



HAL
open science

DNA Polymerase Sequences of New World Monkey Cytomegaloviruses: Another Molecular Marker with Which To Infer Platyrrhini Systematics

Samantha James, Damien Donato, Jean-Francois Pouliquen, Manuel Ruiz-García, Anne Lavergne, Vincent Lacoste

► **To cite this version:**

Samantha James, Damien Donato, Jean-Francois Pouliquen, Manuel Ruiz-García, Anne Lavergne, et al.. DNA Polymerase Sequences of New World Monkey Cytomegaloviruses: Another Molecular Marker with Which To Infer Platyrrhini Systematics. *Journal of Virology*, 2018, 92 (18), pp.e00980-18. 10.1128/JVI.00980-18 . pasteur-02197592

HAL Id: pasteur-02197592

<https://riip.hal.science/pasteur-02197592v1>

Submitted on 3 Apr 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **DNA polymerase sequences of New World monkey cytomegaloviruses: another molecular**
2 **marker to infer Platyrrhini systematics**

3
4 *Samantha James,^{a,b} Damien Donato,^a Jean-François Pouliquen,^a Manuel Ruiz-García,^c Anne*
5 *Lavergne,^{a,b} Vincent Lacoste^{a,b,d,#}*

6
7 ^aLaboratoire des Interactions Virus-Hôtes, Institut Pasteur de la Guyane, Cayenne, French Guiana

8 ^bUniversité de la Guyane, Ecole doctorale 587 « Diversités, Santé et Développement en
9 Amazonie », Cayenne, French Guiana

10 ^cLaboratorio de Genética de Poblaciones - Biología Evolutiva, Unidad de Genética,
11 Departamento de Biología, Facultad de Ciencias, Pontificia Universidad Javeriana, Bogotá, DC,
12 Colombia

13 ^dDépartement de Virologie, Institut Pasteur, Paris, France

14
15 **Running head:** New World monkey cytomegaloviruses

16
17 # Address correspondence to Vincent Lacoste, vincent.lacoste@pasteur.fr

18
19 **Keywords:** *Cytomegalovirus*, CMV, New World monkeys, evolution, phylogeny.

20
21 **Word counts:** Abstract **245 words** Importance **123 words** Main text **4036 words**

22
23 **Figures:** **3** **Tables:** **4**

24 **ABSTRACT**

25 Over the past few decades, a large number of studies have identified herpesvirus sequences from
26 many mammalian species around the world. Among the different non-human primate species
27 tested so far for cytomegaloviruses, only a few were from the New World. Seeking to identify
28 cytomegalovirus homologues in New World monkeys (NWMs), we carried out molecular
29 screening of 244 blood DNA samples from 20 NWM species from Central and South America.
30 Our aim was to reach a better understanding of their evolutionary processes within the Platyrrhini
31 parvorder. Using polymerase chain reaction amplification with degenerate consensus primers
32 targeting highly conserved amino acid motifs of the herpesvirus *DNA polymerase* gene, we
33 characterized novel viral sequences from 12 species belonging to seven genera representative of
34 the three NWM families. BLAST searches, pairwise nucleotide and amino acid sequence
35 comparisons, as well as phylogenetic analyses confirmed that they all belonged to the
36 *Cytomegalovirus* genus. Previously determined host taxa allowed us to demonstrate a good
37 correlation between the distinct monophyletic clades of viruses with those of the infected
38 primates at the genus level. In addition, the evolutionary branching points that separate NWM
39 CMVs were congruent with the divergence dates of their hosts at the genus level. These results
40 significantly expand our knowledge of the host range of this viral genus and strongly support the
41 co-speciation between these viruses and their hosts. In this respect, we propose that NWM CMV
42 *DNA polymerase* sequences could serve as a reliable molecular marker to infer Platyrrhini
43 phylogenetics.

44 **IMPORTANCE**

45 Investigating evolutionary processes between viruses and non-human primates has led to the
46 discovery of a high number of herpesviruses. No study published so far on primate
47 cytomegaloviruses has extensively studied New World monkeys (NWMs) at the subspecies,
48 species, genus and family levels. The present study sought to identify cytomegalovirus
49 homologues in NWMs and decipher their evolutionary relationships. This led us to characterize
50 novel viruses in 12 of the 20 primate species tested representative of the three NWM families.
51 The identification of distinct viruses in these primates not only significantly expands our
52 knowledge of the host range of this viral genus, but has also shed light on their evolutionary
53 history. Phylogenetic analyses and molecular dating of the sequences obtained support a virus–
54 host coevolution.

55 INTRODUCTION

56 New World monkeys of tropical forests from Central to South America belong to the Platyrrhini
57 parvorder (1). They first appeared in the Neotropics in the late Eocene or early Oligocene and
58 have subsequently evolved into a broad diversity of families, subfamilies and genera (**Figure 1**)
59 (2). To shed light on their phylogeny and evolution, NWMs have been extensively studied
60 through morphological, bio-geographical, behavioral and molecular data (2-16). Over the last few
61 decades, contrasting hypotheses have been proposed, presumably due to different markers and
62 the presence of polymorphisms in the features considered. Agreement on the main clades of
63 NWMs has been reached using different approaches, revealing a unique phylogenetic
64 arrangement of Platyrrhini with three monophyletic families: Pitheciidae, Atelidae and Cebidae
65 (**Table 1, Figure 1**) (3, 4, 10, 11, 13-15). Nevertheless, the relationships between them continue
66 to be debated. Through the analysis of intergeneric and intrageneric relationships, intrafamily
67 relationships have also been studied in depth. By incorporating all the available data, major
68 advances have been made and many taxonomic controversies have been clarified (5). Therefore,
69 the Pitheciidae family is composed of *Callicebus*, *Pithecia*, *Chiropotes* and *Cacajao*, the Atelidae
70 family of *Alouatta*, *Ateles*, *Brachyteles* and *Lagothrix* while the Cebidae family is divided into
71 *Cebuella*, *Mico*, *Callithrix*, *Callimico*, *Saguinus*, *Leontopithecus*, *Samiri*, *Cebus*, *Sapajus* and
72 *Aotus* genera. However, relationships between or within some subfamilies and/or genera remain
73 under discussion. Among Cebidae, the phylogenetic position of the Aotinae subfamily remains
74 unclear (14). Indeed, molecular data did not allow determining if Aotinae is a sister clade of
75 Callitrichinae or if, alternatively, Aotinae, Saimiriinae and Cebinae are sisters to Callitrichinae (3,
76 11, 13-15). Moreover, the number of platyrrhine genera is also still under discussion, such as the
77 division of *Cebus* into the *Sapajus* (tufted capuchins) and *Cebus* (untufted capuchins) genera (3,
78 16-18). As a result, neither the diversity nor the taxonomy of NWMs are fully known. To

79 appreciate the details of Platyrrhini's evolution, much work still needs to be done at various
80 taxonomic levels.

81
82 Viruses of the genus *Cytomegalovirus* belong to the *Betaherpesvirinae* subfamily within
83 the *Herpesviridae* family, order *Herpesvirales* (19). Eight cytomegaloviruses are recognized as
84 species by the International Committee for Taxonomy of Viruses (ICTV) according to the latest
85 master species list (MSL# 32) released on March 12, 2018
86 (<https://talk.ictvonline.org/files/master-species-lists/m/msl/7185>). *Human betaherpesvirus 5*
87 (HHV5), commonly referred to in the literature as human cytomegalovirus (HCMV), is the
88 cytomegalovirus (CMV) type species. So far, cytomegaloviruses have been characterized only
89 from primates. Since the initial description of cytomegalovirus in African green monkeys in
90 1957, whose current species name is *Cercopithecine betaherpesvirus 5* (CeHV5), natural
91 infections by such viruses have been described in several Old World monkey (OWM) species,
92 including baboons, macaques, colobuses, chimpanzees, gorillas and others (20-31). In contrast,
93 cytomegaloviruses of New World monkeys (NWMs) are represented by only three viral entities
94 from *Aotus trivirgatus* (northern owl monkey), *Saimiri sciureus* (common squirrel monkey) and
95 *Cebus* spp. (capuchin) animals, despite the wide diversity of platyrrhines (32-34). It is presumed
96 that all primate species harbor CMVs following a co-speciation process, but data supporting this
97 assumption are scarce. The most extensive analyses of primate CMVs conducted to date are those
98 of Leendertz *et al.* (26) and Anoh *et al.* (30). These studies have demonstrated, through
99 phylogenetic analyses, a species-specific distribution of these viruses. This species specificity
100 indicates a long-term coevolution of CMVs with their natural hosts. The identification of two
101 clades, each composed of chimpanzee and gorilla CMVs, suggests that they have coevolved

102 following a horizontal transmission event between these great apes millions of years ago (26).
103 Nevertheless, interspecies transmissions in the wild are rare events (26, 28, 30, 31).

104
105 With the exception of the three above-mentioned CMVs of NWMs, there has been little
106 prior organized effort to discover cytomegaloviruses in neotropical primates. The number of
107 NWM species tested to date therefore accounts for only a tiny part of their diversity. We thought
108 that additional investigations on a larger number of species were required. We therefore
109 addressed the possible presence of CMVs in different NWM species from which we previously
110 partially characterized Epstein-Barr-like viruses (35, 36). The purpose was to gain greater insight
111 into the distribution and diversity of CMVs infecting the Platyrrhine primates. Furthermore,
112 based on the coevolution observed between OWMs and their specific CMVs, we wished to
113 determine whether NWM CMV sequences could help decipher evolutionary relationships of their
114 host species (26, 30). Given that multiple molecular markers of mitochondrial and nuclear DNAs
115 are available, host species can be characterized along with their viruses, allowing progress to be
116 made on their respective pattern of diversification. Here, we report finding sequences of
117 cytomegaloviruses in different NWM species and achieve a better understanding of the
118 evolutionary processes between these viruses and their Platyrrhini hosts.

119
120 **RESULTS**
121 To look for the presence of CMV-like viruses in our collection of NWMs, we attempted to
122 amplify a fragment of the highly conserved herpesvirus *DNA polymerase* gene from the PBMC
123 DNA of each wild-caught primate using the PCR conditions as described previously (27, 35). A
124 total of 244 samples from 20 different species of the three NWM families were tested (**Table 1**).
125 DNA samples from 12 species scored positive after the nested PCRs (nPCRs) (**Table 1**). No

126 primate belonging to the *Saguinus* and *Lagothrix* genera scored positive. Indeed, no amplification
127 was observed in any samples from the three tamarin species (*Saguinus midas*, *S. labiatus* and *S.*
128 *oedipus*) and the four woolly monkey subspecies (*Lagothrix lagotricha cana*, *L. l. lagotricha*, *L.*
129 *l. lugens* and *L. l. poeppigii*). We then used different pairs of consensus-degenerate and specific
130 PCR primers to obtain longer sequences of the *DNA polymerase* gene from each positive animal
131 (**Table 2, Figure 2**). The concatenated nucleotide sequences generated are between 448 and
132 2,026 bp in size depending on the viral strain (**Table 1**).

133
134 BLAST searches demonstrated that all sequences identified belonged to the
135 *Cytomegalovirus* genus and revealed the presence of 12 distinct sequences. Four sequences were
136 identified twice in the *Saimiri boliviensis boliviensis*, *Aotus nancymae*, *Pithecia pithecia* and
137 *Alouatta seniculus* species while the sequence in the *Alouatta macconnelli* was identified in five
138 individuals (**Table 1**). Virus names and abbreviations were given to the 12 distinct viruses (**Table**
139 **1**): the viruses were named after the host species (for *Saimiri*, after the subspecies) followed by
140 three upper case letters corresponding to the viral genus (CMV for *Cytomegalovirus*) to which
141 they were then assigned the Arabic numeral 1, as has already been done by us and others (26, 35).

142
143 To obtain a full vision of the genetic diversity of these new CMV sequences, pairwise
144 sequence comparisons were made on the 447-bp/149-aa fragments of the *DNA polymerase* gene
145 common to all primate CMVs. All 12 sequences obtained differed from each other at the
146 nucleotide level. Sequences that were identified in different specimens of the same primate
147 species, e.g., AnanCMV1, AsenCMV1, AmacCMV1, SbolCMV1 and PpitCMV1, were 100%
148 identical with the exception of the two PpitCMV1 sequences showing 99.6% nucleotide identity
149 (**Table 3**). For clarity, comparison of the percentage of identity between the different newly

150 identified NWM CMVs have been reported by grouping viral sequences at the host genus level
151 (**Table 3**). Overall, the new sequences exhibited among themselves and the other available NWM
152 CMV sequences from 71.1% (AnanCMV1 vs. ApanCMV1) to 99.6% (CalbCMV1 vs. CebHV1
153 or *Cebus* sp. herpesvirus) nucleotide identity and from 79.6% (CcapCMV1 vs. ApanCMV1 and
154 AnanCMV1 vs. ApalCMV1) to 100% (AmacCMV1 vs. AsenCMV1 as well as SbolCMV1 vs.
155 SalbCMV1) amino acid identity (**Table 3**). Viruses infecting NWMs of the same genus presented
156 more than 92% of nucleotide and amino acid identities (**Table 3**). Comparison between CMVs of
157 different NWM genera ranged from 71.1% (*Aotus* CMVs vs. *Ateles* CMVs) to 88.7% (*Sapajus*
158 CMVs vs. *Cebus* CMVs) at the nucleotide level and from 79.6% between *Aotus* and *Alouatta*
159 CMVs and *Cebus* and *Ateles* CMVs to 99.3% between *Sapajus* and *Cebus* CMVs at the amino
160 acid level. NWM CMV sequences exhibited 59.8% (*Cebus* CMVs vs. *Macaca* CMVs) to 72.7%
161 (*Aotus* CMVs vs. *Colobus* CMVs) nucleotide sequence identities and 61.9 to 68.7% amino acid
162 sequence identities with OWM CMVs. The level of nucleotide and amino acid sequence identity
163 with CMVs of Hominidae, except HHV5, ranged from 61 to 72% and from 66 to 72.8%,
164 respectively. Identities with HHV5 ranged from 61.6% (*Alouatta* CMVs) to 70.9% (*Aotus*
165 CMVs) at the nucleotide level and from 66% (*Alouatta* CMVs) to 72.8% (*Aotus* CMVs) at the
166 amino acid level.

167
168 All phylogenetic analyses, performed on nucleotide or amino acid sequences between the
169 newly characterized CMV sequences and those of other primate CMVs available in the databases
170 clearly placed the new sequences in a monophyletic lineage of NWM viruses in the
171 *Cytomegalovirus* genus. The phylogenetic analysis presented in **Figure 3** is based on amino acid
172 sequences. The NWM CMV lineage diverged from the OWM CMV lineage with a posterior
173 probability value of 1. Remarkably, considering the OWM CMV lineage, the phylogenetic tree

174 formed two major monophyletic groups consisting of Hominoidea and Cercopithecoidea viruses,
175 respectively. Within the Cercopithecoidea, the *Colobus* CMV sequences are the basal taxa with
176 the formation of additional clades, comprising Asian *Macaca*, *Cercopithecus/Chlorocebus* and
177 *Papio/Mandrillus/Cercocebus* taxa.

178
179 Considering NWM CMVs, analyses demonstrated the existence of five distinct lineages
180 supported by high posterior probability values. The phylogenetic relationships between the
181 different NWM CMVs were correlated with the families and genera to which the infected
182 primates belong. The only exception was the hierarchical branching order of the
183 *Aotus/Saimiri/Cebus* genera within the Cebidae, which was not supported. Thus, viruses of *Aotus*
184 (*AoHV1*, *AvocCMV1* and *AnanCMV1*) all grouped together in a monophyletic clade as did
185 those of *Saimiri* (*SaHV4*, *SscHV*, *SalbCMV1* and *SbolCMV1*), *Cebus/Sapajus* (*CebHV1*,
186 *CebusHV*, *CalbCMV1*, *CcapCMV1* and *SapeCMV1*), *Alouatta* (*ApalCMV1*, *AmacCMV1* and
187 *AsenCMV1*) and *Pithecia* (*PpitCMV1*). Furthermore, viruses from *Alouatta* species were related
188 to *ApanCMV1* from *Ateles* in a monophyletic clade of viruses infecting Atelidae monkeys with a
189 posterior probability of 1, while those of *Saimiri*, *Aotus* and *Cebus/Sapajus* belonged to a
190 monophyletic clade of Cebidae, which was supported by a posterior probability of 0.91.

191
192 To explore the cospeciation hypothesis, a time calibration analysis was performed on our
193 data set. Cytomegaloviruses identified in NWMs diverged from those of OWMs at about 32.45
194 MYA (95% HPD: 17.76–52.33) (**Table 4**). In the New World clade, three major groups were
195 identified corresponding to viruses hosted by (i) Pitheciidae, (ii) Atelidae and (iii) Cebidae. The
196 Pitheciidae viruses diverged from those of the two other groups 22.33 MYA (95% HPD: 16.25–
197 28.21) and intraspecific divergence of *PpitCMV1* occurred 1.15 MYA. The divergence between

198 Atelidae and Cebidae viruses is estimated at 17.16 MYA (95% HPD: 10.29–24.25). Within the
199 group of Atelidae viruses, the divergence between ApanCMV1 identified in *Ateles paniscus* with
200 those identified in *Alouatta* spp. (ApalCMV1, AmacCMV1 and AsenCMV1) is estimated at 8.29
201 MYA (95% HPD: 2.57–14.82). Viruses identified in the three different *Alouatta* species
202 (*palliata*, *macconnelli* and *seniculus*) diverged at 3.11 MYA (95% HPD: 0.47–6.82) (ApalCMV1
203 vs. AmacCMV1 /AsenCMV1) while AmacCMV1 diverged from AsenCMV1 1.33 MYA (95%
204 HPD: 0.02–3.45). Within the Cebidae group, the divergence between viruses of Aotinae, Cebinae
205 and Saimiriinae occurred around 11.54 MYA (95% HPD: 5.82–17.89). Within the group of
206 cytomegaloviruses hosted by the different *Aotus* species, AoHV1 identified in *A. trivirgatus*
207 diverged from the others about 3.98 MYA (95% HPD: 0.8–8.25) while AnanCMV1 diverged
208 from AvocCMV1 2.15 MYA (95% HPD: 0.2–4.88). Within the group of viruses identified in the
209 Cebinae, CcapCMV1 from *C. capucinus* diverged from the others 3.26 MYA (95% HPD: 0.59–
210 6.94). SapCMV1 identified in *Sapajus apella* diverged from the other viruses detected in *Cebus*
211 (*albifrons* and *Cebus* sp.) 2.09 MYA (0.32–4.68). Finally, within the group of viruses identified
212 in the *Saimiri* genus, SsciCMV1/SaHV4 detected in *S. sciureus* diverged from those hosted by *S.*
213 *boliviensis* and *S. albigena*, SbolCMV1 and SalbCMV1, 2.68 MYA (95% HPD: 0.28–6.14).

214

215 **DISCUSSION**

216 This study is the largest conducted to date to molecularly characterize CMVs in NWMs in terms
217 of species diversity. It has partially characterized 12 cytomegaloviruses from 12 distinct species
218 belonging to seven genera and three NWM families. BLAST searches of the *Cytomegalovirus*
219 sequences identified further revealed that all but one were new viral sequences close to, but
220 distinct from, already published CMV sequences from *Aotus trivirgatus*, *Saimiri sciureus* and
221 *Cebus* sp. The only exception was the viral sequence from *C. albifrons* showing 99.6% identity at

222 the nucleotide level to CebusHV (AF292067) and CebHV1 (JQ264772), both identified from
223 unspecified *Cebus* spp. These three viral sequences were therefore considered to correspond to
224 the same viral species. In addition, the newly identified viral sequences are completely host-
225 specific, with no identification of cross-species transmission in our sample. Observations on
226 sequence comparisons, phylogenetic analysis and host-specificity of the sequences reported in
227 this paper are close to species demarcation criteria outlined in the 9th ICTV report for formal
228 recognition of new herpesvirus species (https://talk.ictvonline.org/ictv-reports/ictv_9th_report/)
229 (37).

230
231 By refining the degeneracy of the PCR primers used to screen the sample collection, we
232 were able to specifically target and identify CMV sequences, even though some of the primates
233 tested were coinfecting with lymphocryptoviruses (35). Indeed, we formerly identified 17 EBV-
234 related viruses from 15 NWM species belonging to seven genera and three families from the
235 same collection of samples (35, 36). These new combinations of screening primers are therefore
236 good molecular tools to be used for future studies. Nevertheless, on the 20 NWM species tested,
237 we did not characterize any CMV sequence from our collection of *Saguinus* spp. and *Lagothrix*
238 spp. samples. Considering the relatively small sampling size of most species belonging to these
239 two genera, with the exception of *Saguinus midas*, it is conceivable that we missed a CMV-like
240 virus from them. Nevertheless, for the other primate species tested, the sampling size was
241 equivalent or even smaller and we identified CMV sequences for almost all of them. More
242 strikingly, despite the large sample size of *Saguinus midas* screened (54 individuals) and the
243 different PCR approaches used (different combinations of primers with different levels of
244 degeneracy and PCR cycling conditions), no PCR product was identified. The negativity of the
245 *Saguinus* and *Lagothrix* genera for CMV-related viruses can be explained by a lack of primer

246 matching or by a loss of CMV viruses during evolution within these genera. Likewise, in our
247 former studies of EBV-related sequences, we were unsuccessful in amplifying EBV sequences
248 from individuals of the *Aotus* and *Alouatta* genera (35, 36). Taken together, these results
249 highlight the need for more in-depth analyses of a representative sample from these and other
250 species of these genera to clarify this point. Moreover, for part of the positive samples, we were
251 unsuccessful in generating longer sequences of the *DNA polymerase* gene. Whether this is due to
252 the low quality/low amount of the remaining DNA, a low viral load, or reflects technical
253 difficulties, *i.e.* an inadequate level of degeneracy of the primers designed for some of these
254 viruses, is not clear. Nonetheless, sequence data generated here were sufficient to gain insight on
255 their genetic relationships.

256

257 Pairwise nucleotide and amino acid sequence comparisons demonstrated that the viral
258 sequences analyzed present different levels of genetic diversity among them (**Table 3**); the
259 smallest divergences were detected when viral sequences from primates belonging to the same
260 genus were analyzed. Phylogenetic analyses showed that CMV sequences are grouped according
261 to the primate genus from which they were detected. Thereafter, the phylogenetic clustering and
262 diversification follow those proposed for NWM species, corroborating the hypothesis of joint
263 evolution of the viruses with the speciation of their hosts (3, 5). Comparatively, analyses of
264 NWM EBV sequences have fallen short of achieving a completely resolved phylogeny (35).
265 While a clear co-speciation can be seen in the terminal branchings within major lineages
266 according to the primate subfamilies, the phylogenetic relationships between them are not
267 concordant with the current interpretations of their hosts' pattern of diversification at the family
268 level. In addition, for OWMs, there is a similar incongruence between the *Lymphocryptovirus*
269 phylogeny and that of the corresponding host lineages (38, 39). One can therefore argue that,

270 within the *Herpesviridae* family, *DNA polymerase* sequences from viruses of the
271 *Cytomegalovirus* genus are a better molecular marker than those of the *Lymphocryptovirus* genus
272 to test hypotheses of herpesvirus/primate co-evolution. On the basis of the available data, our
273 analysis nevertheless has two limitations regarding viruses of Cebidae that do not perfectly reflect
274 current interpretations of their hosts' diversification pattern. While viral sequences of Cebidae
275 segregate into three well-supported clades, each corresponding to the host genus from which they
276 have been identified, *i.e.* *Cebus/Sapajus*, *Saimiri* and *Aotus*, the relationships between the three
277 clades are not phylogenetically supported (**Figure 3**). The second limitation concerns
278 SapeCMV1, identified from *Sapajus apella*, which, phylogenetically, falls within the group of
279 *Cebus* viruses (**Figure 3**). Nevertheless, pairwise sequence comparison of SapeCMV1 with the
280 viral sequences identified from *Cebus* spp. shows that the nucleotide divergence of SapeCMV1 is
281 over the maximum 8% observed for viral sequences identified from NWM species of the same
282 genus (**Table 3**). These combined results (on SapeCMV1 and the other *Cebus* viruses) do not, for
283 the moment, make it possible to confidently separate the *Cebus* genus into two genera as
284 observed on analyses of *Alu* elements and on phylogenomics (3, 17). However, our virus results
285 agree quite well with the mitogenomics findings obtained in recent studies, where *Sapajus* is a
286 taxon within *Cebus* (18). These limits should be resolved by screening an extensive taxon
287 sampling of the different *Sapajus* spp. as well as of Callithrichinae for the presence of
288 cytomegaloviruses.

289
290 Finally, these data support virus–host coevolution in terms of branching order as well as
291 divergence time. Indeed, for each NWM genus tested, the estimated timing of diversification of
292 viruses is in agreement with host sequence divergence date estimates from previously published
293 studies (**Table 4**) (3, 10, 15). Nevertheless, dates obtained at superior taxonomic levels are more

294 recent than those based on different types of data sets or models. These discrepancies can be
295 attributed to the fact that a majority of NWM taxa remain to be tested and that no CMV sequence
296 is available for numerous genera. This therefore limits the significance of our estimates for the
297 major primate lineages for the moment and emphasizes the need for further studies.

298
299 Here, we conclusively expand our knowledge of the viral diversity, distribution and
300 evolutionary relationships of NWM CMVs. Even if the evolutionary history of these viruses is
301 not fully resolved, and in spite of the limitations mentioned above, these results strongly support
302 the co-evolution hypothesis of these new viruses with their hosts. In light of these data, we
303 propose that *CMV DNA polymerase* sequences could serve as a genetic marker to define the
304 evolutionary links of their host species. Indeed, despite the number of studies conducted over the
305 past few decades and the fast-growing number of host DNA sequence data sets, an unifying
306 consensus of the evolutionary hierarchy of NWMs has not been fully reached, partly because not
307 all phylogenies from these data sets agree but also due to a large proportion of missing data for
308 some taxa (2-15). The search for and identification of *CMV DNA polymerase* sequences
309 therefore seems to be an alternative to help solve this issue. Given the high prevalence rate of
310 CMVs in wild primates, their spread through close contact with infectious bodily fluids and their
311 persistence for the lifelong of the host, cytomegalovirus sequences, if present, should be easily
312 obtained through our PCR approach (20, 21, 40-42). Considering the number of all presently
313 known NWM species, we only tested here a fraction of their diversity. This suggests that a great
314 number of cytomegaloviruses remains to be identified in this important group of primates. These
315 results argue for a wider and more systematic sampling and exploration of NWMs to evaluate the
316 presence of CMVs and to confirm the usefulness of those sequences as a new molecular tool to
317 infer the systematics of Platyrrhini.

318

319 MATERIALS AND METHODS

320 **Sample collection.** The collection of blood DNA samples has been described in detail elsewhere
321 (**Table 1**) (35, 36, 43-45). In brief, we tested a total of 244 DNA samples from 20 NWM species
322 (26 subspecies) belonging to the three families and to six of the seven subfamilies according to
323 Schneider and Sampaio (5). All samples were previously genetically identified by phylogenetic
324 analysis of mtDNA genes, including cytochrome *c* oxidase subunit I (*COXI*) and/or cytochrome
325 *b* (*CytB*) (35).

326

327 **Initial screening of samples.** Molecular screening was done by semi-nested polymerase
328 chain reaction amplification with degenerate consensus primers targeting highly conserved amino
329 acid (aa) motifs of the herpesvirus *DNA polymerase* gene (**Table 2**). To maximize the chances of
330 amplifying CMV-like sequences, the primers of Rose *et al.* were refined from the alignment of all
331 primate cytomegalovirus DNA polymerase sequences available in the databases (**Table 2**) (46).
332 CMV3F1, CMV3F2 and CMV3R1 primers were designed and used in place of QAHNA, VYGA
333 and GDTD1B, respectively, while the sense primer DFASA was kept as it was. Two different
334 combinations of primers (DFASA/ CMV3R1 or CMV3F1/ CMV3R1) were used on each DNA
335 sample in separate reactions for the first-round PCR (**Figure 2**). In the second-round PCR,
336 CMV3F2/CMV3R1 primers were used. PCR analyses were performed at an annealing
337 temperature of 60°C, with an elongation time of 30 s, for 35 cycles. All amplicons of
338 approximately the expected size were purified, cloned by TA cloning and sent for sequencing to
339 Genewiz®, Takeley, UK (<https://www.genewiz.com/>).

340

341 **Partial DNA polymerase gene amplification.** To obtain the nucleotide sequence upstream
342 of the CMV3F2 motif, a degenerate primer (CMV3R2) was derived from the complementary
343 sequences of the small fragments and used in an nPCR amplification with the DFASA or
344 CMV3F1 primer pools using the initial PCR products as templates (**Table 2, Figure 2**). Then, to
345 generate longer segments of the *DNA polymerase* gene for each newly identified virus, we tried
346 to obtain upstream and downstream sequences using different sets of consensus degenerate and
347 species-specific primers designed using the primate CMV DNA polymerase sequence alignment
348 (**Table 2, Figure 2**). Overlapping amplicons were generated, cloned and sequenced as described
349 above. Each sequence corresponds to at least three independent clones sequenced on both strands.
350 Contig sequences were then assembled using MEGA 5.05 software (47).

351
352 **Phylogenetic analysis.** Raw sequences were analyzed and edited in MEGA 5.05 (47).
353 Sequences were confirmed as CMV by homology analysis using the NCBI BLAST search tool
354 (48). Then multiple sequence alignments were constructed using ClustalW with all other
355 previously published primate CMV sequences and alignments were checked manually.
356 Sequences were translated into amino acids and both nucleotide and amino acid sequences were
357 checked for irregularities. Hypervariable regions were removed before performing analyses.
358 Sequence identity was calculated using uncorrected p-distances. Phylogenetic trees were inferred
359 from the aligned amino acid sequences. The JTT + G model was selected as the best-fitted model
360 of amino acid evolution under corrected Akaike information criteria (AICc) using MEGA 5.05
361 and used for the Bayesian approach (49), which was performed with Mr. Bayes 3.2.2 to infer
362 phylogenetic relationships (50). Markov Chain Monte Carlo (MCMC) simulations were run for
363 10,000,000 generations, with four simultaneous chains, using a sample frequency of 500 and a
364 burn-in of 25,000. Majority rule consensus trees were obtained from the output. Validation of the

365 inference was assessed based on the standard deviation of split frequencies, which was less than
366 the expected threshold value of 0.01 (calculated value of 0.002).

367
368 **Time calibration.** Divergence times between clades were calculated using a relaxed
369 Bayesian molecular clock model with an uncorrelated lognormal rate of distribution, as
370 implemented in BEAST version 1.7.4 (51). A monophyletic constraint was imposed for the nodes
371 used to calibrate evolutionary rates. Two calibration points were applied as normal prior to
372 constraining the age of the Platyrrhini and Homo-Pan clades, the time of the most recent common
373 ancestor (tMRCA) of Platyrrhini to 23.5 MYA (stdev = 3.0) and for Homo-Pan to 6.5 MYA
374 (stdev = 0.8) (52-54). These calibration points are based on the fossil date (55). The amino acid
375 substitution model was the same as the one described above. A Yule process of speciation was
376 used as the tree prior. Results were obtained for 10,000,000 generations with the first 2,500,000
377 discarded as burn-in and parameter values sampled every 100 generations. The effective sample
378 size for parameter estimates and convergence were checked using Tracer version 1.5.0 software
379 (56). The final tree with divergence estimates and their 95% highest posterior densities was
380 computed in TreeAnnotator v1.4.5 (51).

381
382 **Accession numbers.** The sequences reported in this paper have been deposited in the
383 GenBank database under accession numbers **KU963225** to **KU963240**.

384

385 **ACKNOWLEDGMENTS**

386 Warm thanks go to Benoît de Thoisy for providing NWM DNA samples from French Guiana.
387 S.J. was supported by a grant from the Université de la Guyane, Ecole doctorale 587 “Diversités,
388 Santé et Développement en Amazonie” and by a grant from the Collectivité Territoriale de la
389 Guyane. This study was funded by a European commission "REGPOT-CT-2011-285837-
390 STRonGer" grant within the FP7 and an "Investissement d’Avenir" grant managed by the Agence
391 Nationale de la Recherche (CEBA, Ref. ANR-10-LABX-25-01). It was also supported by grant
392 numbers 1203-09-11239 (Colciencias) and 120108-E0102141 (Fondo para la Acción Ambiental)
393 to M.R-G. The funders had no role in study design, data collection and interpretation, or the
394 decision to submit the work for publication.

395

396 **COMPETING INTERESTS**

397 The authors declare that they have no competing financial interests.

398

399 **AUTHORS’ CONTRIBUTION**

400 A.L., M.R-G. and V.L. contributed to the study design. M.R-G. collected some of the samples.
401 S.J., D.D., J.-F.P., A.L. and V.L. performed the molecular and genetic analyses. S.J., A.L. and
402 V.L. analyzed the data and wrote the article. All authors participated in its final writing and
403 editing.

404

405 **ETHICAL APPROVAL**

406 Study based on samples collected several years earlier. Biological material from French Guiana
407 was collected in 1994–1995 along the Sinnamary River, Petit Saut Hydroelectric dam, under the
408 supervision of veterinarians of “Faune Sauvage” team led by Jean-Christophe Vié (57). Blood

409 sampling from live animals was carried out in accordance with French animal care regulations
410 and the laws of France. The other samples were collected directly from animals killed, in the
411 field, by indigenous hunters for their own purposes, with the full consent of the hunters and in
412 accordance with the laws of Brazil, Colombia, Mexico, Peru, Guatemala and Argentina.

413 **REFERENCES**

- 414 1. Groves CP. 2005. Order Primates, p 111-184. *In* Wilson DE, Reeder DM (ed), Mammal
415 species of the World: a taxonomic and geographic reference, Third ed, vol 2. The Johns
416 Hopkins University Press, Baltimore, MD.
- 417 2. Rylands AB, Mittermeier RA. 2009. The diversity of the New World Primates
418 (Platyrrhini): an annotated taxonomy., p 23 - 54. *In* Garber PA, Estrada A, Bicca-Marques
419 JC, Heymann EW, Stier KB (ed), South American Primates - Comparative Perspectives in
420 the Study of Behavior, Ecology, and Conservation doi:10.1007/978-0-387-78705-3.
421 Springer-Verlag New York, New York.
- 422 3. Perelman P, Johnson WE, Roos C, Seuanez HN, Horvath JE, Moreira MA, Kessing B,
423 Pontius J, Roelke M, Rumpler Y, Schneider MP, Silva A, O'Brien SJ, Pecon-Slattery J.
424 2011. A molecular phylogeny of living primates. *PLoS Genet* 7:e1001342.
- 425 4. Schneider H, Canavez FC, Sampaio I, Moreira MA, Tagliaro CH, Seuanez HN. 2001. Can
426 molecular data place each neotropical monkey in its own branch? *Chromosoma* 109:515-
427 23.
- 428 5. Schneider H, Sampaio I. 2015. The systematics and evolution of New World primates - A
429 review. *Mol Phylogenet Evol* 82 Pt B:348-57.
- 430 6. Schneider H, Sampaio I, Harada ML, Barroso CM, Schneider MP, Czelusniak J, Goodman
431 M. 1996. Molecular phylogeny of the New World monkeys (Platyrrhini, primates) based on
432 two unlinked nuclear genes: IRBP intron 1 and epsilon-globin sequences. *Am J Phys*
433 *Anthropol* 100:153-79.
- 434 7. Schneider H, Schneider MP, Sampaio I, Harada ML, Stanhope M, Czelusniak J, Goodman
435 M. 1993. Molecular phylogeny of the New World monkeys (Platyrrhini, primates). *Mol*
436 *Phylogenet Evol* 2:225-42.

- 437 8. Rylands AB, Mittermeier RA, Silva Jr. JS. 2012. Neotropical primates: taxonomy and
438 recently described species and subspecies. *Int Zoo Yearb* 46:11-24.
- 439 9. Rylands AB, Schneider H, Langguth A, Mittermeier RA, Groves CP, Rodrigues-Luna E.
440 2000. An assessment of the diversity of new World primates. *Neotrop Primates* 8:61-93.
- 441 10. Opazo JC, Wildman DE, Prychitko T, Johnson RM, Goodman M. 2006. Phylogenetic
442 relationships and divergence times among New World monkeys (Platyrrhini, Primates).
443 *Mol Phylogenet Evol* 40:274-80.
- 444 11. Wildman DE, Jameson NM, Opazo JC, Yi SV. 2009. A fully resolved genus level
445 phylogeny of neotropical primates (Platyrrhini). *Mol Phylogenet Evol* 53:694-702.
- 446 12. Singer SS, Schmitz J, Schwiegk C, Zischler H. 2003. Molecular cladistic markers in New
447 World monkey phylogeny (Platyrrhini, Primates). *Mol Phylogenet Evol* 26:490-501.
- 448 13. Ray DA, Xing J, Hedges DJ, Hall MA, Laborde ME, Anders BA, White BR, Stoilova N,
449 Fowlkes JD, Landry KE, Chemnick LG, Ryder OA, Batzer MA. 2005. Alu insertion loci
450 and platyrrhine primate phylogeny. *Mol Phylogenet Evol* 35:117-26.
- 451 14. Osterholz M, Walter L, Roos C. 2009. Retropositional events consolidate the branching
452 order among New World monkey genera. *Mol Phylogenet Evol* 50:507-13.
- 453 15. Jameson Kiesling NM, Yi SV, Xu K, Gianluca Sperone F, Wildman DE. 2015. The tempo
454 and mode of New World monkey evolution and biogeography in the context of
455 phylogenomic analysis. *Mol Phylogenet Evol* 82 Pt B:386-99.
- 456 16. Alfaro JW, Silva JD, Jr., Rylands AB. 2012. How different are robust and gracile capuchin
457 monkeys? An argument for the use of *sapajus* and *cebus*. *Am J Primatol* 74:273-86.
- 458 17. Martins AM, Jr., Amorim N, Carneiro JC, de Mello Affonso PR, Sampaio I, Schneider H.
459 2015. Alu elements and the phylogeny of capuchin (*Cebus* and *Sapajus*) monkeys. *Am J*
460 *Primatol* 77:368-75.

- 461 18. Ruiz-Garcia M, Castillo MI, Luengas-Villamil K. 2016. It is misleading to use *Sapajus*
462 (robust capuchins) as a genus? : A review of the evolution of the capuchins and suggestions
463 on their systematics, p 209-268. *In* Ruiz-Garcia M, Shostell JM (ed), *Molecular population*
464 *genetics, evolutionary biology and biological conservation of the Neotropical Primates.*
465 *Nova Science Publisher, Inc., New York, USA.*
- 466 19. Davison AJ, Eberle R, Ehlers B, Hayward GS, McGeoch DJ, Minson AC, Pellett PE,
467 Roizman B, Studdert MJ, Thiry E. 2009. The order Herpesvirales. *Arch Virol* 154:171-7.
- 468 20. Blewett EL, White G, Saliki JT, Eberle R. 2001. Isolation and characterization of an
469 endogenous cytomegalovirus (BaCMV) from baboons. *Arch Virol* 146:1723-38.
- 470 21. Blewett EL, Lewis J, Gadsby EL, Neubauer SR, Eberle R. 2003. Isolation of
471 cytomegalovirus and foamy virus from the drill monkey (*Mandrillus leucophaeus*) and
472 prevalence of antibodies to these viruses amongst wild-born and captive-bred individuals.
473 *Arch Virol* 148:423-33.
- 474 22. Asher DM, Gibbs CJ, Jr., Lang DJ, Gajdusek DC, Chanock RM. 1974. Persistent shedding
475 of cytomegalovirus in the urine of healthy Rhesus monkeys. *Proc Soc Exp Biol Med*
476 145:794-801.
- 477 23. Malherbe H, Harwin R. 1957. Seven viruses isolated from the vervet monkey. *Br J Exp*
478 *Pathol* 38:539-41.
- 479 24. Black PH, Hartley JW, Rowe WP. 1963. Isolation of a cytomegalovirus from African green
480 monkey. *Proc Soc Exp Biol Med* 112:601-5.
- 481 25. Wroblewska Z, Gilden D, Devlin M, Huang ES, Rorke LB, Hamada T, Furukawa T,
482 Cummins L, Kalter S, Koprowski H. 1979. Cytomegalovirus isolation from a chimpanzee
483 with acute demyelinating disease after inoculation of multiple sclerosis brain cells. *Infect*
484 *Immun* 25:1008-15.

- 485 26. Leendertz FH, Deckers M, Schempp W, Lankester F, Boesch C, Mugisha L, Dolan A,
486 Gatherer D, McGeoch DJ, Ehlers B. 2009. Novel cytomegaloviruses in free-ranging and
487 captive great apes: phylogenetic evidence for bidirectional horizontal transmission. *J Gen*
488 *Virol* 90:2386-94.
- 489 27. Lacoste V, Mauclore P, Dubreuil G, Lewis J, Georges-Courbot MC, Rigoulet J, Petit T,
490 Gessain A. 2000. Simian homologues of human gamma-2 and betaherpesviruses in
491 mandrill and drill monkeys. *J Virol* 74:11993-9.
- 492 28. Murthy S, Couacy-Hymann E, Metzger S, Nowak K, De Nys H, Boesch C, Wittig R, Jarvis
493 MA, Leendertz FH, Ehlers B. 2013. Absence of frequent herpesvirus transmission in a
494 nonhuman primate predator-prey system in the wild. *J Virol* 87:10651-9.
- 495 29. Prepens S, Kreuzer KA, Leendertz F, Nitsche A, Ehlers B. 2007. Discovery of
496 herpesviruses in multi-infected primates using locked nucleic acids (LNA) and a bigenic
497 PCR approach. *Virol J* 4:84.
- 498 30. Anoh AE, Murthy S, Akoua-Koffi C, Couacy-Hymann E, Leendertz FH, Calvignac-
499 Spencer S, Ehlers B. 2017. Cytomegaloviruses in a Community of Wild Nonhuman
500 Primates in Tai National Park, Cote D'Ivoire. *Viruses* 10.
- 501 31. Seimon TA, Olson SH, Lee KJ, Rosen G, Ondzie A, Cameron K, Reed P, Anthony SJ, Joly
502 DO, Karesh WB, McAloose D, Lipkin WI. 2015. Adenovirus and herpesvirus diversity in
503 free-ranging great apes in the Sangha region of the Republic Of Congo. *PLoS One*
504 10:e0118543.
- 505 32. Daniel MD, Melendez LV, King NW, Barahona HH, Fraser CE, Garcia FG, Silva D. 1973.
506 Isolation and characterization of a new virus from owl monkeys: herpesvirus aotus type 3.
507 *Am J Phys Anthropol* 38:497-500.

- 508 33. Daniel MD, Melendez LV, King NW, Fraser CE, Barahona HH, Hunt RD, Garcia FG,
509 Trum BF. 1971. Herpes virus aotus: a latent herpesvirus from owl monkeys (*Aotus*
510 *trivirgatus*) isolation and characterization. *Proc Soc Exp Biol Med* 138:835-45.
- 511 34. Rangan SR, Chaiban J. 1980. Isolation and characterization of a cytomegalovirus from the
512 salivary gland of a squirrel monkey (*Saimiri sciureus*). *Lab Anim Sci* 30:532-40.
- 513 35. Lavergne A, de Thoisy B, Pouliquen JF, Ruiz-Garcia M, Lacoste V. 2011. Partial
514 molecular characterisation of New World non-human primate lymphocryptoviruses. *Infect*
515 *Genet Evol* 11:1782-9.
- 516 36. de Thoisy B, Pouliquen JF, Lacoste V, Gessain A, Kazanji M. 2003. Novel gamma-1
517 herpesviruses identified in free-ranging new world monkeys (golden-handed tamarin
518 [*Saguinus midas*], squirrel monkey [*Saimiri sciureus*], and white-faced saki [*Pithecia*
519 *pithecia*]) in French Guiana. *J Virol* 77:9099-105.
- 520 37. Pellett PE, Davison AJ, Eberle R, B. E, Hayward GS, Lacoste V, Minson AC, Nicholas J,
521 Roizman B, Studdert MJ, Wang F. 2011. Order *Herpesvirales*, p 99-107. *In* King AMQ,
522 Adams MJ, Carstens EB, Lefkowitz EJ (ed), *Virus Taxonomy: Ninth report of the*
523 *International Committee on Taxonomy of Viruses*. Elsevier Inc., Oxford, United Kingdom.
- 524 38. Ehlers B, Ochs A, Leendertz F, Goltz M, Boesch C, Matz-Rensing K. 2003. Novel simian
525 homologues of Epstein-Barr virus. *J Virol* 77:10695-9.
- 526 39. Ehlers B, Spiess K, Leendertz F, Peeters M, Boesch C, Gatherer D, McGeoch DJ. 2010.
527 Lymphocryptovirus phylogeny and the origins of Epstein-Barr virus. *J Gen Virol* 91:630-
528 42.
- 529 40. Mocarski ES, Shenk T, Griffiths PD, Pass RF. 2013. Cytomegaloviruses, p 1960 - 2014. *In*
530 Knipe DM, Howley PM (ed), *Field's Virology*, vol 2. Lippincott Williams & Wilkins,
531 Philadelphia.

- 532 41. Cannon MJ, Hyde TB, Schmid DS. 2011. Review of cytomegalovirus shedding in bodily
533 fluids and relevance to congenital cytomegalovirus infection. *Rev Med Virol* 21:240-55.
- 534 42. Reeves M, Sinclair J. 2008. Aspects of human cytomegalovirus latency and reactivation.
535 *Curr Top Microbiol Immunol* 325:297-313.
- 536 43. de Thoisy B, Vogel I, Reynes JM, Pouliquen JF, Carme B, Kazanji M, Vie JC. 2001.
537 Health evaluation of translocated free-ranging primates in French Guiana. *Am J Primatol*
538 54:1-16.
- 539 44. Ruiz-Garcia M, Castillo MI, Vasquez C, Rodriguez K, Pinedo-Castro M, Shostell J,
540 Leguizamon N. 2010. Molecular phylogenetics and phylogeography of the white-fronted
541 capuchin (*Cebus albifrons*; Cebidae, Primates) by means of mtCOII gene sequences. *Mol*
542 *Phylogenet Evol* 57:1049-61.
- 543 45. Ruiz-Garcia M, Pinedo-Castro MO. 2010. Molecular systematics and phylogeography of
544 the genus *Lagothrix* (Atelidae, Primates) by means of the mitochondrial COII gene. *Folia*
545 *Primatol (Basel)* 81:109-28.
- 546 46. Rose TM, Strand KB, Schultz ER, Schaefer G, Rankin GW, Jr., Thouless ME, Tsai CC,
547 Bosch ML. 1997. Identification of two homologs of the Kaposi's sarcoma-associated
548 herpesvirus (human herpesvirus 8) in retroperitoneal fibromatosis of different macaque
549 species. *J Virol* 71:4138-44.
- 550 47. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular
551 evolutionary genetics analysis using maximum likelihood, evolutionary distance, and
552 maximum parsimony methods. *Mol Biol Evol* 28:2731-9.
- 553 48. Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL. 2008. NCBI
554 BLAST: a better web interface. *Nucleic Acids Res* 36:W5-9.

- 555 49. Darriba D, Taboada GL, Doallo R, Posada D. 2011. ProtTest 3: fast selection of best-fit
556 models of protein evolution. *Bioinformatics* 27:1164-5.
- 557 50. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L,
558 Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic
559 inference and model choice across a large model space. *Syst Biol* 61:539-42.
- 560 51. Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012. Bayesian phylogenetics with
561 BEAUti and the BEAST 1.7. *Mol Biol Evol* 29:1969-73.
- 562 52. Poux C, Douzery EJ. 2004. Primate phylogeny, evolutionary rate variations, and divergence
563 times: a contribution from the nuclear gene IRBP. *Am J Phys Anthropol* 124:1-16.
- 564 53. Kay RF, Fleagle JG, Mitchell TR, Colbert M, Bown T, Powers DW. 2008. The anatomy of
565 *Dolichocebus gaimanensis*, a stem platyrrhine monkey from Argentina. *J Hum Evol*
566 54:323-82.
- 567 54. Vignaud P, Durringer P, Mackaye HT, Likius A, Blondel C, Boisserie JR, De Bonis L,
568 Eisenmann V, Etienne ME, Geraads D, Guy F, Lehmann T, Lihoreau F, Lopez-Martinez N,
569 Mourer-Chauvire C, Otero O, Rage JC, Schuster M, Viriot L, Zazzo A, Brunet M. 2002.
570 Geology and palaeontology of the Upper Miocene Toros-Menalla hominid locality, Chad.
571 *Nature* 418:152-5.
- 572 55. Tavaré S, Marshall CR, Will O, Soligo C, Martin RD. 2002. Using the fossil record to
573 estimate the age of the last common ancestor of extant primates. *Nature* 416:726-9.
- 574 56. Rambaut A, Drummond AJ. 2009. Tracer. Version 1.5.0.
- 575 57. Vie JC. 1999. Wildlife rescues - The case of the Petit Saut hydroelectric dam in French
576 Guiana. *Oryx* 33:115-126.

- 577 58. Blewett EL, Sherrod CJ, Texier JR, Conrad TM, Dittmer DP. 2015. Complete Genome
578 Sequences of *Mandrillus leucophaeus* and *Papio ursinus* Cytomegaloviruses. *Genome*
579 *Announc* 3.
- 580 59. Johnson G, Dick D, Ayers M, Petric M, Tellier R. 2003. Detection and species-level
581 identification of primate herpesviruses with a comprehensive PCR test for human
582 herpesviruses. *J Clin Microbiol* 41:1256-8.
- 583 60. Marsh AK, Willer DO, Ambagala AP, Dzamba M, Chan JK, Pilon R, Fournier J,
584 Sandstrom P, Brudno M, MacDonald KS. 2011. Genomic sequencing and characterization
585 of cynomolgus macaque cytomegalovirus. *J Virol* 85:12995-3009.
- 586 61. Russell JN, Marsh AK, Willer DO, Ambagala AP, Dzamba M, Chan JK, Pilon R, Fournier
587 J, Brudno M, Antony JM, Sandstrom P, Evans BJ, MacDonald KS. 2016. A novel strain of
588 cynomolgus macaque cytomegalovirus: implications for host-virus co-evolution. *BMC*
589 *Genomics* 17:277.
- 590 62. Swanson R, Bergquam E, Wong SW. 1998. Characterization of rhesus cytomegalovirus
591 genes associated with anti-viral susceptibility. *Virology* 240:338-48.
- 592 63. Rivaller P, Kaur A, Johnson RP, Wang F. 2006. Genomic sequence of rhesus
593 cytomegalovirus 180.92: insights into the coding potential of rhesus cytomegalovirus. *J*
594 *Virol* 80:4179-82.
- 595 64. Kouzarides T, Bankier AT, Satchwell SC, Weston K, Tomlinson P, Barrell BG. 1987.
596 Sequence and transcription analysis of the human cytomegalovirus DNA polymerase gene.
597 *J Virol* 61:125-33.
- 598 65. Davison AJ, Dolan A, Akter P, Addison C, Dargan DJ, Alcendor DJ, McGeoch DJ,
599 Hayward GS. 2003. The human cytomegalovirus genome revisited: comparison with the
600 chimpanzee cytomegalovirus genome. *J Gen Virol* 84:17-28.

- 601 66. Dunn W, Chou C, Li H, Hai R, Patterson D, Stolc V, Zhu H, Liu F. 2003. Functional
602 profiling of a human cytomegalovirus genome. Proc Natl Acad Sci U S A 100:14223-8.
- 603 67. Murphy E, Yu D, Grimwood J, Schmutz J, Dickson M, Jarvis MA, Hahn G, Nelson JA,
604 Myers RM, Shenk TE. 2003. Coding potential of laboratory and clinical strains of human
605 cytomegalovirus. Proc Natl Acad Sci U S A 100:14976-81.
- 606

607 **FIGURE LEGEND**

608

609 **FIG.1. Diagram representation of Platyrrhini taxa in descending order down to genus**
610 **level.** Adapted from Perelman et al. (3). Black and gray boxes represent NWM genera tested for
611 CMVs. Black boxes correspond to NWM genera from which CMV sequences have been
612 characterized while gray boxes represent NWM genera from which no CMV sequence has been
613 obtained in the present study.

614

615 **FIG. 2. Relative position and orientation of the PCR primers used.** The different
616 combinations of primers used in nested or semi-nested PCRs are represented above the
617 Herpesviral *DNA polymerase* gene sequence in different grey boxes. Primers XxxR1 and XxxR2,
618 represented by dotted arrows, correspond to the different antisense specific primers used in a
619 degenerate (CMV2F1 or CMV2F2) - nondegenerate nPCR assay (see **Table 2**). Bars below the
620 sequence represent the different nested PCR products expected. The pairs of primers on the left-
621 hand side of the bars indicate those used for the first-round PCR, while those on the right-hand
622 side correspond to the one used in the nested PCR reactions. The sequences of the
623 oligonucleotide primers are given in **Table 2**.

624

625 **FIG. 3. Phylogenetic analysis of primate cytomegalovirus sequences.** The phylogenetic tree
626 was derived from the partial amino acid sequences of the *DNA polymerase* gene (149 aa) of 50
627 representatives of primate cytomegaloviruses using the Bayesian method with the JTT + G model
628 of amino acid evolution. Sequences generated in this study are in boldface. Posterior probabilities
629 of the Bayesian analysis (>0.9) are shown next to each node. The scale bar indicates amino acid
630 substitutions per site. The major clades representing Old World and New World primate families,

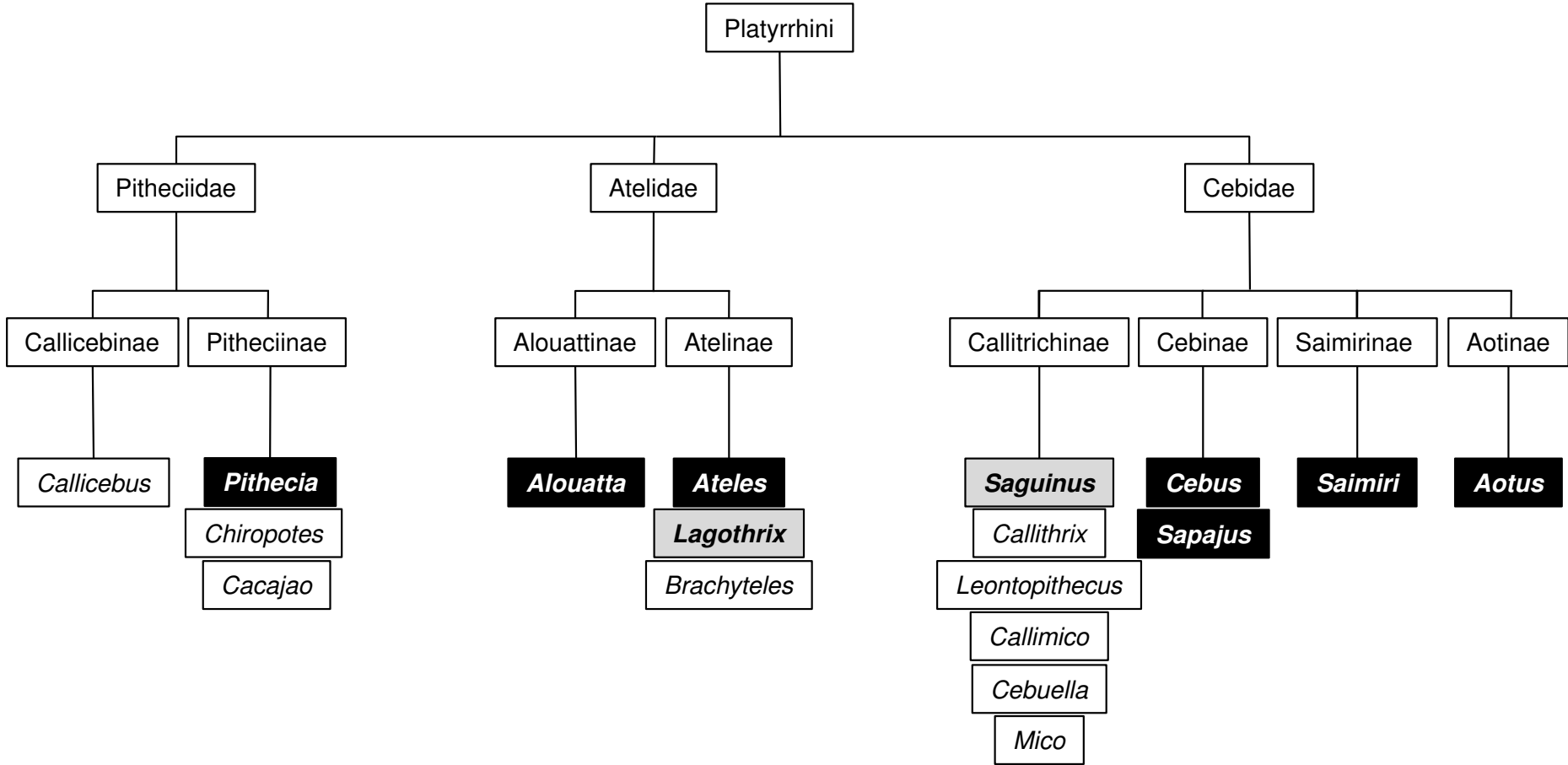
631 super-families (for OWM) and parvorders are labeled on the right-hand side of the figure. The
632 virus names are associated with their accession numbers. Viruses of Cercopithecidae comprise
633 those of the chacma (BaCMV NC_027016 (20, 58)) and olive (BaCMV AF387664 (21))
634 baboons, the moustached guenon (CceCMV1 AY728178), the agile mangabey (CagCMV1
635 AY608713), the drill (MndCMV AF282941 and MndCMV AF387665 strain OCOM6-2 (21,
636 27)), the mandrill (MndCMV AY129399), the African green monkey (CeHV5 AY117754,
637 CeHV5 FJ483969 strain Colburn, CeHV5 FJ483969 strain 2715, CaeCMV AF292066,
638 VervetCMV AY049066 strain CSG and CeHV3 AY049065 (21, 59)), the cynomolgus
639 (MfasCMV1 JN227533 strain Ottawa, MfasCMV1 AY728171 and MfasCMV KP796148 strain
640 Mauritius (60, 61)) and rhesus (McHV3 AF033184 and McHV3 DQ120516 isolate CMV 180.92
641 (62, 63)) macaques and the mantled guereza (CgueCMV1.1 AY129397 and CgueCMV1.2
642 EU118147 (29)). Viruses of Hominidae comprise those of the Bornean orangutan (PpygCMV1.1
643 AY129396), the human (HHV5 M14709 strain AD169, HHV5 NC_006273 strain Merlin, HHV5
644 AY315197 strain Towne and HHV5 AC146905 isolate Toledo (64-67)), the Western gorilla
645 (GgorCMV2.1 FJ538490 (26)) and the common chimpanzee (PnHV2 AF480884 strain
646 Heberling, PtroCMV1.1 FJ538485 and PtroCMV AF292063 (26, 65)). Regarding viruses of New
647 World monkeys, in addition to the ones describe in the present study, viruses of Cebidae
648 comprise those of the capuchin monkey (CebHV1 JQ264772 and CebusHV AF292067 from
649 *Cebus* sp.), the common squirrel monkey (SaHV4 FJ483967 and SscHV AF292065) and the
650 three-striped night monkey (AoHV1 FJ483970).

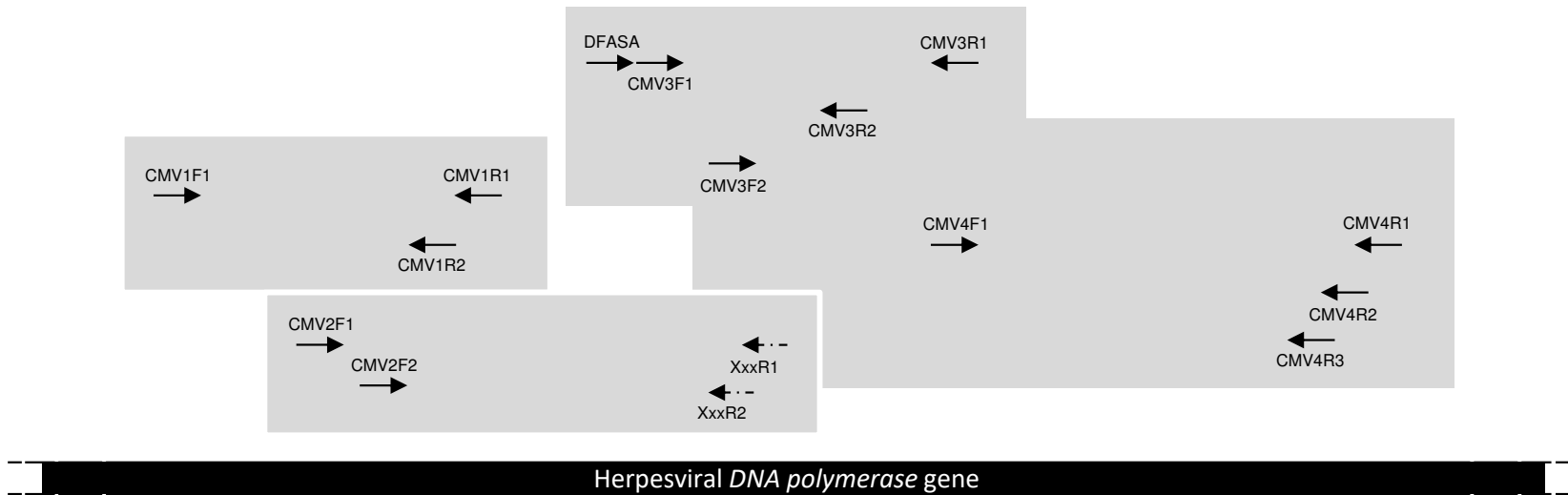
Parvorder

Family

Subfamily

Genus





DFASA/CMV3R1 or CMV3F1/CMV3R1 ————— CMV3F2/CMV3R1

DFASA/CMV3R1 ————— DFASA/CMV3R2

CMV3F1/CMV3R1 ————— CMV3F1/CMV3R2

CMV1F1/CMV1R1 ————— CMV1F1/CMV1R2

CMV2F1/XxxR1 ————— CMV2F2/XxxR2

CMV3F2/CMV4R1 ————— CMV3F2/CMV4R2

CMV3F2/CMV4R1 ————— CMV3F2/CMV4R3

CMV3F2/CMV4R1 ————— CMV4F1/CMV4R1

CMV3F2/CMV4R1 or CMV4F1/CMV4R1 ————— CMV4F1/CMV4R2

CMV3F2/CMV4R1 or CMV4F1/CMV4R1 ————— CMV4F1/CMV4R3

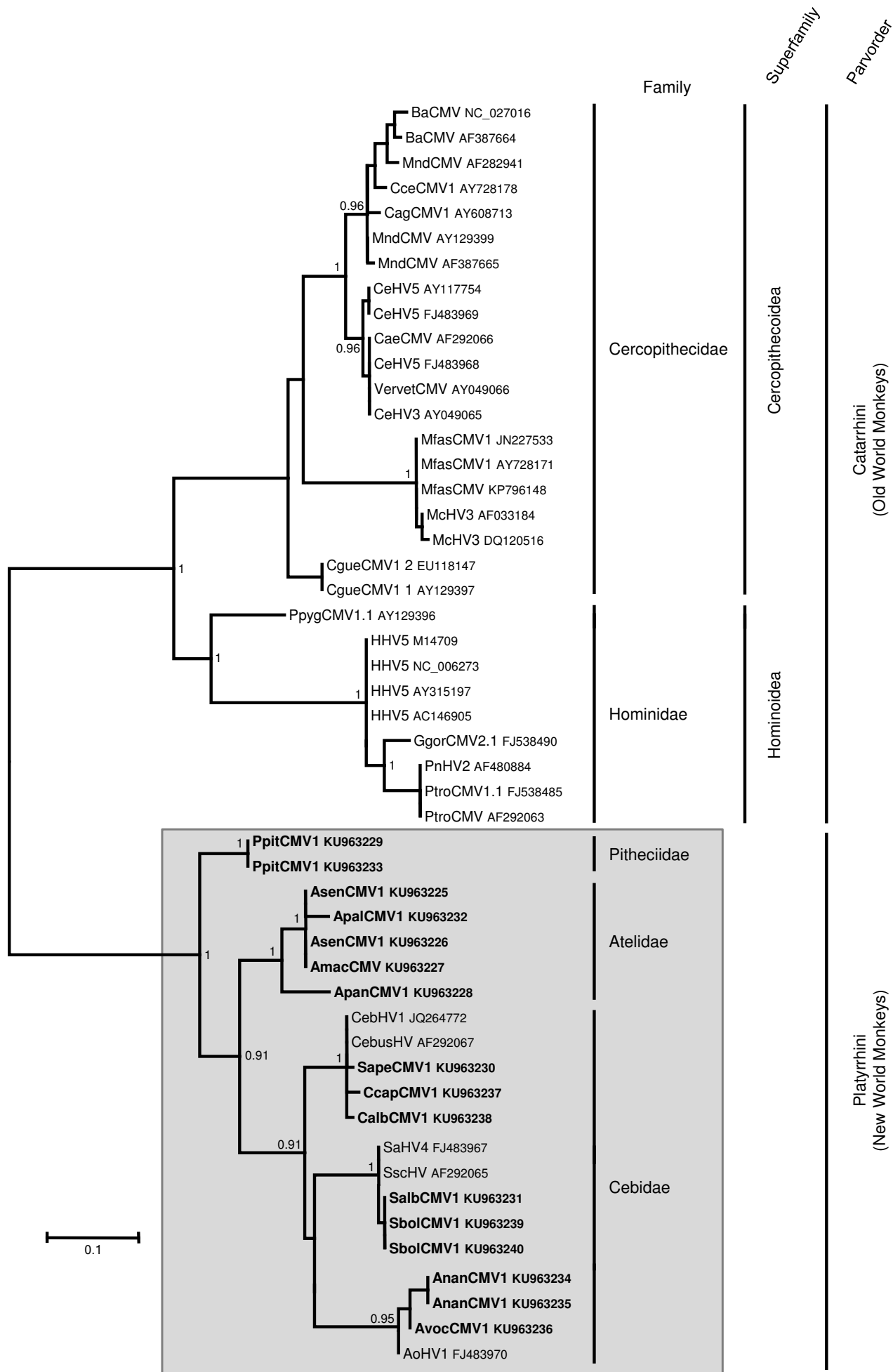


TABLE 2. Oligonucleotide primers used for Cytomegalovirus *DNA polymerase* consensus and specific PCRs

Oligonucleotide	Orientation ^a	Location ^b	5' -> 3' sequence ^c	CMV sequence amplified
<i>DNA pol degenerate primers</i>				
CMV1F1	+	721 - 749	GAC AAG AAG TTG ACN ACN TTY GGN TGG TG	AmacCMV1,AsenCMV1, ApanCMV1
CMV1R1	-	1534 - 1559	ACG CCG GCY TCR TAR TGR AAR TTD AT	AmacCMV1,AsenCMV1, ApanCMV1
CMV1R2	-	1480 - 1504	CGT CCT GHA CRC ART AYT TNC CNA C	AmacCMV1,AsenCMV1, ApanCMV1
CMV2F1	+	1330 - 1352	GAY ATG TAY CCN GTS TGY ATG GC	AmacCMV1, AsenCMV1, ApanCMV1, PpitCMV1, SapeCMV1, SalbCMV1
CMV2F2	+	1372 - 1394	TAC AAR YTV AAY ACB ATG GCS GA	AmacCMV1, AsenCMV1, ApanCMV1, PpitCMV1, SapeCMV1, SalbCMV1
DFASA ^d	+	1768 - 1793	GTG TTC GAC TTY GCN AGY YTN TAY CC	All
CMV3F1	+	1795 - 1817	TCH ATY ATY ATG GCN CAY AAY CT	All
CMV3F2	+	2062 - 2090	ACG TGC AAT TCT TTY TAY GGB TTY ACN GG	All
CMV3R1	-	2269 - 2303	CGA TAG CAC ACA AAC ACR CTR TCN GTR TCN CCR TA	All
CMV3R2	-	2125 - 2147	CCG ATD CGN GTR ATR CTR GCC GC	All
CMV4F1	+	2266 - 2291	ATC TAY G GK GAC ACS GAY AGY GTS TT	AmacCMV1, AsenCMV1, ApalCMV1, AnanCMV1, ApanCMV1
CMV4R1	-	2782 - 2801	GCC GCY ARN CGY TTD ATG AC	ApalCMV1, AnanCMV1, ApanCMV1
CMV4R2	-	2477 - 2498	CGC ACC ARR TCR ACN CCY TTC A	AmacCMV1, AsenCMV1, ApalCMV1, AnanCMV1, ApanCMV1
CMV4R3	-	2419 - 2444	ATA TAC CGY TTY TTR CAG ATC ATC AT	AmacCMV1, AsenCMV1, ApalCMV1, AnanCMV1, ApanCMV1
<i>DNA pol antisense specific primers (in combination with CMV2F1 or CMV2F2)^e</i>				
PitR1	-	1958 - 1978	TGC GCT GAG CAA CCC ATT TAG	PpitCMV1
PitR2	-	1912 - 1932	ACG CAC CTC CGA CTT CAC AAA	PpitCMV1
AotR1	-	2019 - 2039	TTG TCG AGC AGC GTC CTC TTG	AnanCMV1, AvocCMV1
AotR2	-	1884 - 1904	ACC GTA CCG TTT TCG AAG TTA	AnanCMV1, AvocCMV1
SapR1	-	2106 - 2126	GCG ACT GGC AAA CAC GGT AAC	SapeCMV1
CapR2	-	1994 - 2016	GGG ATC TGT GCA ATC TTT CAT GG	CalbCMV1, CcapCMV1
CapR3	-	1939 - 1961	CGG GTC AAC AAT TCA GAA AGC AC	SapeCMV1, CalbCMV1, CcapCMV1
AteR1	-	1994 - 2016	CGG ATC TCT GCA ATT TTT CAT GG	ApanCMV1
AteR2	-	1935 - 1957	TCA GCA GTT CGG ACA ACA CTG AA	ApanCMV1
SaiR1	-	1948 - 1969	CCA CCC ACT TCG TTA GCA GCT C	SbolCMV1, SalbCMV1
SaiR2	-	2155 - 2175	ACG CGC GGT GTC TTG TAA CAT	SbolCMV1, SalbCMV1
AloR1	-	1959 - 1979	TTC CGT TGA GCC ACC CAT TTA	AmacCMV1, AsenCMV1, ApalCMV1
AloR2	-	1932 - 1954	GCA ATT CCG ATA GCA CTG AGG AA	AmacCMV1, AsenCMV1, ApalCMV1

^a +, sense; -, antisense.^b Positions relative to ATG of the *DNA polymerase* gene of AoHV1, accession number FJ483970.^c Positions of degeneracy follow the International Unit Base codes.^d Degenerate oligonucleotide primer described in Rose *et al.* (45).^e For clarity, all antisense specific primers are indicated as XxxR1 for R1 primers and XxxR2 for R2 primers in Figure 2.

TABLE 3. Nucleotide and amino acid identities between the novel Cytomegaloviruses and all other NHP Cytomegaloviruses and HCMV^a

Viruses of		% identity ^b with														
		<i>Sapajus</i>		<i>Cebus</i>		<i>Saimiri</i>		<i>Aotus</i>		<i>Pithecia</i>		<i>Alouatta</i>		<i>Ateles</i>		
		(SapeCMV1)		(CalbCMV1, CcapCMV1)		(SbolCMV1, SalbCMV1)		(AvocCMV1, AnanCMV1)		(PpitCMV1)		(AmacCMV1, AsenCMV1, ApalCMV1)		(ApanCMV1)		
		Nucleotide	Amino acid	Nucleotide	Amino acid	Nucleotide	Amino acid	Nucleotide	Amino acid	Nucleotide	Amino acid	Nucleotide	Amino acid	Nucleotide	Amino acid	
Cd ^c	Cn	<i>Sapajus</i>			87.6 - 88.7	98.0 - 99.3	76.1 - 77.0	89.1	74.3 - 75.4	85.7	72.9 - 73.1	83.7	74.5 - 74.7	83.0 - 85.0	75.9	80.3
		<i>Cebus</i> ^d	87.6 - 88.7	98.0 - 99.3	95.7 - 99.6	98.0 - 99.3	72.2 - 77.2	87.1 - 88.4	74.0 - 75.4	85.0 - 87.1	72.0 - 72.5	83.0 - 84.4	74.3 - 75.4	81.6 - 83.7	74.5 - 74.7	79.6 - 81.0
		<i>Saimiri</i> ^e	76.1 - 77.4	89.1	75.2 - 77.7	87.1 - 87.8	97.1 - 97.7	99.3 - 100	74.3 - 74.9	85.71	74.7 - 75.6	82.31	74.7 - 76.1	85.0 - 85.7	75.2 - 75.6	83.67
		<i>Aotus</i> ^f	74.3 - 75.4	85.7 - 86.4	74.0 - 75.6	85.0 - 87.8	74.0 - 74.7	85.7 - 86.4	94.8 - 96.2	96.6 - 98.6	74.0 - 74.7	83.7 - 85.0	71.8 - 74.8	79.6 - 82.3	71.1 - 72.2	81.63
Pd	Pn	<i>Pithecia</i>	72.9 - 73.1	83.7	72.0 - 72.5	83.0 - 83.7	74.72	82.3 - 83.3	74.0 - 74.7	84.4 - 85.0	99.6	100	72.9 - 75.2	83.0 - 85.0	75.2 - 75.4	84.35
Ad	Al	<i>Alouatta</i>	74.5 - 74.7	82.3 - 83.0	74.3 - 75.4	81.6 - 83.0	74.9 - 76.1	85.0 - 85.7	71.8 - 73.1	79.6 - 81.6	72.9 - 75.2	83.0 - 85.0	92.1 - 98.9	97.3 - 100	82.2 - 83.5	91.2 - 92.5
	At	<i>Ateles</i>	75.9	80.3	74.5 - 74.7	79.6 - 80.3	75.4 - 75.6	83.67	71.1 - 71.6	81.63	75.2 - 75.4	84.35	82.2 - 83.5	91.2 - 92.5	-	-
Hd ^g		<i>Homo</i>	64.1 - 64.6	67.35	64.3 - 65.9	66.7 - 68.0	63.2 - 64.3	67.35	70.0 - 70.9	72.1 - 72.8	65.2 - 66.1	70.07	61.6 - 63.7	65.99	62.1 - 63.0	67.35
	Hn	<i>Pan</i>	61.85	66.67	63.2 - 63.7	66.0 - 66.7	63.2 - 63.4	66.67	71.6 - 72.0	71.4 - 72.1	66.1 - 66.6	70.07	61.2 - 62.7	66.67	62.08	68.03
		<i>Gorilla</i>	62.98	67.35	63.7 - 63.9	66.7 - 68.0	62.3 - 62.5	67.35	71.33	72.1 - 72.8	62.98	70.75	61.0 - 62.1	65.99	60.95	68.03
	Po	<i>Pongo</i>	64.11	69.39	63.4 - 64.1	69.4	65.5 - 65.7	68.03	70.4 - 70.7	71.43	66.14	69.39	62.5 - 63.4	68.0 - 68.7	62.53	68.71
Ce ^g		<i>Macaca</i>	62.1 - 63.4	63.3 - 64.0	59.8 - 61.6	61.9 - 63.3	62.1 - 63.2	64.6 - 65.3	65.9 - 66.6	65.3 - 67.4	63.4 - 63.9	64.6 - 65.3	60.3 - 62.1	64.6 - 66.7	62.5 - 63.0	64.6 - 65.3
		<i>Papio</i>	64.8 - 66.4	64.0 - 64.6	62.1 - 63.4	61.9 - 63.3	62.1 - 62.3	66.0 - 66.7	68.2 - 69.1	64.0 - 65.3	64.6 - 64.8	65.3 - 66.0	62.3 - 63.4	66.0 - 66.7	62.53	65.3
		<i>Mandrillus</i>	64.1 - 65.7	64.0 - 66.0	61.8 - 63.4	61.9 - 64.6	61.4 - 63.0	66.7 - 68.0	67.5 - 68.8	64.6 - 67.4	64.3 - 65.9	66.0 - 68.0	60.9 - 64.6	66.0 - 68.0	61.8 - 62.7	65.3 - 67.4
		<i>Cerco/Chloro</i>	63.7 - 65.5	64.6 - 66.0	61.6 - 68.8	63.3 - 65.3	60.9 - 63.7	66.7 - 68.0	67.5 - 69.3	65.3 - 68.0	63.7 - 65.5	66.7 - 68.7	61.8 - 65.0	66.7 - 68.7	60.5 - 63.2	66.0 - 68.0
Co		<i>Colobus</i>	63.9 - 64.1	65.3	63.9	63.3 - 65.3	63.7 - 64.6	68.7	72.2 - 72.7	68.0	66.1 - 66.4	66.7	63.4 - 66.1	67.4 - 68.0	62.5 - 62.8	66.7

^a Numbers refer to values obtained in comparison with the 447-bp fragment of the conserved *DNA polymerase* gene that is available for all viruses.

^b Sequences identified from specimens from the same primate species showing 100% nucleotide identity e.g. SbolCMV1, AnanCMV1, AmacCMV1 and AsenCMV1 are not included.

^c Abbreviations, Cd, Cebidae; Cn, Cebinae; Sn, Saimiriinae; An: Aotinae; Pd, Pitheciidae; Pn, Pitheciinae; Ad, Atelidae; Al, Alouattinae; At, Atelinae; Hd, Homnidae; Hn, Homninae; Po, Pongidae; Ce, Cercopithecidae; Cr, Cercopithecinae; Co, Colobinae.

^d Nucleotide and amino acid identities of viruses of *Cebus* rely on the sequences generated in this study as well as on sequences of CebHV1_{JQ264772} and CebusHV_{AF292067}.

^e Nucleotide and amino acid identities of viruses of *Saimiri* rely on the sequences generated in this study as well as on sequences of SaHV4_{FJ483967} and SscHV_{AF292065}.

^f Nucleotide and amino acid identities of viruses of *Aotus* rely on the sequences generated in this study as well as on the sequence of AoHV1_{FJ483970}.

^g Homnidae and Cercopithecidae viral sequences used to calculate nucleotide and amino acid identities correspond to those shown in Figure 3. Their GenBank accession numbers and associated publications are all reported in the figure and its legend.

TABLE 4. Estimates of Platyrrhini divergence time from CMV *DNA polymerase* sequence data and comparison with other estimates

Node	This study	Perelman et al. (3)	Jameson Kiesling et al. (15)
Catarrhini / Platyrrhini	32.45 ^a (17.76 - 52.33) ^b	43.47 (38.55 - 48.36)	37.72 (36.04 - 42.07)
Pitheciidae / Atelidae+Cebidae	22.33 (16.25 - 28.21)	24.82 (20.55 - 29.25)	25.51 (25.14 - 26.36)
Atelidae / Cebidae	17.16 (10.29 - 24.25)	22.76 (18.14 - 27.08)	24.04 (22.6 - 25.29)
Atelinae / Alouattinae	8.29 (2.57 - 14.82)	16.13 (10.52 - 21.35)	15.29 (13.29 - 17.99)
Within <i>Alouatta</i>	3.11 (0.47 - 6.82)	ND	5.14 (3.65 - 6.8)
Within Cebidae	11.54 (5.82 - 17.89)	19.95 (15.66 - 24.03)	20.86 (18.48 - 22.86)
Within <i>Aotus</i>	3.98 (0.8 - 8.25)	5.54 (3.20 - 7.85)	4.39 (3.12 - 5.75)
Within <i>Cebus</i>	3.26 (0.59 - 6.94)	6.00 (3.13 - 9.35)	5.19 (3.69 - 6.78)
Within <i>Saimiri</i>	2.68 (0.28 - 6.14)	2.24 (1.05 - 3.73)	0.97 (0.51 - 1.45)

^a times in units of millions of years

^b values given in the parentheses indicate 95% HPD in millions of years.