genome (A/Puerto
452 Rico/8/1934) in https://www.biologicscorp.com/tools/CAICalculator. For full construct sequences, please refer to
453 Table S1.

454 **Table 2. Previously characterised mutations in NS1 that reverse ZAP/TRIM25-mediated IAV attenuation**

NS1 target	Abrogating mutation	Mechanism disrupted	Reference
TRIM25	L95/S99A	RIG-I ubiquitination/ vRNP binding	(50)
ZAP	E96/E97A	Can no longer prevent TRIM25-mediated ubiquitination of ZAP	(51)
ZAP	PKQK107-110A	Can no longer prevent TRIM25-mediated ubiquitination of ZAP	(51)
Interferon	R38K41A	NS1 can no longer bind RNA	(49)

455

456 **Virus rescues.** Viruses were rescued as previously described (68,69). HEK293T cells in 6 well plates
457 were transfected in Opti-MEM Reduced Serum Medium (Thermo Fisher) with 250ng each of the 8
458 plasmids corresponding with each segment of viral genome (or 7 plasmids for mock) and 4 µl
459 Lipofectamine 2000 transfection reagent (Thermo Fisher) per well. pDUAL plasmids were used for all

460 segments, which are bi-directional; vRNA is under a polI promoter, and in the opposite orientation,
461 positive sense RNA is under a polII promoter. After 24 hours, medium was replaced with serum free
462 medium containing 1 µg/ml tosyl phenylalanyl chloromethyl ketone (TPCK)-treated trypsin (Sigma) and
463 0.14% BSA (w/v) (Sigma) (viral growth medium). At 48-72 hours post-transfection, supernatants were
464 collected and passaged, either (1) onto MDCK cells in viral growth medium, with supernatants collected
465 48 hours later; or (2) 100 µl was inoculated into the allantoic fluid of an embyronated hen's egg at days
466 11-12 development. A packaging mutant, '4c6c' (multiple nucleotide changes in segment 4 and
467 segment 6, encoding the haemagglutinin and neuraminidase viral surface proteins respectively) for use
468 as a control for RNA:PFU ratio assays was propagated in MDCK cells rather than eggs to avoid
469 reversion (33,41), as was an NS1 (segment 8) R38K41A mutant (defective in regulating interferon) (49).

470 **Egg infections.** Dekalb White fertilised eggs were purchased from Henry Stewart (Norfolk, UK) and
471 incubated at 37.5ºC, 50% relative humidity until infection at day 11-12 (unless stated). Eggs were
472 candled to confirm viability. Under sterile conditions, an egg punch was used to create two small holes
473 in the shell – one in the air cavity and one in the locality of the allantoic fluid. 100 µl virus was injected
474 into the allantoic fluid and both holes were sealed with sticky tape. For preparation of virus stocks,
475 infected eggs were left for 48 hours prior to allantoic fluid harvest, which was spun down for 5 minutes
476 at 800g to pellet cellular debris, aliquoted and used as neat virus stock; presence of virus was confirmed
477 by HA assay (70). For comparison of viral titres between days 6 and 10, eggs were inoculated with 100
478 PFU of virus and were left for 24 hours prior to harvesting of allantoic fluid for virus titration. For kill
479 curves, eggs were inoculated at day 10 with 0.001 – 1000 PFU virus and incubated until day 14 unless
480 embryonic death was observed before that time. At experimental endpoints, eggs were chilled overnight
481 and embryonic death was confirmed by membrane disruption.

482 **Virus titrations.** Viral titres were quantified by plaque assays; 800 µl of ten-fold serial dilutions of viral
483 stocks/ experiment supernatants/ allantoic fluid were incubated on sub-confluent MDCKs in 6-well
484 plates for an hour, and then 2 ml overlay comprising 1:1 viral growth medium: 2.4% cellulose (Sigma)
485 was added. Plates were incubated without disturbance for 48 hours, after which time 1 ml 10% formalin
486 was added to each well and cells were fixed for at least 20 minutes. Overlays were removed and cells
487 were stained with 0.1% toluidine blue (Sigma).

488 **Virus infections.** Cells were infected at either a high multiplicity of infection (MOI) of 10 (single cycle
489 infections) or low MOI of 0.01-0.1 (cell line dependent, multi-cycle infections) for 24 or 48 hours
490 respectively unless otherwise stated. Viruses were added to cells and incubated for 1 hour at 37°C, and
491 then inoculum was removed, cells were washed with serum free medium, and then either serum free
492 medium (single cycle infection) or viral growth medium (multi-cycle infection) was added. Infected cells
493 were incubated at 37°C for the duration of the experiment.

494 **Generation of lentivirus vectors expressing ZAPL and ZAPS.** In order to reconstitute isoform-
495 specific ZAP expression in ZAP -/- A549 cells, lentiviruses expressing ZAPL or ZAPS isoforms under a
496 polII promoter were produced using a pSCRPSY (accession number KT368137.1) vector-based system
497 (a kind gift from the laboratory of Prof Sam Wilson). Both the A549 Cas9+ ZAP -/- cells and pSCRPSY
498 carry puromycin resistance, and so a blasticidin S-resistance gene was cloned into the NheI site of
499 pSCRPSY. pSCRPSY ZAPL and pSCRPSY ZAPS constructs were produced by cloning a cDNA of
500 the open reading frame of ZAPL (accession number NM_020119.4) or ZAPS (accession number
501 NM_024625.4) into the dual SfiI sites of pSCRPSY. For lentivirus production, the growth media on
502 HEK293Ts at 80% confluency in a T175 flask was replaced with 25ml of Opti-MEM. 6 µg PMD2.G
503 (Addgene, 12259), 18µg PsPAX2 (Addgene, 12260) and 12 µg of pSCRPSY (pSCRPSY, pSCRPSY-
504 ZAPL or pSCRPSY-ZAPS) were added to 450µl of Opti-MEM. 72µl of Lipofectamine 2000 was added
505 to 450µl of Opti-MEM. Diluted plasmids and Lipofectamine were mixed and incubated for 30 minutes at
506 room temperature. The transfection mixture was added to the flask and cells were then incubated at
507 37ºC, 5% CO_2 for 18h. Opti-MEM was then replaced with 25ml of DMEMsupplemented with 10% FBS.
508 After a further 48 hours, supernatant containing P0 lentivirus was collected and centrifuged to remove
509 cell debris. Due to low titre of ZAPL-expressing lentivirus, 75 ml of cell culture supernatant containing
510 lentivirus was concentrated by ultracentrifugation using a SW32Ti rotor in a Beckman Coulter Optima
511 Max-E ultracentrifuge at 50,000 g for 2 hours. The viral pellet was then resuspended in 1 ml of 10%
512 FCS-containing D-MEM. Lentivirus was aliquoted and stored at -80ºC.

513 **Generation of A549 cells expressing specific ZAP isoforms.** A549 ZAP-/- cells at 70% confluency
514 in a 6 well plate were transduced with ZAPL- or ZAPS-expressing pSCRPSY lentiviruses. 1ml of
515 lentivirus (pSCRPSY empty vector control, concentrated ZAPL, or ZAPS) or 10% FBS D-MEM as
516 control, supplemented with polybrene (Merck, TR-1003-G) at 8 µg/ml, was added to each well of a 6
517 well plate. Spinoculation was performed (500 *g*, 1 hour). 48 hours after transduction, the inoculum was

518 removed and cells were passaged in 10% FBS D-MEM supplemented with 50 µg/ml puromycin (Gibco)
519 to select for transduced cells. Expression of the appropriate ZAP isoforms was confirmed by western
520 blotting and confocal imaging (described below).

521 **Western blotting.** Samples for western blotting were collected in 2X homemade Laemmli buffer (71),
522 boiled for 10 minutes, cooled and loaded onto 10% pre-cast gels (Bio-Rad) for separation by SDS-
523 PAGE. Proteins were transferred onto nitrocellulose membranes (Fisher Scientific) using wet transfer
524 (100V for 90 minutes on ice) and membranes were blocked using 5% milk (Morrisons, Edinburgh, UK)
525 for at least 30 minutes. Membranes were washed three times with TBS-T (Tris-buffered-saline/0.1%
526 Tween 20 (Sigma)) and incubated overnight at 4°C with primary antibodies in 2% BSA/TBS-T (0.1%).
527 Commercial antibodies were used for β-tubulin (clone YL1/2, Bio-Rad, Watford, UK; 1:5000), β-actin
528 (Abcam ab8227; 1:2000), ZAP (Invitrogen PA5-31650; 1:1000), TRIM25 (Abcam ab167154; 1:1000)
529 and KHNYN (Insight Bio sc-514168; 1:1000). In-house rabbit polyclonal antibodies raised against NP
530 and PB2 are previously described (72) and were both diluted 1:1000. Membranes were then washed 3
531 times and incubated with species-specific secondary antibodies conjugated to Alexafluor-680 or -800
532 (Thermo Fisher) diluted 1:5000 in 2% BSA/TBS-T (0.1%) for 90 minutes. Membranes were washed 3
533 times and visualised using a LICOR Odyssey Fc imaging system.

534 **Confocal imaging.** Immunofluorescence was used to confirm ZAP expression in lentivirus-transduced
535 A549 cells, and presence of viral NP in chicken enteroids. For imaging of A549 cells, A549 pSCRPSY-
536 transduced cells at 95% confluency on coverslips in 24 well plates were fixed using 500 µl of 3.7%
537 neutral buffered formalin (CellPath) for 15 minutes, and then washed with PBS. Cells were
538 permeabilised using 500 µl of 0.02% Triton X100 for 5 minutes and blocked for 30 minutes in 2% FCS.
539 Anti-ZC3HAV1 Antibody (16820-1-AP) used 1:500 was diluted in 2% FCS/ PBS and incubated on cells
540 overnight at 4°C. 200 µl of 1:1000 AlexaFluor-488 anti-rabbit secondary antibody (Thermo Fisher,
541 A32731) was added for 1 hour. Cell nuclei were stained using 1:5000 Hoechst (Thermo Fisher) in PBS
542 for 10 minutes. Between addition of each reagent, cells were washed 3 times using 500 µl of PBS. All
543 images were acquired with Zeiss LSM 710 confocal microscope fitted with 40X/1.4 oil-immersion
544 objective lens. Controls for background signal were mock infected cells and infected cells stained using
545 secondary antibodies only.

546 For organoid imaging, 50 organoids in a 24 well plate were fixed in 4% neutral buffered formalin for 45
547 minutes and washed with PBS. Organoids were then permeabilised and blocked with 5% v/v goat serum
548 (Sigma) in permeabilisation solution (0.5% v/v bovine serum albumin and 0.1% v/v Saponin (Merck,
549 Glasgow, UK) in PBS) for 30 minutes. Antibodies were diluted in permeabilisation solution and
550 incubated with organoids (mouse monoclonal NP antibody AA5H clone (Abcam, Cambridge, UK
551 ab20343) used 1:500 overnight at 4°C, phalloidin-647 (Thermo Fisher A22287) used 1:50 and
552 incubated for 1 hour at room temperature). 200ul of 1:1000 AlexaFluor-488 anti-mouse secondary
553 antibody (Thermo Fisher A32723) was added for 1 hour. Cell nuclei were stained using 1:5000 Hoechst
554 in PBS for 10 minutes. Between addition of each reagent, cells were washed 3x using 500 µl of PBS.
555 All images were acquired with Zeiss LSM 880 confocal microscope fitted with 40X/1.4 oil-immersion
556 objective lens. Secondary-only and isotype (mouse IgG2b) controls were used.

557 **Reverse transcription (RT)-quantitative (q)PCR**. RT-qPCR was performed as previously described
558 (67). RNA was extracted from cells using Qiagen QiAmp Viral RNA mini kits. Nucleic acid extracts were
559 treated using RQ1 RNAse-free DNAse (Promega) according to manufacturer's instructions. qPCR was
560 performed in technical triplicates using BioLine (London, UK) SensiFAST One Step RT-qPCR Lo-Rox
561 kit (BIO-82020), with primers targeting a conserved region of segment 1 or segment 5 of IAV PR8 strain
562 (**Table S2B**). Cycling conditions were 45°C for 10 minutes, then 40 cycles of 95°C for 10 seconds then
563 60°C for 30 seconds. Melt curves were performed between 50°C-99°C with one degree increments at
564 second intervals.

565 **Urea-PAGE RNA gels.** IAV virions equivalent to 10^{10} PFU from egg stocks were semi-purified by
566 centrifugation through a buffered 25% sucrose cushion (w/v) (100 mM NaCl, 10 mM Tris-HCl [pH 7],
567 1mM EDTA) using a SW32Ti rotor in a Beckman Coulter Optima Max-E ultracentrifuge at 280,000 *g* for
568 90 minutes at 4°C. RNAs were extracted from virions using an RNEasy mini extraction kit (Qiagen)
569 following resuspension of the pellets directly into 350 µl of buffer RLT. Extracted RNA was separated
570 by PAGE using homemade 5% urea polyacrylamide gels in 1X Tris-borate-EDTA (TBE) buffer (89mM
571 Tris-borate, 2mM EDTA [pH 8.3]) run for 6 hours at 120 V. Viral gene segments were visualised by
572 silver staining using the Silver Stain Plus Kit (Bio-Rad) according to manufacturer's instructions and
573 imaged using a Samsung Xpress C480FW scanner.

574 **Bisulphite conversion.** RNA was extracted from ~5x10^8 IAV virions (egg virus stock) using a Qiagen
575 RNEasy mini extraction kit, followed by Promega RQ1 DNAse treatment. Bisulphite conversion
576 reactions were performed on extracted egg stock RNAs or *in vitro* generated genomic RNA (gRNA;
577 negative sense) transcripts using the Zymo (Freiburg, Germany) RNA bisulphite conversion kit (EZ RNA
578 Methylation Kit, RS001). In the reaction, cytosines are converted to uracils unless protected by 5-
579 methylcytosine (5-mc) methylation. Following conversion, RNA libraries for sequencing were prepared
580 using the Takara (Saint-Germain-en-Laye, France) low input RNAseq kit and sequenced on a MiSeq
581 Micro v2 150PE flowcell at Edinburgh Genomics (Edinburgh UK). Sequencing data handling and
582 analyses were performed using the Galaxy platform (73). Adapter sequences
583 (AGATCGGAAGAGCACACGTCTGAACTCCAGTCA and AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT)
584 were trimmed from reads using cutadapt and output sequences were joined using fastq-join. Sequences
585 were aligned to an *in silico* converted PR8 segment 1 (where all cytosines in the negative sense were
586 replaced with thymines) using Bowtie2 (74) and resultant BAM datasets were converted to tabular
587 pileup format to allow for calculation of the proportion of retained cytosines (indicating presence of 5-
588 mc).

589 Specific regions selected for further analysis were amplified by RT-PCR from the converted RNA. RNAs
590 were reverse transcribed using SuperScript III (Invitrogen) with random hexamer primers according to
591 manufacturer's instructions. PCRs were performed using specific primers (**Table S2C**) and Q5 High-
592 Fidelity Polymerase (NEB, Hitchin, UK) according to manufacturer's instructions. Cycling conditions
593 were 98°C for 30 seconds, then 45 cycles of 98°C for 10 seconds, 58°C for 21 seconds then 72°C for
594 30 seconds with a final extension of 72°C for 2 minutes. Amplicons were sent for sequencing using the
595 Amplicon-EZ service (Genewiz, Leipzig, Germany). Sequences were analysed as above except that
596 adapter trimming was performed using the primer sequences.

597 **HEK Blue assay.** To assay type I IFN production, A549 cells in 24 well plates were infected at an MOI
598 of 10 as described above but in the absence of trypsin. At 10 hours post infection, supernatants were
599 harvested and viruses were inactivated by exposure to 120 mJ/cm^2 of UV for 10 minutes in a UVP CL-
600 1000 UV crosslinker. 20 µl of inactivated sample or type I IFN standard (Abcam, 500, 50 or 5 pg/ µl)
601 was added to HEK-Blue^{TM} IFNα/β cells (InvivoGen hkb-ifnab) at 4x10^4 cells/ well in 96 well format. HEK-
602 Blue cells have a gene for secreted alkaline phosphatase incorporated under an ISG54 promoter (that
603 is robustly induced in the presence of type I IFN). Cells were incubated with supernatant from A549

604 infections for 24 hours. HEK-Blue cell supernatants were then collected and added to QuantiBlue
605 reagent (InvivoGen) according to manufacturer's instructions. After 15-30 minutes, when a colour
606 change was visible in the positive control (titrated IFN), absorbance was read at 620 nm, as a correlate
607 of IFN concentration.

608 **Minigenome assays.** Minigenome assays were performed as previously described (54). Culture
609 medium on 70-90% confluent HEK293T cells in 24 well plates was replaced with Opti-MEM. Cells were
610 then transfected with 50 ng/well (unless stated) of plasmids encoding IAV segments 1, 2, 3, and 5
611 (sufficient for viral ribonucleoprotein (vRNP) formation) with positive orientation under a PolII promoter
612 (the same pDUAL plasmid constructs used for virus rescue were used here), along with 20 ng/ well of
613 a reporter construct expressing an IAV-like vRNA with a luciferase coding sequence in place of an IAV
614 open reading frame. At 48 hours post-transfection, cells were harvested in 100 µl 5X reporter lysis buffer
615 (Promega, Chilworth, UK) and refrigerated to allow cell debris to settle to the bottom of the tube. 60 µl
616 supernatant was added to 25 µl 600 µM beetle luciferin (Promega) in a white walled 96 well plate, and
617 luminescence was measured on a GloMax luminometer (Promega). Each data point was represented
618 by four technical repeats, and each assay was undertaken for at least four biological repeats.

619 ***In vitro* transcription assays.** PCR amplicons were generated for the production of positive or negative
620 sense, full-length segment 1 RNA by the addition of a T7 promoter sequence to the forward or reverse
621 primer, respectively (**Table S2D**). PCRs were performed using Q5 DNA polymerase kit (NEB) with the
622 appropriate primer pair and 25 ng of plasmid as template (segment 1 of PR8, CDLR, or CpGH). Cycling
623 conditions were 95°C for 5 minutes, then 30 cycles of 95°C for 30 seconds, 50°C for 30 seconds then
624 72°C for 2 minutes, with a final extension of 72°C for 5 minutes. Amplicons were purified by agarose
625 gel extraction using the MinElute Gel Extraction kit (Qiagen) and quantified using the Qubit dsDNA BR
626 Assay kit (Thermo Fisher). RNA transcripts were generated with 40 ng of these amplicons as template
627 using the MEGAscript T7 Transcription kit (Thermo Fisher) in half reaction volumes according to
628 manufacturer's instructions for the indicated times. Following incubation, transcript concentrations were
629 determined using the Qubit RNA HS Assay kit (Thermo Fisher).

630 ***In vitro* translation assays.** *In vitro* generated RNA transcripts above were used templates for *in vitro*
631 translation assays using a Rabbit Reticulocyte Lysate kit (Promega) at the indicated concentrations.
632 Reactions were performed at 1/5th volume with the inclusion of Transcend Biotin-Lysyl-tRNA (Promega)

633 to allow antibody independent detection of all translated products. Samples were assayed by western
634 blot (as above) with the exception that membranes were blocked with 5% BSA/TBS for 1 hour,
635 incubated with IRDye 800CW Streptavidin (LICOR) for 1 hour, washed 3 times and visualised using the
636 LICOR Odyssey Fc imaging system.

637 **Northern blotting.** RNA probes and positive control full-length, negative sense transcripts were
638 generated by first producing PCR amplicons of the target sequence with the addition of a T7 promoter
639 sequence to the forward or reverse primer, (**Table S2E**). PCRs were performed using Q5 DNA
640 polymerase kit (NEB) with the appropriate primer pair and 25 ng of plasmid as template (Seg1-PR8 or
641 Seg5-PR8). Cycling conditions were 95°C for 5 minutes, then 30 cycles of 95°C for 30 seconds, 50°C
642 for 30 seconds then 72°C for 2 minutes, with a final extension of 72°C for 5 minutes. Amplicons were
643 purified by agarose gel extraction using the MinElute Gel Extraction kit (Qiagen) and quantified using
644 the Qubit dsDNA BR Assay kit. RNA transcripts were generated with 40 ng of these amplicons as
645 template using the MEGAscript T7 Transcription kit (Thermo Fisher) in half reaction volumes for 4 hours
646 according to manufacturer's instructions except for the inclusion of Biotin-11-UTP (Cambridge
647 BioScience) at a molar ratio of 2:3 with unlabelled UTP for the generation of labelled probes. Following
648 incubation, transcript concentrations were determined using the Qubit RNA HS Assay kit (Thermo
649 Fisher).

650 Parental A549 Cas9+ or A549 ZAP-/- cells were infected at a high MOI as described above and cell
651 pellets were collected at 8 hours post infection. RNA was extracted from cell pellets using the RNeasy
652 Mini Kit (Qiagen). RNA gels were run, blotted and hybridised using reagents from the NorthernMax –
653 Gly Kit (Thermo Fisher) unless otherwise indicated. Briefly, RNA samples were run on 1% Agarose-LE
654 gels at low voltage for 4 hours along with 1ng each of positive control positive sense segment 1 and
655 segment 5 transcripts and RiboRuler High Range RNA ladder (Thermo Fisher). Nucleic acids were
656 blotted on to Nytran SuPerCharge nylon membranes (Cytiva) by overnight downward transfer using the
657 Turbo Blotter System apparatus (Cytiva). Membranes were baked at 68°C for 10 minutes and
658 crosslinked by exposure to 1200 (x100) J/cm2 twice in CL-1000 UV crosslinker. Membranes were pre-
659 hybridised in ULTRAhyb Ultrasensitive Hybridisation Buffer for 30 minutes at 68°C and hybridised
660 overnight at 68°C with biotinylated probes at a concentration of 0.1nM. Membranes were washed once
661 in 2xSCC, 0.1%SDS at room temperature for 15 minutes and three times with 0.1xSCC, 0.1%SDS at
662 68°C for 15 minutes each. Membranes were blocked in 5% dried milk powder in TBS-T for 4 hours at

663 room temperature, washed six times with TBS-T and incubated with HRP-conjugated Streptavidin
664 (Proteintech) in 5%BSA/TBS-T for 30 minutes. Membranes were washed four times in TBS-T,
665 incubated briefly with LumiGLO Peroxidase Chemiluminescent Substrate (Insight Bio) and visualised
666 by exposure to Amersham High Performance Chemiluminescence film (Cytiva) for 5 minutes.

667 **Mouse infections.** Mouse experiments were approved by the Roslin Institute Animal Welfare and
668 Ethical Review Board under Home Office project license PF795AF60. Five-week-old female BALB/c
669 mice, maintained in groups of six, were purchased from Harlan (Oxon, UK) and housed for one week
670 prior to infections. For infection experiments, mice were inoculated intranasally with 200 PFU virus
671 (either PR8, seg1-CDLR, seg1-CpGH or mock) under brief isofluorane anaesthesia. Mice were weighed
672 daily until day 5, when mice were euthanised by terminal isofluorane anaesthesia. Left lungs were
673 harvested in 2 ml Eppendorf (Stevenage, UK) Safe-Lock micro test tubes containing 500 µl DMEM and
674 a stainless-steel carbide bead (Qiagen), and snap frozen on dry ice until virus titration. For titration,
675 lungs were thawed on wet ice and fragmented using a Qiagen TissueLyser II at 200 oscillations/minute
676 for 2 minutes or until tissue was visibly homogenised. Suspended tissue was then serially diluted and
677 titrated by plaque assay. For challenge infections, groups of ten six-week old mice were inoculated with
678 20 PFU virus (either PR8, CDLR, CpGH or mock) under brief isofluorane anaesthesia, except for the
679 PR8 group which comprised six mice. Mice were weighed and monitored for clinical signs until 10 days
680 post-infection. At 20 days post-infection, tail bleeds were performed. At 21 days, 5 mice from each
681 group (or 6 from PR8) were culled by terminal isofluorane anaesthesia and cardiac puncture. Sera were
682 collected. The remaining mice were challenged with 200 PFU PR8 WT virus (due to licensing
683 constraints inoculation then challenge with the same PR8 WT virus was not possible). Mice were
684 monitored until cull at 5 days post-challenge by terminal isofluorane anaesthesia and cardiac puncture,
685 with sera and left lung collected as described above for ELISA and plaque assays respectively.

686 **Mouse serum ELISA.** To test mouse sera cross-reactivity with IAV virions, virions were semi-purified
687 by overlaying 5 mL of clarified PR8, CDLR or CpGH virus stocks on 25% buffered sucrose (w/v) cushion
688 (100 mM NaCl, 10 mM Tris-HCl pH7, 1 mM EDTA) in a Beckman Coulter Optima Max-E ultracentrifuge
689 with a SW32Ti rotor at 60,000 *g* for 2 hours at 4°C. Viral pellets were resuspended in 500 µL PBS.
690 Virions equivalent to an input of 2 x 10^6 PFU were diluted in 100 µL of 0.1 M carbonate-bicarbonate
691 buffer (pH 9.6) (Sigma-Aldrich) and immobilised on 96-well ELISA plates (Greiner Bio-One) overnight
692 at 4°C. Plates were washed twice with PBS-Tween 20 (0.05% v/v) and blocked with 1% BSA/PBS at

693 room temperature for 2 hours. Mouse sera diluted 1/100 in 1% BSA/PBS were then applied as primary
694 antibody along with a 5-fold dilution series of a positive control serum designated as 100 arbitrary units
695 (AU)/µl and plates were incubated at room temperature with agitation for 2 hours. Plates were washed
696 six times with PBS-T and incubated with HRP-conjugated goat anti-mouse IgG (H+L) (1:2000) (Bio-
697 Rad) in 1% BSA/PBS for 1 hour with agitation. Plates were washed thrice with PBS-T and incubated
698 with 2,2'-azino-bis(3'-ethylbenzothiazoline-6-sulfonic acid) substrate (SLS) after which the colorimetric
699 change was arrested by adding 70 µL/well 1%SDS. Optical density was read at 405 nm on a Cytation
700 3 Cell Imaging Multi-Mode Reader (Agilent) and analysed in Bioek Gen5 (Agilent).

701 **Dot blotting with mouse sera.** To test mouse sera cross-reactivity with IAV proteins, MDCK cells were
702 infected at MOI of 10 with PR8 strain for 8 hours, and infected cell lysates were harvested in 2X Laemmli
703 buffer. A single lysate long well was run out by SDS-PAGE and transferred to a nitrocellulose membrane
704 as described above. Membranes were placed in a strip-blot and individual lanes were incubated with
705 day 20 sera from individual mice diluted 1:100 in 2% BSA/PBS for 1 hour with agitation. Western blotting
706 then proceeded as described above.

707 **Chicken enteroid infections.** Chicken enteroids were prepared as described previously (58). Briefly,
708 small intestinal tissue was removed from Hy-Line Brown embryos supplied by the National Avian
709 Research Facility, Edinburgh, UK (Home Office Project Licenses PE263A4FA) and cut into 5 mm
710 sections. Tissue was digested and villi were isolated and cultured. At day 2 of culture, approximately
711 200 enteroids were infected with 10^5 virus particles/ enteroid for 1 hour, washed 3 times in PBS and
712 then incubated for 48 hours in the presence of 2 µg/mL TPCK trypsin for virus quantification by plaque
713 assay, confocal imaging, or arrayed gene expression analyses.

714 **Fluidigm analysis of innate immune response gene expression in infected chicken enteroids.** At
715 48 hpi, enteroids were suspended in RNALater and RNA was isolated using Tru-Seq total RNA Sample
716 Preparation v2 kit (Illumina, Great Abingdon, UK) according to manufacturer's instructions. cDNA was
717 synthesized from 100 ng of RNA using High Capacity Reverse Transcription kit (Life Technologies, UK)
718 according to manufacturer's instructions with a random hexamer primer and oligo(dT). cDNA was pre-
719 amplified as previously described (75) and analysed through quantitative PCR (qPCR) with the
720 microfluidic 96.96 Dynamic array (Fluidigm, San Francisco, CA) performed on the BioMark HD
721 instrument (BioMark, Boise, ID). Raw Cq values were processed with GenEx6 and GenEx Enterprise

722 (MultiD Analyses AB), with correction for primer efficiency and reference gene normalisation. Reference
723 genes used were TATA box binding protein (TBP), Tubulin alpha chain (TUBA8B), beta-actin (ACTB),
724 beta-glucuronidase (GUSB), glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and ribosomal
725 28S (r28S).

726 **Serial passage.** Viruses were serially passaged ten times at an MOI of 0.01 in A549 cells. Viruses were
727 titred by plaque assay after each passage. Biological duplicates were performed with viruses rescued
728 from both eggs and from MDCK cells. At the end of ten passages, viruses were sequenced. RNA was
729 extracted from culture supernatants using Viral RNA mini kits (Qiagen) according to manufacturer's
730 instructions. Extracted RNAs were reverse transcribed using SuperScript III (Invitrogen) with the IAV
731 vRNA specific Uni12 primer (AGCAAAAGCAGG) (76) according to manufacturer's instructions. PCR
732 primers were designed to amplify approximately 400 base fragments with approximately 60 base
733 overlaps across the genome (**Table S2F**). Additional primers were designed specific to the segment 1
734 and segment 5 recoded regions as appropriate. Four sets of primer pools each containing 10 non-
735 overlapping sets of primer pairs were generated (1A-3E odd, 1B-3F even, 3G-8B odd and 4A-8C even)
736 with each primer at a final concentration of 5 µM. PCRs were performed using specific primers and Q5
737 High-Fidelity Polymerase according to manufacturer's instructions with primers at a concentration of
738 500 nM each in the final reactions. Cycling conditions were 98°C for 30 seconds, then 45 cycles of 98°C
739 for 20 seconds, 55°C for 20 seconds then 72°C for 30 seconds with a final extension of 72°C for 2
740 minutes. The 4 multiplexed amplicon sets for each sample were purified using a PureLink PCR
741 Purification kit, quantified using a Qubit dsDNA BR assay kit (Invitrogen) and pooled at equimolar ratios.
742 Amplicon pools were sent for sequencing using the Amplicon-EZ service (Genewiz). Sequencing data
743 handling and analyses were performed using the Galaxy platform (73). Primer sequences were trimmed
744 from reads using cutadapt and output sequences were joined using fastq-join. Sequences were aligned
745 to the PR8 reference genome or genome sequences with recoded segments 1 or 5, as appropriate,
746 using Bowtie2 (74). Variants and coverage levels in the resultant BAM datasets were analysed using
747 iVar (77) variants with tabular outputs.

748

749 **ACKNOWLEDGEMENTS**

750 We are grateful to the labs of Prof Sam Wilson (MRC-University of Glasgow Centre for Virus Research,
751 UK) for provision of A549 ZAP-/- cell lines and pSCRPSY lentivirus vector, Prof Gracjan Michlewski
752 (University of Edinburgh, UK/ International Institute of Molecular and Cell Biology, Poland) for the
753 provision of TRIM25-/- HEK293 cells, Dr Chad Swanson (King's College London, UK) for A549 KHNYN-
754 /- cells and Dr Sascha Trapp (French National Institute for Agriculture, Food and Environment, France)
755 for the provision of CLEC213 cells. We are indebted to the staff in the Roslin Biological Research Facility
756 for technical support for mouse infection experiments and mouse line maintenance, to the staff of the
757 Central Services Unit, Facilities team, Bioimaging facilities, and to the National Avian Research Facility
758 for supply of eggs in the Roslin Institute, University of Edinburgh.

759 We are grateful to Christina Vrettou (The Roslin Institute, UK) for technical assistance, to Dr James
760 Glover (The Roslin Institute, UK) for critical feedback on the manuscript and to Dr Edward Hutchinson
761 (MRC-University of Glasgow Centre for Virus Research, UK) for helpful discussions.

762

763 **FINANCIAL DISCLOSURE STATEMENT**

764 FG, LV, PD and EG are supported by a BBSRC Institute Strategic Programme grant (BB/P013740/1).
765 BT was supported by a Microbiology Society Harry Smith Vacation studentship (GA002550), Wellcome
766 Trust ISSF3 funding and a Carnegie Trust Undergraduate Vacation Scholarship (VAC011866). FG and
767 PD are supported by BBSRC grant no BB/S00114X/1. OD is supported by a Roslin Studentship Award.
768 LT is supported by an EASTBio Studentship Award. PS is supported by Wellcome Investigator Award
769 (WT103767MA). LV and PD are also supported by the European Union's Horizon 2020 research and
770 innovation programme under grant agreement no. 727922 (DELTA-FLU) and SS under grant
771 agreement no. 731014 (VetBioNet). TN was supported by an iCase studentship from the BBSRC
772 (BB/MO14819). EG is also supported by a Wellcome Trust/ Royal Society Sir Henry Dale Fellowship
773 (211222_Z_18_Z). The funders had no role in study design, data collection and analysis, decision to
774 publish, or preparation of the manuscript. For the purpose of open access, the author has applied a CC
775 BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

776

777 **COMPETING INTERESTS**

None declared.

FIGURES

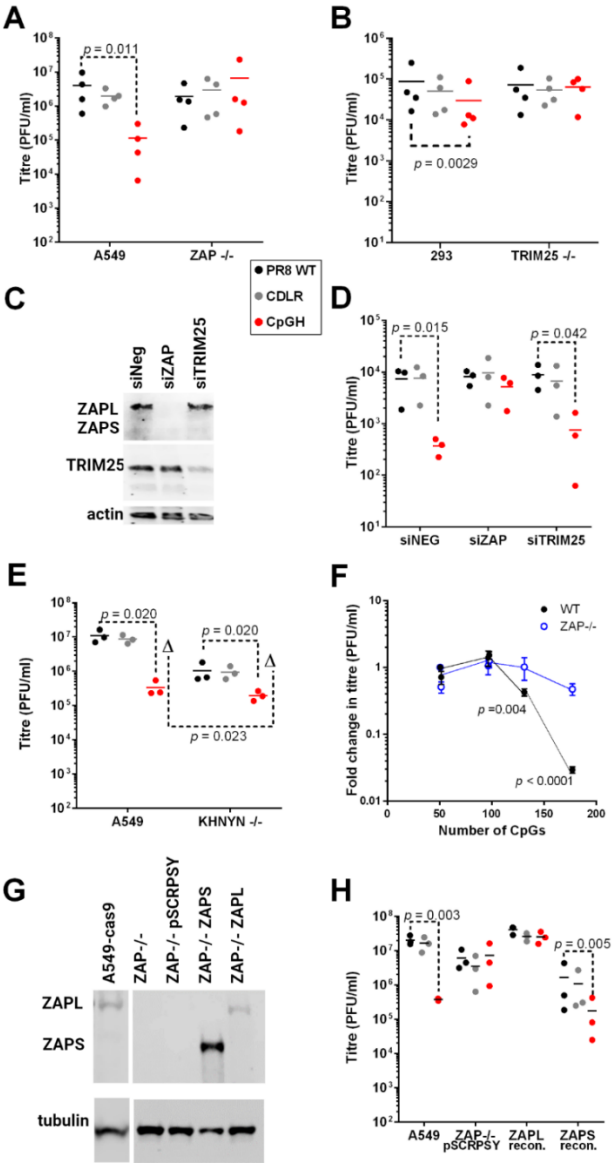


Figure 1. ***Influenza A viruses with CpG enrichment in segment 1 have replication defects mediated by ZAP.*** WT PR8, synonymously permuted control virus (CDLR) and CpG enriched (CpGH) viruses were used to infect various cell types at low MOI for 48 hours. Synonymous recoding was performed in segment 1. Viruses were grown in A549 and paired ZAP -/- cells (**A**) or in 293 and paired TRIM25 -/- cells (**B**). **C.** siRNA knockdown of ZAP or TRIM25 in NHDF primary cells was confirmed by western blot. **D.** Impact of ZAP or TRIM25 knockdown on replication of the seg1 virus panel in NHDFs. **E.** Seg1 virus panel was grown in A549 and paired KHNYN -/- cells. **F.** The number of CpGs introduced into the CpGH construct was depleted by a third (80 CpGs added, giving a segment total of 131) or by two thirds (46 CpGs added, giving a segment total of 97). When 46 CpGs were added, this was done either by retaining the 46 CpGs added to the 3' end of RNA (positive orientation), or by spacing the 46 CpGs added across the segment. This panel of CpG mutants, along with CDLR and CpGH, were titred in WT A549s and paired ZAP-./- cells as in (A). **G.** ZAPL/ ZAPS expression was reconstituted in A549 cells using lentivirus expression constructs, and isoform-specific expression was confirmed by western blotting. **H.** A549 ZAP-/- cells were reconstituted for isoform-specific ZAPL or ZAPS expression, and titres of seg1 virus panel were determined. For validation of KO phenotypes, please refer to **Fig. S1**. Confocal visualisation of ZAP reconstitution is presented in **Fig. S2**.

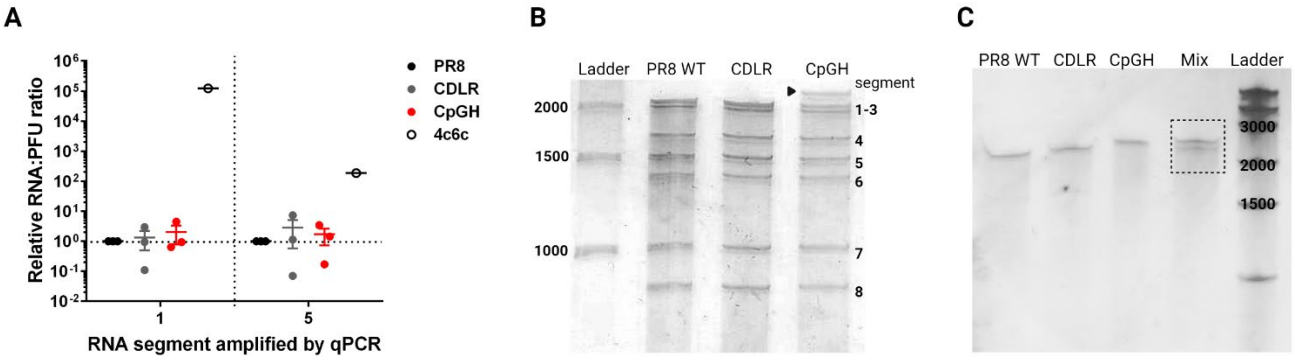
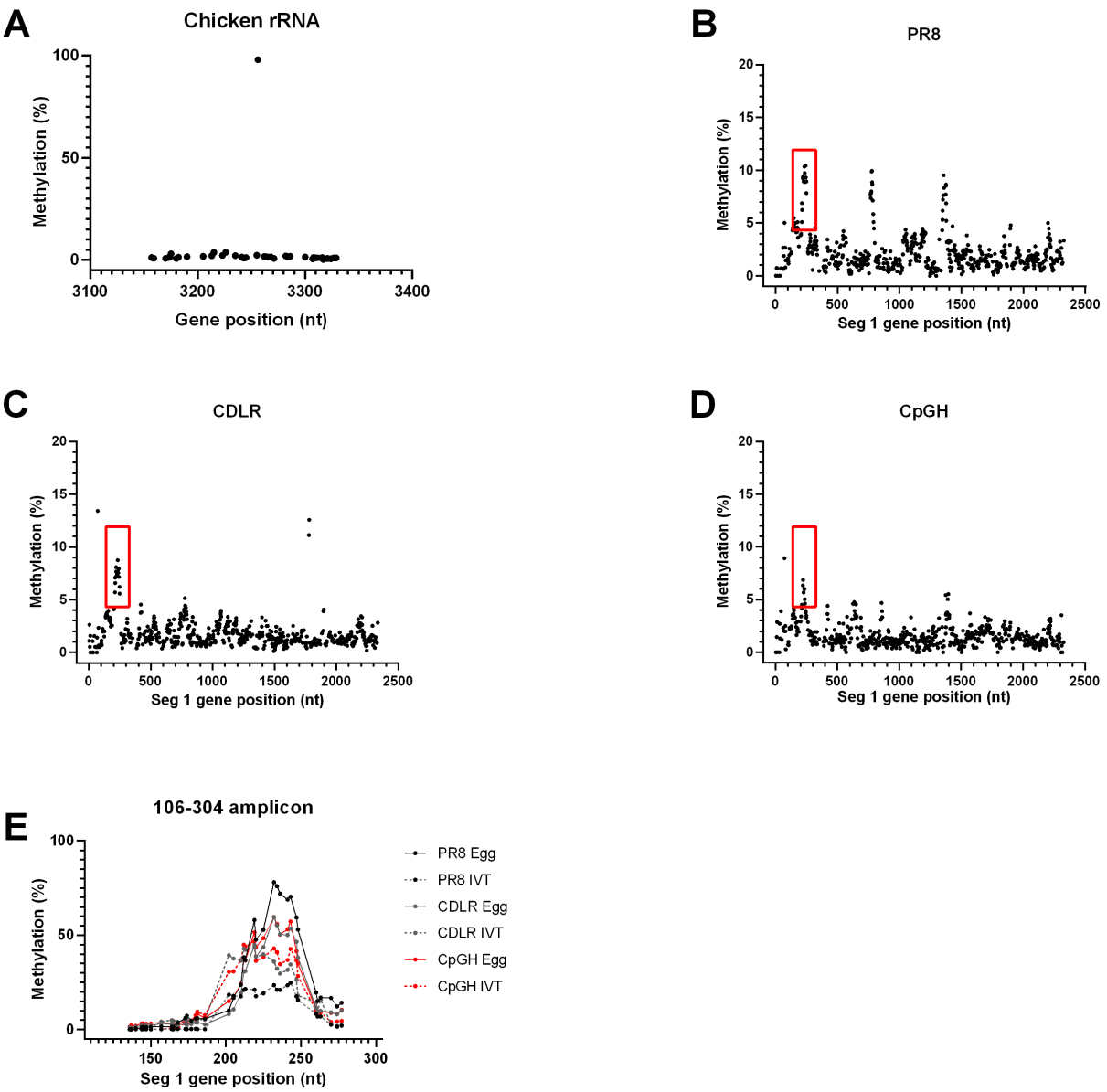


Figure 2. CpG enrichment does not affect packaging but may alter electrophoretic mobility. A. Virus stocks of the seg1 virus panel were titred by plaque assay, and then the relative copy numbers of segment 1 and segment 5 were determined by qPCR. RNA:PFU ratios were then calculated. A '4c6c' mutant virus (54) was used as a packaging mutant control. To visualise packaged viral RNAs, Viral stocks were purified by ultracentrifugation and subjected to urea-PAGE, thereby separating the 8 segments of the viral genome **(B)**. A shift in the electrophoretic mobility of segment 1 of the CpGH virus was observed (large black arrow), and in order to determine whether this was due to RNA modification, RNAs generated through *in vitro* transcription (i.e. a cell-free system, free from RNA modifying enzymes) were run under the same conditions **(C)**. For 'mix' samples, equimolar combinations of CDLR and CpGH constructs were combined and run to demonstrate whether a mobility shift was apparent in CpGH constructs.



806

807 **Figure 3. CpG enrichment of viral RNA does not alter the methylation profile.** Segment 1 modified viruses were processed
808 through bisulphite conversion reactions and then deep sequenced to determine the frequency of methylation in unmodified versus
809 CpG-enriched transcripts. **A.** As virus reconstitution was performed in embryonated hens' eggs, chicken ribosomal RNA was
810 sequenced from viral stocks as a positive control due to a known methylation site in this RNA. **B.** 5-methylcytosine signal intensity
811 across PR8 segment 1 vRNA. **C.** 5-methylcytosine signal intensity across CDLR segment 1 vRNA. **D.** 5-methylcytosine signal
812 intensity across CDLR segment 1 vRNA. **E.** 5-methylcytosine signal intensity across PCR amplicons of either vRNA (PR8 egg/
813 CDLR egg/ CpGH egg) or *in vitro* transcribed (IVT) RNA, across segment 1 positions 106-304. This region was selected due to
814 the spike in signal observed across segment 1 vRNAs for the segment 1 virus panel (B-D, red boxes). For PR8 whole genome
815 methylation plots please refer to **Fig. S4**.

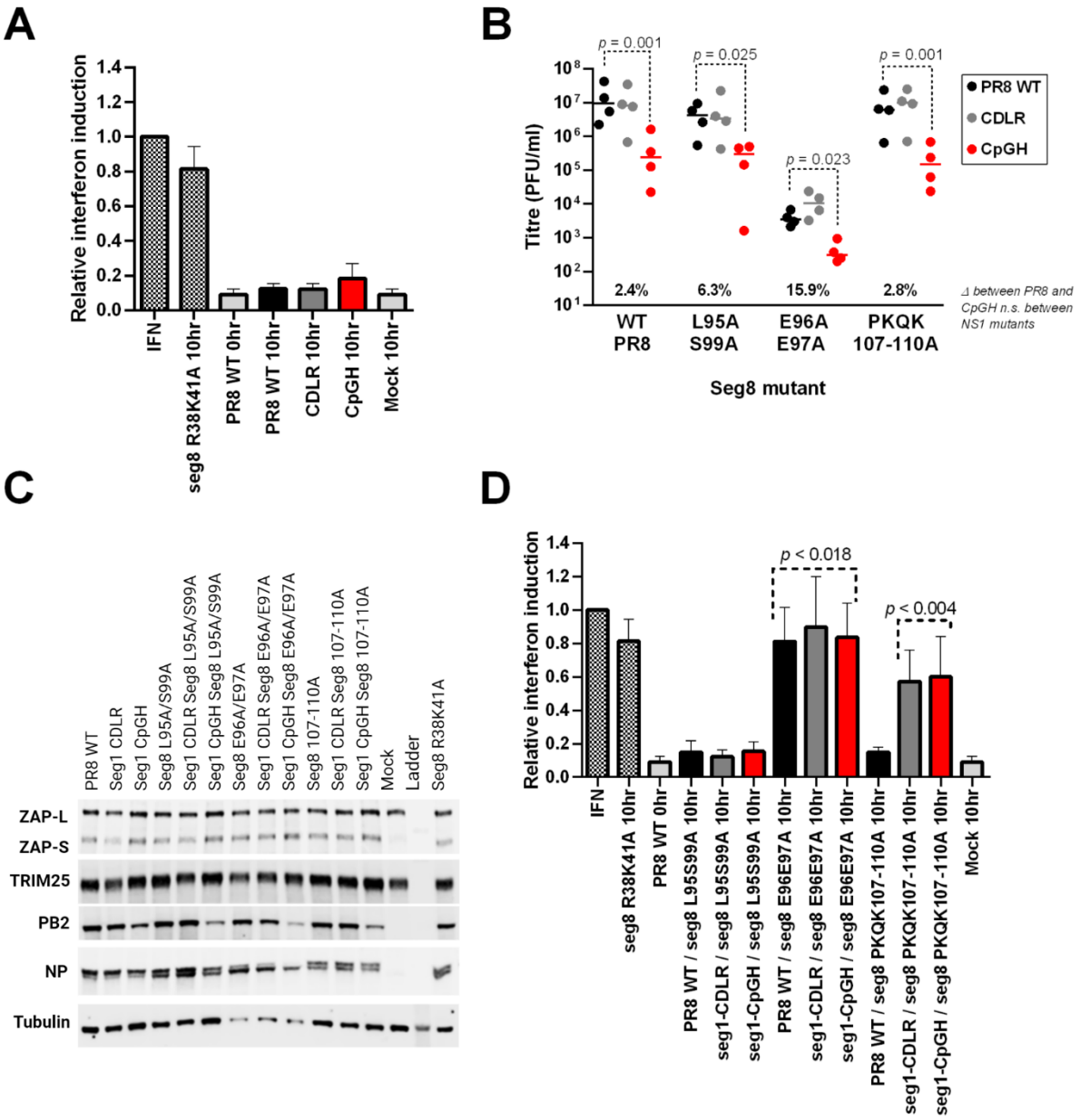


Figure 4. Interaction of CpG-enriched viruses with the IFN pathway. A. A549 cells were infected with WT PR8, CDLR or CpGH viruses at MOI 10 for 24 hours, and then type I IFN induction was quantified by HEK-Blue assay. Titrated IFN was used as a standard, and an IAV mutated in segment 8 to inhibit IFN induction (R38K41A) was used as a positive control. **B.** IAV WT PR8, CDLR or CpGH viruses were further mutated in segment 8 to abrogate the ZAP/TRIM25 targeting activity of NS1 encoded on this segment. Mutations introduced into segment 8 were L95S99A, reportedly defective in TRIM25 targeting (50), E96/97A, reportedly defective in ZAP targeting (51), and PKQK107-110A, reportedly defective in ZAP targeting (51). Seg1/8 mutant viruses were used to infect A549 cells at MOI 0.01, for 48h, and viral titres were quantified. **C.** Western blots examining ZAP and TRIM25 expression in the same infections. **D.** The seg1/8 mutant virus panel was used to infect A549 cells at high MOI for 10 hours, and then IFN induction was quantified by HEK-Blue assay (A640 read-out).

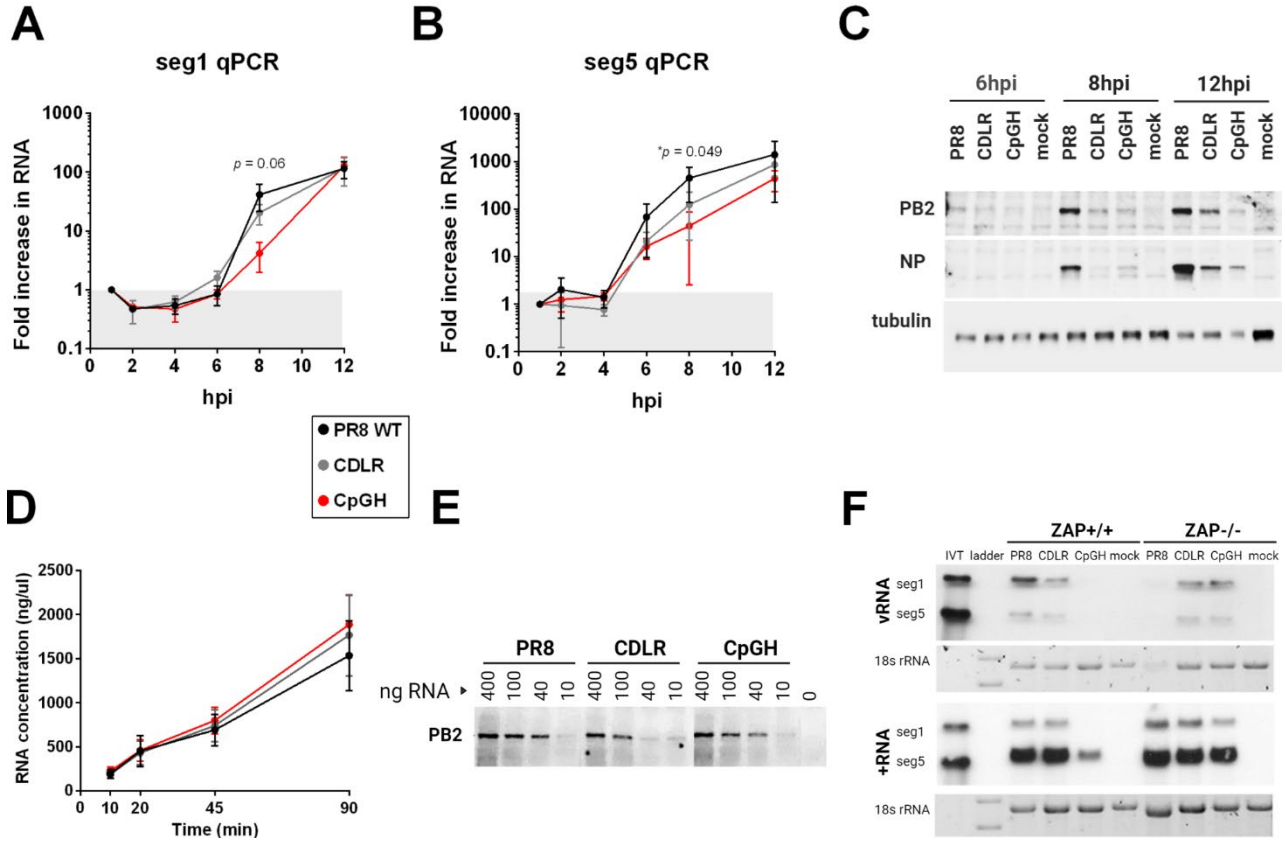


Figure 5. CpG enrichment in the IAV genome results in reduced transcriptional levels during virus replication. Transcription and translation were investigated in the context of infection using high MOI time courses. The virus panel was analysed for segment 1 (**A**) or segment 5 (**B**) RNA production. **C.** For assessment of protein production, PB2 (segment 1) and NP (segment 5) were assayed by western blot. **D.** Equal amounts of segment 1 WT (PR8), permuted control (CDLR) or CpG-enriched (CpGH) amplicons under T7 promoters were used in *in vitro* transcription assays to assess whether synonymous recoding impacted transcriptional efficiency in a cell free system. **E.** The efficiency of translation of RNAs was then assessed by titrating the synthesised RNAs in *in vitro* translation assays. **F**. A549 WT or ZAP-/- cells were infected with either WT PR8 or CpGH virus at an MOI of 15 for 8 hours, after which time RNA was harvested and probed by northern blotting using specific primers to detect segment 1 and segment 5 vRNA or positive polarity RNA.

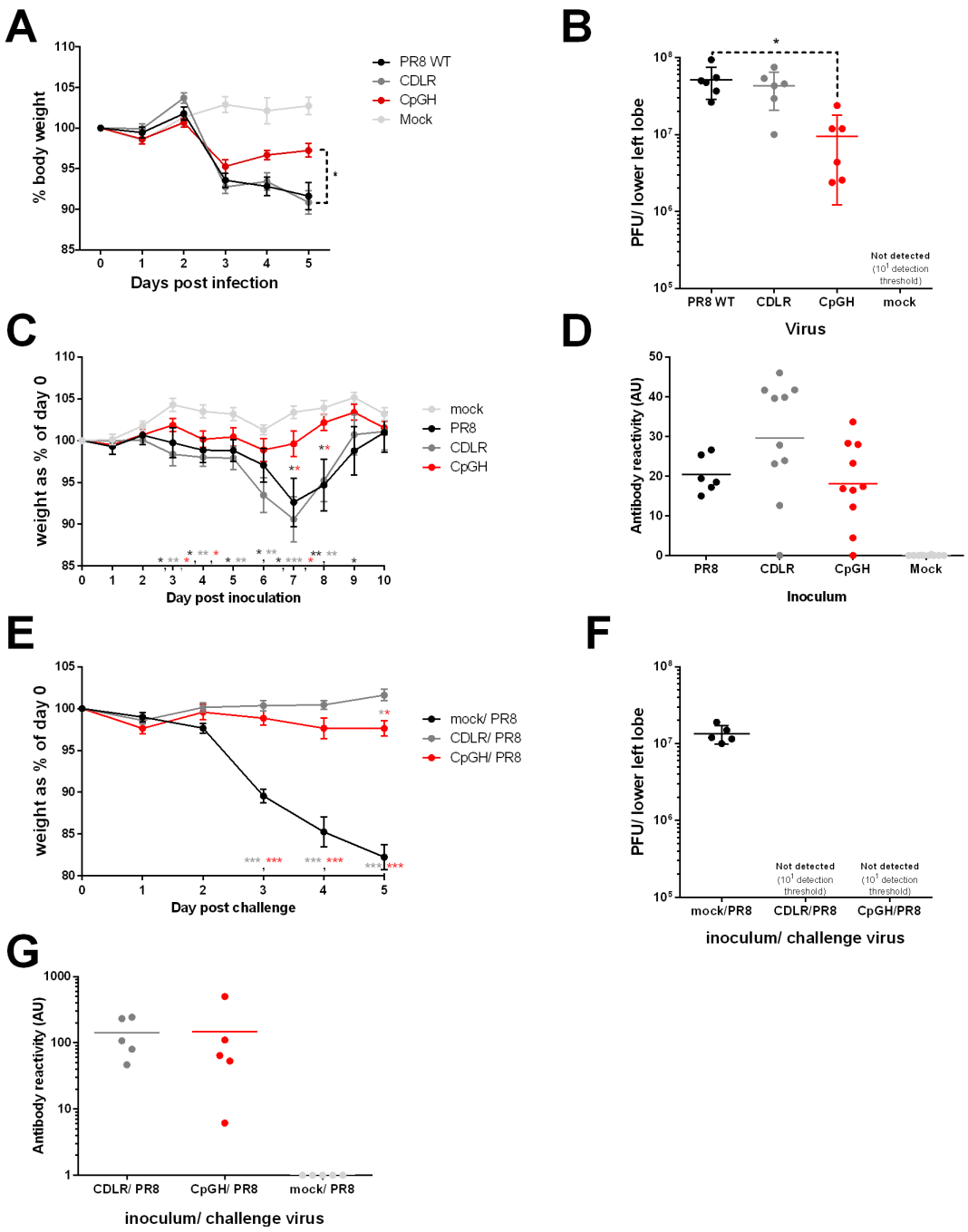


Figure 6. CpGH IAV is attenuated in mice and conveys protection from challenge with a lethal dose of WT virus. Groups of six, 6 week-old female BALB/c mice were infected with 200 PFU of either WT PR8 virus, CDLR permuted control, CpGH virus or mock infected (*n* = 2) for 5 days; weight change measured daily (**A**). On day 5, lower left lung lobes were harvested and the presence of infectious virus was quantified (**B**). To determine the protective effect of CpGH virus infection, groups of 6-10 mice were infected with a low dose (20 PFU) of either PR8 (n = 6), CDLR control (n = 10), CpGH virus (n = 10) or were mock infected (n = 10) and weighed daily for 10 days until full recovery (**C**). At day 20 post-exposure, tail bleeds were used to collect sera for ELISA to quantify anti-IAV antibody levels (**D**). The next day, five mice exposed to either CDLR or CpGH virus, or mock exposed, were challenged with a potentially lethal dose of PR8 WT virus (200 PFU) and weighed daily for 5 days (**E**) at which time mice were culled and lower left lung lobes were collected for virus titration (**F**) and antibody levels in sera were quantified by ELISA (**G**). Sera were also tested for IAV protein binding using western blotting (**Fig. S5**).

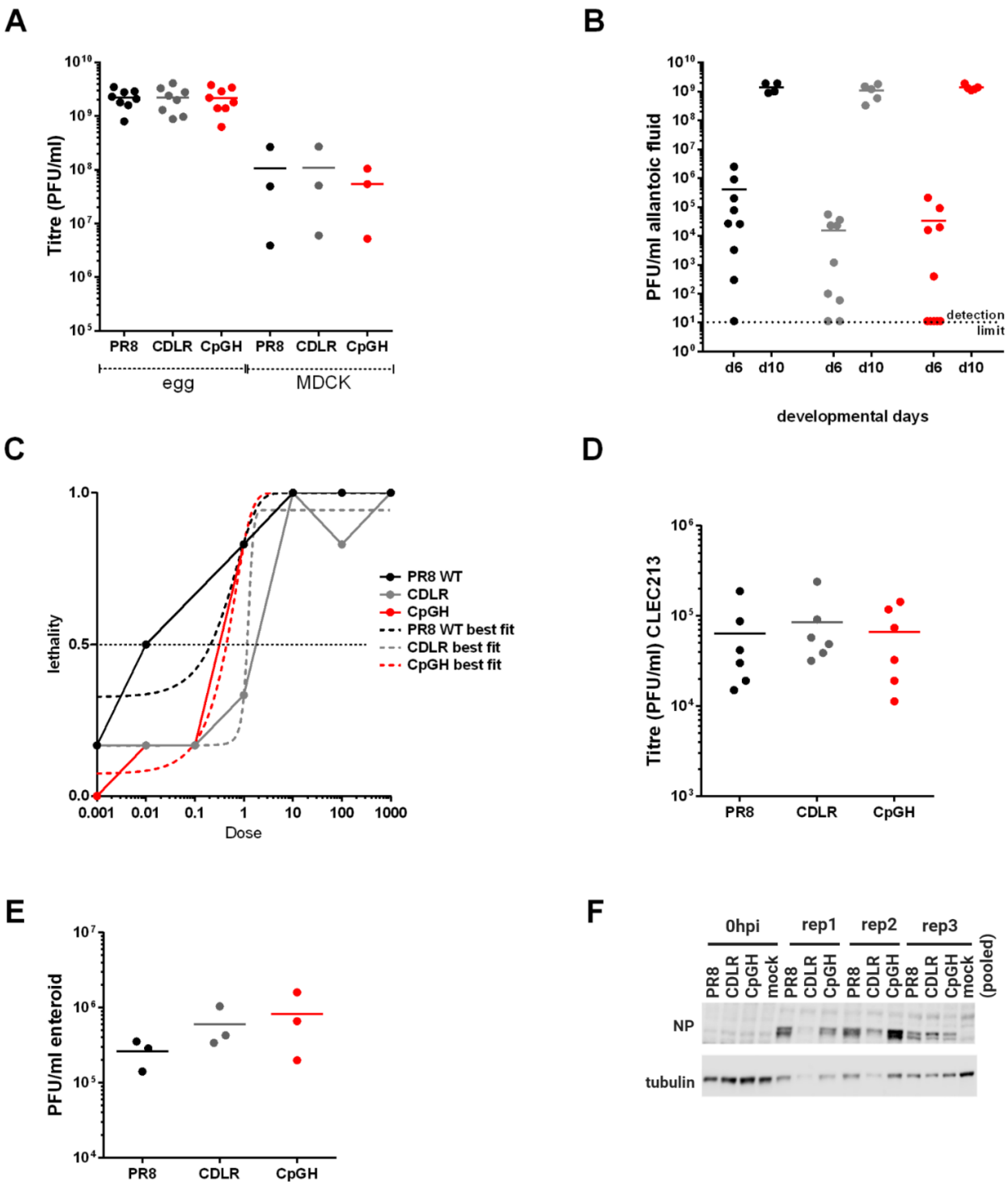


Figure 7. A ZAP-sensitive attenuated CpG-enriched IAV replicates equivalently to WT virus in vaccine propagation systems. A. PR8 WT, CDLR and CpGH viruses were grown in embryonated hens' eggs or in MDCK cells and titred. **B.** Embroyonated hens' eggs were inoculated with 100 PFU virus at developmental day 5 or day 9, and viral titre in allantoic fluid was assessed 24 hours later. **C.** Lethality of the virus panel in eggs was fit to nonlinear regression curves for LD50 calculations. **D.** The virus panel was grown in chicken epithelial cell line CLEC213 and viral titres were quantified by plaque assay. **E.** The virus panel was used to infect chicken enteroids at an MOI of 10 for 24 hours, and the production of infectious virus, and viral protein production **(F)** was assayed. For confocal imaging of infected enteroids, please refer to **Fig. S8**. For innate gene expression analyses from infected enteroids, please refer to **Fig. S9**.

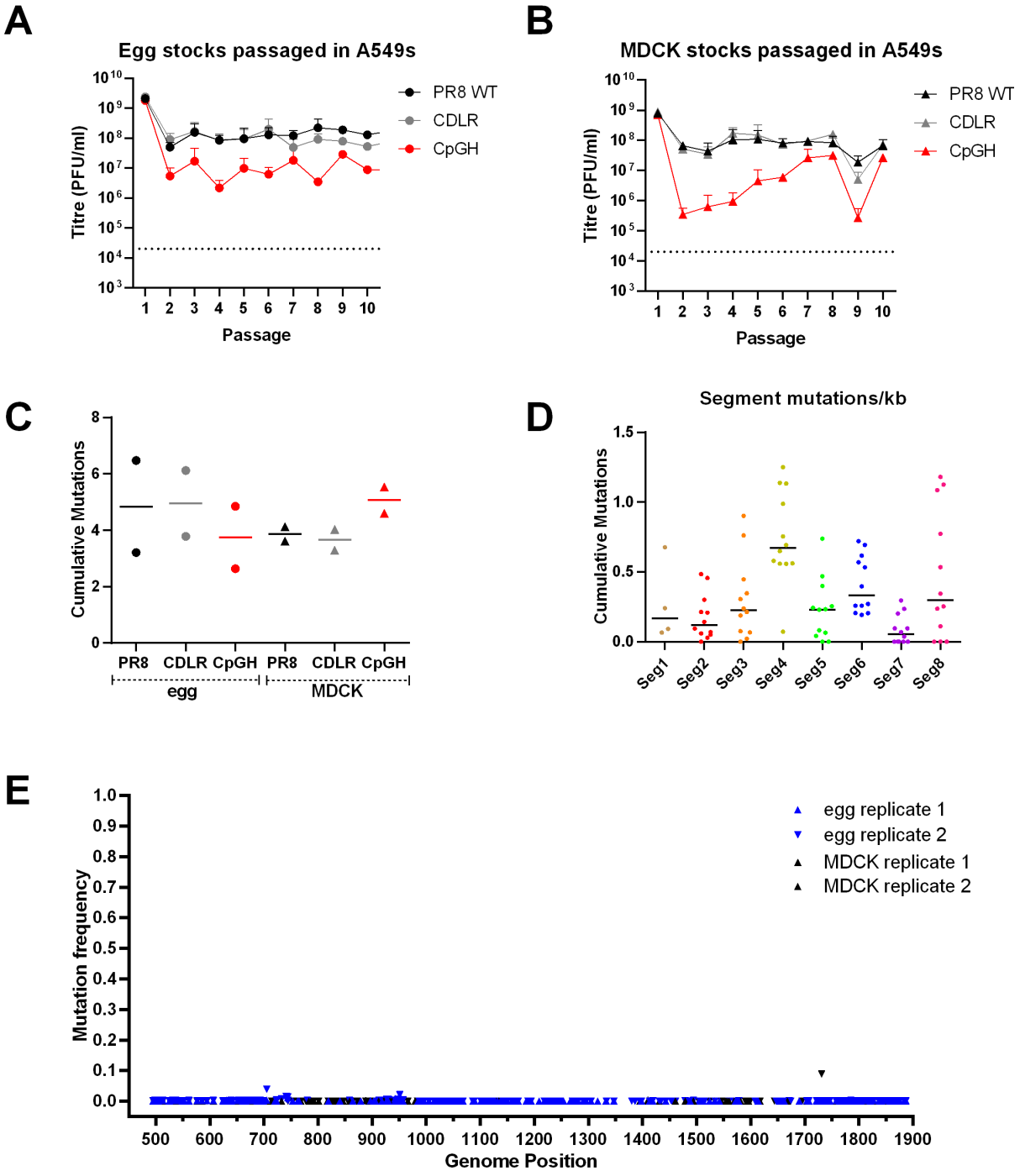


Figure 8. Serial passage of CpG modified IAV does not enable deselection of CpG motifs. Egg **(A)** or MDCK **(B)** derived virus stocks were passaged at an MOI of 0.01 ten times in A549 cells, and titred after each passage (*n* = 2). After ten passages, viruses were whole genome sequenced to test for reversion or epistatic compensatory mutations. (**C**). Total cumulative genome mutations after ten passages were calculated, with proportional representation of mutations that had not become fixed (e.g. if 10% of sequences contained a mutation at a given position, this scored as 0.1 mutations) calculated for egg virus stock passages and MDCK virus stock passages. (**D**). Cumulative mutations were calculated for each segment for each of the three viruses in the panel across four replicates (**E**). Mutation frequency at recoded positions across segment 1 for CpGH virus. The mutation frequencies of the four serially passaged replicates are overlaid. No mutations at CpG sites were observed. For sequence coverage, please refer to **Fig. S10** and for full genome mutational profiles to **Fig. S11**.

866 **SUPPLEMENTARY TABLES AND FIGURES**

867 **Table S1. Segment 1 modified PR8 sequences.**

868 >A/human/PuertoRico/8_seg1_wildtype

869 AGCGAAAGCAGGTCAATTATATTCAATATGGAAAGAATAAAAGAACTAAGAAATCTAATGTCGCAGTCTCGCACCCGCGAG
870 ATACTCACAAAAACCACCGTGGACCATATGGCCATAATCAAGAAGTACACATCAGGAAGACAGGAGAAGAACCCAGCACTT
871 AGGATGAAATGGATGATGGCAATGAAATATCCAATTACAGCAGACAAGAGGATAACGGAAATGATTCCTGAGAGAAATGAG
872 CAAGGACAAACTTTATGGAGTAAAATGAATGATGCCGGATCAGACCGAGTGATGGTATCACCACTGGCTGTGACATGGTG
873 GAATAGGAATGGACCAATAACAAATACAGTTCATTATCCAAAAATCTACAAAACTTATTTTGAAAGAGTCGAAAGGCTAAAG
874 CATGGAACCTTTGGCCCTGTCCATTTTAGAAACCAAGTCAAAATACGTCGGAGAGTTGACATAAATCCTGGTCATGCAGAT
875 CTCAGTGCCAAGGAGGCACAGGATGTAATCATGGAAGTTGTTTTCCCTAACGAAGTGGGAGCCAGGATACTAACATCGGA
876 ATCGCAACTAACGATAACCAAAGAGAAGAAAGAAGAACTCCAGGATTGCAAAATTTCTCCTTTGATGGTTGCATACATGTTG
877 GAGAGAGAACTGGTCCGCAAAACGAGATTCCTCCCAGTGGCTGGTGGAACAAGCAGTGTGTACATTGAAGTGTTGCATTT
878 GACTCAAGGAACATGCTGGGAACAGATGTATACTCCAGGAGGGGAAGTGAGGAATGATGATGTTGATCAAAGCTTGATTAT
879 TGCTGCTAGGAACATAGTGAGAAGAGCTGCAGTATCAGCAGATCCACTAGCATCTTTATTGGAGATGTGCCACAGCACACA
880 GATTGGTGGAATTAGGATGGTAGACATCCTTAGGCAGAACCCAACAGAAGAGCAAGCCGTGGATATATGCAAGGCTGCAA
881 TGGGACTGAGAATTAGCTCATCCTTCAGTTTTGGTGGATTCACATTTAAGAGAACAAGCGGATCATCAGTCAAGAGAGAGG
882 AAGAGGTGCTTACGGGCAATCTTCAAACATTGAAGATAAGAGTGCATGAGGGATATGAAGAGTTCACAATGGTTGGGAGAA
883 GAGCAACAGCCATACTCAGAAAAGCAACCAGGAGATTGATTCAGCTGATAGTGAGTGGGAGAGACGAACAGTCGATTGCC
884 GAAGCAATAATTGTGGCCATGGTATTTTCACAAGAGGATTGTATGATAAAAGCAGTCAGAGGTGATCTGAATTTCGTCAATA
885 GGGCGAATCAGCGATTGAATCCTATGCATCAACTTTTAAGACATTTTCAGAAGGATGCGAAAGTGCTTTTTCAAAATTGGGG
886 AGTTGAACCTATCGACAATGTGATGGGAATGATTGGGATATTGCCCGACATGACTCCAAGCATCGAGATGTCAATGAGAGG
887 AGTGAGAATCAGCAAAATGGGTGTAGATGAGTACTCCAGCACGGAGAGGGTAGTGGTGAGCATTGACCGTTTTTTGAGAA
888 TCCGGGACCAACGAGGAAATGTACTACTGTCTCCCGAGGAGGTCAGTGAAACACAGGGAACAGAGAAACTGACAATAACT
889 TACTCATCGTCAATGATGTGGGAGATTAATGGTCCTGAATCAGTGTTGGTCAATACCTATCAATGGATCATCAGAAACTGGG
890 AAACTGTTAAAATTCAGTGGTCCCAGAACCCTACAATGCTATACAATAAAATGGAATTTGAACCATTTCAGTCTTTAGTACCT
891 AAGGCCATTAGAGGCCAATACAGTGGGTTTGTAAGAACTCTGTTCCAACAAATGAGGGATGTGCTTGGGACATTTGATACC
892 GCACAGATAATAAAACTTCTTCCCTTCGCAGCCGCTCCACCAAAGCAAAGTAGAATGCAGTTCTCCTCATTTACTGTGAATG
893 TGAGGGGATCAGGAATGAGAATACTTGTAAGGGGCAATTCTCCTGTATTCAACTATAACAAGGCCACGAAGAGACTCACAG
894 TTCTCGGAAAGGATGCTGGCACTTTAACTGAAGACCCAGATGAAGGCACAGCTGGAGTGGAGTCCGCTGTTCTGAGGGGA
895 TTCCTCATTCTGGGCAAAGAGAACAAGAGATATGGGCCAGCACTAAGCATCAATGAACTGAGCAACCTTGCGAAAGGAGA
896 GAAGGCTAATGTGCTAATTGGGCAAGGAGACGTGGTGTTGGTAATGAAACGGAAACGGGACTCTAGCATACTTACTGACA
897 GCCAGACAGCGACCAAAAGAATTCGGATGGCCATCAATTAGTGTCGAATAGTTTAAAAACGACCTTGTTTCTACT

898 > A/human/PuertoRico/8_seg1_CDLR

899 AGCGAAAGCAGGTCAATTATATTCAATATGGAAAGAATAAAAGAACTAAGAAATCTAATGTCGCAGTCTCGCACCCGCGAG
900 ATACTCACAAAAACCACCGTGGACCATATGGCCATAATCAAGAAGTACACATCAGGAAGACAGGAGAAGAACCCAGCACTT
901 AGGATGAAATGGATGATGGCAATGAAATATCCAATTACAGCAGACAAGAGGATAACGGAAATGATTCCTGAGAGAAATGAG

902 CAAGGACAAACTTTATGGAGTAAAATGAATGATGCCGGATCAGACCGAGTGATGGTATCACCACTGGCTGTGACATGGTG
903 GAATAGGAATGGACCAATAACAAATACAGTTCATTATCCAAAAATCTACAAAACTTATTTTGAAAGAGTCGAAAGGCTAAAG
904 CATGGAACCTTTGGCCCTGTCCATTTTAGAAACCAAGTCAAAATACGTCGGAGAGTTGACATAAATCCTGGTCATGCAGAT
905 CTCAGTGCCAAGGAGGCTCAGGATGTAATCATGGAAGTGGTGTTCCCAAACGAGGTTGGAGCAAGAATCCTCACTTCTGA
906 GTCGCAACTAACAATAACTAAAGAGAAAAAAGAGGAGCTACAAGATTGCAAAATATCACCTTTGATGGTTGCATACATGTTG
907 GAGAGAGAACTGGTCCGAAAGACGAGATTCCTACCCGTGGCCGGTGGGACCAGCAGTGTATACATAGAAGTGTTGCATTT
908 GACTCAAGGGACGTGCTGGGAACAGATGTATACACCAGGGGGGGAAGTGAGAAATGATGATGTTGACCAGAGCTTGATAA
909 TTGCCGCGAGGAACATTGTGAGAAGAGCCGCCGTATCGGCAGATCCTCTAGCATCCTTATTGGAAATGTGTCACAGCACA
910 CAGATTGGTGGAATAAGAATGGTTGACATCCTTAGACAGAATCCCACAGAAGAGCAGGCTGTGGATATTTGCAAGGCTGCC
911 ATGGGGCTGAGAATCAGTTCATCATTTAGTTTTGGTGGATTCACTTTCAAGAGAACAAGCGGGTCATCAGTCAAGAGAGAG
912 GAAGAAGTGCTAACGGGGAATCTCCAAACATTGAAAATAAGAGTTCATGAGGGCTATGAAGAGTTCACGATGGTTGGGAG
913 GAGAGCAACCGCCATACTTAGAAAAGCAACGAGGAGATTGATACAACTGATAGTGAGTGGCAGAGACGAGCAATCTATTG
914 CAGAAGCAATAATTGTGGCTATGGTGTTTTCCCAGGAAGATTGTATGATTAAAGCAGTCAGAGGAGACCTGAACTTTGTTAA
915 TAGGGCAAATCAACGATTGAACCCTATGCATCAACTTTTAAGACATTTTCAAAAGGATGCAAAAGTGCTTTTTCAAAATTGG
916 GGAGTTGAACCAATCGACAATGTGATGGGCATGATAGGAATTTTGCCAGACATGACTCCAAGCATCGAGATGTCAATGAGA
917 GGTGTCAGAATCAGCAAAATGGGTGTAGATGAATACTCGAGCACTGAGAGGGTAGTGGTCAGCATTGACCGATTTTTGAGA
918 ATTCGGGATCAACGAGGGAATGTTCTTCTGTCACCGGAAGAGGTGAGTGAAACTCAGGGAACTGAGAAACTGACAATAAC
919 ATATTCCTCATCAATGATGTGGGAGATTAATGGCCCAGAATCAGTATTGGTGAATACATATCAATGGATCATCAGAAACTGG
920 GAGACTGTTAAAATTCAGTGGTCTCAGAATCCCACTATGCTATACAATAAAATGGAGTTCGAGCCCTTCCAATCTTTAGTGC
921 CTAAGGCCATTAGAGGCCAGTACAGTGGATTCGTAAGAACACTGTTCCAACAGATGAGAGATGTACTTGGGACATTTGATA
922 CAGCACAGATTATAAAACTGCTTCCATTTGCAGCCGCTCCACCAAAGCAAAGTAGAATGCAGTTCTCCTCATTTACTGTGAA
923 TGTGAGGGGATCAGGAATGAGAATACTTGTAAGGGGCAATTCTCCTGTATTCAACTATAACAAGGCCACGAAGAGACTCAC
924 AGTTCTCGGAAAGGATGCTGGCACTTTAACTGAAGACCCAGATGAAGGCACAGCTGGAGTGGAGTCCGCTGTTCTGAGGG
925 GATTCCTCATTCTGGGCAAAGAGAACAAGAGATATGGGCCAGCACTAAGCATCAATGAACTGAGCAACCTTGCGAAAGGA
926 GAGAAGGCTAATGTGCTAATTGGGCAAGGAGACGTGGTGTTGGTAATGAAACGGAAACGGGACTCTAGCATACTTACTGA
927 CAGCCAGACAGCGACCAAAAGAATTCGGATGGCCATCAATTAGTGTCGAATAGTTTAAAAACGACCTTGTTTCTACT

928 > A/human/PuertoRico/8_seg1_CpG-high

929 AGCGAAAGCAGGTCAATTATATTCAATATGGAAAGAATAAAAGAACTAAGAAATCTAATGTCGCAGTCTCGCACCCGCGAG
930 ATACTCACAAAAACCACCGTGGACCATATGGCCATAATCAAGAAGTACACATCAGGAAGACAGGAGAAGAACCCAGCACTT
931 AGGATGAAATGGATGATGGCAATGAAATATCCAATTACAGCAGACAAGAGGATAACGGAAATGATTCCTGAGAGAAATGAG
932 CAAGGACAAACTTTATGGAGTAAAATGAATGATGCCGGATCAGACCGAGTGATGGTATCACCACTGGCTGTGACATGGTG
933 GAATAGGAATGGACCAATAACAAATACAGTTCATTATCCAAAAATCTACAAAACTTATTTTGAAAGAGTCGAAAGGCTAAAG
934 CATGGAACCTTTGGCCCTGTCCATTTTAGAAACCAAGTCAAAATACGTCGGAGAGTTGACATAAATCCTGGTCATGCAGAT
935 CTCAGTGCAAAAGAAGCTCAAGATGTGATCATGGAAGTAGTTTTTCCGAATGAAGTCGGCGCGCGAATACTAACGTCGGAA
936 TCGCAACTGACGATAACGAAAGAGAAAAAAGAAGAACTTCAGGATTGCAAAATTTCGCCATTAATGGTCGCGTATATGTTG
937 GAACGCGAACTCGTTCGAAAAACGCGATTTCTTCCGGTTGCCGGAGGAACGAGCAGCGTGTACATAGAAGTATTACATTTG
938 ACGCAAGGAACGTGTTGGGAACAAATGTATACGCCGGGCGGCGAAGTTCGAAACGATGACGTAGATCAAAGTTTGATTAT

939 AGCGGCTCGAAATATCGTGCGACGCGCGGCGGTTTCCGCCGATCCGCTCGCGTCATTATTGGAAATGTGTCACAGCACGC
940 AGATAGGAGGAATCCGAATGGTCGACATTCTACGGCAAAATCCGACGGAAGAACAAGCAGTCGACATTTGCAAGGCCGCG
941 ATGGGACTGCGAATCAGTTCGTCTTTTAGTTTTGGCGGATTCACGTTCAAACGTACGAGTGGATCGTCAGTGAAGCGCGAA
942 GAAGAAGTTCTGACGGGAAATCTTCAAACATTGAAAATCCGAGTGCATGAAGGATACGAAGAATTCACTATGGTCGGTCGA
943 CGCGCGACAGCGATACTACGAAAAGCGACGCGTCGATTAATTCAGCTGATAGTTAGCGGACGGGACGAACAGTCGATCGC
944 CGAAGCGATCATCGTAGCGATGGTGTTTTCGCAAGAAGATTGCATGATAAAAGCCGTGCGAGGCGATCTAAATTTCGTAAA
945 TCGAGCGAATCAACGATTGAATCCGATGCATCAACTTTTGCGACATTTTCAGAAAGACGCGAAAGTGCTTTTTCAAAATTGG
946 GGAGTCGAACCGATCGACAACGTAATGGGAATGATAGGTATTTTACCCGACATGACGCCGAGCATCGAAATGTCGATGCG
947 CGGCGTTCGAATCAGCAAAATGGGCGTCGACGAATATTCAAGCACAGAGCGAGTCGTTGTAAGCATTGATCGTTTTTTACG
948 AATACGCGATCAACGAGGTAACGTTCTTCTTTCGCCGGAAGAAGTAAGTGAAACGCAAGGTACGGAGAAACTTACGATCAC
949 GTATTCGTCGTCAATGATGTGGGAGATCAACGGTCCAGAATCCGTCTTGGTAAACACGTATCAATGGATCATCCGAAATTG
950 GGAAACAGTGAAAATTCAGTGGTCGCAAAATCCGACGATGCTTTACAACAAAATGGAATTCGAACCGTTTCAATCATTAGTT
951 CCAAAAGCGATTCGCGGTCAATATAGTGGATTTGTTCGAACGCTGTTTCAACAAATGCGTGACGTACTCGGAACGTTCGAC
952 ACCGCGCAAATTATTAAACTTCTACCGTTCGCAGCCGCTCCACCAAAGCAAAGTAGAATGCAGTTCTCCTCATTTACTGTGA
953 ATGTGAGGGGATCAGGAATGAGAATACTTGTAAGGGGCAATTCTCCTGTATTCAACTATAACAAGGCCACGAAGAGACTCA
954 CAGTTCTCGGAAAGGATGCTGGCACTTTAACTGAAGACCCAGATGAAGGCACAGCTGGAGTGGAGTCCGCTGTTCTGAGG
955 GGATTCCTCATTCTGGGCAAAGAGAACAAGAGATATGGGCCAGCACTAAGCATCAATGAACTGAGCAACCTTGCGAAAGG
956 AGAGAAGGCTAATGTGCTAATTGGGCAAGGAGACGTGGTGTTGGTAATGAAACGGAAACGGGACTCTAGCATACTTACTG
957 ACAGCCAGACAGCGACCAAAAGAATTCGGATGGCCATCAATTAGTGTCGAATAGTTTAAAAACGACCTTGTTTCTACT

958 > A/human/PuertoRico/8_seg1_5'-CpGH

959 AGCGAAAGCAGGTCAATTATATTCAATATGGAAAGAATAAAAGAACTAAGAAATCTAATGTCGCAGTCTCGCACCCGCGAG
960 ATACTCACAAAAACCACCGTGGACCATATGGCCATAATCAAGAAGTACACATCAGGAAGACAGGAGAAGAACCCAGCACTT
961 AGGATGAAATGGATGATGGCAATGAAATATCCAATTACAGCAGACAAGAGGATAACGGAAATGATTCCTGAGAGAAATGAG
962 CAAGGACAAACTTTATGGAGTAAAATGAATGATGCCGGATCAGACCGAGTGATGGTATCACCACTGGCTGTGACATGGTG
963 GAATAGGAATGGACCAATAACAAATACAGTTCATTATCCAAAAATCTACAAAACTTATTTTGAAAGAGTCGAAAGGCTAAAG
964 CATGGAACCTTTGGCCCTGTCCATTTTAGAAACCAAGTCAAAATACGTCGGAGAGTTGACATAAATCCTGGTCATGCAGAT
965 CTCAGTGCAAAAGAAGCTCAAGATGTGATCATGGAAGTAGTTTTTCCGAATGAAGTCGGCGCGCGAATACTAACGTCGGAA
966 TCGCAACTGACGATAACGAAAGAGAAAAAAGAAGAACTTCAGGATTGCAAAATTTCGCCATTAATGGTCGCGTATATGTTG
967 GAACGCGAACTCGTTCGAAAAACGCGATTTCTTCCGGTTGCCGGAGGAACGAGCAGCGTGTACATAGAAGTATTACATTTG
968 ACGCAAGGAACGTGTTGGGAACAAATGTATACGCCGGGCGGCGAAGTTCGAAACGATGACGTAGATCAAAGTTTGATTAT
969 AGCGGCTCGAAATATCGTGCGACGCGCGGCGGTTTCCGCCGATCCGCTCGCGTCATTATTGGAAATGTGTCACAGCACGC
970 AGATAGGAGGAATCCGAATGGTCGACATTCTACGGCAAAATCCGACGGAAGAACAAGCAGTCGACATTTGCAAGGCCGCG
971 ATGGGACTGCGAATCAGTTCGTCTTTTAGTTTTGGCGGATTCACGTTCAAACGTACGAGTGGATCGTCAGTGAAGCGCGAA
972 GAAGAAGTTCTGACGGGAAATCTTCAAACATTGAAAATCCGAGTGCATGAAGGATACGAAGAATTCACTATGGTCGGTCGA
973 CGCGCGACAGCGATACTACGAAAAGCGACGCGTCGATTAATTCAGCTGATAGTTAGCGGACGGGACGAACAGTCGATCGC
974 CGAAGCGATCATCGTAGCGATGGTGTTTTCGCAAGAAGATTGCATGATAAAAGCCGTGCGAGGCGATCTAAATTTCGTAAA
975 TCGAGCGAATCAACGATTGAATCCGATGCATCAACTTTTAAGACATTTTCAGAAGGATGCGAAAGTGCTTTTTCAAAATTGG

976 GGAGTTGAACCTATCGACAATGTGATGGGAATGATTGGGATATTGCCCGACATGACTCCAAGCATCGAGATGTCAATGAGA
977 GGAGTGAGAATCAGCAAAATGGGTGTAGATGAGTACTCCAGCACGGAGAGGGTAGTGGTGAGCATTGACCGTTTTTTGAG
978 AATCCGGGACCAACGAGGAAATGTACTACTGTCTCCCGAGGAGGTCAGTGAAACACAGGGAACAGAGAAACTGACAATAA
979 CTTACTCATCGTCAATGATGTGGGAGATTAATGGTCCTGAATCAGTGTTGGTCAATACCTATCAATGGATCATCAGAAACTG
980 GGAAACTGTTAAAATTCAGTGGTCCCAGAACCCTACAATGCTATACAATAAAATGGAATTTGAACCATTTCAGTCTTTAGTAC
981 CTAAGGCCATTAGAGGCCAATACAGTGGGTTTGTAAGAACTCTGTTCCAACAAATGAGGGATGTGCTTGGGACATTTGATA
982 CCGCACAGATAATAAAACTTCTTCCCTTCGCAGCCGCTCCACCAAAGCAAAGTAGAATGCAGTTCTCCTCATTTACTGTGAA
983 TGTGAGGGGATCAGGAATGAGAATACTTGTAAGGGGCAATTCTCCTGTATTCAACTATAACAAGGCCACGAAGAGACTCAC
984 AGTTCTCGGAAAGGATGCTGGCACTTTAACTGAAGACCCAGATGAAGGCACAGCTGGAGTGGAGTCCGCTGTTCTGAGGG
985 GATTCCTCATTCTGGGCAAAGAGAACAAGAGATATGGGCCAGCACTAAGCATCAATGAACTGAGCAACCTTGCGAAAGGA
986 GAGAAGGCTAATGTGCTAATTGGGCAAGGAGACGTGGTGTTGGTAATGAAACGGAAACGGGACTCTAGCATACTTACTGA
987 CAGCCAGACAGCGACCAAAAGAATTCGGATGGCCATCAATTAGTGTCGAATAGTTTAAAAACGACCTTGTTTCTACT

988 > A/human/PuertoRico/8_seg1_3'-CpGH

989 AGCGAAAGCAGGTCAATTATATTCAATATGGAAAGAATAAAAGAACTAAGAAATCTAATGTCGCAGTCTCGCACCCGCGAG
990 ATACTCACAAAAACCACCGTGGACCATATGGCCATAATCAAGAAGTACACATCAGGAAGACAGGAGAAGAACCCAGCACTT
991 AGGATGAAATGGATGATGGCAATGAAATATCCAATTACAGCAGACAAGAGGATAACGGAAATGATTCCTGAGAGAAATGAG
992 CAAGGACAAACTTTATGGAGTAAAATGAATGATGCCGGATCAGACCGAGTGATGGTATCACCACTGGCTGTGACATGGTG
993 GAATAGGAATGGACCAATAACAAATACAGTTCATTATCCAAAAATCTACAAAACTTATTTTGAAAGAGTCGAAAGGCTAAAG
994 CATGGAACCTTTGGCCCTGTCCATTTTAGAAACCAAGTCAAAATACGTCGGAGAGTTGACATAAATCCTGGTCATGCAGAT
995 CTCAGTGCCAAGGAGGCACAGGATGTAATCATGGAAGTTGTTTTCCCTAACGAAGTGGGAGCCAGGATACTAACATCGGA
996 ATCGCAACTAACGATAACCAAAGAGAAGAAAGAAGAACTCCAGGATTGCAAAATTTCTCCTTTGATGGTTGCATACATGTTG
997 GAGAGAGAACTGGTCCGCAAAACGAGATTCCTCCCAGTGGCTGGTGGAACAAGCAGTGTGTACATTGAAGTGTTGCATTT
998 GACTCAAGGAACATGCTGGGAACAGATGTATACTCCAGGAGGGGAAGTGAGGAATGATGATGTTGATCAAAGCTTGATTAT
999 TGCTGCTAGGAACATAGTGAGAAGAGCTGCAGTATCAGCAGATCCACTAGCATCTTTATTGGAGATGTGCCACAGCACACA
1000 GATTGGTGGAATTAGGATGGTAGACATCCTTAGGCAGAACCCAACAGAAGAGCAAGCCGTGGATATATGCAAGGCTGCAA
1001 TGGGACTGAGAATTAGCTCATCCTTCAGTTTTGGTGGATTCACATTTAAGAGAACAAGCGGATCATCAGTCAAGAGAGAGG
1002 AAGAGGTGCTTACGGGCAATCTTCAAACATTGAAGATAAGAGTGCATGAGGGATATGAAGAGTTCACAATGGTTGGGAGAA
1003 GAGCAACAGCCATACTCAGAAAAGCAACCAGGAGATTGATTCAGCTGATAGTGAGTGGGAGAGACGAACAGTCGATTGCC
1004 GAAGCAATAATTGTGGCCATGGTATTTTCACAAGAGGATTGTATGATAAAAGCAGTCAGAGGTGATCTGAATTTCGTCAATA
1005 GGGCGAATCAGCGATTGAATCCTATGCATCAACTTTTGCGACATTTTCAGAAAGACGCGAAAGTGCTTTTTCAAAATTGGG
1006 GAGTCGAACCGATCGACAACGTAATGGGAATGATAGGTATTTTACCCGACATGACGCCGAGCATCGAAATGTCGATGCGC
1007 GGCGTTCGAATCAGCAAAATGGGCGTCGACGAATATTCAAGCACAGAGCGAGTCGTTGTAAGCATTGATCGTTTTTTACGA
1008 ATACGCGATCAACGAGGTAACGTTCTTCTTTCGCCGGAAGAAGTAAGTGAAACGCAAGGTACGGAGAAACTTACGATCACG
1009 TATTCGTCGTCAATGATGTGGGAGATCAACGGTCCAGAATCCGTCTTGGTAAACACGTATCAATGGATCATCCGAAATTGG
1010 GAAACAGTGAAAATTCAGTGGTCGCAAAATCCGACGATGCTTTACAACAAAATGGAATTCGAACCGTTTCAATCATTAGTTC
1011 CAAAAGCGATTCGCGGTCAATATAGTGGATTTGTTCGAACGCTGTTTCAACAAATGCGTGACGTACTCGGAACGTTCGACA
1012 CCGCGCAAATTATTAAACTTCTACCGTTCGCAGCCGCTCCACCAAAGCAAAGTAGAATGCAGTTCTCCTCATTTACTGTGAA

1013 TGTGAGGGGATCAGGAATGAGAATACTTGTAAGGGGCAATTCTCCTGTATTCAACTATAACAAGGCCACGAAGAGACTCAC
1014 AGTTCTCGGAAAGGATGCTGGCACTTTAACTGAAGACCCAGATGAAGGCACAGCTGGAGTGGAGTCCGCTGTTCTGAGGG
1015 GATTCCTCATTCTGGGCAAAGAGAACAAGAGATATGGGCCAGCACTAAGCATCAATGAACTGAGCAACCTTGCGAAAGGA
1016 GAGAAGGCTAATGTGCTAATTGGGCAAGGAGACGTGGTGTTGGTAATGAAACGGAAACGGGACTCTAGCATACTTACTGA
1017 CAGCCAGACAGCGACCAAAAGAATTCGGATGGCCATCAATTAGTGTCGAATAGTTTAAAAACGACCTTGTTTCTACT

1018 > A/human/PuertoRico/8_seg1_CpGM

1019 AGCGAAAGCAGGTCAATTATATTCAATATGGAAAGAATAAAAGAACTAAGAAATCTAATGTCGCAGTCTCGCACCCGCGAG
1020 ATACTCACAAAAACCACCGTGGACCATATGGCCATAATCAAGAAGTACACATCAGGAAGACAGGAGAAGAACCCAGCACTT
1021 AGGATGAAATGGATGATGGCAATGAAATATCCAATTACAGCAGACAAGAGGATAACGGAAATGATTCCTGAGAGAAATGAG
1022 CAAGGACAAACTTTATGGAGTAAAATGAATGATGCCGGATCAGACCGAGTGATGGTATCACCACTGGCTGTGACATGGTG
1023 GAATAGGAATGGACCAATAACAAATACAGTTCATTATCCAAAAATCTACAAAACTTATTTTGAAAGAGTCGAAAGGCTAAAG
1024 CATGGAACCTTTGGCCCTGTCCATTTTAGAAACCAAGTCAAAATACGTCGGAGAGTTGACATAAATCCTGGTCATGCAGAT
1025 CTCAGTGCCAAGGAGGCGCAGGATGTAATCATGGAAGTTGTTTTCCCGAACGAAGTGGGAGCCCGGATACTAACATCTGA
1026 ATCCCAACTAACGATAACGAAAGAGAAGAAAGAAGAACTCCAGGATTGCAAAATTTCGCCTTTGATGGTTGCGTACATGTT
1027 GGAGAGAGAACTCGTCCGCAAAACTCGATTCCTCCCAGTGGCTGGCGGAACAAGCAGCGTGTACATTGAAGTGTTGCATT
1028 TGACGCAAGGAACATGCTGGGAACAGATGTATACGCCAGGAGGGGAAGTGAGGAATGATGACGTTGATCAAAGCTTGATT
1029 ATTGCTGCTCGGAACATAGTGAGAAGAGCTGCGGTATCAGCAGATCCACTAGCGTCTTTATTGGAGATGTGCCACAGCAC
1030 GCAGATTGGTGGAATTAGGATGGTCGACATCCTTAGGCAGAACCCAACGGAAGAGCAAGCTGTCGATATATGCAAGGCTG
1031 CGATGGGACTGAGAATTAGCTCGTCCTTCAGTTTTGGCGGATTCACATTTAAGAGAACAAGCGGATCATCAGTCAAGCGAG
1032 AGGAAGAGGTGCTTACGGGCAATCTTCAAACATTGAAGATACGAGTGCATGAGGGATACGAAGAGTTCACAATGGTTGGG
1033 AGAAGAGCGACAGCCATACTCCGAAAAGCAACCAGGCGATTGATTCAGCTGATAGTGAGCGGGAGAGACGAACAGTCGAT
1034 TGCCGAAGCGATAATTGTGGCCATGGTATTTTCGCAAGAGGATTGTATGATAAAAGCAGTCCGAGGTGATCTGAATTTCGT
1035 CAATAGGGCCAATCAGCGATTGAATCCTATGCATCAACTTTTACGACATTTTCAGAAGGATGCGAAAGTGCTTTTTCAAAAT
1036 TGGGGAGTCGAACCTATTGACAACGTGATGGGAATGATTGGGATATTGCCCGACATGACTCCAAGCATTGAGATGTCGAT
1037 GAGAGGAGTGCGAATCAGCAAAATGGGTGTAGACGAGTACTCCAGCACGGAGAGGGTCGTGGTGAGCATTGACCGTTTTT
1038 TGCGAATCAGGGACCAAAGGGGAAACGTACTACTGTCTCCCGAGGAGGTCAGTGAAACGCAGGGAACAGAGAAACTGAC
1039 AATAACGTACTCATCGTCAATGATGTGGGAGATTAACGGTCCTGAATCAGTGTTGGTCAATACGTATCAATGGATCATCCGA
1040 AACTGGGAAACCGTTAAAATTCAGTGGTCCCAGAACCCTACGATGCTATACAATAAAATGGAATTCGAACCGTTTCAGTCTT
1041 TAGTACCTAAGGCCATTCGAGGCCAATACAGTGGGTTTGTAAGAACGCTGTTCCAACAAATGAGGGACGTGCTCGGGACA
1042 TTTGATACTGCGCAGATAATAAAACTTCTTCCCTTCGCAGCTGCTCCACCGAAGCAAAGTAGAATGCAGTTCTCCTCATTTA
1043 CTGTGAATGTGAGGGGATCAGGAATGAGAATACTTGTAAGGGGCAATTCTCCTGTATTCAACTATAACAAGGCCACGAAGA
1044 GACTCACAGTTCTCGGAAAGGATGCTGGCACTTTAACTGAAGACCCAGATGAAGGCACAGCTGGAGTGGAGTCCGCTGTT
1045 CTGAGGGGATTCCTCATTCTGGGCAAAGAGAACAAGAGATATGGGCCAGCACTAAGCATCAATGAACTGAGCAACCTTGC
1046 GAAAGGAGAGAAGGCTAATGTGCTAATTGGGCAAGGAGACGTGGTGTTGGTAATGAAACGGAAACGGGACTCTAGCATAC
1047 TTACTGACAGCCAGACAGCGACCAAAAGAATTCGGATGGCCATCAATTAGTGTCGAATAGTTTAAAAACGACCTTGTTTCTA
1048 CT

1049

1050 **Table S2. Primers used for mutagenesis and sequencing.**

1051

1052 **Table S2A.** Primers used for site directed mutagenesis of IAV PR8 segment 8.

Target	F primer	R primer
Seg8 L95S99A	TGACATGACTGCTGAGGAAATGGCAAGG	CCTTGCCATTTCCTCAGCAGTCATGTCA
Seg8 E9697A	ATGACTCTTGCGGCAATGTCAAGGG	CCCTTGACATTGCCGCAAGAGTCAT
Seg8 P107A	CATGCTCATAGCCAAGCAGAAAG	CTTTCTGCTTGGCTATGAGCATG
Seg8 P107K108A	GCTCATAGCCGCGCAGAAAGTGG	CCACTTTCTGCGCGGCTATGAGC
Seg8 P107K108Q109A	CATAGCCGCGGCGAAAGTGGCAG	CTGCCACTTTCGCCGCGGCTATG
Seg8 P107K108Q109K110A	AGCCGCGGCGGCAGTGGCAGGC	GCCTGCCACTGCCGCCGCGGCT

1053

1054 **Table S2B.** Primers for amplification of influenza A virus A/Puerto Rico/8/1934 (PR8) viral transcripts by qPCR.

Target	Forward primer	Reverse primer
PR8 segment 1	GGCTAATGTGCTAATTGGGCAAG	GGCTGTCAGTAAGTATGCTAGAG
PR8 segment 5	ATCATGGCGTCTCAAGGCACCAA	GCTCTGATTTCAGTGGCATTCTGG

1055

1056 **Table S2C.** Primers for amplification of specific bisulphite converted RNA regions.

Target	Forward primer	Reverse primer
PR8 106-304	CATATAACCATAATCAAAAAATACACATC	GTGATATTATTATTTGGTTTGATTTGG
Chicken rRNA 3163-3307	GTGATTTTTGTTTAGTGTTTTGAATG	AAATCTCATTCATCCATTCATACAC

1057

1058 **Table S2D.** Primers for generation of *in vitro* transcription amplicon templates.

Target	Forward primer	Reverse primer
T7 Seg1 -sense	AGCGAAAGCAGGTCAATTATATTC	TAATACGACTCACTATAGGGAGTAGAAA CAAGGTCGTTTTTAAACTA
T7 Seg1 +sense	TAATACGACTCACTATAGGGAGCGAAA GCAGGTCAATTATATTC	AGTAGAAACAAGGTCGTTTTTAAACTA
T7 Seg5 -sense	AGCAAAAGCAGGGTAGATAATCA	TAATACGACTCACTATAGGGAGTAGAAA CAAGGGTATTTTTCTTTAATTGT
T7 Seg5 +sense	TAATACGACTCACTATAGGGAGCAAAA GCAGGGTAGATAATCA	AGTAGAAACAAGGGTATTTTTCTTTAATT GT

1059

1060 **Table S2E.** Primers for Northern blotting probe amplicon templates.

Target	Forward primer	Reverse primer
PR8 segment 1 -sense probe	TAATACGACTCACTATAGGGGCCGCTC CACCAAAGCAAAG	TTCGACACTAATTGATGGCC
PR8 segment 1 +sense probe	GCCGCTCCACCAAAGCAAAG	TAATACGACTCACTATAGGGTTCGACAC TAATTGATGGCC
PR8 segment 5 -sense probe	TAATACGACTCACTATAGGGCAAAAGC AGGGTAGATAATC	TCACCATTATTAGCTTGGCG
PR8 segment 5 +sense probe	CAAAAGCAGGGTAGATAATC	TAATACGACTCACTATAGGGTCACCATT ATTAGCTTGGCG

1061

1062 **Table S2F.** Primers for overlapping full genome sequencing.

Target	Forward primer	Reverse primer
Seg1 amplicon A	AGCGAAAGCAGGTCAATTAT	CCATGCTTTAGCCTTTCGAC
Seg1 amplicon B	ACATGGTGGAATAGGAATGG	TGCAACACTTCAATGTACAC
Seg1 amplicon C	CAAAACGAGATTCCTCCCAG	TGAAGATTGCCCGTAAGCAC
Seg1 amplicon D	AAGGCTGCAATGGGACTGAG	AGCACTTTCGCATCCTTCTG
Seg1 amplicon E	GCGATTGAATCCTATGCATC	TTGTAGGGTTCTGGGACCAC
Seg1 amplicon F	ATGGATCATCAGAAACTGGG	GGACTCCACTCCAGCTGTGC
Seg1 amplicon G	GCCGCTCCACCAAAGCAAAG	TTCGACACTAATTGATGGCC
Seg2 amplicon A	GAAAGCAGGCAAACCATTTG	TCATAGGTCTGTCGGCCTTG
Seg2 amplicon B	AACGATGGAGGTTGTTCAGC	CCCTGGGGTTGCAATTGCTC
Seg2 amplicon C	CAATGACCAAAGATGCTGAG	CTTGCTCTCAAACATATACCC
Seg2 amplicon D	AATCAGCCCGAATGGTTCAG	GTCGACTCCGGCTTGAATCC
Seg2 amplicon E	TGGTCTTCAATCCTCTGACG	GGGTTTGCTCCCACAGTTTC
Seg2 amplicon F	ATGCCATAGAGGTGACACAC	AATTGCAGCACCTTTGGTAC
Seg2 amplicon G	CCCATTTGTCAGCCATAAAG	GAAGGACAAGCTAAATTCAC
Seg3 amplicon A	AGCGAAAGCAGGTACTGATC	ACTTCTCTCCTTGTTACTCC
Seg3 amplicon B	GGCTGAGAAACCAAAGTTTC	GAATCCATCCACATAGGCTC
Seg3 amplicon C	TTGCCGACCAAAGTCTCCCG	TCATTCTCAATGTCCTGCAG
Seg3 amplicon D	GTTAAACCACACGAAAAGGG	ATGTATTCTGTGGCTCTGCAG
Seg3 amplicon E	ACATTGCAAGCATGAGAAGG	GAGGTTCCATTTGTTCTCAC
Seg3 amplicon F	AAGAAGTGCCATAGGCCAGG	CTCAATTGCTTCATATAGCC
Seg3 amplicon G	ATGATTGAAGCTGAGTCCTC	TTTTGGACAGTATGGATAGC
Seg4 amplicon A	CAACCAAAATGAAGGCAAAC	GGCCATGAGCTTTCTTTGGG

Seg4 amplicon B	GAGCTCAGTGTCATCATTCG	AATATTATTGTGTCTCCGGG
Seg4 amplicon C	TAAGAGATCAAGCTGGGAGG	AGCCTGATCCCTGTTCATTC
Seg4 amplicon D	ACTGGAATGATAGATGGATG	CATTTCTTACACTTTCCATGC
Seg4 amplicon E	GGATTTCTGGACATTTGGAC	ATTCTGCACTGCAAAGATCC
Seg5 amplicon A	CAAAAGCAGGGTAGATAATC	TCACCATTATTAGCTTGGCG
Seg5 amplicon B	GAGAGAACTCATCCTTTATG	TTCCGGCTCTCTCTCACTTG
Seg5 amplicon C	GAAATTTCAAACTGCTGCAC	TTGGAAGCAATTTGAACTCC
Seg5 amplicon D	CCAAGAGGGAAGCTTTCCAC	GTACTCCTCTGCATTGTCTC
Seg6 amplicon A	TGGATCAATCTGTCTGGTAG	ACAGGGCAGCTCATTAAGGCC
Seg6 amplicon B	TTCAAGTGGGACTGTTAAGG	GGTATCAGGGTAACAGGAAC
Seg6 amplicon C	GAGTTGAATGCACCTAATTC	CCGCTATACCCTGACCAATC
Seg6 amplicon D	GTGTATGTTGATGGAGCAAAC	GAACAGACTACTTGTCAATG
Seg7 amplicon A	GATGAGTCTTCTAACCGAGG	TGCCACTTCAGTGGTCACAG
Seg7 amplicon B	CCAAATAACATGGACAAAGC	GCTAGGATGAGTTCCAATGG
Seg7 amplicon C	GCTATGGAGCAAATGGCTGG	TTTTTACTCCAGCTCTATGC
Seg8 amplicon A	CATAATGGATCCAAACACTG	TCTCCAGCCGGTCAAAAATC
Seg8 amplicon B	AGAAGAATCCGATGAGGCAC	TTACTGCTTCTCCAAGCGAATC
Seg8 amplicon C	GACTCTAATATTGCTAAGGG	AAGCTGAAACGAGAAAGTTC
Seg1 CDLR amplicon B	ACATGGTGGAATAGGAATGG	TGCAACACTTCTATGTATAC
Seg1 CDLR amplicon C	AAAGACGAGATTCCTACCCG	TGGAGATTCCCCGTTAGCAC
Seg1 CDLR amplicon D	AAGGCTGCCATGGGGCTGAG	AGCACTTTTGCATCCTTTTG
Seg1 CDLR amplicon E	ACGATTGAACCCTATGCATC	TAGTGGGATTCTGAGACCAC
Seg1 CDLR amplicon F	ATGGATCATCAGAAACTGGG	GGACTCCACTCCAGCTGTGC
Seg1 CpGH amplicon B	ACATGGTGGAATAGGAATGG	TAATACTTCTATGTACACGC
Seg1 CpGH amplicon C	AAAAACGCGATTTCTTCCGG	TGAAGATTTCCCGTCAGAAC
Seg1 CpGH amplicon D	AAGGCCGCGATGGGACTGCG	AGCACTTTCGCGTCTTTCTG
Seg1 CpGH amplicon E	ACGATTGAATCCGATGCATC	TCGTCGGATTTTGCGACCAC
Seg1 CpGH amplicon F	ATGGATCATCCGAAATTGGG	GGACTCCACTCCAGCTGTGC

1063

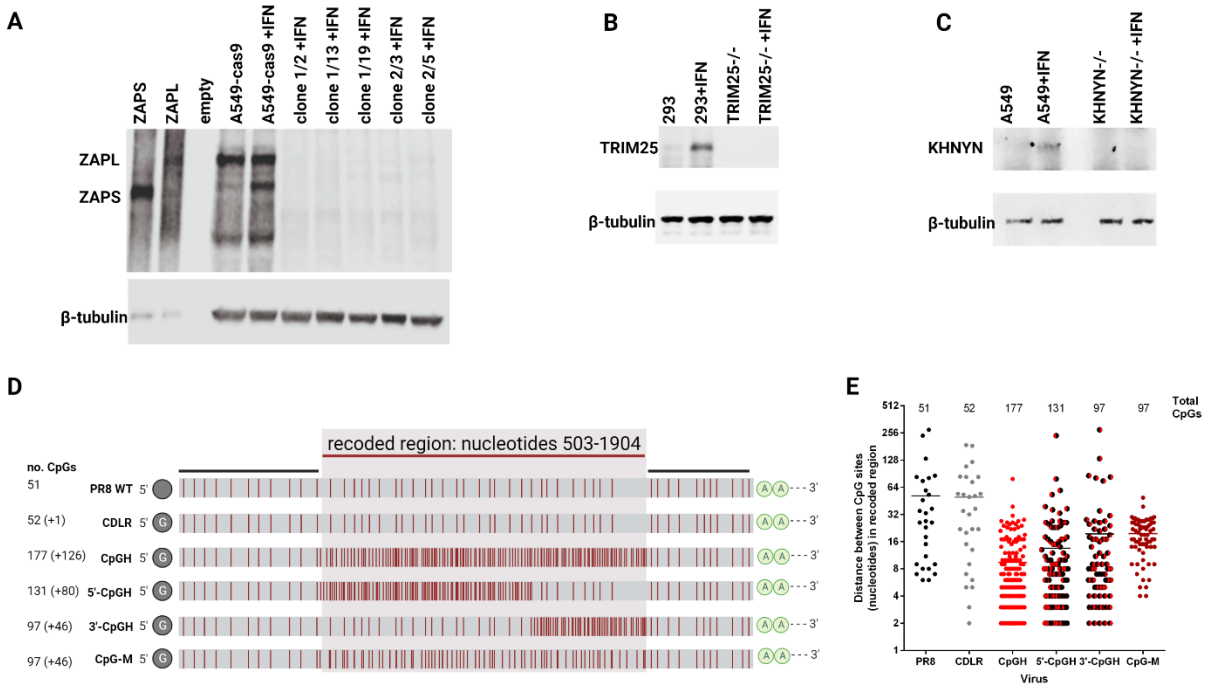
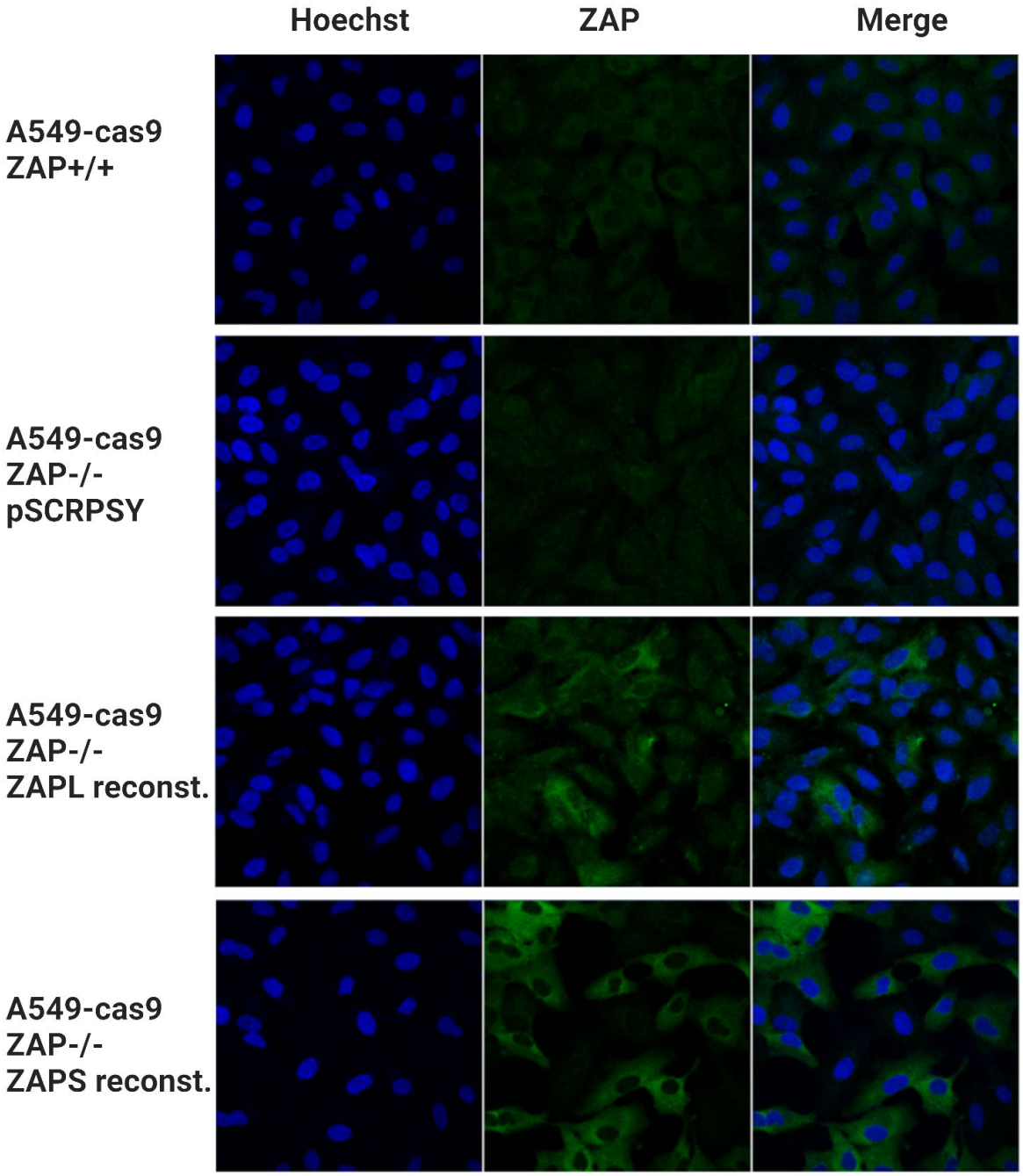


Figure S1. Validation of ZAP pathway inhibition assays. A. ZAP -/- A549 cells (37) were validated by western blotting. **B.** TRIM25 -/- 293 cells (38) were validated by western blotting. **C.** KHNYN -/- A549 cells (18) were validated by western blotting. **D.** Distribution of CpGs in segment 1 of recoded viruses. **E.** Distance between CpG sites in recoded region of segment 1.

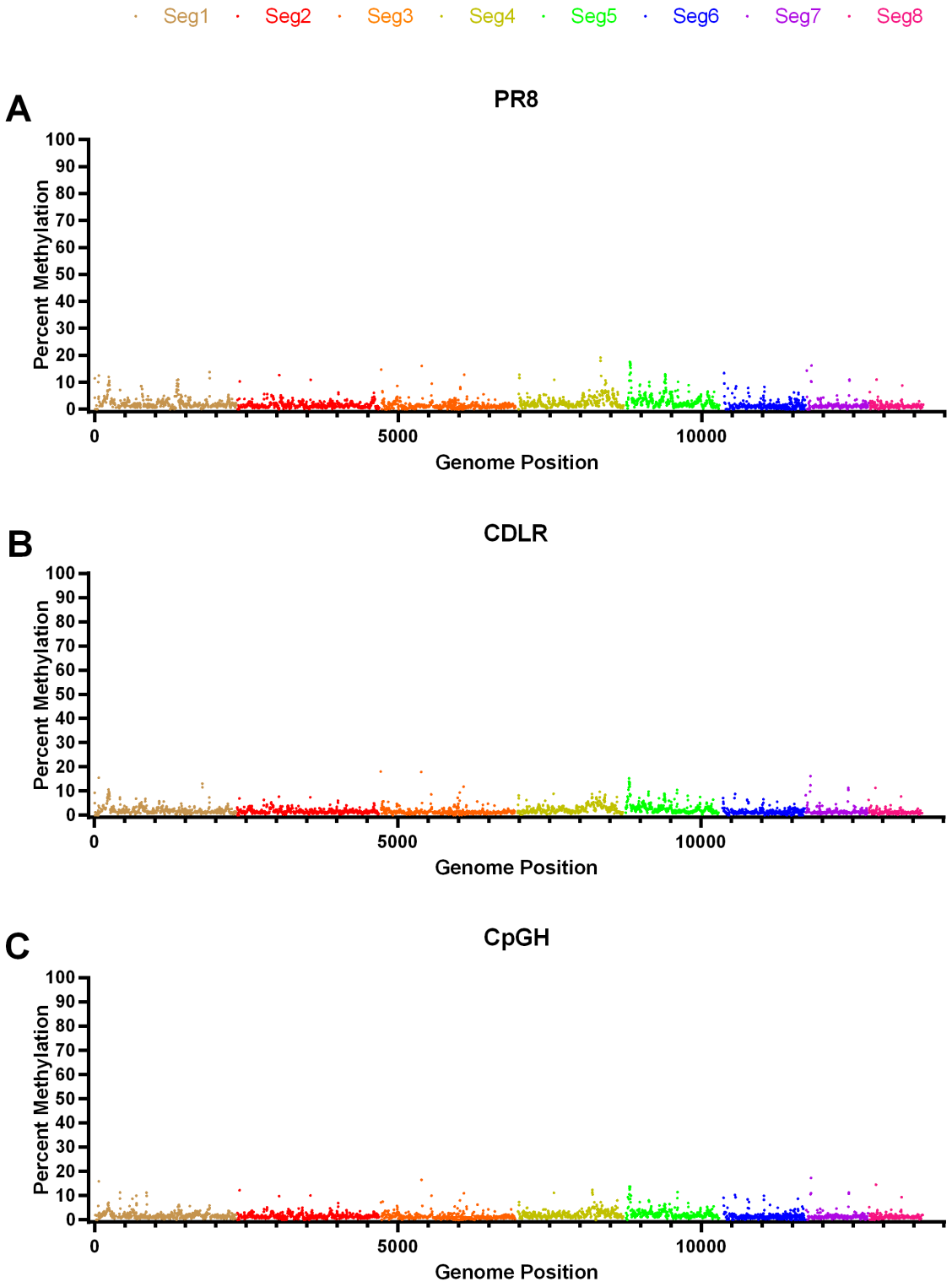
1069



1070

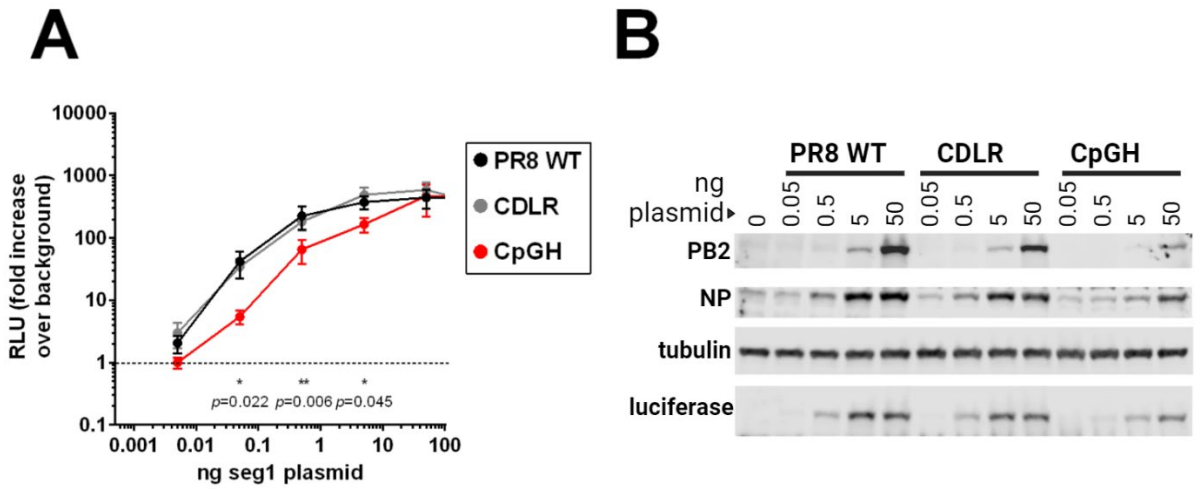
1071 **Figure S2. Validation of ZAP expression reconstitution in ZAP-/- A549 cells delivered by lentivirus vector and examined**

1072 **by confocal imaging.**



1073

1074 **Figure S3. Bisulphite conversion does not identify any conserved m5c methylation sites in the IAV genome, and CpG**
1075 **enrichment does not introduce any such methylation events.** Methylation signal identified across the genome of wildtype
1076 PR8 **(A)**, CDLR **(B)** and CpGH **(C)**. While the segment 1 sequence was different in A-C, the genome sequence across segments
1077 2-8 was identical, and so for these segments, the data are representative of 3 biological repeats.



1078

1079 **Figure S4. Minireplicon assays confirmed decreased transcript and protein production associated with CpG enrichment**
1080 **of PR8 segment 1.** Minireplicon assays that reconstitute the viral polymerase were performed using a luciferase reporter (**A**,
1081 luciferase reporter signal; **B**, protein signal by western blot).

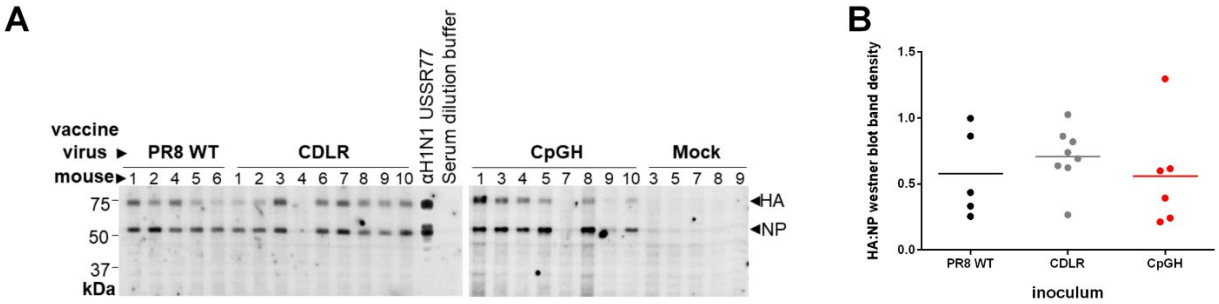
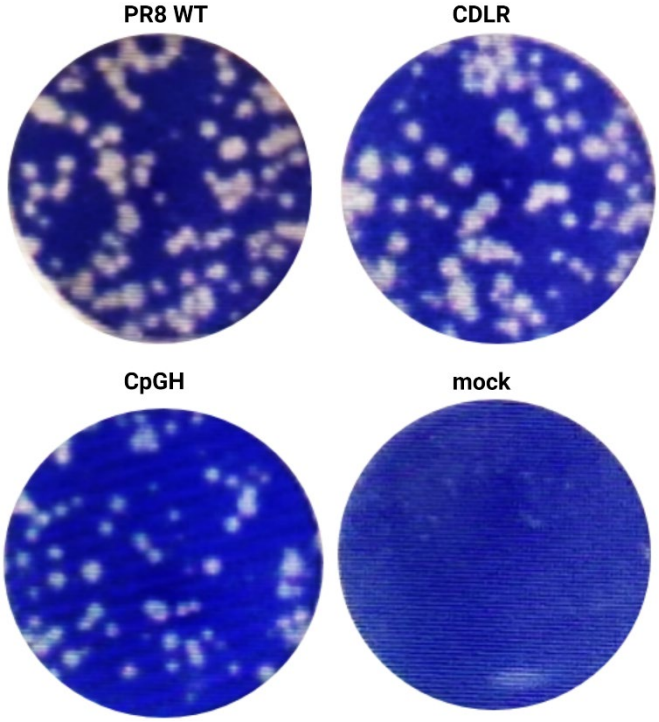


Figure S5. Sera from mice infected with 20 PFU of CpGH virus contain antibodies that are IAV-specific. 6-week old female BALB/c mice were infected with wildtype PR8 (n = 6), CDLR (n = 10) or CpGH (n = 10) viruses. After 20 days, tail bleeds were performed and sera harvested from all mice. **A.** For mice yielding sufficient sera, serum was diluted 1:100 in 2% BSA/ PBS and cross-reacted with cell lysate from MDCK cells infected with wildtype PR8 virus at MOI of 10 for 8 hours. Whole anti-IAV antibody (αH1N1 USSR77) was used as a positive control. Variable cross-reactivity to HA and NP proteins was observed for the different virus inocula, but this variability was consistent across viruses (**B**).



1089

1090 **Figure S6. Plaque phenotype of synonymously recoded viruses.** WT PR8, CDLR and CpGH viruses were rescued in
1091 embryonated hens' eggs and titred by plaque assay in MDCK cells. Representative images taken from plaque assays performed
1092 at the same time.

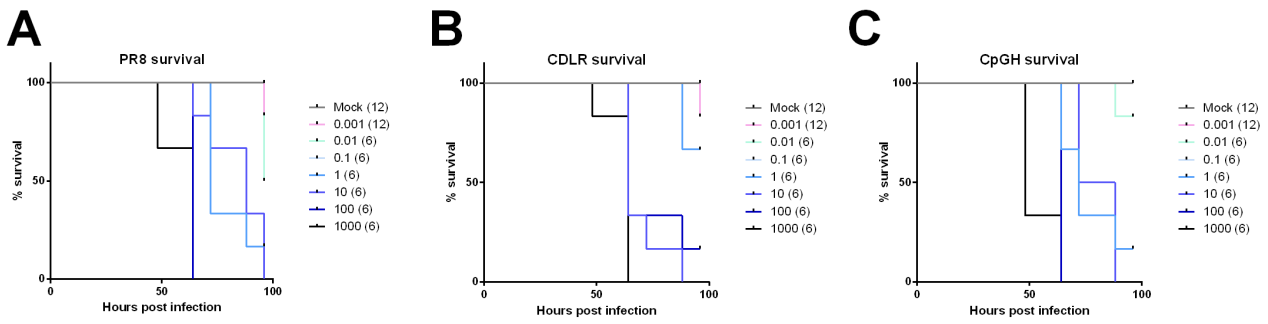


Figure S7. Survival of chicken embryos infected at day 10 post fertilisation. Embryonated hens' eggs at developmental day 10 were inoculated with ten-fold serial dilutions of virus stock (**A,** PR8; **B,** CDLR; **C,** CpGH) and survival was monitored by candling every 12 hours for 96 hours. The number of eggs inoculated at each dose is indicated in brackets.

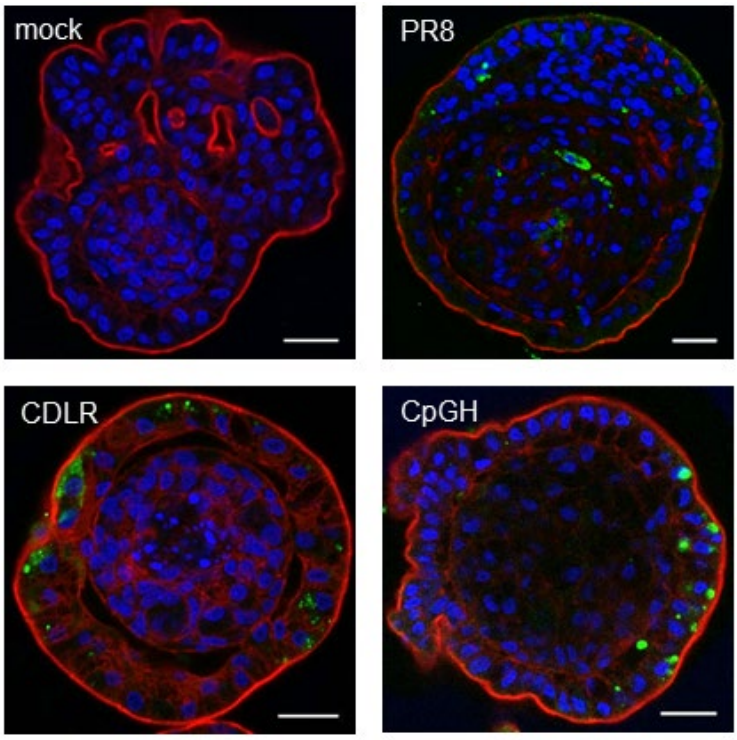


Figure S8. Confocal imaging of infected chicken enteroids. WT PR8, CDLR and CpGH viruses were used to infect chicken enteroids at 10^5 PFU/enteroid for 48 hours, and viral protein production was assessed visually using confocal microscopy. Representative z-axis projections of chicken enteroid whole-mount stained to detect cell nuclei (DAPI, blue), F-actin-expressing brush border (red) and virus nucleoprotein (green). Scale bar: 20 µm.

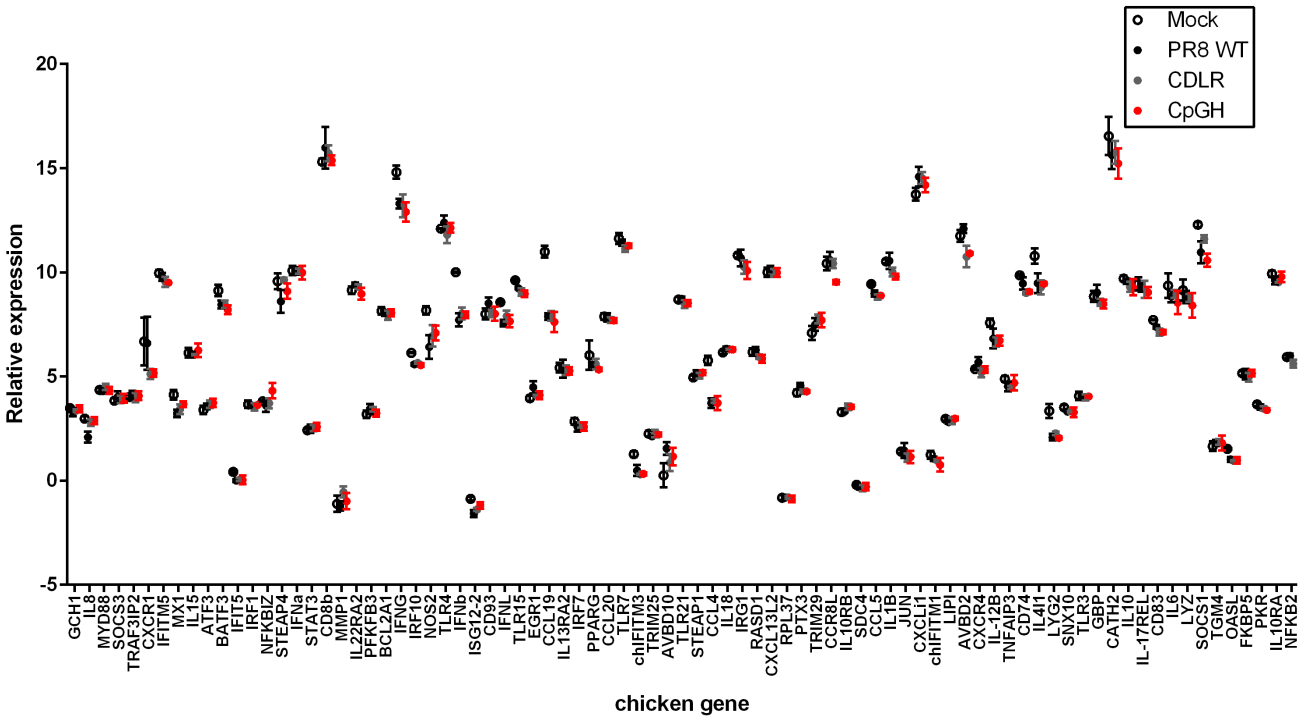
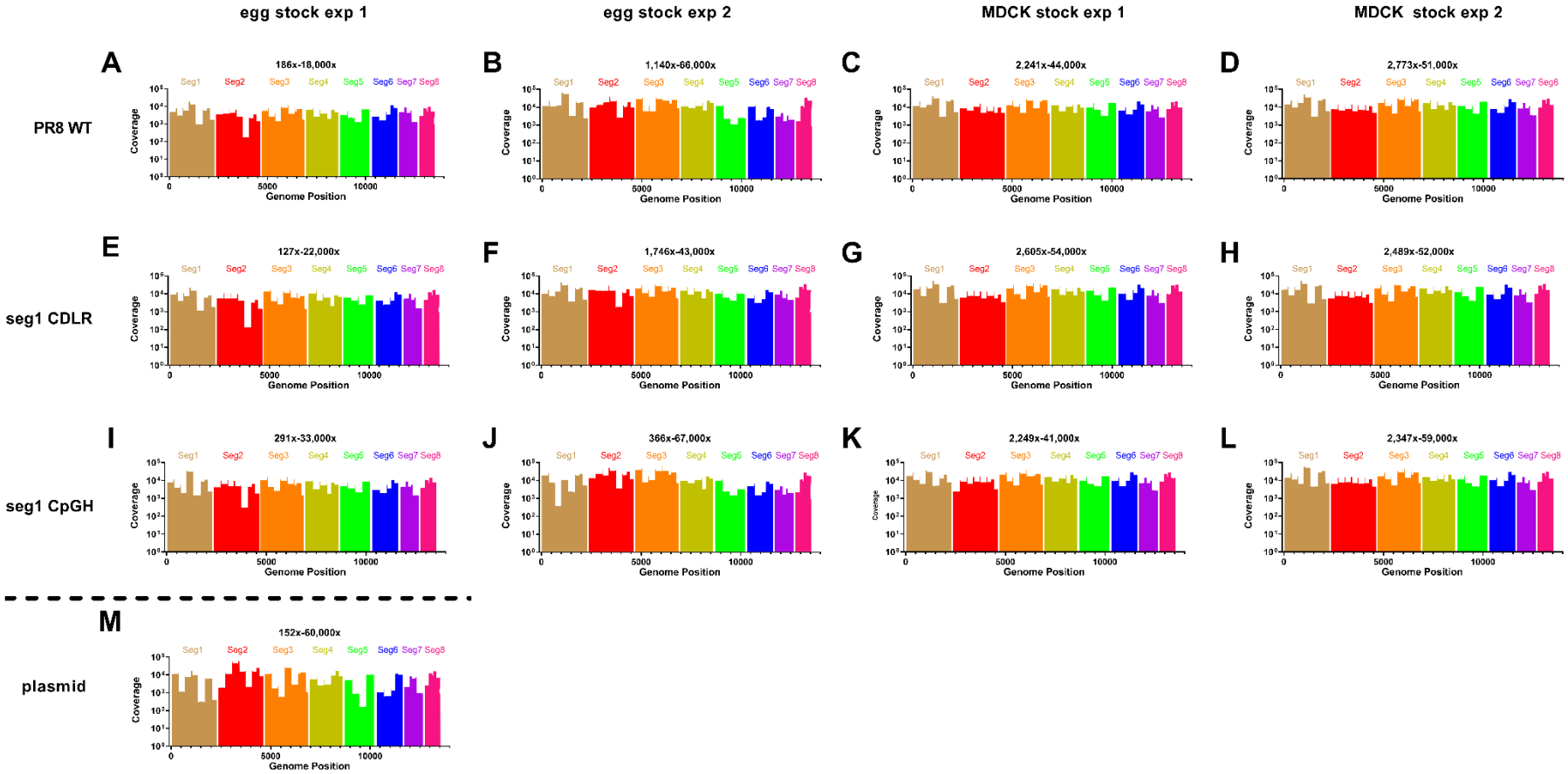


Figure S9. Innate immune gene expression levels of WT PR8, CDLR and CpGH viruses in chicken organoids.

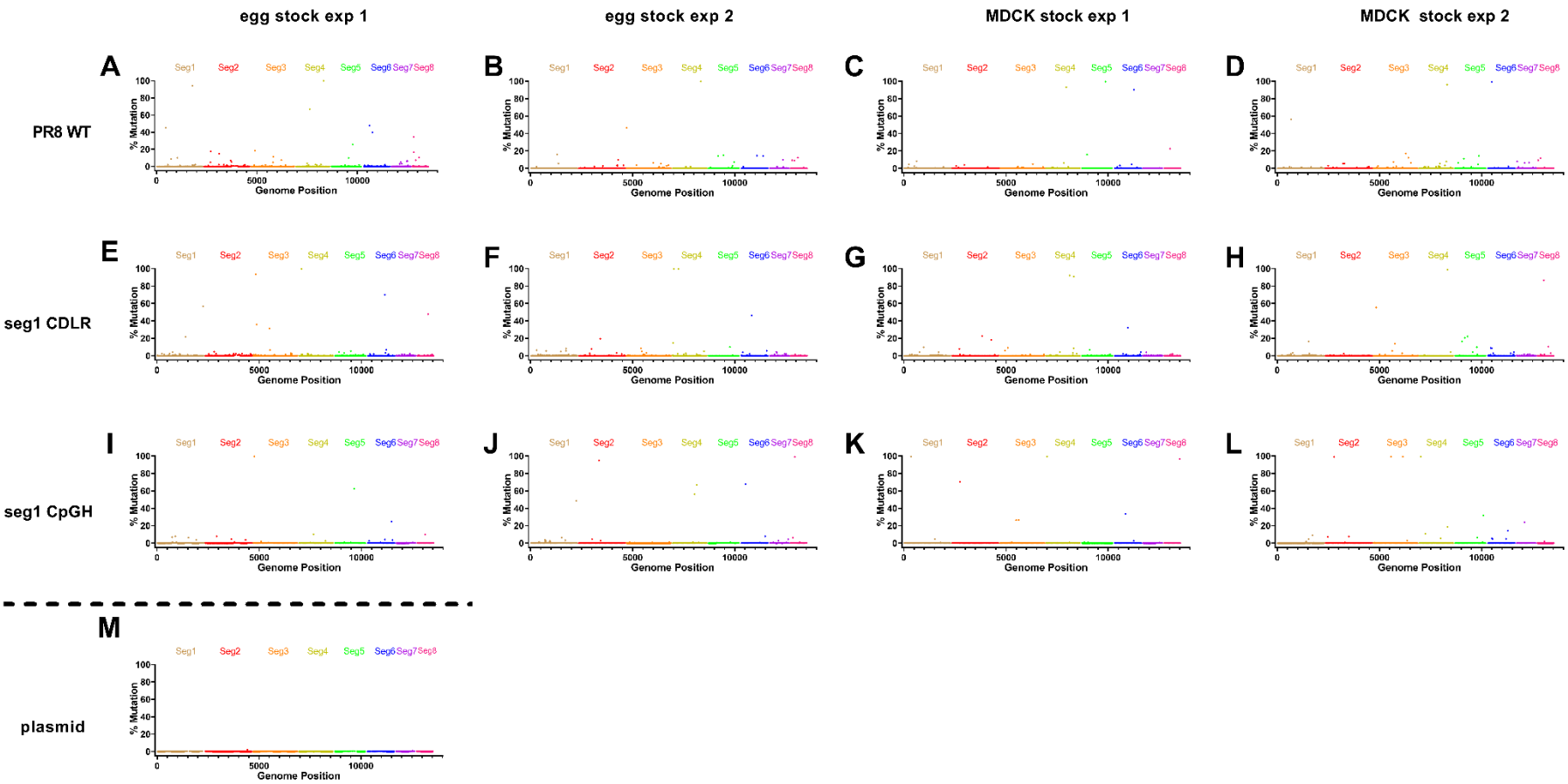
1104



1105

1106 **Figure S10. Sequence coverage depth for deep sequenced serially passaged viruses recoded in segment 1, over the complete genome. A-D**, PR8 wildtype virus; **E-H**, CDLR; **I-L**, CpGH; **M,** input

1107 plasmid.



1108

1109 **Figure S11. Full genome mutation traces for all serially passaged viruses. A-D**, PR8 wildtype virus; **E-H**, CDLR; **I-L**, CpGH; **M,** input plasmid.

1110 **REFERENCES**

1111 1. Simmonds, P., Xia, W., Baillie, J.K. and McKinnon, K. (2013) Modelling mutational and selection
1112 pressures on dinucleotides in eukaryotic phyla –selection against CpG and UpA in cytoplasmically
1113 expressed RNA and in RNA viruses. *BMC Genomics*, **14**.
1114 2. Rima, B.K. and McFerran, N.V. (1997) Dinucleotide and stop codon frequencies in single-stranded
1115 RNA viruses. *Journal of General Virology*, **78**, 2859-2870.
1116 3. Karlin, S., Doerfler, W. and Cardon, L.R. (1994) Why is CpG suppressed in the genomes of virtually
1117 all small eukaryotic viruses but not in those of large eukaryotic viruses? *Journal of Virology*, **68**,
1118 2889-2897.
1119 4. Takata, M.A., Gonçalves-Carneiro, D., Zang, T.M., Soll, S.J., York, A., Blanco-Melo, D. and Bieniasz,
1120 P.D. (2017) CG dinucleotide suppression enables antiviral defence targeting non-self RNA. *Nature*,
1121 **550**, 124-127.
1122 5. Gaunt, E.R. and Digard, P. (2021) Compositional biases in RNA viruses: Causes, consequences and
1123 applications. *WIREs RNA*, e1679.
1124 6. Li, M.M.H., Aguilar, E.G., Michailidis, E., Pabon, J., Park, P., Wu, X., de Jong, Y.P., Schneider, W.M.,
1125 Molina, H., Rice, C.M. *et al.* (2019) Characterization of Novel Splice Variants of Zinc Finger Antiviral
1126 Protein (ZAP). *Journal of Virology*, **93**, e00715-00719.
1127 7. Chen, S., Xu, Y., Zhang, K., Wang, X., Sun, J., Gao, G. and Liu, Y. (2012) Structure of N-terminal
1128 domain of ZAP indicates how a zinc-finger protein recognizes complex RNA. *Nature structural &*
1129 *molecular biology*, **19**, 430-435.
1130 8. Todorova, T., Bock, F.J. and Chang, P. (2014) PARP13 regulates cellular mRNA post-transcriptionally
1131 and functions as a pro-apoptotic factor by destabilizing TRAILR4 transcript. *Nat Commun*, **5**, 5362.
1132 9. Gao, G., Guo, X. and Goff, S.P. (2002) Inhibition of Retroviral RNA production by ZAP, a CCCH-type
1133 zinc finger protein. *Science*, **297**, 1703-1706.
1134 10. Kmiec, D., Lista, M.J., Ficarelli, M., Swanson, C.M. and Neil, S.J.D. (2021) S-farnesylation is essential
1135 for antiviral activity of the long ZAP isoform against RNA viruses with diverse replication strategies.
1136 *PLoS Pathog*, **17**, e1009726.
1137 11. Hayakawa, S., Shiratori, S., Yamato, H., Kameyama, T., Kitatsuji, C., Kashigi, F., Goto, S., Kameoka, S.,
1138 Fujikura, D., Yamada, T. *et al.* (2011) ZAPS is a potent stimulator of signaling mediated by the RNA
1139 helicase RIG-I during antiviral responses. *Nature immunology*, **12**, 37-44.
1140 12. Guo, X., Carroll, J.-W.N., MacDonald, M.R., Goff, S.P. and Gao, G. (2004) The Zinc Finger Antiviral
1141 Protein Directly Binds to Specific Viral mRNAs through the CCCH Zinc Finger Motifs. *Journal of*
1142 *Virology*, **78**, 12781.
1143 13. Li, M.M.H., Lau, Z., Cheung, P., Aguilar, E.G., Schneider, W.M., Bozzacco, L., Molina, H., Buehler, E.,
1144 Takaoka, A., Rice, C.M. *et al.* (2017) TRIM25 enhances the antiviral action of zinc-finger antiviral
1145 protein (ZAP). *PLoS Pathogens*, **13**, e1006145.
1146 14. Zheng, X., Wang, X., Tu, F., Wang, Q., Fan, Z. and Gao, G. (2017) TRIM25 is required for the antiviral
1147 activity of zinc finger antiviral protein. *Journal of Virology*, **91**, e00088-00017.
1148 15. Zhu, Y., Chen, G., Lv, F., Wang, X., Ji, X., Xu, Y., Sun, J., Wu, L., Zheng, Y.T. and Gao, G. (2011) Zinc-
1149 finger antiviral protein inhibits HIV-1 infection by selectively targeting multiply spliced viral mRNAs
1150 for degradation. *Proc Natl Acad Sci U S A*, **108**, 15834-15839.
1151 16. Ficarelli, M., Wilson, H., Pedro Galão, R., Mazzon, M., Antzin-Anduetza, I., Marsh, M., Neil, S.J.D.
1152 and Swanson, C.M. (2019) KHNYN is essential for the zinc finger antiviral protein (ZAP) to restrict
1153 HIV-1 containing clustered CpG dinucleotides. *Elife*, **8**, e46767.
1154 17. Ficarelli, M., Antzin-Anduetza, I., Hugh-White, R., Firth, A.E., Sertkaya, H., Wilson, H., Neil, S.J.D.,
1155 Schulz, R. and Swanson, C.M. (2020) CpG dinucleotides inhibit HIV-1 replication through zinc finger
1156 antiviral protein (ZAP)-dependent and -independent mechanisms. *Journal of Virology*, **94**, e01337-
1157 01319.
1158 18. Gonçalves-Carneiro, D., Takata, M.A., Ong, H., Shilton, A. and Bieniasz, P.D. (2021) Origin and
1159 evolution of the zinc finger antiviral protein. *PLoS Pathogens*, **17**, e1009545.
1160 19. Goossens, K.E., Karpala, A.J., Ward, A. and Bean, A.G. (2014) Characterisation of chicken ZAP.
1161 *Developmental and comparative immunology*, **46**, 373-381.
1162 20. Gomez, P.L. and Robinson, J.M. (2018) Vaccine Manufacturing. *Plotkin's Vaccines*, 51-60.e51.

1163 21. Gonçalves-Carneiro, D., Mastrocola, E., Lei, X., DaSilva, J., Chan, Y.F. and Bieniasz, P.D. (2022)
1164 Rational attenuation of RNA viruses with zinc finger antiviral protein. *Nature Microbiology*, **7**, 1558-
1165 1567.
1166 22. Grohskopf, L.A., Alyanak, E., Ferdinands, J.M., Broder, K.R., Blanton, L.H., Talbot, H.K. and Fry, A.M.
1167 (2021) Prevention and Control of Seasonal Influenza with Vaccines: Recommendations of the
1168 Advisory Committee on Immunization Practices, United States, 2021-22 Influenza Season.
1169 *Morbidity and mortality weekly report. Recommendations and reports*, **70**, 1-28.
1170 23. Singanayagam, A., Zambon, M., Lalvani, A. and Barclay, W. (2018) Urgent challenges in
1171 implementing live attenuated influenza vaccine. *The Lancet Infectious Diseases*, **18**, e25-e32.
1172 24. Gould, P.S., Easton, A.J. and Dimmock, N.J. (2017) Live Attenuated Influenza Vaccine contains
1173 Substantial and Unexpected Amounts of Defective Viral Genomic RNA. *Viruses*, **9**.
1174 25. Ran, F.A., Hsu, Patrick D., Lin, C.-Y., Gootenberg, Jonathan S., Konermann, S., Trevino, A.E., Scott,
1175 David A., Inoue, A., Matoba, S., Zhang, Y. *et al.* (2013) Double Nicking by RNA-Guided CRISPR Cas9
1176 for Enhanced Genome Editing Specificity. *Cell*, **154**, 1380-1389.
1177 26. Idoko-Akoh, A., Taylor, L., Sang, H.M. and McGrew, M.J. (2018) High fidelity CRISPR/Cas9 increases
1178 precise monoallelic and biallelic editing events in primordial germ cells. *Scientific Reports*, **8**.
1179 27. Manini, I., Trombetta, C.M., Lazzeri, G., Pozzi, T., Rossi, S. and Montomoli, E. (2017) Egg-
1180 Independent Influenza Vaccines and Vaccine Candidates. **5**, 18.
1181 28. Milián, E. and Kamen, A.A. (2015) Current and Emerging Cell Culture Manufacturing Technologies
1182 for Influenza Vaccines. *BioMed Research International*, **2015**, 504831.
1183 29. Chen, J.-R., Liu, Y.-M., Tseng, Y.-C. and Ma, C. (2020) Better influenza vaccines: an industry
1184 perspective. *Journal of Biomedical Science*, **27**.
1185 30. Hegde, N.R. (2015) Cell culture-based influenza vaccines: A necessary and indispensable investment
1186 for the future. *Hum Vaccin Immunother*, **11**, 1223-1234.
1187 31. Fodor, E., Devenish, L., Engelhardt, O.G., Palese, P., Brownlee, G.G. and García-Sastre, A. (1999)
1188 Rescue of Influenza A Virus from recombinant DNA. *Journal of Virology*, **73**, 9679-9682.
1189 32. Hoffmann, E., Neumann, G., Kawaoka, Y., Hobom, G. and Webster, R.G. (2000) A DNA transfection
1190 system for generation of influenza A virus from eight plasmids. *Proceedings of the National*
1191 *Academy of Sciences*, **97**, 6108.
1192 33. Gog, J.R., Afonso, E.D.S., Dalton, R.M., Leclercq, I., Tiley, L., Elton, D., von Kirchbach, J.C., Naffakh,
1193 N., Escriou, N. and Digard, P. (2007) Codon conservation in the influenza A virus genome defines
1194 RNA packaging signals. *Nucleic Acids Research*, **35**, 1897-1907.
1195 34. Wise, H., Barbezange, C., Jagger, B.W., Dalton, R.M., Gog, J.R., Curran, M.D., Taubenberger, J.K.,
1196 Anderson, E.C. and Digard, P. (2011) Overlapping signals for translational regulation and packaging
1197 of influenza A virus segment 2. *Nucleic Acids Research*, **39**, 7775-7790.
1198 35. Dadonaite, B., Gilbertson, B., Knight, M.L., Trifkovic, S., Rockman, S., Laederach, A., Brown, L.E.,
1199 Fodor, E. and Bauer, D.L.V. (2019) The structure of the influenza A virus genome. *Nature*
1200 *Microbiology*, **4**, 1781-1789.
1201 36. Pinto, R.M., Lycett, S., Gaunt, E. and Digard, P. (2020) Accessory Gene Products of Influenza A Virus.
1202 *Cold Spring Harb Perspect Med*, **1**, a038380.
1203 37. Shaw, A.E., Rihn, S.J., Mollentze, N., Wickenhagen, A., Stewart, D.G., Orton, R.J., Kuchi, S., Bakshi, S.,
1204 Collados, M.R., Turnbull, M.L. *et al.* (2021) The antiviral state has shaped the CpG composition of
1205 the vertebrate interferome to avoid self-targeting. *PLoS biology*, **19**, e3001352.
1206 38. Choudhury, N.R., Heikel, G., Trus, I., Dos Santos Pinto, R.M., Trubitsyna, M., Gaunt, E., Digard, P.
1207 and Michlewski, G. (2021) TRIM25 inhibits Influenza A infection by destabilising its mRNA and is
1208 redundant for the RIG-I pathway. *bioRxiv*, 2021.2009.2013.460052.
1209 39. Lin, Y.-T., Chiweshe, S., McCormick, D., Raper, A., Wickenhagen, A., DeFillipis, V., Gaunt, E.,
1210 Simmonds, P., Wilson, S.J. and Grey, F. (2020) Human cytomegalovirus evades ZAP detection by
1211 suppressing CpG dinucleotides in the major immediate early genes. *PLoS Pathogens*,
1212 2020.2001.2007.897132.
1213 40. Li, X., Gu, M., Zheng, Q., Gao, R. and Liu, X. (2021) Packaging signal of influenza A virus. *Virology*
1214 *Journal*, **18**, 36.

1215 41. Wise, H., Gaunt, E., Ping, J., Holzer, B., Jasim, S., Lycett, S.J., Murphy, L., Livesey, A., Brown, R.,
1216 Smith, N. *et al.* (2019) An alternative AUG codon that produces an N-terminally extended form of
1217 the influenza A virus NP is a virulence factor for a swine-derived virus. *bioRxiv*, 738427.
1218 42. Moore, L.D., Le, T. and Fan, G. (2013) DNA Methylation and Its Basic Function.
1219 *Neuropsychopharmacology*, **38**, 23-38.
1220 43. Goll, M.G., Kirpekar, F., Maggert, K.A., Yoder, J.A., Hsieh, C.L., Zhang, X., Golic, K.G., Jacobsen, S.E.
1221 and Bestor, T.H. (2006) Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2.
1222 *Science*, **311**, 395-398.
1223 44. Brzezicha, B., Schmidt, M., Makałowska, I., Jarmołowski, A., Pieńkowska, J. and Szweykowska-
1224 Kulińska, Z. (2006) Identification of human tRNA:m5C methyltransferase catalysing intron-
1225 dependent m5C formation in the first position of the anticodon of the pre-tRNA Leu (CAA). *Nucleic*
1226 *Acids Research*, **34**, 6034-6043.
1227 45. Cooper, D.N. and Krawczak, M. (1989) Cytosine methylation and the fate of CpG dinucleotides in
1228 vertebrate genomes. *Human Genetics*, **83**, 181-188.
1229 46. Khabar, K.S.A. and Young, H.A. (2007) Post-transcriptional control of the interferon system.
1230 *Biochimie*, **89**, 761-769.
1231 47. Atasheva, S., Frolova Elena, I. and Frolov, I. (2014) Interferon-Stimulated Poly(ADP-Ribose)
1232 Polymerases Are Potent Inhibitors of Cellular Translation and Virus Replication. *Journal of Virology*,
1233 **88**, 2116-2130.
1234 48. García-Sastre, A., Egorov, A., Matassov, D., Brandt, S., Levy, D.E., Durbin, J.E., Palese, P. and Muster,
1235 T. (1998) Influenza A Virus Lacking the NS1 Gene Replicates in Interferon-Deficient Systems.
1236 *Virology*, **252**, 324-330.
1237 49. Talon, J., Horvath Curt, M., Polley, R., Basler Christopher, F., Muster, T., Palese, P. and García-
1238 Sastre, A. (2000) Activation of Interferon Regulatory Factor 3 Is Inhibited by the Influenza A Virus
1239 NS1 Protein. *Journal of Virology*, **74**, 7989-7996.
1240 50. Koliopoulos, M.G., Lethier, M., van der Veen, A.G., Haubrich, K., Hennig, J., Kowalinski, E., Stevens,
1241 R.V., Martin, S.R., Reis e Sousa, C., Cusack, S. *et al.* (2018) Molecular mechanism of influenza A NS1-
1242 mediated TRIM25 recognition and inhibition. *Nature communications*, **9**, 1820-1820.
1243 51. Tang, Q., Wang, X. and Gao, G. (2017) The Short Form of the Zinc Finger Antiviral Protein Inhibits
1244 Influenza A Virus Protein Expression and Is Antagonized by the Virus-Encoded NS1. **91**, e01909-
1245 01916.
1246 52. Pereira, C.F., Wise, H.M., Kurian, D., Pinto, R.M., Amorim, M.J., Gill, A.C. and Digard, P. (2018)
1247 Effects of mutations in the effector domain of influenza A virus NS1 protein. *BMC Research Notes*,
1248 **11**, 673.
1249 53. Guo, X., Ma, J., Sun, J. and Gao, G. (2007) The zinc-finger antiviral protein recruits the RNA
1250 processing exosome to degrade the target mRNA. *PNAS*, **104**, 151-156.
1251 54. Wise, H.M., Foeglein, A., Sun, J., Dalton, R.M., Patel, S., Howard, W., Anderson, E.C., Barclay, W.S.
1252 and Digard, P. (2009) A complicated message: Identification of a novel PB1-related protein
1253 translated from influenza A virus segment 2 mRNA. *J Virol*, **83**, 8021-8031.
1254 55. Pérez Rubio, A. and Eiros, J.M. (2018) Cell culture-derived flu vaccine: Present and future. *Hum*
1255 *Vaccin Immunother*, **14**, 1874-1882.
1256 56. Esnault, E., Bonsergent, C., Larcher, T., Bed'hom, B., Vautherot, J.-F., Delaleu, B., Guigand, L.,
1257 Soubieux, D., Marc, D. and Quéré, P. (2011) A novel chicken lung epithelial cell line:
1258 Characterization and response to low pathogenicity avian influenza virus. *Virus Research*, **159**, 32-
1259 42.
1260 57. Hinshaw, V.S., Webster, R.G., Naeve, C.W. and Murphy, B.R. (1983) Altered tissue tropism of
1261 human-avian reassortant influenza viruses. *Virology*, **128**, 260-263.
1262 58. Nash, T.J., Morris, K.M., Mabbott, N.A. and Vervelde, L. (2021) Inside-out chicken enteroids with
1263 leukocyte component as a model to study host–pathogen interactions. *Communications Biology*, **4**,
1264 377.
1265 59. Saitou, N. and Nei, M. (1986) Polymorphism and evolution of influenza A virus genes. *Molecular*
1266 *Biology and Evolution*, **3**, 57-74.

1267 60. Magor, K.E., Miranzo Navarro, D., Barber, M.R.W., Petkau, K., Fleming-Canepa, X., Blyth, G.A.D. and
1268 Blaine, A.H. (2013) Defense genes missing from the flight division. *Developmental and comparative*
1269 *immunology*, **41**, 377-388.
1270 61. Maassab, H.F. and DeBorde, D.C. (1985) Development and characterization of cold-adapted viruses
1271 for use as live virus vaccines. *Vaccine*, **3**, 355-369.
1272 62. Jang, Y.H. and Seong, B.-L. (2012) Principles underlying rational design of live attenuated influenza
1273 vaccines. *Clin Exp Vaccine Res*, **1**, 35-49.
1274 63. Maassab, H.F. and Bryant, M.L. (1999) The development of live attenuated cold-adapted influenza
1275 virus vaccine for humans. *Reviews in Medical Virology*, **9**, 237-244.
1276 64. de Wit, E., Spronken, M.I., Bestebroer, T.M., Rimmelzwaan, G.F., Osterhaus, A.D. and Fouchier, R.A.
1277 (2004) Efficient generation and growth of influenza virus A/PR/8/34 from eight cDNA fragments.
1278 *Virus Res*, **103**, 155-161.
1279 65. Simmonds, P. (2012) SSE: a nucleotide and amino acid sequence analysis platform. *BMC Research*
1280 *Notes*, **5**, 50.
1281 66. Yamayoshi, S., Watanabe, M., Goto, H. and Kawaoka, Y. (2016) Identification of a Novel Viral
1282 Protein Expressed from the PB2 Segment of Influenza A Virus. *J Virol*, **90**, 444-456.
1283 67. Gaunt, E., Wise, H.M., Zhang, H., Lee, L.N., Atkinson, N.J., Nicol, M.Q., Highton, A.J., Klenerman, P.,
1284 Beard, P.M., Dutia, B.M. *et al.* (2016) Elevation of CpG frequencies in influenza A genome
1285 attenuates pathogenicity but enhances host response to infection. *Elife*, **5**, e12735-e12735.
1286 68. Hutchinson E. C., Curran M. D., Read E. K., Gog J. R. and Digard, P. (2008) Mutational Analysis of cis-
1287 Acting RNA Signals in Segment 7 of Influenza A Virus. *Journal of Virology*, **82**, 11869-11879.
1288 69. Wise HM., Foeglein, A., Sun, J., Dalton Rosa, M., Patel, S., Howard, W., Anderson E. C., Barclay W. S.
1289 and Digard, P. (2009) A Complicated Message: Identification of a Novel PB1-Related Protein
1290 Translated from Influenza A Virus Segment 2 mRNA. *Journal of Virology*, **83**, 8021-8031.
1291 70. Hirst , G.K. (1942) The quantitative determination of influenza virus and antibodies by means of red
1292 cell agglutination. *Journal of Experimental Medicine*, **75**, 49-64.
1293 71. Cold Spring Harbor. (2006) Laemmli sample buffer (2X). pdb.rec10424.
1294 72. Carrasco, M., Amorim, M.J. and Digard, P. (2004) Lipid raft-dependent targeting of the influenza A
1295 virus nucleoprotein to the apical plasma membrane. *Traffic (Copenhagen, Denmark)*, **5**, 979-992.
1296 73. Afgan, E., Baker, D., Batut, B., van den Beek, M., Bouvier, D., Čech, M., Chilton, J., Clements, D.,
1297 Coraor, N., Grüning, B.A. *et al.* (2018) The Galaxy platform for accessible, reproducible and
1298 collaborative biomedical analyses: 2018 update. *Nucleic Acids Research*, **46**, W537-W544.
1299 74. Langmead, B., Trapnell, C., Pop, M. and Salzberg, S.L. (2009) Ultrafast and memory-efficient
1300 alignment of short DNA sequences to the human genome. *Genome Biology*, **10**, R25.
1301 75. Borowska, D., Kuo, R., Bailey, R.A., Watson, K.A., Kaiser, P., Vervelde, L. and Stevens, M.P. (2019)
1302 Highly multiplexed quantitative PCR-based platform for evaluation of chicken immune responses.
1303 *PLoS ONE*, **14**, e0225658.
1304 76. Hoffmann, E., Stech, J., Guan, Y., Webster, R.G. and Perez, D.R. (2001) Universal primer set for the
1305 full-length amplification of all influenza A viruses. *Arch Virol*, **146**, 2275-2289.
1306 77. Grubaugh, N.D., Gangavarapu, K., Quick, J., Matteson, N.L., De Jesus, J.G., Main, B.J., Tan, A.L., Paul,
1307 L.M., Brackney, D.E., Grewal, S. *et al.* (2019) An amplicon-based sequencing framework for
1308 accurately measuring intrahost virus diversity using PrimalSeq and iVar. *Genome Biology*, **20**, 8.