1 **Structural variation turnovers and defective genomes: key drivers for the *in***
2 ***vitro* evolution of the large double-stranded DNA koi herpesvirus (KHV)**
3
4

5 Nurul Novelia Fuandila[1], Anne-Sophie Gosselin-Grenet[2], Marie-Ka Tilak[1], Sven M Bergmann[3],
6 Jean-Michel Escoubas[4], Sandro Klafack[5], Angela Mariana Lusiastuti[6], Munti Yuhana[7], Anna-
7 Sophie Fiston-Lavier[1,8], Jean-Christophe Avarre[1*], Emira Cherif[1*]
8 [1]ISEM, Univ Montpellier, CNRS, IRD, Montpellier, France
9 [2]DGIMI, Univ Montpellier, INRAE, Montpellier, France
10 [3]Institute of Infectology, Friedrich-Loeffer-Institut, Federal Research Institute for Animal Health,
11 Greifswald-Insel Riems, Germany
12 [4]IHPE, Univ Montpellier, CNRS, Ifremer, UPVD, Montpellier, France
13 [5]Department of Experimental Animal Facilities and Biorisk Management, Friedrich-Loeffer-
14 Institut, Federal Research Institute for Animal Health, Greifswald-Insel Riems, Germany
15 [6]Fish Health Laboratory, Research Institute for Freshwater Aquaculture and Fisheries
16 Extension, Bogor, Indonesia
17 [7]Faculty of Fisheries and Marine Sciences, Bogor Agricultural University, Indonesia
18 [8]Institut Universitaire de France (IUF)
19 * corresponding authors: jean-christophe.avarre@ird.fr; emira.cherif@ird.fr

20

21 **Abstract**

22 Structural variations (SVs) constitute a significant source of genetic variability in virus genomes.
23 Yet knowledge about SV variability and contribution to the evolutionary process in large double-
24 stranded (ds)DNA viruses is limited. Cyprinid herpesvirus 3 (CyHV-3), also commonly known as
25 koi herpesvirus (KHV), has the largest dsDNA genome within herpesviruses. This virus has
26 become one of the biggest threats to common carp and koi farming, resulting in high morbidity
27 and mortalities of fishes, serious environmental damage, and severe economic losses. A
28 previous study analyzing CyHV-3 virulence evolution during serial passages onto carp cell
29 cultures suggested that CyHV-3 evolves, at least *in vitro*, through an assembly of haplotypes
30 that alternatively become dominant or under-represented. The present study investigates the
31 SV diversity and dynamics in CyHV-3 genome during 99 serial passages in cell culture using,
32 for the first time, ultra-deep whole-genome and amplicon-based sequencing. The results
33 indicate that KHV polymorphism mostly involves SVs. These SVs display a wide distribution
34 along the genome and exhibit high turnover dynamics with a clear bias towards inversion and

35   deletion events. Analysis of the pathogenesis-associated ORF150 region in ten intermediate cell

36   passages highlighted mainly deletion, inversion and insertion variations that deeply altered the

37   structure of ORF150. Our findings indicate that SV turnovers and defective genomes represent

38   key drivers in the viral population dynamics and *in vitro* evolution of KHV. Thus, the present

39   study can contribute to the basic research needed to design safe live-attenuated vaccines,

40   classically obtained by viral attenuation after serial passages in cell culture.

41

42

43   Keywords

44   KHV, virus evolution, virulence, virus attenuation, structural variations, defective genome, carp

45

46

47   Introduction

48   Viruses have a remarkable ability to adapt to the complex and hostile host immune and

49   physiological constraints. Such capability is directly associated with the viral population's genetic

50   diversity, and deep characterization of this diversity is the cornerstone of our understanding of

51   virus evolutive response to a new cellular environment. The starkly evident examples are RNA

52   viruses. These viruses, such as HIV, hepatitis C, and influenza, display high mutation rates

53   which generate significant polymorphism levels, allowing the viral population to quickly adapt to

54   newly infected cellular environments and evolve vaccine and antiviral drug resistance (Lauring

55   and Andino 2010; Loiseau et al., 2020). However, the genetic diversity has been thoroughly

56   characterized in only a handful of viruses, mainly targeting SNPs (Single Nucleotide

57   Polymorphisms) in RNA viruses (Sanjuán and Domingo-Calap 2016). Although the mutation

58   rate of large double-stranded (ds) DNA viruses is up to four folds lower than that of RNA

59   viruses, due to the use of high-fidelity proofreading polymerases, and most SNPs in dsDNA

60   viruses are neutral and at low frequency, SNP-based approaches were chosen to analyze the

61   genetic diversity in Human cytomegalovirus (HCMV in the species *Human betaherpesvirus 5,*

62   HHV-5), Autographa californica multiple nucleopolyhedrovirus (AcMNPV) and Human

63   herpesvirus 2 (HHV-2), for example (Renzette et al., 2015; Chateigner et al., 2015; Akhtar et al.,

64   2019).

65   Structural variations (SVs) play a key role in viral evolutionary processes. Genome

66   rearrangements such as deletions, insertions, duplications and inversions can lead to defective

67   viral genomes (DVGs) (O'Hara et al., 1984; Molenkamp et al., 2000; Vignuzzi and López 2019).

68 Preben von Magnus first identified DVGs in the late 40's as incomplete influenza viruses that
69 can interfere with the wild-type virus replication (Vignuzzi and López 2019). Since then, the role
70 of DVGs in antiviral immunity, viral persistence and their negative impact on virus replication
71 and production has been established (Bull et al., 2003; Li et al., 2011; Vignuzzi and López 2019;
72 Loiseau et al., 2020). Nowadays, DVGs have been described in most RNA viruses and to a
73 lesser extent in dsDNA viruses (Vignuzzi and López 2019; Loiseau et al., 2020). Despite the
74 critical role of SVs in virus infection dynamics, the knowledge about structural variation diversity,
75 and their evolutionary impact in viral populations, especially those with large dsDNA, is limited.
76 The large dsDNA Cyprinid herpesvirus 3 (CyHV-3), more commonly known as koi herpesvirus
77 (KHV), is one of the most virulent viruses of fish. It is a lethally infectious agent that infects
78 common carp and koi (*Cyprinus carpio*) at all stages of their life (Hedrick et al., 2000; Haenen et
79 al., 2004). KHV infections are usually associated with high morbidities and mortalities (up to
80 95%), resulting in serious environmental damages and severe economic losses (Sunarto et al.,
81 2011; Rakus et al., 2013). This threatening virus had a rapid worldwide spread due to global fish
82 trade and international ornamental koi exhibitions (Gotesman et al., 2013). Classified within the
83 family *Alloherpesviridae*, genus *Cyprinivirus*, CyHV-3 is the subject of an increasing number of
84 studies and has become the archetype of alloherpesviruses (Boutier et al., 2015). Despite this
85 "status", only 19 isolates have been entirely sequenced so far (source: NCBI) since the release
86 of the first complete genome sequences in 2007 (Aoki et al., 2007). Such a low number of full
87 genomes impairs large-scale phylogenomic studies (Gao et al., 2018). On the other hand, KHV
88 infections have been shown to be the result of haplotype mixtures, both *in vivo* and *in vitro*
89 (Hammoumi et al, 2016; Klafack et al, 2019). If mixed-haplotype infections probably represent
90 an additional source of diversification for KHV (Renner and Szpara, 2018), they make genomic
91 comparisons more challenging.

92

93 KHV has the largest genome among all known herpesviruses, with a size of approximately 295
94 kb and 156 predicted open reading frames (ORFs) (Aoki et al. 2007). Several studies focusing
95 on the analysis of viral ORFs have shown the implication of some of them in KHV virulence
96 (Boutier et al., 2015; Fuchs et al. 2014). KHV isolates are known to carry mutations in ORFs
97 that are likely to alter gene functions, and these mutations may vary from virus to virus (Gao et
98 al., 2018) and even within viruses (Hammoumi et al., 2016). Aoki et al (2007) hypothesized that
99 virulent KHV would have arisen from a wild-type ancestor by loss of gene function. However,
100 nearly 15 years later, this hypothesis has still not been tested, probably because of the lack of
101 extensive genomic comparisons. SVs may play a key role in this gene function loss, as recently

102 shown by Klafack et al (2019). These authors conducted a comparative study of a cell culture-
103 propagated isolate that suggested that CyHV-3 evolves through an assemblage of haplotypes
104 whose composition changes within cell passages. This study revealed a deletion of 1,363 bp in
105 the ORF150 of the majority of haplotypes after 78 passages (P78), which was not detected after
106 99 passages. Furthermore, experimental infections showed that the virus passaged 78 times
107 was much less virulent compared to the original wild-type on the one hand and slightly less
108 virulent compared to the same virus passaged 99 times (P99), highlighting the potentially
109 important role of the ORF150 in the virulence of KHV. Besides, this study demonstrated that
110 haplotype assemblages evolve very rapidly along successive *in vitro* cell passages during

Commenté [lc1]: Assemblies

111 infectious cycles, and raised many questions regarding the mechanisms leading to such rapid
112 gene loss and gain *in vitro*.
113 The present study sought to characterize the SV diversity and dynamics in the KHV genome
114 using viruses propagated onto cell cultures. First, P78 and P99 whole virus genomes were
115 sequenced using ultra-deep long-read sequencing, a first with KHV. Then, the obviously
116 pathogenesis-associated ORF150 region (~5 kb) was sequenced in ten intermediate successive
117 cell passages through an Oxford nanopore® amplicon-based sequencing approach to gain
118 insights into the gene loss and gain mechanisms.
119119

120 Material and methods

121 Extraction of high molecular weight DNA from P78 and P99 cell culture passages
122 The virus isolate used in this study was the same as that previously described in Klafack et al.
123 (2019), *i.e.* an isolate collected from an infected koi in Taiwan (KHV-T) and passed 99 times
124 onto common carp brain (CCB) cells. Considering previous results, a special focus was made
125 on passages 78 (P78) and 99 (P99). Genomic DNA was extracted from cell cultures stored at -
126 80°C, using the MagAttract HMW DNA Kit (Qiagen). Each frozen culture was thawed quickly in
127 a 37°C water bath, equilibrated to room temperature (25°C) and divided into 12 cell culture
128 aliquots of 250 µL. Tubes were centrifuged at 3,000× g for 1 minute and supernatants were
129 transferred into new 2-mL tubes containing 200 µL of proteinase K and RNase A solution. DNA
130 was subsequently extracted according to the manufacturer's recommendations and eluted in
131 200 µL distilled water provided in the kit. The 12 replicates of each sample were pooled together
132 and evaporated at room temperature using a vacuum concentrator, to reach a final volume of
133 around 60 µL. Concentrated DNA was quantified by fluorometry (Qbit, ThermoFisher Scientific)

Commenté [lc2]: Qubit

134   and its quality was evaluated by spectrophotometry (Nanodrop 2100) and agarose gel

135   electrophoresis. The final concentration of P78 and P99 was 14.4 and 2.6 ng·µL^{-1}, respectively.

136136

137   Quantitative PCR assays

138   Quantitative PCR (qPCR) was applied to evaluate cellular and viral DNA ratio. Two sets of

139   primers were used: primers targeting the ORF150 of CyHV-3 (GenBank #AP008984.1, KHV-J,

140   nt   259,965-260,110:   5'-GAGCGAGGAACTCTACACAAC-3'   and   5'-

141   GGTAAGGGTAAAGCAGACCATC-3') and primers targeting the glucokinase gene of *Cyprinus*

142   *carpio* (GenBank #AF053332.2, nt 225-293: 5'-ACTGCGAGTGGAGACACAT-3' and 5'-

143   TCAGGTGTGGAGGGGACAT-3'). Amplification reactions contained 1 µL of 2X SYBR Green I

144   Master mix (Roche), 200 nM of each primer, and 1 µl of template DNA in a final volume of 10

145   µL. Amplifications were carried out in a LightCycler 480 (Roche) and cycling conditions

146   consisted in an initial denaturation at 95°C for 5 min followed by 45 cycles of amplification at

147   95°C for 10 sec, annealing at 60°C for 20 sec and elongation at 72°C for 10 sec with a single

148   fluorescence measurement. After amplification, a melting step was applied, which comprised a

149   denaturation at 95°C for 5 sec, a renaturation at 65°C for 60 sec and a heating step from 65 to

150   97°C with a ramp of 0.1°C per second and a continuous fluorescence acquisition. Specificity of

151   amplification was verified by visual inspection of the melting profiles, and the ratio between

152   cellular and viral DNA was estimated as $2^{-\Delta Cq}$, assuming that each primer pair has an

153   amplification efficiency close to 2 and that each amplicon is present as a single copy per

154   genome.

155155

156   Genomic library preparation and MinION sequencing

157   High-quality genomic DNA from the two samples (P78 and P99) was sequenced using Oxford

158   Nanopore technology®. DNA libraries were prepared using the Ligation Sequencing Kit (SQK-

159   LSK109) according to the manufacturer's instructions (Oxford Nanopore®). A total input amount

160   of 48 µL (corresponding to 692 ng of P78 and 124 ng of P99 high molecular weight DNA) was

161   used for sequencing library preparation. DNA was first end-repaired using NEBNext FFPE DNA

162   Repair Mix and NEBNext Ultra II End repair, and then cleaned up with Agencourt AMPure XP

163   beads (Beckman Coulter Inc) at a 1:1 bead to DNA ratio. Sixty-one µL of clean-up elution were

164   transferred into a new 1.5-mL tube for subsequent adapter ligation. Adapter ligation was

165   achieved using NEBNext Quick T4 DNA ligase adapter mix (AMX), ligation buffer (LNB), Long

166   Fragment Buffer (LFB) and Elution Buffer (EB), following the provider's recommendations. The

167   quantity of the retained DNA fragments was measured again by Qbit fluorometry. The final

---

Commenté [lc3]: Maybe replace or precise « MinION » by « Whole Genome Sequencing MinION »

Commenté [lc4]: Qubit

168 amount of P78 and P99 DNA was 346 ng and 94 ng, respectively. Each library was directly
169 sequenced on R9.4.1 flow cells using a MinION sequencing device. Sequencing runs were
170 controlled with MinKNOW version 0.49.3.7 and operated for about 30 hours.

171 171

172 Amplicon-based MinIon sequencing

173 To specifically investigate ORF150 region, a fragment of ~4.3 kb encompassing the whole
174 ORF150 of CyHV-3 (257,103-261,345 according to GenBank #AP008984) was sequenced
175 from various intermediate passages of KHV-T (P10, P20, P30, P40, P50, P70, P78, P80, P90
176 and P99). A ligase-free protocol was used to limit the risk of potential artifacts linked with
177 sample preparation, e.g. the creation of chimeric sequences during the end-repair or ligation
178 steps (White et al., 2017). Moreover, the inversion found in the middle of the reads excluded the
179 formation of *in silico* chimeras, i.e., chimeras resulting from the basecaller when two molecules
180 are sequenced in the same pore that undergoes fast reloading (Martin and Legget, 2021).
181 Genomic DNA was extracted from cell cultures stored at -80°C, using the Nucleospin virus kit
182 (Macherey-Nagel). Purified DNA was subsequently used for PCR amplification with the
183 following primers: 5'-TGGGCGCAATCAAGATGT-3' (F) and 5'-TGAAGTGTCAGGGTCAGAGT-
184 3' (R). PCR was performed with GoTaq G2 DNA Polymerase (Promega) in a final volume of 40
185 µL containing 1 µL of total genomic DNA, 10 µL of Taq Buffer, 5 µL of dNTPs 2 mM, 2.5 µL of
186 $MgCl_2$ and 1.25 µL of each primer (10 µM). Cycling conditions were as follows: initial
187 denaturation at 95°C for 10 min, amplification with 40 cycles of 95°C for 10 sec, 60°C for 20
188 sec, 72°C for 3 min, and final extension at 72°C for 5 min. PCR products were purified using 1X
189 Agencourt AMPure XP beads, tested for purity using the NanoDrop™ One spectrophotometer
190 (ThermoFisher Scientific), and quantified fluorometrically using the Qubit dsDNA High sensitivity
191 kit. DNA libraries were prepared using the Rapid Barcoding kit (SQK-RBK004), following the
192 manufacturer's instructions. For each sample, 400 ng of purified amplicon were adjusted with
193 nuclease-free water to a total volume of 7.5 µL and supplemented with 2.5 µL of Fragmentation
194 Mix RB01-4 (one for each sample). Four barcoded samples were combined with an equimolar
195 ratio by mixing 2.5 µL of each sample in a total volume of 10 µL. Pooled libraries were
196 sequenced on 2 R9.4.1 flow cells for 24 hours and sequencing runs were controlled with
197 MinKNOW version 0.49.3.7.

198 DNA sequence analysis

199 For each sample, bases from raw FAST5 files with a *pass* sequencing tag were recalled using

---

Commenté [lc5]: Question: Did you well used one flow cell per sample ?

Commenté [lc6]: VH: For me, it is not exactly an amplicon but more a « target region » , what do you think ?

Commenté [lc7]: Did you really performed the sequencing for these ten passages, because you precised line 194 using « Four barcodes » and two flow cells (So 8 samples and not ten) ?

Commenté [lc8]: VH: For me, this sentence is not a Method, and should be find in the result or in the Discussion part..

200   the high-accuracy model of ONT Guppy basecalling software version 4.0.15 which improves the

201   basecalling accuracy. The obtained fastq files were filtered to keep reads with a length ≥ 2 kb.

202   Sequencing depth was calculated for each sample using the *plotCoverage* tool implemented in

203   deepTools2.0 tool suite (Ramírez et al*.,* 2016). Sequencing coverage was assessed with the

204   bamCoverage tool from the same tool suite and normalized using the RPGC (reads per genome

205   coverage) method (Figure 1).

206206

207   Structural variant detection

208   To detect structural variants (SVs) in the P78 and P99 whole genomes, a mapping step followed

209   by BAM filtering was performed. Two aligners were used to map the raw long-reads against the

210   KHV-J AP008984.1 reference genome: minimap2 (Li, 2018) and NGMLR (Sedlazeck et al.,

211   2018). BAM files were then filtered using the option '-F' of Samtools view (Li et al., 2018) with

212   the flag "4" to keep only mapped reads and with the flag "0x800" to remove chimeric reads

213   (inconsistent/supplmental mapping). For P78 and P99, 99.33% and 97.77% of reads were

214   mapped, respectively, before the filtering steps. Chimeric reads represented 28.05% and

215   17.74% of the mapped reads in P78 and P99, respectively. The resulting filtered BAM files from

216   each mapper were used as input data for SV caller, Sniffles (Sedlazeck et al., 2018). Only SVs

217   ≥ 30 bp and supported by at least 10 reads were kept in the final VCF files. A cross-validation

218   step was performed using SURVIVOR (Jeffares et al., 2017) by extracting common SVs from

219   each mapper/caller combination for each sample. Although the KHV-J reference genome used

220   for the mapping is phylogenetically close to the KHV-T isolate, some genetic diversity exists

221   (Klafack et al. 2017). Hence, to exclude inter-isolate SVs, a pairwise comparison between P78

222   and P99 was made.

223   The distribution of the different SVs along with P78 and P99 genomes was assessed by

224   estimating their occurrences using a 5 kb sliding window. SNPs and Indels variants were called

225   in P78 and P99 by *medaka_variant* implemented in medaka (1.4.4) using KHV-J AP008984.1

226   as a reference genome. To detect structural variants in the amplified region (257,103-261,345)

227   of P10 to P99 samples, a size filtering step using *guppyplex* was added to the steps described

228   above. Only reads from 1.5 kb to 8 kb were used for the analysis.

229229

230   Results

231   Main features of sequencing data for P78 and P99

232   A total of 4,900,000 and 2,293,830 long-reads were obtained for P78 and P99, respectively.

Commenté [lc9]: VH: This step is not really clear. The basecalling with Guppy could have been done with High-accuracy parameters during the sequencing process.

Commenté [lc10]: VH: Could you precise why you selected this length please ?

Commenté [lc11]: VH: Did you performed a mapping process to evaluate the sequencing depth, if so, maybe precise the mapper at this moment? (As Line 210 « minimap2… »)

Commenté [lc12]: VH: Did these steps concern the length filtering ?

233   After filtering, 462,982 long-reads with an average length of 4.96 kb were retained for P78, and
234   418,034 reads with an average length of 7.06 kb for P99 (Table 1, Table S1). The ratio between
235   cellular and viral DNA was 34800 and 33200 for P78 and P99, and the percent of mapped reads
236   99.33% and 97.77%, respectively. 100% of the sampled bases from the P78 genome had at
237   least 5,000 overlapping reads and 100% of the sampled bases from the P99 genome had at
238   least 7,500 overlapping reads (Figure S2). Both P78 and P99 genomes were entirely covered
239   by the sequencing data (Figure 1).

240240
241241
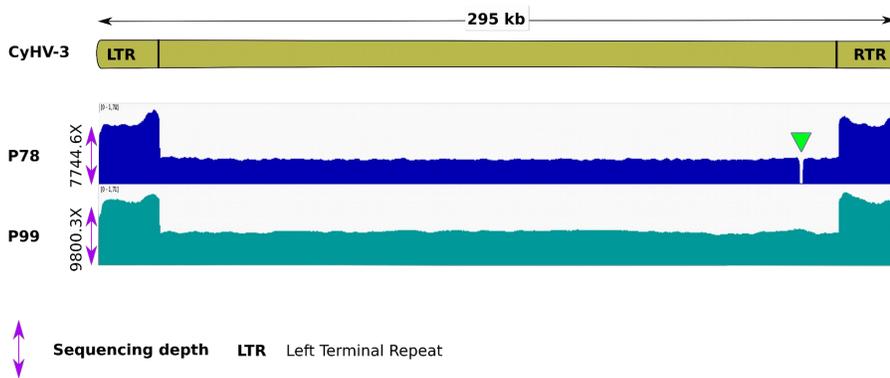
242
243243

244   **Figure 1.** Normalized sequencing coverage for P78 and P99 samples using the RPGC (reads
245   per genome coverage) method. Both P78 and P99 genomes were totally covered by the
246   sequencing. For P78, the coverage break (green triangle) corresponds to the 1.3 kb deletion.

247247
248248

249   **Table 1** Main features of reads obtained for each genome after the read length filtering step.
250   min_len= minimum length, avg_len= average length, max_len= maximum length.

Sample	reads	min_len	avg_len	max_len	N50	Q20(%)	Q30(%)
P78	462982	2000	4960.8	91494	6091	59.53	17.71
P99	418034	2000	7061.4	176512	8845	58.27	19.22

251   N50= the length for which all sequenced reads of that length or greater sum to 50% of the set's
252   total size Q20=1 in 100 probability of an incorrect base call, Q30= 1 in 1000 probability of an
253   incorrect base call. Here 59.53% of P78 bases=Q20 and 58.27% of P99 bases=Q20.

254254
255255

Commenté [lc13]: The ratio is based on PCR results ? If so maybe the sentence need to be move to the end of the paragraph ?

Commenté [lc14]: VH: Perhaps, it could be interesting to add the position of the ORF150 on the CyHV-3 genome schema ..

Commenté [lc15]: VH: These Q score values are surprising, they seems to correspond to Illumina data. For the MinION we use usually Q7 or Q9 ..

256   SV distribution in P78 and P99

257   For P78, the mapper/caller combination minimap2/Sniffles detected 731 structural variations

258   (SVs), and the combination NGMLR/Sniffles detected 460 SVs (Table S2). For P99, the

259   combination minimap2/Sniffles detected 210 SVs and NGMLR/Sniffles detected 397 SVs (Table

260   S2). Independently from the mapper/caller combination that was used, P78 showed more SVs

261   than P99 (Table S2). After the cross-validation step (extracting common SVs from each mapper/

262   caller combination), 236 and 87 SVs were kept for P78 and P99, respectively. For comparison,

263   the number of short SVs (with a size < 30 nt) amounted to 57 and 77 for P78 and P99,

264   respectively.

265   In both samples, inversions and deletions were the most prevalent SVs. In P78, inversions

266   represented 75% of the events and deletions 20% (Figure 2.A). In P99, 80% of the SVs were

267   inversions and 11% were deletions (Figure 2.A). Inversions were found along the entire

268   genome, with the highest number detected within the [70-75 kb] window in P78 and within the

269   [235-245 kb] and [270-275 kb] windows in P99 (Figure 2.B). In P78, deletions were mainly

270   detected within [10-40 kb], [70-100 kb] and [245-270 kb] windows (Figure 2.B). In P99, the two

271   major deletions were found within [10-20 kb] and [205-215 kb] regions (Figure 2.B). The

272   frequencies of these SVs were low and did not exceed 1% of the total reads (with a few

273   exceptions, Table S2). In spite of such low frequencies, it is interesting to note that the most

274   frequent SVs were located in ORFs potentially involved in DNA replication and encapsidation,

275   e.g. ORF33, 46, 47, 55 (Aoki et al, 2007; Table S2).

276   Altogether, these results highlight high SV turnover dynamics during the *in vitro* infection cycles

277   (from 78 passages to 99) with a clear bias towards inversion and deletion events.

278   Dynamics and impacts of SVs in ORF150 region

279   Taking advantage of the high-resolution SV detection provided by the long-read sequencing, we

280   looked for the SV events around the potential virulence-linked ORF150 in P78 and P99 (nt

281   257,103-261,345 according to AP008984.1). Results confirmed that P99 had a reference-like

282   profile with an unmodified ORF150. In P78, the deletion (nt 258,154-259,517; D258153) was

283   found in 6,902 reads (100% of the reads), whereas the reference haplotype was also detected

284   in 30 reads, representing 0.44% of the total 6,734 supporting reads (Figure 3, Tables S3).

285   Surprisingly, 26 reads revealed a haplotype as yet unidentified (INV258153), consisting of an

286   inversion of the same length (1,363 bp) and at the same breakpoints as the deletion. The

287   inverted haplotype (INV258153) in P78 deeply altered the ORF150, by inverting the first 1200

288   bp of the ORF and 160 bp of the 5'UTR in the middle of the ORF (Figure 3).

289   In order to trace the unexpected dynamics of gain and loss of the full ORF150 along passages,

Commenté [lc16]: VH: on Line217 you indicated that you kept only > or = 30. Could you precise this information on line 263 ?

Commenté [lc17]: VH: I do not correctly understand this sentence ? Why a bias ? Could you develop a little ?

290 we searched for the SV turnovers during 10 intermediate passages (P10, P20, P30, P40, P50,

291 P70, P78, P80, P90, P99). This analysis revealed the presence of haplotype D258153 at low

292 frequency (from 0.05 to 0.15% of the reads) in passages P10 to P40 and a strong increase in its

293 frequency at P50 (88.7% of the reads) (Figure 4). The frequency of the haplotype D258153

294 reached a maximum at P78 (100% of the reads) then dropped quickly at P80 (30.7% of the

295 reads) to stabilize at low frequency (0.31% of readings) at P90., as during the first 40 passages

296 (Figure 4, Table S3). Interestingly, shorter deletions of 119 and 881 bp were observed near the

297 5' end of the ORF150 in P40 and P80, respectively, at low frequencies (0.42 % in P40 and

298 0.18% in P80) (Figure 4, Table S3). The haplotype D258153 completely disappeared at

299 passage 99 (Figure 4, Table S3).

300 This analysis also evidenced several other SVs that alter the structure of ORF150 and of its

301 upstream region, including the beginning of ORF149 (Figure 5). Besides the large deletion,

302 inversions and insertions were also observed in the ORF149-ORF150 region. Inversions were

303 at a low frequency (between 0.01% and 0.53% of the supporting reads) in all passages except

304 for P70, P90 and P99. P10 and P40 showed the lowest and the highest inversion frequencies,

305 respectively (Figure 5, Table S3). A large insertion of about 1 kb appeared in P50 and P70 at

306 moderate frequencies (14,34% and 16.01% of the supporting reads, respectively) to disappear

307 in P78 and re-appear at a lower frequency (6,79% of the supporting reads) in P80. The

308 consensus sequence of this insertion corresponds to the fragment 259,517-260,477 of the KHV

309 genome, with an identity of about 90%. In P90, an intriguing inverted-duplicated haplotype was

310 observed at a low proportion (0.054% of the supporting reads). Surprisingly, P99 exhibited a

311 unique reference-like, SV-free haplotype (Figure 5, Table S3). All the variations deeply impacted

312 the structure of ORF150 - and sometimes that of ORF149 as well - by shrinking or increasing its

313 size, causing the ORF149 and ORF150 fusion, inverting the ORF150 sequences and

314 duplicating the ORF150 with the deleted, inserted, inverted and inverted-duplicated haplotypes

315 (Figure 5).

316316

317317

318318

319319

320320

321321

322322

> **Commenté [lc18]:** VH: You precise in Material investigated only the ORF150 (Line173) so how did you obtained the ORF149 ? Not clear enough ..
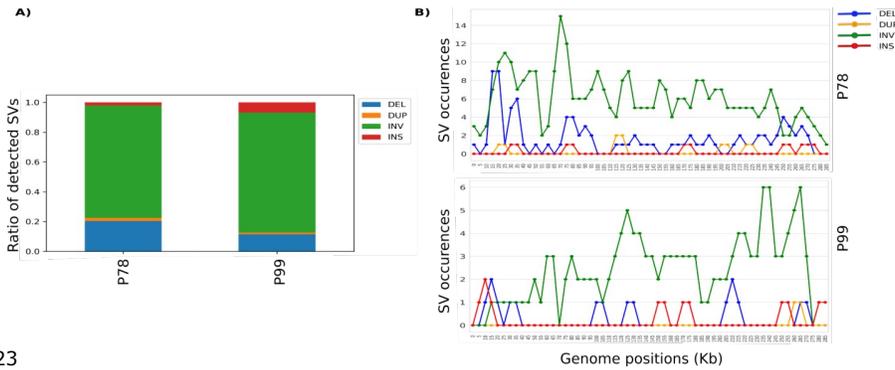
323323

324 **Figure 2.** The ratio and distribution of SVs detected in P78 and P99 genomes. **A)** The
325 frequency of each type of SV detected in P78 and P99. **B)** The occurrences and distribution of
326 each SV type in the P78 and P99 genomes using a 5kb window.

327327

328328

329329

330 Figure 3. Impact of structural variations D258153 and INV258153 on the P78 genomic
331 **Figure 3**. Impact of structural variations D258153 and INV258153 on the P78 genomic
332 structure. The inversion is highlighted by an inverted arrow compared to the reference
333 haplotype.

334334

335335

336336

337337

338338

339339

Commenté [lc19]: VH: it is not the same scale for P78 and P99 and it could be interesting to see the region of ORF150 at the top of the graph B.

Commenté [lc20]: VH: The scale indicate the length of ORF150 only so it is not really clear with the presence of ORF149. Maybe you could replace the scale or replace the scale with ORF150

Percentage of deleted vs reference haplotypes
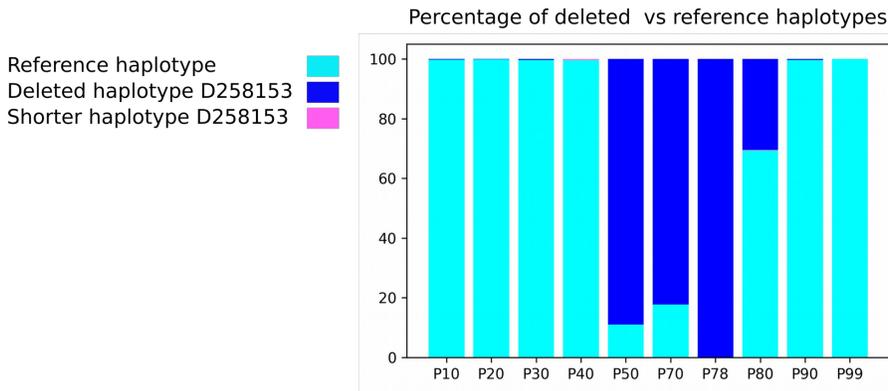
Figure 4. The prevalence of the deleted haplotype during the 10 intermediate passages (P10, P20, P30, P40, P50, P70, P78, P80, P90, P99)
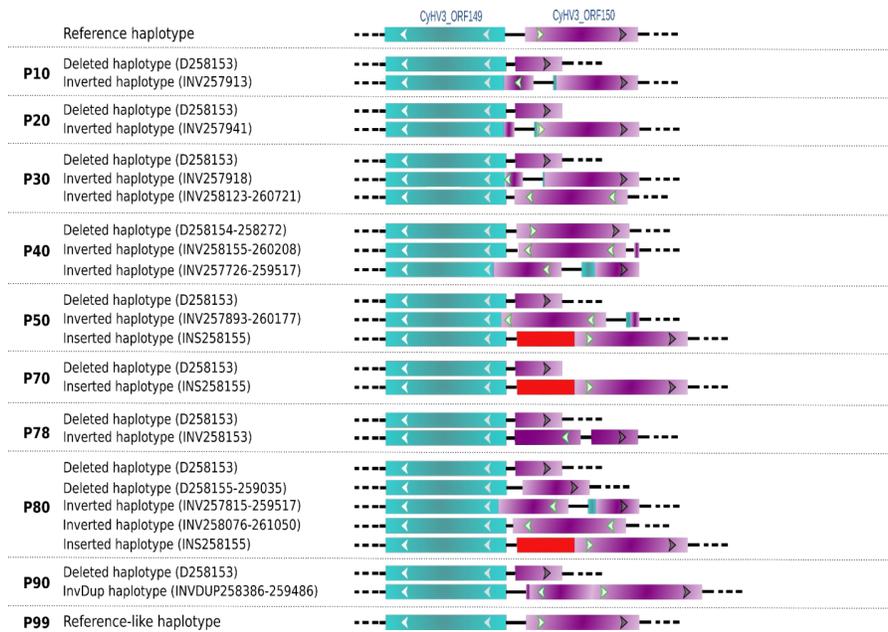
Figure 5. Impact of SV dynamics on the ORF149-ORF150 structure in the successive passages P10, P20, P30, P40, P50, P70, P78, P80, P90 and P99. The inversion is highlighted by an inverted arrow compared to the reference haplotype. Red blocks correspond to an inserted

348   sequence.
349   InvDup = Inverted-duplicated haplotype.
350350

351   Discussion

352   SVs significantly impact the adaptation of viruses to their natural host and environment (Pérez-
353   Losada et al., 2015). Yet the role of SV diversity and dynamics in large DNA viruses is barely
354   known. Ultra-deep long-read sequencing opens unprecedented ways to gain insights into these
355   untapped viral genome polymorphisms. The present study started to tackle the impact of SVs in
356   the evolution of the large dsDNA KHV during cell culture serial passages using ultra-deep
357   whole-genome and amplicon-based sequencing. The sequence data showed a wide distribution
358   of various SVs along the genome associated with high SV turnover dynamics during the *in vitro*
359   infection cycles and a clear bias towards inversion and deletion events. Analysis of the
360   pathogenicity-associated ORF150 region in ten serial passages mainly highlighted deletions,
361   inversions and insertions that deeply altered the structure of ORF150.

362   Serial passages of viruses in cell culture may lead to the accumulation of mutations and gene
363   disruptions (Spatz 2010; Colgrove et al., 2014). These mutations can modify viral adaptation
364   and increase or decrease virulence (Boutier et al., 2017; López-Muñoz et al., 2021; Vancsok et
365   al., 2017). In the case of KHV, a previous work using short-read sequencing showed that 99
366   consecutive *in vitro* passages onto CCB cells resulted in the accumulation of less than 60 small
367   variations (<100 nt) (Klafack et al., 2019). It also showed that the haplotype composition can
368   quickly vary along with infection cycles of KHV *in vitro*. The present study unexpectedly
369   highlighted a high number of structural variations: 87 for P99 and 236 for P78. In contrast, the
370   accumulation of small variations was consistent with what had been observed with short-read
371   sequencing (Klafack et al, 2019). These findings illustrate that long-read sequencing is highly
372   suitable for genome-wide comparisons of viruses. Most importantly, they revealed a hidden
373   source of virus diversification, which had never been reported so far for KHV. They also
374   confirmed that P78 consists of a mixture of undeleted and deleted haplotypes revealing that the
375   undeleted haplotype did not correspond to the native one but to an inverted version of it. We
376   experimentally validated this inversion in P78 by PCR followed by sequencing. These structural
377   variations form a 'mosaic' of viral subpopulations that seem to result from multiple
378   rearrangement events, mainly inversions and deletions and, to a lesser extent, insertions and
379   duplications. Such sub-viral variants often lead to Defective Viral Genomes (DVGs) (Vignuzzi
380   and López 2019). Because of their negative impact on viral replication, some forms of DVGs
381   have been extensively studied, and three pathogenesis-related functions have been well-

382 described: interference with viral replication, immunostimulation, and viral persistence (Marriott
383 and Dimmock 2010; Vignuzzi and López 2019). DVGs play a role in viral production interference
384 by accumulating at higher rates than the full-length viral genomes and consequently interfere
385 with viral replication by taking up the polymerase activity and competing for structural proteins
386 (Calain and Roux 1995; Portner and Kingsbury 1971). In addition, DVGs can act as primary
387 stimuli and trigger antiviral immunity by inducing the expression of some interleukins and pro-
388 inflammatory cytokines (for review, see Vignuzzi and López 2019; Xiao et al, 2021). However,
389 the implication of DVGs for virus persistence is complex and still unclear. Nevertheless, a
390 combination of viral replication interfering and cycle asynchrony between full-length viruses and
391 DVGs seems to establish chronical infection (for review, see Vignuzzi and López 2019).
392 The observed dynamics of structural variations in the KHV genomes during cell culture
393 passages where selection pressures are virtually non-existent may indicate an accumulation of
394 DVGs impacting the pathogenicity. Indeed, Klafack et al. (2019) showed that serial passages
395 significantly attenuated this infectious virus. This attenuation was associated with a truncated
396 KHV genome bearing a 1.3-kb deletion that removes the majority of the ORF150. A very recent
397 study showed that partial or complete depletion of ORF150 leads to a clear attenuation of the
398 virus. It also brought *in vivo* evidence that ORF150 may down-regulate the inflammatory
399 response of carp to enhance viral proliferation, thus confirming a key role of ORF150 product in
400 KHV virulence (Klafack et al., 2022 submitted). ORF150 seems to belong to the RING family
401 genes. In aquatic viruses, RING family genes have been reported to be involved in virus
402 latency, replication, and host protein degradation (Shekar and Venugopal 2019; Wang et al.,
403 2021).
404 As previously observed, the most significant difference affecting all viral haplotypes between
405 P78 and P99 was around the ORF150. For this reason, we focused on this interesting region to
406 determine the passage number at which the presence of the deletion inversion first arose, by
407 sequencing selected representative passages (P10, P20, P30, P40, P50, P70, P80 and P90).
408 We found that deleted and inverted haplotypes appeared as soon as P10 and that their
409 proportion varied along the successive passages. The most striking feature was the rapid and
410 total disappearance of the deletion between P8 and P99, raising many questions regarding the
411 mechanisms that led to the clearance of this major haplotype.
412 The rapid SV turnover of DNA viruses, including herpesviruses, likely involves recombination
413 (Szpara and VanDoorslaer, 2021; Renner and Szpara, 2018; Cudini et al., 2019; Wilkinson and
414 Weller, 2003; Kolb et al., 2017; Tomer et al., 2019), which is often linked to replication and DNA
415 repair, as well as errors during viral genome replication (Kulkarni and Fortunato 2011; Xiaofei

Commenté [lc21]: P78

416 and Kowalik 2014). In the present case, the pretty good conservation of breakpoints around
417 ORF150 may be a sign of homologous recombination. However, whether this recombination
418 occurs within the same genomic entities or between different viruses remains open. It would be
419 interesting to assess the involvement of each of these mechanisms in generating the observed
420 structural diversity. The multiple mechanisms of DNA virus evolution beyond single nucleotide
421 substitutions likely confer KHV a high level of evolutionary adaptability.

422 Classically, the generation of live-attenuated vaccines is achieved by passaging the virus in cell
423 culture under different conditions (in different host species or at lower or higher temperatures),
424 in order to induce mutation accumulation that supports viral adaptation to the specific conditions
425 and provides viral attenuation (Minor, 2015, Hanley 2011). With the exception of the OPV polio
426 vaccine viruses (Kew et al., 2005), the exact mechanisms by which these mutations lead to
427 attenuated phenotypes are usually poorly characterized (Lauring et al., 2010). However, live-
428 attenuated viruses can revert to virulent phenotypes either by reversions (as shown here
429 between P78 and P99), introduction of compensatory mutations, or recombination with viruses
430 belonging to the same genus (Cann et al., 1984; Bull et al., 2018; Muslin et al., 2019).
431 Additionally, the combination of multiple live-attenuated viruses may result in competition or
432 facilitation between individual vaccine viruses, resulting in undesirable increases in virulence or
433 decreases in immunogenicity (Hanley 2011; Pereira-Gomez et al., 2021). Recently, genetic
434 engineering has led to many novel approaches to generate live-attenuated virus vaccines that
435 contain modifications to prevent reversion to virulence (Yeh et al., 2020) and improve
436 interferences among multiple vaccine strains (Pereira-Gomez et al., 2021).

437 Conclusion

438 Our findings confirm that CyHV-3 can evolve rapidly during infectious cycles in cell culture, and
439 SVs are a major component in the evolutionary process of this virus. SVs are extremely
440 dynamic under *in vitro* controlled conditions, and it would now be interesting to evaluate their
441 dynamics *in vivo*. The present study also contributes to the basic research on the mechanisms
442 underlying attenuation, and may have important outcomes for the design of safe live-attenuated
443 vaccine formulations.

444 Acknowledgements

457 Conflict of interest disclosure

458 The authors declare they have no conflict of interest relating to the content of this article.

459 Authors' contributions

460 **Nurul Novelia Fuandila:** Investigation; Methodology; Formal analysis; Writing – original draft;

461 Data Validation

462 **Anne-Sophie Gosselin-Grenet:** Investigation; Methodology; Data curation; Data Validation;
463 Writing – original draft
464 464
465 **Marie-Ka Tilak:** Investigation; Methodology ; Data Validation
466 466
467 **Sven M Bergmann**: Resources; Writing – review and editing
468 468
469 **Sandro Klafack:** Resources; Writing – review and editing
470 470
471 **Angela Lusiastuti:** Methodolgy; Writing – review and editing
472 472
473 **Munti Yuhanna:** Methodolgy; Writing – review and editing

474 **Jean-Michel Escoubas:** Methodology; Funding acquisition; Writing – review and editing

475 **Anna-Sophie Fiston-Lavier:** Writing – review and editing

476 **Jean-Christophe Avarre:** Conceptualization; Data curation; Methodology; Funding acquisition;
477 Data Validation; Writing – original draft

478 **Emira Cherif:** Conceptualization; Data curation; Formal analysis; Software; Writing – original
479 draft;

480 480

481 Data availability

482 Raw sequences (fastq files) were stored in the public Sequence Read Archive (SRA) repository

483 and can be accessed under the bioproject PRJNA511566.

484 The original genetic data yielded by genotyping (in vcf format) and additional metadata are

485 available in a publicly-available OSF repository:

486 https://osf.io/3c2ag/?view_only=c9cc68a2cc3943138373eda1ed05ed25

487 References

488 Akhtar, L. N., Bowen, C. D., Renner, D. W., Pandey, U., Della Fera, A. N., Kimberlin, D. W., ... &
489 Szpara, M. L. (2019). Genotypic and phenotypic diversity of herpes simplex virus 2 within the
490 infected   neonatal   population.   MSphere,   4(1),   e00590-18.
491 https://doi.org/10.1128/mSphere.00590-18

492 Aoki, T., Hirono, I., Kurokawa, K., Fukuda, H., Nahary, R., Eldar, A., ... & Hedrick, R. P. (2007).
493 Genome sequences of three koi herpesvirus isolates representing the expanding distribution of
494 an emerging disease threatening koi and common carp worldwide. Journal of virology, 81(10),
495 5058-5065.

496 Boutier, M., Ronsmans, M., Rakus, K., Jazowiecka-Rakus, J., Vancsok, C., Morvan, L., ... &
497 Vanderplasschen, A. (2015). Cyprinid herpesvirus 3: an archetype of fish alloherpesviruses.
498 Advances in Virus Research, 93, 161-256.

499 Boutier, Maxime; Gao, Yuan; Vancsok, Catherine; Suárez, Nicolás M.; Davison, Andrew J.;
500 Vanderplasschen, Alain (2017): Identification of an essential virulence gene of cyprinid
501 herpesvirus 3. In: *Antiviral research* 145, S. 60–69. DOI: 10.1016/j.antiviral.2017.07.002.

502 Bull, J. C., Godfray, H. C. J., & O'Reilly, D. R. (2003). A few-polyhedra mutant and wild-type
503 nucleopolyhedrovirus remain as a stable polymorphism during serial coinfection in Trichoplusia
504 ni.   Applied   and   environmental   microbiology,   69(4),   2052-2057.
505 https://doi.org/10.1128/AEM.69.4.2052-2057.2003

506 Bull, J. J., Smithson, M. W., & Nuismer, S. L. (2018). Transmissible viral vaccines. Trends in
507 microbiology, 26(1), 6-15.

508 Calain, P., & Roux, L. Functional characterisation of the genomic and antigenomic promoters of
509 Sendai virus. Virology 212, 163–173 (1995).

510 Cann, A. J., Stanway, G., Hughes, P. J., Minor, P. D., Evans, D. M. A., Schild, G. T., & Almond,

511   J. (1984). Reversion to neurovirulence of the live-attenuated Sabin type 3 oral pollovirus
512   vaccine. Nucleic acids research, 12(20), 7787-7792. https://doi.org/10.1093/nar/12.20.7787

513   Cann, A. J., Stanway, G., Hughes, P. J., Minor, P. D., Evans, D. M. A., Schild, G. T., & Almond,
514   J. W. (1984). Reversion to neurovirulence of the live-attenuated Sabin type 3 oral pollovirus
515   vaccine. Nucleic acids research, 12(20), 7787-7792.

516   Chateigner, A., Bézier, A., Labrousse, C., Jiolle, D., Barbe, V., & Herniou, E. A. (2015). Ultra
517   deep sequencing of a baculovirus population reveals widespread genomic variations. Viruses,
518   7(7), 3625-3646. https://doi.org/10.3390/v7072788

519   Colgrove, R., Diaz, F., Newman, R., Saif, S., Shea, T., Young, S., ... & Knipe, D. M. (2014).
520   Genomic sequences of a low passage herpes simplex virus 2 clinical isolate and its plaque-
521   purified derivative strain. Virology, 450, 140-145. DOI: 10.1016/j.virol.2013.12.014.

522   Cudini, J., Roy, S., Houldcroft, C. J., Bryant, J. M., Depledge, D. P., Tutill, H., ... & Breuer, J.
523   (2019). Human cytomegalovirus haplotype reconstruction reveals high diversity due to
524   superinfection and evidence of within-host recombination. Proceedings of the National Academy
525   of Sciences, 116(12), 5693-5698. https://doi.org/10.1073/pnas.1818130116

526   Fuchs, Walter; Granzow, Harald; Dauber, Malte; Fichtner, Dieter; Mettenleiter, Thomas C.
527   (2014): Identification of structural proteins of koi herpesvirus. In: *Archives of virology* 159 (12),
528   S. 3257–3268. DOI: 10.1007/s00705-014-2190-4.

529   Gao, Y., Suárez, N. M., Wilkie, G. S., Dong, C., Bergmann, S., Lee, P. Y. A., ... & Boutier, M.
530   (2018). Genomic and biologic comparisons of cyprinid herpesvirus 3 strains. Veterinary
531   research, 49(1), 1-11.

532   Gotesman, M., Kattlun, J., Bergmann, S. M., & El-Matbouli, M. (2013). CyHV-3: the third
533   cyprinid   herpesvirus.   Diseases   of   aquatic   organisms,   105(2),   163-174.
534   https://doi.org/10.3354/dao02614

535   Haenen, O. L. M., Way, K., Bergmann, S. M., & Ariel, E. (2004). The emergence of koi
536   herpesvirus and its significance to European aquaculture. Bulletin of the European Association
537   of Fish Pathologists, 24(6), 293-307.

538   Hammoumi, Saliha; Vallaeys, Tatiana; Santika, Ayi; Leleux, Philippe; Borzym, Ewa; Klopp,
539   Christophe; Avarre, Jean-Christophe (2016): Targeted genomic enrichment and sequencing of
540   CyHV-3 from carp tissues confirms low nucleotide diversity and mixed genotype infections. In:
541   *PeerJ*, e2516. DOI: 10.7717/peerj.2516.

542   Hanley, K. A. (2011). The double-edged sword: How evolution can make or break a live-
543   attenuated virus vaccine. Evolution: Education and Outreach, 4(4), 635-643.

544   Hedrick, R. P., Gilad, O., Yun, S., Spangenberg, J. V., Marty, G. D., Nordhausen, R. W., ... &
545   Eldar, A. (2000). A herpesvirus associated with mass mortality of juvenile and adult koi, a strain
546   of common carp. Journal of Aquatic Animal Health, 12(1), 44-57.

547 Jeffares, D. C., Jolly, C., Hoti, M., Speed, D., Shaw, L., Rallis, C., ... & Sedlazeck, F. J. (2017).
548 Transient structural variations have strong effects on quantitative traits and reproductive
549 isolation in fission yeast. Nature communications, 8(1), 1-11. doi:10.1038/NCOMMS14061

550 Kew, O. M., Sutter, R. W., de Gourville, E. M., Dowdle, W. R., & Pallansch, M. A. (2005).
551 Vaccine-derived polioviruses and the endgame strategy for global polio eradication. Annu. Rev.
552 Microbiol., 59, 587-635.

553 Klafack, S., Fiston-Lavier, A. S., Bergmann, S. M., Hammoumi, S., Schröder, L., Fuchs, W., ... &
554 Avarre, J. C. (2019). Cyprinid herpesvirus 3 evolves in vitro through an assemblage of
555 haplotypes that alternatively become dominant or under-represented. Viruses, 11(8), 754.
556 https://doi.org/10.3390/v11080754

557 Klafack, S., Schröder, L., Jin, Y., Lenk, M., Lee, P.Y., Fuchs, W, Avarre, J. C., Bergmann, S. M.
558 (2022). Development of an attenuated vaccine against Koi Herpesvirus Disease (KHVD)
559 suitable for oral administration and immersion. (submitted, NPJ Vaccines)

560 Klafack, S., Wang, Q., Zeng, W., Wang, Y., Li, Y., Zheng, S.C., Kempter, J., Lee, P.Y., &
561 Bergmann, S. M. (2017). Genetic Variability of Koi Herpesvirus In vitro—A Natural Event?.
562 Front. Microbiol., 8, 982. https://doi.org/10.3389/fmicb.2017.00982

563 Kolb, A. W., Lewin, A. C., Moeller Trane, R., McLellan, G. J., & Brandt, C. R. (2017).
564 Phylogenetic and recombination analysis of the herpesvirus genus varicellovirus. BMC
565 genomics, 18(1), 1-17. https://doi.org/10.1186/s12864-017-4283-4

566 Kulkarni, A. S., & Fortunato, E. A. (2011). Stimulation of homology-directed repair at I-SceI-
567 induced DNA breaks during the permissive life cycle of human cytomegalovirus. Journal of
568 virology, 85(12), 6049-6054. https://doi.org/10.1128/JVI.02514-10

569 Lauring, A. S., & Andino, R. (2010). Quasispecies theory and the behavior of RNA viruses.
570 PLoS pathogens, 6(7), e1001005. https://doi.org/10.1371/journal.ppat.1001005

571 Lauring, A. S., Jones, J. O., & Andino, R. (2010). Rationalizing the development of live
572 attenuated virus vaccines. Nature biotechnology, 28(6), 573-579.
573 https://doi.org/10.1038/nbt.1635
574 Lecompte, L., Peterlongo, P., Lavenier, D., & Lemaitre, C. (2020). SVJedi: Genotyping structural
575 variations with long reads. Bioinformatics, 36(17), 4568-4575.
576 https://doi.org/10.1093/bioinformatics/btaa527

577 Li, D., Lott, W. B., Lowry, K., Jones, A., Thu, H. M., & Aaskov, J. (2011). Defective interfering
578 viral particles in acute dengue infections. PloS one, 6(4), e19447.
579 https://doi.org/10.1371/journal.pone.0019447

580 Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics, 34(18),
581 3094-3100. doi: 10.1093/bioinformatics/bty191

582 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., ... & Durbin, R. (2009). The
583 sequence alignment/map format and SAMtools. Bioinformatics, 25(16), 2078-2079.
584 https://doi.org/10.1093/bioinformatics/btp352

585 Loiseau, V., Herniou, E. A., Moreau, Y., Lévêque, N., Meignin, C., Daeffler, L., ... & Gilbert, C.
586 (2020). Wide spectrum and high frequency of genomic structural variation, including
587 transposable elements, in large double-stranded DNA viruses. Virus evolution, 6(1), vez060
588 https://doi.org/10.1093/ve/vez060

589 López-Muñoz, A. D., Rastrojo, A., Martín, R., & Alcamí, A. (2021). Herpes simplex virus 2 (HSV-
590 2) evolves faster in cell culture than HSV-1 by generating greater genetic diversity. PLoS
591 pathogens, 17(8), e1009541.

592 Marriott, A. C., and Dimmock, N. J. (2010) 'Defective Interfering Viruses and Their Potential as
593 Antiviral Agents', Reviews in Medical Virology, 20: 51–62.

594 Martin, S., & Leggett, R. M. (2021). Alvis: a tool for contig and read ALignment VISualisation
595 and chimera detection. BMC bioinformatics, 22(1), 1-10. https://doi.org/10.1186/s12859-021-
596 04056-0

597 Minor, P. D. (2015). Live attenuated vaccines: Historical successes and current challenges.
598 Virology, 479, 379-392.

599 Molenkamp, R., Rozier, B. C., Greve, S., Spaan, W. J., & Snijder, E. J. (2000). Isolation and 600
characterization of an arterivirus defective interfering RNA genome. Journal of virology, 74(7), 601
3156-3165. https://doi.org/10.1128/JVI.74.7.3156-3165.2000

602 Muslin, C., Mac Kain, A., Bessaud, M., Blondel, B., & Delpeyroux, F. (2019). Recombination in
603 enteroviruses, a multi-step modular evolutionary process. Viruses, 11(9), 859.

604 O'Hara, P. J., Nichol, S. T., Horodyski, F. M., & Holland, J. J. (1984). Vesicular stomatitis virus 605
defective interfering particles can contain extensive genomic sequence rearrangements and 606 base
substitutions. Cell, 36(4), 915-924. https://doi.org/10.1016/0092-8674(84)90041-2

607 Pereira-Gómez, M., Carrau, L., Fajardo, Á., Moreno, P., & Moratorio, G. (2021). Altering 608
Compositional Properties of Viral Genomes to Design Live-Attenuated Vaccines. Frontiers in 609
Microbiology, 1706. https://doi.org/10.3389/fmicb.2021.676582

610 Pérez-Losada, M., Arenas, M., Galán, J. C., Palero, F., & González-Candelas, F. (2015). 611
Recombination in viruses: mechanisms, methods of study, and evolutionary consequences. 612
Infection, Genetics and Evolution, 30, 296-307.

613 Portner, A., & Kingsbury, D. W. Homologous interference by incomplete Sendai virus particles:
614 changes in virus-specifc ribonucleic acid synthesis. J. Virol. 8, 388–394 (1971).

615 Rakus, K., Ouyang, P., Boutier, M., Ronsmans, M., Reschner, A., Vancsok, C., ... &
616 Vanderplasschen, A. (2013). Cyprinid herpesvirus 3: an interesting virus for applied and

617   fundamental research. Veterinary Research, 44(1), 1-16. https://doi.org/10.1186/1297-9716-44-
618   85

619 Ramírez, F., Ryan, D. P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A. S., ... & Manke, T. 620
(2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic 621 acids
research, 44(W1), W160-W165. doi:10.1093/nar/gkw257.

622   Renner, D. W., & Szpara, M. L. (2018). Impacts of genome-wide analyses on our understanding
623   of human herpesvirus diversity and evolution. Journal of virology, 92(1), e00908-17.

624 Renzette, N., Pokalyuk, C., Gibson, L., Bhattacharjee, B., Schleiss, M. R., Hamprecht, K., ... & 625
Kowalik, T. F. (2015). Limits and patterns of cytomegalovirus genomic diversity in humans. 626
Proceedings of the National Academy of Sciences, 112(30), E4120-E4128. 627
https://doi.org/10.1073/pnas.1501880112

628   Sanjuán, R., & Domingo-Calap, P. (2016). Mechanisms of viral mutation. Cellular and molecular
629   life sciences, 73(23), 4433-4448. https://doi.org/10.1007/s00018-016-2299-6

630 Sedlazeck, F. J., Rescheneder, P., Smolka, M., Fang, H., Nattestad, M., Von Haeseler, A., & 631
Schatz, M. C. (2018). Accurate detection of complex structural variations using single-molecule 632
sequencing. Nature methods, 15(6), 461-468. doi:10.1038/s41592-018-0001-7

633 Shekar, M., & Venugopal, M. N. (2019). Identification and characterization of novel double zinc 634
fingers encoded by putative proteins in genome of white spot syndrome virus. Archives of   635 virology,
164(4), 961-969.

636 Spatz, S. J. (2010). Accumulation of attenuating mutations in varying proportions within a high 637
passage very virulent plus strain of Gallid herpesvirus type 2. Virus research, 149(2), 135-142. 638
DOI:10.1016/j.virusres.2020.01.007

639 Sunarto, A., McColl, K. A., Crane, M. S. J., Sumiati, T., Hyatt, A. D., Barnes, A. C., & Walker, P. 640
J. (2011). Isolation and characterization of koi herpesvirus (KHV) from Indonesia: identification 641 of a
new genetic lineage. Journal of fish diseases, 34(2), 87-101. https://doi.org/10.1111/j.1365- 642
2761.2010.01216.x

643 Szpara, M.L., & Van Doorslaer, K. Mechanisms of dna virus evolution. In Dennis H. Bamford 644
and Mark Zuckerman, editors, Encyclopedia of Virology (Fourth Edition),pages 71–78. 645
Academic Press, Oxford, fourth edition edition, 2021. 3 https://doi.org/10.1016/B978-0-12- 646
809633-8.20993-X

647 Te Yeh, M., Bujaki, E., Dolan, P. T., Smith, M., Wahid, R., Konz, J., ... & Andino, R. (2020).   648
Engineering the live-attenuated polio vaccine to prevent reversion to virulence. Cell host & 649
microbe, 27(5), 736-751. https://doi.org/10.1016/j.chom.2020.04.003

650 Tomer, E., Cohen, E. M., Drayman, N., Afriat, A., Weitzman, M. D., Zaritsky, A., & Kobiler, O. 651
(2019). Coalescing replication compartments provide the opportunity for recombination between 652
coinfecting herpesviruses. The FASEB journal, 33(8), 9388-9403.

653   https://doi.org/10.1096/fj.201900032R

654 Vancsok, C., Peñaranda, M. M. D., Raj, V. S., Leroy, B., Jazowiecka-Rakus, J., Boutier, M., ... &
655 Vanderplasschen, A. F. (2017). Proteomic and functional analyses of the virion transmembrane 656 proteome of cyprinid herpesvirus 3. Journal of Virology, 91(21), e01209-17. DOI: 657 10.1128/JVI.01209-17

658   Vignuzzi, M., & López, C. B. (2019). Defective viral genomes are key drivers of the virus–host
659   interaction. Nature microbiology, 4(7), 1075-1087. https://doi.org/10.1038/s41564-019-0465-y

660 Wang, Z. H., Ke, F., Zhang, Q. Y., & Gui, J. F. (2021). Structural and functional diversity among 661 five ring finger proteins from Carassius auratus herpesvirus (CaHV). Viruses, 13(2), 254. DOI: 662 10.3390/v13020254

663   White, R., Pellefigues, C., Ronchese, F., Lamiable, O., & Eccles, D. (2017). Investigation of
664   chimeric   reads   using   the   MinION.   F1000Research,   6.
665   https://doi.org/10.12688/f1000research.11547.2

666   Wilkinson, D., & Weller, S. (2003). The role of DNA recombination in herpes simplex virus DNA
667   replication. IUBMB life, 55(8), 451-458. https://doi.org/10.1080/15216540310001612237

668   Wilkinson, Dianna E., Weller, Sandra K. (2003): The role of DNA recombination in herpes
669   simplex   virus   DNA   replication.   In: *IUBMB   life*   55   (8),   S.   451–458.   DOI:
670   10.1080/15216540310001612237

671 Xiao, Y., Lidsky, P. V., Shirogane, Y., Aviner, R., Wu, C. T., Li, W., ... & Andino, R. (2021). A   672 defective viral genome strategy elicits broad protective immunity against respiratory viruses.   673 Cell, 184(25), 6037-6051.

674   Xiaofei, E., & Kowalik, T.F. The dna damage response induced by infection with human
675   cytomegalovirus   and   other   viruses.   Viruses,   6(5)   :2155–2185,   2014.
676   https://doi.org/10.3390/v6052155

677

678

679

680

681

682

683

684