Phylogenetic analysis of enteroviruses from non-human primates reveals two new species within the genus *Enterovirus* and inter-species recombination

Corentin Aubé[1], Paulina Cruz de Casas[1], Matthieu Prot[2], Artem Baidaliuk[2], Marie-Claire Endegue Zanga[3], Etienne Simon-Lorière[2], Nolwenn Jouvenet[1], Serge A. Sadeuh-Mba[3,4], Maël Bessaud[1, 5]

[1] *Institut Pasteur, Université de Paris Cité, CNRS UMR 3569, Virus sensing and signaling Unit. 25-28, rue du Dr Roux 75015 Paris. France*

[2] *Institut Pasteur, Université Paris Cité, CNRS UMR2000, Evolutionary Genomics of RNA Viruses Unit. 25-28, rue du Dr Roux 75015 Paris. France*

[3] *Centre Pasteur du* Cameroun, Virology Laboratory, *451, avenue 2005 - Yaoundé 2. Cameroon*

[4] Salisbury Animal Health Laboratory, Maryland Department of Agriculture, 27722 Nanticoke Rd, Salisbury, MD 21801, United States of America

[5] *Laboratoire associé au Centre national de référence pour les entérovirus & paréchovirus - 25-28, rue du Dr Roux 75015 Paris. France*

Abstract

To date, 15 species have been described within the genus *Enteroirus.* Previous studies suggested the existence of another species comprising strains isolated from the stools specimens of non-human primates (NHPs) in Central Africa. Moreover, numerous full-length or partial genomic sequences of NHP enteroviruses (EVs) can be found in GenBank without being properly classified. To our knowledge, no comprehensive synthesis of NHP EV data exists, leaving genetic relationships between strains across independent studies unclear. To address these gaps, we sequenced the complete genome of four NHP EVs from our stool collection and conducted an extensive search of NHP EV sequences in GenBank to perform a comprehensive phylogenetic analysis. Our analyses revealed two new species tentatively named *Enterovirus mbel* and *Enterovirus noa*, which contain at least 6 and 2 virus types, respectively. We also identified new virus types within the known species EV-J.

Phylogenetic analyses strongly suggest interspecies recombination events between NHP EVs in the non-structural region of the genome, challenging the long-held view that recombination is confined within narrowly defined subsets of EVs belonging to the same species. We also performed the first comprehensive comparative analysis of full length human and NHP EV genomes, focusing on GC content, dinucleotide frequencies and codon-usage bias. GC content emerged as the most robust host-associated marker: all NHP-associated virus types within species *E. alphacoxsackie* and *E. betacoxsackie* displayed GC % below 47 %, whereas human-derived virus types exhibited GC % above 47 %. Dinucleotide frequency, Effective Number of Codons (ENC) and Relative Synonymous Codon Usage highlighted distinct codon-bias clusters that mirror the phylogenetic relationships between EVs but only partially correlate with their respective hosts of origin.

This work enhances our understanding of EVs circulating in NHPs and paves the way for future research aiming at understanding the mechanisms underlying host-adaptation among EVs.

Keywords

Enterovirus, *Picornaviridae*, Recombination, non-human primates, GC content, host-adaptation

66 **INTRODUCTION**

67 Enteroviruses (EVs) are small, non-enveloped viruses with an icosahedral capsid. Their
68 genome is a single-stranded and positive-sense RNA molecule of approximately 7,200-8,500
69 nucleotides (nt) in length. It contains two untranslated regions (5'UTR and 3'UTR) flanking a
70 large open reading frame (ORF) that encodes a polyprotein, which is subsequently cleaved into
71 three precursors (P1, P2 and P3). P1 gives rise to the four capsid proteins (VP1 to VP4) while
72 P2 and P3 are cleaved into different proteins involved in viral replication. A short additional
73 upstream ORF (uORF) encoding a single protein has been described in the genome of some
74 EVs (Lulla et al., 2019).

75 The classification of EVs has undergone significant revisions since their discovery.
76 Initially, EVs were classified into different groups based on their phenotypic properties (e.g.
77 tropism and pathogenicity in mice) and physicochemical characteristics (density of the virus
78 particles, sensibility to acid pHs) (Bessaud et al., 2018). Within each group, virus isolates were
79 grouped into serotypes, based on *in vitro* cross-neutralization assays. In the 2000s, genetic
80 criteria replaced phenotypic and serological criteria as the basis for classification (M. S. Oberste
81 et al., 1999). Enteroviruses are now classified into species and virus types, depending on their
82 phylogenetic relationships. According to the International Committee for Taxonomy of Viruses
83 guidelines (ICTV), EVs belong to the same species if they share at least 70% amino acid (aa)
84 identity in the complete ORF and in the 2C+3CD genes, and at least 60% aa identity in the
85 capsid (P1 region). The *Enterovirus* genus currently comprises 15 species: *Enterovirus*
86 *alphacoxsackie* (EV-A), *Enterovirus betacoxsackie* (EV-B), *Enterovirus coxsackiepol* (EV-C),
87 *Enterovirus deconjuncti* (EV-D) *Enterovirus eibovi* (EV-E), *Enterovirus fitauri* (EV-F)
88 *Enterovirus geswini* (EV-G), *Enterovirus hesimi* (EV-H), *Enterovirus idromi* (EV-I),
89 *Enterovirus jesimi* (EV-J), *Enterovirus krodeni* (EV-K), *Enterovirus lesimi* (EV-L), *Enterovirus*
90 *alpharhino* (RV-A), *Enterovirus betarhino* (RV-B) and *Enterovirus cerhino* (RV-C), each of

which comprise multiple virus types based on the VP1-encoding sequence. Virus type within a species are defined by >75% nt identity and >85%-88% aa identity in this region (Brown et al., 2009; M. Steven Oberste et al., 1999).

EVs infect a wide range of mammals, including humans, non-human primates (NHPs), small ruminants, pigs, cattle, camels and rodents. While each EV appears adapted to a limited set of host species that constitutes its natural reservoirs, host barriers can be overcome, as demonstrated by numerous documented cross-species transmissions between humans and animals, bidirectionally, or among different animal species (Doté et al., 2024; Fieldhouse et al., 2018; Modiyinji et al., 2025). Notably, EVs do not exhibit clear host-associated phylogenetic clustering, as some species include virus types circulating in different hosts. For instance, some EV-As, EV-Bs and EV-Ds circulate in humans while others circulate in NHPs. Similarly, some EV-Gs circulate in pigs while others circulate in small ruminants.

The first identified NHP EVs were detected in simian cell cultures and laboratory animals in the mid-20th century (Oberste et al., 2002). Since then, new EV types have been discovered in captive, wild, healthy and sick NHPs. Over the past 20 years, next-generation sequencing has increased the number of publicly available NHP EV sequences, however, data remain sparse. Furthermore, many sequences span diverse sub-genomic regions and remain to be properly classified. Moreover, no comprehensive synthesis of all available NHP EV molecular data exists, leaving the genetic relationships between NHP EVs across independent studies fragmented. Additionnally, sequences of NHP EVs sampled in Central Africa have suggested the existence of at least one additional EV species (Sadeuh-Mba et al., 2014), but it has never been formally described. To address these gaps, we conducted an extensive search of NHP EV sequences in GenBank and performed a comprehensive phylogenic analysis. To make this analysis as complete as possible, we also sequenced the complete genome of four NHP EVs from our NHP stool collection and samples from our previous work (Sadeuh-Mba et al., 2014).

116 Furthermore, we compared genetic characteristics, such as GC percentage (GC%), dinucleotide
117 ratios or codon usage (effective number of codons, ENC and relative synonymous codon usage,
118 RSCU) between NHP and human EVs. Our analysis identified two previously uncharacterized
119 EV species and revealed that specific genomic signatures align with phylogeny and may be
120 associated with host-adaptation.

121

122 **METHODS**

123 **Amplicon-based sequencing**

124 EV-B114 strain Z057 was grown in RD cells and RNA was extracted from 250 µL of
125 clarified cell culture supernatant using the High Pure Viral RNA kit (Roche Diagnostics,
126 Meylan, France) following the manufacturer's instructions. Genome amplification was
127 performed by RT-PCR with primers targeting previously described conserved regions of the
128 EV genome (Joffret et al., 2018). Sequencing was performed by Illumina technology and de
129 novo assembly conducted with CLC Genomics Workbench as previously described (Bessaud
130 et al., 2016). As EV-J121 strain CHE20 could not be propagated in cell culture (Sadeuh-Mba
131 et al., 2014), RNA was extracted directly from 140 µL of clarify 20% (Weight/Volume) stool
132 suspension using the QIAamp Viral RNA Mini Kit (Qiagen, France). Genome amplification
133 was performed by RT-PCR with primers targeting the sub-genomic regions previously
134 sequenced (Sadeuh-Mba et al., 2014), as well as a primer targeting the 3' terminal poly A
135 sequence. Sequencing was performed by the Sanger technique.

136

137 **Metatranscriptomic next-generation sequencing**

138 Two strains, RCHM3 and RCHM6, that could not be grown in cell culture were
139 sequenced using a protocol previously described (Gámbaro et al., 2021). After extraction from

140 140 µl of fecal suspension using the QIAamp Viral RNA Mini Kit (Qiagen) followed by Turbo
141 DNAse treatment (Ambion), RNA was purifed with Agencourt RNAClean XP beads. Host
142 rRNA were depleted using custom probes and RNAse H treatment (Matranga et al., 2014).
143 After purification using AMPure RNA clean beads (Beckman Coulter Genomics) and elution
144 in 10 ul of RNase-free water, RNA was retro transcribed using random primers and SuperScript
145 IV (Invitrogen). Second-strand DNA was synthesized using *E. coli* DNA ligase, RNAse H and
146 DNA polymerase (New England Biolabs) and purified using Agencourt AMPure XP beads
147 (Beckman Coulter). Libraries were then prepared using the Nextera XT DNA Library Prep Kit
148 (Illumina), quality controlled using a Bioanalyzer high sensitivity DNA analysis kit (Agilent)
149 and sequenced on an Illumina NextSeq500 (2 × 75 cycles).
150 Adapters and low-quality reads were removed using Trimmomatic v0.39 (Bolger et al.,
151 2014). Trimmed reads were assembled using megahit v1.2.9 (Li et al., 2016) with default
152 parameters. The contigs were queried against the NCBI non-redundant protein database using
153 DIAMOND v2.0.4 (Buchfink et al., 2015), leading to the identification of novel EVs sequences.
154 The EVs genomes were then reconstructed by iterative mapping of the trimmed reads, using
155 clc-assembly-cell v5.1.0. We used Samtools v1.10 (Li et al., 2009) to generate alignment
156 statistics, and ivar v1.0 (Grubaugh et al., 2019) to generate consensus sequences, using a
157 minimum of 3× read depth coverage. The mapping data was visually checked to confirm the
158 accuracy of the obtained genomes using Geneious Prime 2023 (www.geneious.com).
159

160 **Phylogenetic analysis**

161 Multiple sequence alignments were performed with CLC Main Workbench 24.0.2
162 software (CLC bio, Aarhus, Denmark). Phylogenetic trees were inferred by the neighbour-
163 joining (NJ) or maximum likelihood (ML) methods implemented in MEGA X (Kumar et al.,
164 2018). Distance-based phylogenetic trees were reconstructed by the NJ method with the

165 Tajima-Nei model for computing evolutionary distances (Tajima and Nei, 1984). The ML
166 phylogenetic trees were reconstructed by the Tajima-Nei model and Nearest-Neighbor-
167 Interchange method for ML heuristic method. The reliability of tree topologies was estimated
168 by bootstrap analysis with 1,000 pseudoreplicates for NJ trees and 500 pseudoreplicates for ML
169 trees due to computational constraints. Tree representations were generated using MEGA X
170 (version 10.2.6) or FigTree (version 1.4.4) (https://tree.bio.ed.ac.uk/software/figtree/). Links to
171 Nextstrain interactive representations of trees from Figures 2 and 6, Supplementary Figure 1,
172 and tanglegrams are available on GitHub:
173 https://github.com/CorentinAube/NHP_EV_genomic_analysis.
174

175 **Genomic comparison**

176 Prototype genomes of human and NHP EVs were retrieved from the Picornaviridae
177 database (http://www.picornaviridae.com), while more recent sequences were retrieved from
178 the BV-BRC database (Olson et al., 2023). Other partial NHP EV sequences were retrieved
179 from the scientific literature indexed in PubMed. Protein sequences translation, MSA,
180 nucleotide and amino acid identity matrices were calculated with CLC Main Workbench. The
181 GC content, dinucleotide ratios, ENC and RSCU values were calculated using custom Python
182 (version 3.12.12) scripts available on GitHub with also the list of all accession numbers of
183 sequences used (https://github.com/CorentinAube/NHP_EV_genomic_analysis). The formulas
184 for dinucleotide ratio, ENC and RSCU calculations were obtained from Zeng and colleagues
185 (Zeng et al., 2022). The determination of GC3 bias was based on the method published by Butt
186 and colleagues (Butt et al., 2016).
187

188 **Data availability**

189 Nucleotide sequences generated in this study were deposited on the GenBank database
190 (accession numbers PX682287-PX682290). An isolate of the EV-B114 strain Z057 has been
191 shared via the EVAg repository (Romette et al., 2018).
192
193 **RESULTS**

194 **Whole-genome sequencing of four NHP EVs from our stool collection**

195 Four EVs that had only been previously partially sequenced (Sadeuh-Mba et al., 2014),
196 were subjected to whole genome sequencing. Contigs covering nearly the entire viral genome
197 were obtained for three viruses: Z057, CHE20, and RCHM3 (Figure 1A). For RCMH6,
198 approximately 150 nt could not successfully be sequenced at the 5' extremity. All 4 viruses
199 featured the canonical EV genomic organization, with a large ORF flanked by two UTRs.
200 Comparing these sequences with annotated reference EV genomes, canonical polyprotein
201 cleavage sites were identified, suggesting that the polyprotein processing in these viruses is
202 similar to that of other EVs (Supplementary Table 1). Additionally, the genomes of Z057,
203 RCMH3, and RCMH6 contained a putative uORF. This putative uORF overlaps with the main
204 ORF with a frameshift, and its putative start codon is located 133 or 136 nt upstream of the
205 main ORF (Figure 1B). In contrast, the EV-J121 CHE20 genome lacked a uORF, as previously
206 reported for other members of the EV-J species (Lulla et al., 2019). All four genomes also
207 contained a cis-replication element (CRE) within the 2C gene, matching the consensus
208 sequence 5'-GXXXAAAXXXXXXA-3' (Figure 1C) (Cordey et al., 2008).
209

210 **Comprehensive phylogeny of non-human primate enteroviruses**

211 The search for NHP EV sequences in GenBank led to the identification of 336
212 sequences, of which 24 were complete genomes, 91 were VP1 sequences, 103 were VP4/2

213 sequences, 19 were VP2 sequences, 26 were 3D sequences and 29 were sequences covering
214 various regions of the genome, referred to as partial (Table 1).

215 A phylogenetic tree based on the complete VP1 gene was generated using all available
216 NHP EV sequences spanning this region along with representatives of all known human derived
217 EV types (Figure 2 and Supplementary Figure 1). This VP1-based tree clearly delineates the
218 EV species. Non-human primate derived EVs were found in species EV-A, -B, -D, -H, -J and -
219 L. The data for EV-L are very sparce with only one known representative. Species EV-H
220 includes only 6 sequences, of which 4 are epidemiologically related to each other. To determine
221 whether these sequences belong to the same virus type, the 75% nt and 85% aa identity rule
222 (Brown et al., 2009) was applied to their VP1 sequences. As previously described (Nguyen et
223 al., 2014), theses sequences belong to a single virus type (EV-H1), sharing 84.09%-100% nt
224 identity and 96.17%-100% aa identity (Supplementary Figure 2). In contrast to the EV-H and
225 EV-L species, distinct EV-J virus types have been previously described. The capsid-based
226 phylogram accurately delineated the EV-J103, EV-J108 and EV-J122 virus types and confirmed
227 the EV-J121 type, of which CHE20 is the first known representative that was previously
228 characterized based on partial sequence (Sadeuh-Mba et al., 2014) (Figure 3A). In the VP4/VP2
229 region, three strains (CPML8127, CPOV3340 and CPOV3342) were closely related to CHE20
230 (Figure 3C), thus suggesting that they could also be assigned to the EV-J121 virus type. Identity
231 matrices were calculated to assign several untyped EV-J sequences (Figure 3A and 3B,
232 Supplementary Figure 3). NOLA1 and NOLA3 share over 75.41% and 85.31% at nt and aa
233 level, respectively, with an EV-J103 strain and were thus assigned to this virus type. Similarly,
234 the sequence comparison of the partial VP1 sequence of cg5275 with EV-J prototype strains
235 supported its assignment to the EV-J103 type (>77.36% and >85.43 nt and aa identity,
236 respectively, with other EV-J103 viruses, Supplementary Figure 4). As previously reported
237 (Wang et al., 2016) the Sev-nJ strain represent a novel virus type, sharing less than 74.48% and

238 84.27% with other EV-J sequences (Supplementary Figure 3). The CHN-BJ-LXL1 sequence,
239 published 5 years later than the Sev-nJ strain, belonged to the same EV-J type, as it shares
240 82.09% and 92.63% nt and aa identity with Sev-nJ, respectively (Supplementary Figure 3). We
241 propose to assign these two strains to a novel virus type, EV-J123, following current taxonomic
242 nomenclature. Additionally, several partial VP1 and VP4/2 sequences remained to be assigned
243 (Harvala et al., 2014). Based on partial VP1 sequences, the CPOV3336 strain is related to EV-
244 J121 strain CHE20 (Figure 3B), but their nt and aa identities (66.04% and 75.81%,
245 Supplementary Figure 4) indicating that CPOV3336 belongs to a putative distinct virus type
246 (tentatively named EV-J124). A set of short VP4/VP2 sequences (CPOV3337, DJ4139,
247 CPGR7921, CPLM8102 and CPLM8125) were very close that of CPOV3336 (Figure 3C),
248 suggesting that they belong to the newly proposed EV-J124 type.

249 Some viruses previously identified as EV-Js (CHN-BJ-2J, -1J, -8J, -10J and -9J) form a
250 specific group (Figure 3). To determine whether these viruses defined a separated additional
251 new species, we applied the ICTV criteria. The CHN-BJ-2J, 1J, 8J, 10J and 9J sequences share
252 69.41% to 72.74% identity in the ORF and 80.72% to 84.50% in the 2C+3CD genes with other
253 EV-J types (Supplementary Figure 5A and B). However, they share only between 55.36% to
254 59.25% aa identity in the capsid gene (Supplementary Figure 5C), as previously reported (Li et
255 al., 2020). These results support the existence of a separated new species, which we tentatively
256 named *Enterovirus nao* (EV-N), after the Chinese word náo (猱) meaning "monkey" (Couvreur,
257 1896; Loewe, 1993). As previouly suggested (Li et al., 2020), CHN-BJ-1J, 8J, 9J, 10J and CHN-
258 BJ-1-2J form two distinct EV types within this species, as they share less than 72.94% nt and
259 80.99% aa identity (Supplementary Figure 3). We propose naming these types EV-N1 (for the
260 CHN-BJ-1J, 8J, 9J, 10J strains) and EV-N2 (for the CHN-BJ-2J strain).

261 Three EV species, EV-A, -B and -D, include some EVs that originated from humans
262 while others were recovered from NHPs. Within the EV-B species, NHP EVs segregate from

263 human derived EVs and dfine a specific branch (Figure 2). The unique exception to the host
264 association of EV-B types is the case of GAB98 strain, isolated from wild chimpanzees, which
265 is an EV-B107 known to circulate in humans in Asia (Maan et al., 2019; Yamashita et al., 2010).
266 The isolation of strain GAB98 in a chimpanzee was explained as a sporadic human-to-NHP
267 transmission, potentially facilitated by the adjacent or overlapping humans and chimpanzee
268 habitats (Mombo et al., 2017). As previously reported, 4 distinct virus types exist in the NHP
269 specific branch of the EV-B species. Strain Z057, of which a complete genome sequence was
270 obtained in this work (Figure 1), belongs to the EV-B114 virus type (Figure 4), that was first
271 isolated in 1963 from NHP derived cell cultures (Malherbe and Harwin, 1963). Since its
272 discovery, only two EV-B114 EVs have been reported: CPOV3344 (Nguyen et al., 2014) and
273 Z057 (Sadeuh-Mba et al., 2014). They share 82.09%-84.58% nt and 92.88%-93.95% aa
274 identities between each other in the VP1 sequence; thus confirming that they belong to the same
275 virus type (Figure 4). In contrast to EV-B species, the phylogenetic boundaries between NHP
276 and human EVs are less distinct in species EV-A and EV-D (Figure 2). As previously reported,
277 members of the species EV-A fall into four main phylogroups. Clades 1 and 2 consist of EV
278 types that circulate in humans worldwide (Clade 1: EV-A120, EV-A71, CVA7, CVA14 and
279 CVA16; Clade 2: EV-A114, CVA2, CVA3, CVA4, CVA5, CVA6, CVA8, CVA10 and CVA12).
280 Clade 3 inlcudes EVs detected both in human and NHPs (Clade 3: EV-A76, EV-A89, EV-A90,
281 EV-A91, EV-A119, EV-A121 and EV-A125). Finally, clade 4 contains only NHP EV strains
282 (Clade 4: EV-A122, EV-A123 and EV-A124). Similarly, within EV-D, some types are human-
283 specific (EV-D70, EV-D68 and EV-D94), or NHP-specific (EV-D120), while EV-D111 has
284 been detected in both hosts (Sadeuh-Mba et al., 2019). No unidentified or misidentified NHP
285 derived EV-A -B or -D were found in public databases.

286 Notably, 7 sequences spanning the entire main ORF (RCMH3, RCMH6 -both sequenced
287 in this study-, CPML3961, CPML8109, GR2815 and NGR_NHP_1) clustered together in the

288 VP1-based phylogenetic tree and did not align with any known species (Supplementary Figure
289 1). To determine whether these sequences belong to a single new species, we calculated their
290 pairwise nt and aa identities (Supplementary Figure 6). Our analysis showed that the
291 corresponding viruses met the ICTV criteria for classification as the same species: they share
292 over 70% aa identity in the ORF (85.69%-95.02%) and in the 2C+3CD sequences (97.03%-
293 99.69%), as well as over 60% aa identity in the capsid sequences (69.27%-89.82%). Therefore,
294 we propose the creation of a new species, tentavely named *Enterovirus mbel* (EV-M), after the
295 name of *Pterocarpus soyauxii*, a tree native to the tropical forest in Central African (Albreht et
296 al., 2025), known as *mbel* in various Bantu languages spoken in Central and Southern
297 Cameroon, where the first representatives of this species were sampled (Galley, 1964; Tsala,
298 1956). Phylogenetic analyses of the VP1 gene revealed several clusters within this new EV-M
299 species (Figure 5). Based on the established thresholds for discriminating EV types (>75% nt
300 identity and >85% aa identity in the VP1 gene) (Brown et al., 2009), we identified 6 distinct
301 virus types (Supplementary Figure 7): EV-M1 (prototype: RCHM6), EV-M2 (prototype:
302 RCHM3), EV-M3 (prototype: CPML3961), EV-M4 (prototype: CPML8109), EV-M5
303 (prototype: GR2815) and EV-M6 (prototype: NGR_NHP_1). Identity matrices showed that
304 numerous VP1 sequences deposited in GenBank belong to these newly defined EV-M virus
305 types (Supplementary Figure 8 and Figure 5). EV-M1, EV-M2, EV-M4 and EV-M5 virus types
306 gathered sequences that displayed >78.57% nt and >90.62% aa identities between each other.
307 Within the EV-M3 type, all sequences except two (LM1962 and CPOV3334) displayed a nt
308 identity >77.04% with the prototype strain CPML3961, exceeding the 75% threshold. The nt
309 identities of LM1962 and CPOV3334 compared to CPML396 fell slightly below this threshold
310 (73.22% and 74.14%, respectively), but both shared an aa identity >86.0% with the prototype,
311 supporting their classification as EV-M3. Besides the prototype NGR_NHP_1, no additional
312 members of virus type EV-M6 were identified.

313 Shorter VP1 or VP2 sequences, averaging 350 nt in length, have also been published
314 and annotated as belonging to the species EV-J (Mombo et al., 2023, 2021, 2017). However,
315 dendrograms based on these regions revealed that they actually cluster within the EV-M species
316 (Supplementary Figure 9A and 9B). The corresponding sequences are too short to allow
317 accurate typing of the corresponding viruses.

318 Collectively, these results reveal a complex landscape of NHP EV diversity, including
319 two new species (EV-M and EV-N) and two new EV-J types (EV-J123 and EV-J124). This
320 refined classification will facilitate the proper assignation of NHP EV strains to their respective
321 species and virus type.

322 **Phylogenetic relationships in the non-structural genomic regions**

323 Recombination events are frequent among co-circulating EVs, generating new viruses
324 with chimeric genomes that may outcompete parental strains if they confer a fitness advantage.
325 Because recombination constantly reshuffles EV genomes, different genomic regions are not
326 subjected to the same extend of selective pressure and do not follow the same evolutionary
327 trajectories. After clarifying the phylogenetic relationships among the NHP derived EVs in the
328 capsid-encoding region, we investigated their evolutionary history by constructing
329 phylogenetic trees considering 2A2B, 2C, 3A3B, 3C, and 3D regions, using all available full-
330 length ORF sequences (Figure 6). As previously reported (Mombo et al., 2017; Oberste et al.,
331 2007; Sadeuh-Mba et al., 2023), striking intra-species phylogenetic incongruences were
332 observed between trees generated from structural (capsid) and non-structural (2A2B, 2C, 3A3B
333 and 3D) regions. Within the EV-B species, the NHP and human derived EVs formed two distinct
334 clades, consistent with the VP1 phylogeny (Figure 6). Of note, all known EV-Ms clustered with
335 the NHP EV-B clade across all non-structural regions. Consequently, EV-Ms and NHP EV-Bs
336 are phylogenetically indistinguishable from each other across the non-structural part of the
337 genome. Similarly, as previously reported (Oberste et al., 2007; Sadeuh-Mba et al., 2023), EV-

338 A125 strain BA13 did not cluster with other EV-As in the non-structural regions downstream
339 the capsid. Our analyses further revealed that BA13 strain clustered with EV-J108 strains in the
340 2A2B, 2C and 3A3B regions, with the EV-J103 strains in the 3C gene and with the members of
341 the new species EV-N in the 3D gene (Figure 6). Finally, the species EV-D, EV-H, and EV-L
342 exhibited complex phylogenetic relationships. Indeed, they form a single group in the 2A2B
343 region, but segregated into three distinct branches in the 3A3B and 3C based trees. The other
344 regions showed specific pairwise associations: EV-D grouped with EV-H and EV-L in the 2C
345 region and the 3D region, respectively.

346

347 **Comparison of genomic signatures of human and NHP derived EVs**

348 To our knowledge, no comparative genomic analysis of human and NHP derived EV
349 species has been published so far. A previous study, limited to EV-A species, revealed that clade
350 3 viruses displayed distinct relative dinucleotide ratios and codon usage biases compared to
351 human EV-As of clades 1 and 2 (Zeng et al., 2022), suggesting that specific genomic signatures
352 could be associated with the virus' natural history and, particularly its host range. To determine
353 whether NHP derived EVs from different host species shared unique genomic traits, we
354 compared GC content (GC%), dinucleotide ratio, and codon usage (ENC and RSCU analysis)
355 between human and NHP derived EVs. To minimize potential bias in these analyses, we
356 included, where possible, at least the prototype strain and two additional genomes sampled from
357 different countries for each virus type.

358 The GC% ranged from 41% to 49% across all species (Figure 7A). Interestingly, the
359 GC% strongly correlated with the host origin rather than EV species classification, with all NHP
360 virus types featuring a GC% <47%. In particular, NHP-specific EV species (EV-J, EV-H, EV-
361 L, EV-M and EV-N) display low GC%, ranging from 43% to 46.4%, as did the few known EV-

362 D virus types (41.0%-43.5%). Strikingly, virus types segregated with respect to their host of
363 origin within the EV-A and EV-B species: all EV-A virus types of clades 1 and 2, originating
364 from humans, displayed GC%>47%, while NHP-associated clades 3 and 4 had GC%<47%
365 (Figure 7B). A similar clustering was found within EV-B species, with all human derived EV
366 types featuring a GC% >47%, while all NHP derived EV types had a GC% <45% (Figure 7C).
367 Notably, the GC % of EV-C virus types, which are restricted to human hosts, mirrored their
368 phylogenetic relationships: respiratory virus types (EV-C104, EV-C105, EV-C109, EV-C117
369 and EV-C118) exhibited a high GC % (48–48.5%), cultivable enteric virus types (PV1, PV2,
370 PV3, CVA11, CVA13, CVA17, CVA20, CVA21, CVA24, EV-C95, EV-C96, EV-C99 and
371 EV-C102) displayed an intermediate GC % (45–46.5 %), while non-cultivable viruses (CVA1,
372 CVA19, CVA22, EV-C113, EV-C116 and EV-C119) showed a low GC % (43-44 %) (Figure
373 7D).

374 A previous study revealed that EV-As clade 3 viruses display lower CpG and UpA
375 dinucleotide frequencies compared to clades 1 and 2 (Zeng et al., 2022). To determine whether
376 NHP derived EV-As of clade 4 share this trait, we analyzed dinucleotide ratios of EV-A strains
377 representative of the different clades (Figure 8). Our analysis revealed that EV-A viruses
378 belonging to clade 4 had the same low CpG ratio as clade 3 EV-As (Figure 8B). The same
379 observation was made within EV-B species, with NHP derived EV types displaying a CpG ratio
380 substantially lower than that of human derived EV types (NHP EVs<0.42 vs Human EVs>0.56)
381 (Figure 8C). With the exception of EV-H, a low CpG ratio seems to be a common characteristic
382 of NHP derived EVs. Indeed, members of the EV-J, EV-L, EV-M and EV-N species featured
383 CpG ratios<0.5 and comparable to those of NHP derived EV-A and EV-B species (Figure 8A).
384 It is noteworthy that all EV-D virus types had CpG ratio<0.35, irrespective of their host of
385 origin.

386 We also investigated codon usage bias in human and NHP derived EVs. The effective
387 number of codons (ENC), which ranges from 20 (strong bias, with one codon used for each
388 amino-acid) to 61 (no bias, even use of each codon) (Wright, 1990), revealed disctinct patterns.
389 Human EV-As (clades 1 and 2), EV-Bs, respiratory EV-Cs, and NHP derived EV-Hs had ENC
390 greater than 54 (Figure 9A), while clade 3, NHP derived EV-As, NHP derived EV-Bs, cultivable
391 and non-cultivable EV-Cs, EV-Ds, EV-Js and EV-Ms featured ENC values between 50 and 54.
392 The sole known EV-L had an ENC value of 48. Since the GC% at the third codon position
393 (GC3) could introduce a bias in the ENC analysis, we compared the ENC values for each EV
394 species with the expected ENC based on the respective GC3 percentage (Butt et al., 2016)
395 (Supplementary Figure 10). All calculated ENC values fell below the expected curve,
396 confirming that GC3 content did not account for the variations across virus group.

397 We also analyzed the Relative Synonymous Codon Usage (RSCU), which quantifies
398 how frequently each synonymous codon is used in a gene or genome relative to the expected
399 usage. The results, summarized in a cladogram, grouped EV types into three main clusters
400 (Figure 9B). Group A containes only the EV-H1 type. Group B includes all other NHP-specific
401 EV types (except EV-A123), whichever their species, all human cultivable EV-Cs, all EV-Ds
402 and EV-A clades 3 and 4. Group C containes all other human EV-Cs and EV-A123. These results
403 suggest that codon-usage bias varies markedly among EV groups, yet it only partially mirrors
404 the underlying phylogenetic relationships among them.

405 Collectively, these results reveal that dinucleotide ratios, ENC and RSCU values vary
406 among NHP EV species. Remarkably, GC% mirrors phylogenetic relationships between EVs
407 and could serve as a predictive marker of human versus NHP origin of EVs.

408

409 **DISCUSSION**

410 For decades, no standardized nomenclature existed for NHP derived EVs. In the 1980s,
411 it was proposed to designate these viruses under the common name, *Simian enterovirus*,
412 followed by a number corresponding to a given serotype (Kalter, 1982). For a long time, NHP
413 derived EVs were not unassigned to specific species. In the 2000s and 2010s, the generalization
414 of genetic sequencing brought about major changes in EV classification, as phylogenetic
415 relationships replaced phenotypic criteria to delineate EV types and species. Based on their
416 genetic relationships, NHP derived EVs were classified into different species, some of which
417 also included human derived EVs. By considering all NHP EV genome sequences available in
418 public databases and integrating our two new full-genome sequences, our analyses confirmed
419 the existence of a new species (Figure 5 and Supplementary Figure 6) that had been previously
420 suggested by two independent studies (Nguyen et al., 2014; Sadeuh-Mba et al., 2014). We found
421 that this new EV species gathered viruses sampled in three countries, Gabon, Cameroon, and
422 Nigeria (Mombo et al., 2023, 2021, 2017; Nguyen et al., 2014; Sadeuh-Mba et al., 2023). This
423 geographic distribution justifies the name proposed for this species, *Enterovirus mbel*, which
424 refers to an endemic tree of this region, called African padauk in English (Albreht et al., 2025).
425 We also identified another new EV species, comprising a few viruses sampled in China
426 previously misidentified as EV-Js (Li et al., 2020) (Figure 3 and Supplementary Figure 5). Since
427 this species has only been documented in China, we propose naming it *Enterovirus nao,* after
428 the Chinese word referring to a monkey in a poem written 2,500-3,000 years ago (Couvreur,
429 1896; Loewe, 1993). However, it is important to note that the genuine geographic area of
430 circulation of these two new EV species remain to be documented, especially for *E. nao* which
431 was detected in captive animals.

432 The discovery of the uORF in the genome of some EVs a few years ago constituted a
433 drastic change in our understanding of their biology. This uORF was experimentally shown to
434 encode a non-essential protein that enhances virus replication in gut cells (Lulla et al., 2019).

435 Interestingly, the putative uORFs identified in the genomes of EV-B114 Z057, EV-M1 RCMH6,
436 and EV-M3 RCMH3 are respectively located 134 nt, 133 nt, and 136 nt upstream of the AUG
437 start codon of the main ORF. These distances fall below the cutoff set by Lulla et al. for
438 screening for uORF of EV genomic sequences (Lulla et al., 2019). Although functional
439 validation of these candidate uORFs is still lacking, our findings suggest that the prevalence of
440 uORFs across EVs may have been underestimated and warrants re-evaluation using a more
441 permissive distance threshold.

442 The phylogenetic analyses performed on the non-structural part of the genome further
443 confirmed the close genetic relationships between NHP EVs recovered from different species,
444 as previouly documented (Figure 6) (Mombo et al., 2017; Nguyen et al., 2014; Oberste et al.,
445 2007; Sadeuh-Mba et al., 2014). This close genetic relatedness could be explained by
446 convergent evolution, which may have driven EV genomes from different species to become
447 phylogenetically more similar as a consequence of their replication within NHP hosts. However,
448 it unlikely that convergent evolution could uniformly affect the entire non-structural region of
449 the genome, as it would be the case for NHP derived EV-Bs and EV-Ms that cluster together
450 across this region. Furthermore, convergent evolution among NHP derived EVs is not supported
451 by the fact that most known NHP derived EV-As do not cluster with other NHP derived EVs,
452 except for EV-A125 strain BA13, which is phylogenetically close to EV-Js downstream the
453 capsid. If convergent evolution was shaping the evolution of EV-A125, it would be expected to
454 similarly affect other NHP derived EV-As, which seems unlikely. Therefore, these phylogenetic
455 incongruences found in this study are probably due to recombinations, a well-established driver
456 of EV evolution (Muslin et al., 2019).

457 While interspecies recombination has been observed in the 5'UTR (Boros et al., 2012;
458 Muslin et al., 2015; Santti et al., 1999), it has never been documented downstream of the capsid-
459 encoding region among human derived EVs. Even intraspecies recombination is limited to

460 specific subsets of EV types that can recombine with one another but not with members of other
461 subsets. For instance, EV-As of clade 3 do not exchange genetic sequences with EV-As of clades
462 1 and 2 (Oberste et al., 2005; Sadeuh-Mba et al., 2014). Similarly, respiratory EV-Cs and
463 non-cultivable EV-Cs do not recombine with each other or with cultivable EV-Cs (Brouwer et
464 al., 2020). Several factors could explain the apparent inability of some EVs to generate
465 recombinant viruses, such as functional incompatibilities between genomic regions from distant
466 phylogenetic groups, lack of fitness advantages conferred to the recombinant viruses, or low
467 frequency of co-infection due to different host and tissue tropisms. Our analysis provided
468 substantial evidence to the existence of interspecies recombination between some NHP derived
469 EV species, particularly between EV-Ms and EV-Bs. These findings change our understanding
470 of recombination between EVs, which appear to be capable of producing larger leaps than
471 previously believed. This raise fundamental questions about the past and future evolutionary
472 path of these viruses. For instance, it is unknown whether NHP derived EV-Bs acqiured non-
473 structural genes from EV-Ms, or vice-versa, unless both inherited these genes from a yet-
474 unidentified common source. The ability of EV-Ms to recombine with human derived EV-Bs
475 and the potential for such inter-species macro-evolutionary events to produce new human-
476 adapted EV-Ms are also difficult to predict.

477 Previously studies have shown that genomes of clade 3 EV-As, which were detected in
478 both NHPs and humans, had a lower CpG frequency than human specific EV-As and featured
479 a specific codon usage bias that impairs their expression of these EV-A genomes in human cells
480 (Zeng et al., 2022). These observations led us to thypothesize that some genetic characteristics
481 shared by NHP derived EVs may play a role in host adaptation. Nucleotide composition has
482 been previously used to predict the host of novel picorna-like viruses (Kapoor et al., 2010).
483 While our analysis of dinucleotide frequency and codon usage bias across the various EV
484 species circulating in humans and in NHPs did not reveal a clear segregation between human

485 and NHP derived viruses (Figure 9), our results identified a relatively low GC% appears as a
486 signature shared by virus types circulating among NHPs, regardless of their virus species
487 (Figure 7). Unfortunately, the lack of data precluded robust statistical analysis, as some EV
488 types are represented by hundreds of sequences while others, such as EV-L species, are
489 represented by only one or two sequences. Based on genetic and epidemiological data, we and
490 others have previously postulated that EV-Ds known in humans may have recently emerged
491 from zoonotic sources, presumably NHPs (Kono et al., 1981; Miyamura et al., 1986; Sadeuh-
492 Mba et al., 2019). The low GC% values displayed by all EV-D virus types, irrespective of their
493 tissue tropism or host of origin, are consistent with that hypothesis. The relative relatedness of
494 EV-Ds and NHP-specific EV species in the non-structural part of the genome further supporting
495 the hypothesis of a zoonotic origin of EV-Ds. However, host adaptation is not the only driver
496 of GC%, as illustrated by EV-C virus types, which display a broad range of GC% values despite
497 all circulating in humans (Figure 7). In addition, the GC% of human and NHP EV strains falls
498 within the broad range observed in the human transcriptome (average 48%, varying from <40%
499 to >60% among mRNAs) (Courel et al., 2019; Piovesan et al., 2019). Unfortunately, the GC%
500 of transcriptomes of NHP infected with EVs remains uncharacterized to our knowledge. The
501 GC% overlap between EVs and human transcriptome is insufficient to conclude host-species
502 adaptation, and the lack of data prevents comparative analysis that could elucidate potential
503 host-driven GC% adaptation in NHP EV strains. The genomic signature result from a dynamic
504 equilibrium influenced by many factors, of which polymerase fidelity, RNA secondary structure
505 requirements, host translational preferences, immune pressure, and genetic recombination.
506 Furthermore, identifying genomic signatures that reliably distinguish NHP erived EVs from
507 human derived EVs is currently limited by the paucity of fully sequenced strains, our limited
508 knowledge of the natural history of these viruses, and particularly the uncertainty surrounding
509 their host tropisms.

ACKNOWLEDGEMENT

The authors are indebted to Dr. WEI Yu, who shared her knowledge of ancient Chinese poetry, and Dr. Nick J. Knowles, for his feedback regarding enterovirus phylogeny. This work has used the computational and storage services (Maestro cluster) provided by the IT department at Institut Pasteur, Paris.

AUTHORS CONTRIBUTIONS

MB and SASM conceived the study. SASM, PCdC, MP, MLJ, MCEZ, AB and ESL generated the sequencing data. CA, ESL, and MB performed the analyses. CA, NJ and MB wrote the article, which was approved by all co-authors.

FUNDING

This project was funded by the Institut Pasteur through a program of incentive actions (*Actions concertée inter-pasteuriennes*, ACIP-162). CA's salary is funded by a grant of the Gates Foundation (INV-067045). The E.S.-L. laboratory is funded by Institut Pasteur, the INCEPTION program (Investissements d'Avenir grant ANR-16-CONV-0005), the Ixcore foundation for research, the French Government's Investissement d'Avenir programme, Laboratoire d'Excellence 'Integrative Biology of Emerging Infectious Diseases' (grant no. ANR-10-LABX-62-IBEID), the HERA Projects DURABLE (grant no 101102733) and LEAPS (grant no 101094685). The funders of this study had no role in study design, data collection, analysis and interpretation, or writing of the article.

REFERENCES

Albreht, L., Bourland, N., Klitgård, B., Schley, R., Janssens, S.B., Hardy, O.J., 2025. Species delimitation and phylogenomics of economically important African Pterocarpus trees, with an implication for the development of DNA-based species identification tools. Mol Phylogenet Evol 204, 108277. https://doi.org/10.1016/j.ympev.2024.108277

Bessaud, M., Blondel, B., Joffret, M.-L., Mac Kain, A., Delpeyroux, F., 2018. Nomenclature et classification des entérovirus : quelle histoire ! Virologie (Montrouge) 22, 289–303. https://doi.org/10.1684/vir.2018.0756

Bessaud, M., Sadeuh-Mba, S.A., Joffret, M.-L., Razafindratsimandresy, R., Polston, P., Volle, R., Rakoto-Andrianarivelo, M., Blondel, B., Njouom, R., Delpeyroux, F., 2016. Whole Genome Sequencing of Enterovirus species C Isolates by High-Throughput Sequencing: Development of Generic Primers. Front Microbiol 7, 1294. https://doi.org/10.3389/fmicb.2016.01294

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Boros, Á., Pankovics, P., Knowles, N.J., Reuter, G., 2012. Natural interspecies recombinant bovine/porcine enterovirus in sheep. J Gen Virol 93, 1941–1951. https://doi.org/10.1099/vir.0.041335-0

Brouwer, L., Benschop, K.S.M., Nguyen, D., Kamau, E., Pajkrt, D., Simmonds, P., Wolthers, K.C., 2020. Recombination Analysis of Non-Poliovirus Members of the Enterovirus C Species; Restriction of Recombination Events to Members of the Same 3DPol Cluster. Viruses 12, 706. https://doi.org/10.3390/v12070706

Brown, B.A., Maher, K., Flemister, M.R., Naraghi-Arani, P., Uddin, M., Oberste, M.S., Pallansch, M.A., 2009. Resolving ambiguities in genetic typing of human enterovirus species C clinical isolates and identification of enterovirus 96, 99 and 102. J Gen Virol 90, 1713–1723. https://doi.org/10.1099/vir.0.008540-0

Buchfink, B., Xie, C., Huson, D.H., 2015. Fast and sensitive protein alignment using DIAMOND. Nat Methods 12, 59–60. https://doi.org/10.1038/nmeth.3176

Butt, A.M., Nasrullah, I., Qamar, R., Tong, Y., 2016. Evolution of codon usage in Zika virus genomes is host and vector specific. Emerg Microbes Infect 5, e107. https://doi.org/10.1038/emi.2016.106

Cordey, S., Gerlach, D., Junier, T., Zdobnov, E.M., Kaiser, L., Tapparel, C., 2008. The cis-acting replication elements define human enterovirus and rhinovirus species. RNA 14, 1568–1578. https://doi.org/10.1261/rna.1031408

Courel, M., Clément, Y., Bossevain, C., Foretek, D., Vidal Cruchez, O., Yi, Z., Bénard, M., Benassy, M.-N., Kress, M., Vindry, C., Ernoult-Lange, M., Antoniewski, C., Morillon, A., Brest, P., Hubstenberger, A., Roest Crollius, H., Standart, N., Weil, D., 2019. GC content shapes mRNA storage and decay in human cells. Elife 8, e49708. https://doi.org/10.7554/eLife.49708

Couvreur, S., 1896. Chant IX. Kio Koung, in: Cheu King. Texte Chinois Avec Une Double Traduction En Français et En Latin, Une Introduction et Un Vocabulaire. Imprimerie de la Mission catholique, Ho-kien fou, pp. 302–304.

Doté, J., Joffret, M.-L., Beta, B.N., Ait-Ahmed, M., Banga-Mingo, V., Knowles, N.J., Jouvenet, N., MBaïkoua, M.-N., Gouandjika-Vasilache, I., Bessaud, M., 2024.

Characterization of enteroviruses circulating among farm animals and children in Central African Republic. Emerg Microbes Infect 13, 2368212. https://doi.org/10.1080/22221751.2024.2368212

Fieldhouse, J.K., Wang, X., Mallinson, K.A., Tsao, R.W., Gray, G.C., 2018. A systematic review of evidence that enteroviruses may be zoonotic. Emerging Microbes & Infections 7, 1–09. https://doi.org/10.1038/s41426-018-0159-1

Galley, S., 1964. Dictionnaire Fang -Français et Français - Fang. Éditions Henri Messeiller, Neuchatel, Suisse.

Gámbaro, F., Pérez, A.B., Agüera, E., Prot, M., Martínez-Martínez, L., Cabrerizo, M., Simon-Loriere, E., Fernandez-Garcia, M.D., 2021. Genomic surveillance of enterovirus associated with aseptic meningitis cases in southern Spain, 2015–2018. Sci Rep 11, 21523. https://doi.org/10.1038/s41598-021-01053-4

Grubaugh, N.D., Gangavarapu, K., Quick, J., Matteson, N.L., De Jesus, J.G., Main, B.J., Tan, A.L., Paul, L.M., Brackney, D.E., Grewal, S., Gurfield, N., Van Rompay, K.K.A., Isern, S., Michael, S.F., Coffey, L.L., Loman, N.J., Andersen, K.G., 2019. An amplicon-based sequencing framework for accurately measuring intrahost virus diversity using PrimalSeq and iVar. Genome Biol 20, 8. https://doi.org/10.1186/s13059-018-1618-7

Harvala, H., Van Nguyen, D., McIntyre, C., Ahuka-Mundeke, S., Ngole, E.M., Delaporte, E., Peeters, M., Simmonds, P., 2014. Co-circulation of enteroviruses between apes and humans. J Gen Virol 95, 403–407. https://doi.org/10.1099/vir.0.059048-0

Joffret, M.-L., Polston, P.M., Razafindratsimandresy, R., Bessaud, M., Heraud, J.-M., Delpeyroux, F., 2018. Whole Genome Sequencing of Enteroviruses Species A to D by High-Throughput Sequencing: Application for Viral Mixtures. Front Microbiol 9, 2339. https://doi.org/10.3389/fmicb.2018.02339

Kalter, S.S., 1982. Enteric viruses of nonhuman primates. Vet Pathol Suppl 7, 33–43.

Kapoor, A., Simmonds, P., Lipkin, W.I., Zaidi, S., Delwart, E., 2010. Use of nucleotide composition analysis to infer hosts for three novel picorna-like viruses. J Virol 84, 10322–10328. https://doi.org/10.1128/JVI.00601-10

Kono, R., Sasagawa, A., Yamazaki, S., Nakazono, N., Minami, K., Otatsume, S., Robin, Y., Renaudet, J., Cornet, M., Afoakwa, S.N., Mingle, J.A., Obinim, J.K., Huros, A., 1981. Seroepidemiologic studies of acute hemorrhagic conjunctivitis virus (enterovirus type 70) in West Africa. III. Studies with animal sera from Ghana and Senegal. Am J Epidemiol 114, 362–368. https://doi.org/10.1093/oxfordjournals.aje.a113203

Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K., 2018. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol Biol Evol 35, 1547–1549. https://doi.org/10.1093/molbev/msy096

Li, D., Luo, R., Liu, C.-M., Leung, C.-M., Ting, H.-F., Sadakane, K., Yamashita, H., Lam, T.-W., 2016. MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. Methods 102, 3–11. https://doi.org/10.1016/j.ymeth.2016.02.020

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 1000 Genome Project Data Processing Subgroup, 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Li, W., Qiang, X., Qin, S., Huang, Y., Hu, Y., Bai, B., Hou, J., Gao, R., Zhang, X., Mi, Z., Fan, H., Ye, H., Tong, Y., Mao, P., 2020. Virome diversity analysis reveals novel enteroviruses and a human picobirnavirus in stool samples from African green monkeys with diarrhea. Infect Genet Evol 82, 104279. https://doi.org/10.1016/j.meegid.2020.104279

Loewe, M., 1993. Shih ching, in: Early Chinese Texts, Early China Special Monographs Series. The Society for the Study of Early China and The Institute of East Asian Studies, University of California, Berkeley, pp. 415–423.

Lulla, V., Dinan, A.M., Hosmillo, M., Chaudhry, Y., Sherry, L., Irigoyen, N., Nayak, K.M., Stonehouse, N.J., Zilbauer, M., Goodfellow, I., Firth, A.E., 2019. An upstream protein-coding region in enteroviruses modulates virus infection in gut epithelial cells. Nat Microbiol 4, 280–292. https://doi.org/10.1038/s41564-018-0297-1

Maan, H.S., Dhole, T.N., Chowdhary, R., 2019. Identification and characterization of nonpolio enterovirus associated with nonpolio-acute flaccid paralysis in polio endemic state of Uttar Pradesh, Northern India. PLoS One 14, e0208902. https://doi.org/10.1371/journal.pone.0208902

Malherbe, H., Harwin, R., 1963. The cytopathic effects of vervet monkey viruses. S Afr Med J 37, 407–411.

Matranga, C.B., Andersen, K.G., Winnicki, S., Busby, M., Gladden, A.D., Tewhey, R., Stremlau, M., Berlin, A., Gire, S.K., England, E., Moses, L.M., Mikkelsen, T.S., Odia, I., Ehiane, P.E., Folarin, O., Goba, A., Kahn, S.H., Grant, D.S., Honko, A., Hensley, L., Happi, C., Garry, R.F., Malboeuf, C.M., Birren, B.W., Gnirke, A., Levin, J.Z., Sabeti, P.C., 2014. Enhanced methods for unbiased deep sequencing of Lassa and Ebola RNA viruses from clinical and biological samples. Genome Biol 15, 519. https://doi.org/10.1186/PREACCEPT-1698056557139770

Miyamura, K., Tanimura, M., Takeda, N., Kono, R., Yamazaki, S., 1986. Evolution of enterovirus 70 in nature: All isolates were recently derived from a common ancestor. Archives of Virology 89, 1–14. https://doi.org/10.1007/BF01309875

Modiyinji, A.F., Joffret, M.-L., Nombot-Yazenguet, M.P. de M., Endengue Zanga, M.-C., Sadeuh-Mba, S., Njouom, R., Bessaud, M., 2025. Molecular characterization of enteroviruses circulating among pigs and goats in two Central African countries, Cameroon and the Central African Republic. Access Microbiol 7, 000886.v3. https://doi.org/10.1099/acmi.0.000886.v3

Mombo, I.M., Boundenga, L., Suquet, E., Ngoubangoye, B., Maganga, G.D., Leroy, E.M., Charpentier, M.J., Rougeron, V., 2021. Natural infection of free-ranging mandrills (Mandrillus sphinx) by enteroviruses and astroviruses in southern Gabon. Microb Pathog 150, 104659. https://doi.org/10.1016/j.micpath.2020.104659

Mombo, I.M., Lukashev, A.N., Bleicker, T., Brünink, S., Berthet, N., Maganga, G.D., Durand, P., Arnathau, C., Boundenga, L., Ngoubangoye, B., Boué, V., Liégeois, F., Ollomo, B., Prugnolle, F., Drexler, J.F., Drosten, C., Renaud, F., Rougeron, V., Leroy, E., 2017. African Non-Human Primates Host Diverse Enteroviruses. PLOS ONE 12, e0169067. https://doi.org/10.1371/journal.pone.0169067

Mombo, I.M., Ngoubangoye, B., Boundenga, L., Mebaley, T.G.N., Assoghe, A.A., Tsoumbou, T., Bisseye, C., Maganga, G., Leroy, E., 2023. Natural Non-Polio-Enteroviruses Detected in Captive Mandrills (Mandrillus sphinx) and Chimpanzees (Pan troglodytes troglodytes) Housed in a Primate Center in

Franceville, Gabon. Int J Primatol 44, 847–851. https://doi.org/10.1007/s10764-023-00361-y

Muslin, C., Joffret, M.-L., Pelletier, I., Blondel, B., Delpeyroux, F., 2015. Evolution and Emergence of Enteroviruses through Intra- and Inter-species Recombination: Plasticity and Phenotypic Impact of Modular Genetic Exchanges in the 5' Untranslated Region. PLoS Pathog 11, e1005266. https://doi.org/10.1371/journal.ppat.1005266

Muslin, C., Kain, A.M., Bessaud, M., Blondel, B., Delpeyroux, F., 2019. Recombination in Enteroviruses, a Multi-Step Modular Evolutionary Process. Viruses 11, 859. https://doi.org/10.3390/v11090859

Nguyen, D.V., Harvala, H., Ngole, E.M., Delaporte, E., Woolhouse, M.E.J., Peeters, M., Simmonds, P., 2014. High Rates of Infection with Novel Enterovirus Variants in Wild Populations of Mandrills and Other Old World Monkey Species. Journal of Virology 88, 5967–5976. https://doi.org/10.1128/jvi.00088-14

Oberste, M. Steven, Maher, K., Kilpatrick, D.R., Flemister, M.R., Brown, B.A., Pallansch, M.A., 1999. Typing of Human Enteroviruses by Partial Sequencing of VP1. Journal of Clinical Microbiology 37, 1288–1293. https://doi.org/10.1128/jcm.37.5.1288-1293.1999

Oberste, M. S., Maher, K., Kilpatrick, D.R., Pallansch, M.A., 1999. Molecular evolution of the human enteroviruses: correlation of serotype with VP1 sequence and application to picornavirus classification. J Virol 73, 1941–1948. https://doi.org/10.1128/JVI.73.3.1941-1948.1999

Oberste, M.S., Maher, K., Michele, S.M., Belliot, G., Uddin, M., Pallansch, M.A., 2005. Enteroviruses 76, 89, 90 and 91 represent a novel group within the species Human enterovirus A. J Gen Virol 86, 445–451. https://doi.org/10.1099/vir.0.80475-0

Oberste, M.S., Maher, K., Pallansch, M.A., 2007. Complete genome sequences for nine simian enteroviruses. J Gen Virol 88, 3360–3372. https://doi.org/10.1099/vir.0.83124-0

Oberste, M.S., Maher, K., Pallansch, M.A., 2002. Molecular Phylogeny and Proposed Classification of the Simian Picornaviruses. Journal of Virology 76, 1244–1251. https://doi.org/10.1128/jvi.76.3.1244-1251.2002

Olson, R.D., Assaf, R., Brettin, T., Conrad, N., Cucinell, C., Davis, J.J., Dempsey, D.M., Dickerman, A., Dietrich, E.M., Kenyon, R.W., Kuscuoglu, M., Lefkowitz, E.J., Lu, J., Machi, D., Macken, C., Mao, C., Niewiadomska, A., Nguyen, M., Olsen, G.J., Overbeek, J.C., Parrello, B., Parrello, V., Porter, J.S., Pusch, G.D., Shukla, M., Singh, I., Stewart, L., Tan, G., Thomas, C., VanOeffelen, M., Vonstein, V., Wallace, Z.S., Warren, A.S., Wattam, A.R., Xia, F., Yoo, H., Zhang, Y., Zmasek, C.M., Scheuermann, R.H., Stevens, R.L., 2023. Introducing the Bacterial and Viral Bioinformatics Resource Center (BV-BRC): a resource combining PATRIC, IRD and ViPR. Nucleic Acids Res 51, D678–D689. https://doi.org/10.1093/nar/gkac1003

Piovesan, A., Pelleri, M.C., Antonaros, F., Strippoli, P., Caracausi, M., Vitale, L., 2019. On the length, weight and GC content of the human genome. BMC Res Notes 12, 106. https://doi.org/10.1186/s13104-019-4137-z

Romette, J.L., Prat, C.M., Gould, E.A., de Lamballerie, X., Charrel, R., Coutard, B., Fooks, A.R., Bardsley, M., Carroll, M., Drosten, C., Drexler, J.F., Günther, S., Klempa, B.,

Pinschewer, D., Klimkait, T., Avsic-Zupanc, T., Capobianchi, M.R., Dicaro, A., Ippolito, G., Nitsche, A., Koopmans, M., Reusken, C., Gorbalenya, A., Raoul, H., Bourhy, H., Mettenleiter, T., Reiche, S., Batten, C., Sabeta, C., Paweska, J.T., Eropkin, M., Zverev, V., Hu, Z., Mac Cullough, S., Mirazimi, A., Pradel, F., Lieutaud, P., 2018. The European Virus Archive goes global: A growing resource for research. Antiviral Res 158, 127–134. https://doi.org/10.1016/j.antiviral.2018.07.017

Sadeuh-Mba, S., Doté, J., Joffret, M.-L., Chesnais, M., Filhol, T., Gouandjika-Vasilache, I., Jouvenet, N., Bessaud, M., 2023. Animal enteroviruses: a glimpse of a wide evolutionary landscape. Virologie (Montrouge) 27, 22–34. https://doi.org/10.1684/vir.2023.1000

Sadeuh-Mba, S.A., Bessaud, M., Joffret, M.-L., Zanga, M.-C.E., Balanant, J., Ngole, E.M., Njouom, R., Reynes, J.-M., Delpeyroux, F., Rousset, D., 2014. Characterization of Enteroviruses from Non-Human Primates in Cameroon Revealed Virus Types Widespread in Humans along with Candidate New Types and Species. PLOS Neglected Tropical Diseases 8, e3052. https://doi.org/10.1371/journal.pntd.0003052

Sadeuh-Mba, S.A., Joffret, M.-L., Mazitchi, A., Endegue-Zanga, M.-C., Njouom, R., Delpeyroux, F., Gouandjika-Vasilache, I., Bessaud, M., 2019. Genetic and phenotypic characterization of recently discovered enterovirus D type 111. PLoS Negl Trop Dis 13, e0007797. https://doi.org/10.1371/journal.pntd.0007797

Santti, J., Hyypiä, T., Kinnunen, L., Salminen, M., 1999. Evidence of recombination among enteroviruses. J Virol 73, 8741–8749. https://doi.org/10.1128/JVI.73.10.8741-8749.1999

Tajima, F., Nei, M., 1984. Estimation of evolutionary distance between nucleotide sequences. Mol Biol Evol 1, 269–285. https://doi.org/10.1093/oxfordjournals.molbev.a040317

Tsala, T., 1956. Dictionnaire Ewondo - Français. Lyon.

Wang, X., Shao, S., Wang, H., Shen, Q., Yang, S., Zhang, W., 2016. An enterovirus from a captive primate in China. Springerplus 5, 1281. https://doi.org/10.1186/s40064-016-2966-y

Wright, F., 1990. The “effective number of codons” used in a gene. Gene 87, 23–29. https://doi.org/10.1016/0378-1119(90)90491-9

Yamashita, T., Ito, M., Tsuzuki, H., Sakae, K., Minagawa, H., 2010. Molecular identification of enteroviruses including two new types (EV-98 and EV-107) isolated from Japanese travellers from Asian countries. J Gen Virol 91, 1063–1066. https://doi.org/10.1099/vir.0.016014-0

Zeng, L., Chen, M., Wang, M., Zhu, L., Yan, J., Zhang, X., Xu, J., Zhang, S., 2022. Enterovirus A Shows Unique Patterns of Codon Usage Bias in Conventional Versus Unconventional Clade. Front. Cell. Infect. Microbiol. 12. https://doi.org/10.3389/fcimb.2022.941325

764 **Legends of figures and table**

765 **Figure 1. Genome organization of CHE20, Z057, RCHM6, and RCHM3 strains. (A)**
766 Schematic representation of the genomic organization of the four newly sequenced genomes.
767 The uORF is shown in brown (Lulla et al., 2019), the P1 region in purple, the P2 region in
768 green, and the P3 region in blue. **(B)** Alignment of nucleotide sequences of the potential uORF
769 compared to the E7 Wallace strain (highlighted by brown boxes). The start codons of the
770 putative uORF and the main ORF, as well as, the stop codons of the uORF are highlighted. **(C)**
771 Alignment of the CRE sequence compared to the PV3 Leon strain (Cordey et al., 2008). The
772 sequence alignment was performed using the CLC Main Workbench v24 software.

773

774 **Figure 2. Phylogenetic analysis of complete VP1 sequences from non-human primate and**
775 **representative human enteroviruses.** The Neighbor-Joining tree was reconstructed from a
776 full-length VP1 nucleotide sequence alignment. Branches are coloured according to EV species:
777 EV-A (light green), EV-B (dark blue), EV-C (red), EV-D (dark green), EV-H (light blue), EV-J
778 (pink), EV-L (purple), EV-M (brown) and EV-N (orange). Circles indicate NHP derived EV
779 strains. The arrows indicate new genomes sequenced in our study (Figure 1). The tree is rooted
780 with the PV1 Mahoney strain (V01148). Bootstrap values greater than 70% (based on 1,000
781 pseudoreplicates) are shown at the corresponding nodes. The sequence alignment was
782 performed using the CLC Main Workbench v24, the phylogenetic tree was constructed with
783 MEGA X and the visualization was generated with FigTree. The tree is drawn to scale, with
784 branch lengths measured in the number of substitutions per site.

785

786 **Figure 3. Phylogenetic analysis of enterovirus J species.** The maximum-likelihood trees
787 were constructed with all available full capsid (length of alignment 2571 nt) **(A)**, partial VP1

788 (length of alignment 754 nt) **(B)** and partial VP4/2 (length of alignment 429 nt) **(C)** sequences.
789 The trees were rooted with the EV-A124 OM22 strain. The symbol * indicates proposed new
790 type classifications. The colours of the strain names indicate their host origins: *Mandrillus*
791 (green), *Chlorocebus* (orange), *Cercocebus* (brown), *Papio* (blue), *Macaca* (purple) and *Pan*
792 (red). Bootstrap values greater than 70% (based on 500 pseudoreplicates) are shown at the
793 corresponding nodes. The sequence alignments were performed using the CLC Main
794 Workbench v24, the phylogenetic trees were constructed with MEGA X and the visualization
795 was generated with FigTree. The trees are drawn to scale, with branch lengths measured in the
796 number of substitutions per site.

797

798 **Figure 4. VP1-based phylogenetic tree of enterovirus B species.** The maximum-likelihood
799 tree was constructed with all available full NHP EV-B VP1 sequences. The tree is rooted with
800 the EV-B3 SSM strain. The colours of the strain names indicate their host origin: *Mandrillus*
801 (green), *Homo* (black), *Cercopithecus* (grey), and *Pan* (red). Bootstrap values greater than 70%
802 (based on 500 pseudoreplicates) are shown at the corresponding nodes. Percentages of
803 nucleotide (Nt) and amino acid (Aa) identities between EV-B114 strains are indicated. The
804 sequence alignment was performed using the CLC Main Workbench v24, and the phylogenetic
805 tree was constructed with MEGA X. The tree is drawn to scale, with branch lengths measured
806 in the number of substitutions per site.

807

808 **Figure 5. VP1-based phylogenetic tree of the enterovirus M species.** The maximum-
809 likelihood tree was constructed with all available full VP1 sequences. VP1 sequences derived
810 from complete genomes are in bold. The colours of the strain names indicate their host origin:
811 *Mandrillus* (green), *Cercocebus* (brown), *Macaca* (purple) and *Pan* (red). The tree is rooted

812 with the EV-J121 CHE20 strain. Bootstrap values greater than 70% (based on 500
813 pseudoreplicates) are shown at the corresponding nodes. The sequence alignment was
814 performed using the CLC Main Workbench v24, and the phylogenetic tree was constructed with
815 MEGA X. The tree is drawn to scale, with branch lengths measured in the number of
816 substitutions per site.

817

818 **Figure 6. Phylogenetic analysis of non-structural genes derived from complete genomes of**
819 **NHP and representative human EVs.** The maximum-likelihood trees were constructed from
820 the 2A2B, 2C, 3A3B, 3C and 3D regions of all complete NHP and representative human EV
821 genomes. Branches are coloured according to EV species: EV-A (light green), EV-B (dark blue),
822 EV-C (red), EV-D (dark green), EV-H (light blue), EV-J (pink), EV-L (purple), EV-M (brown)
823 and EV-N (orange). The colours of the strain names indicate their host origin: *Macaca* (purple),
824 *Pan* (red), *Papio* (light blue), *Chlorocebus* (orange), *Mandrillus* (green), *Cercocebus* (brown)
825 and *Homo* (black). The trees are rooted with the PV1 Mahoney strain (V01148). Bootstrap
826 values greater than 70% (based on 500 pseudoreplicates) are shown at the corresponding nodes.
827 The sequence alignments were performed using the CLC Main Workbench v24, the
828 phylogenetic trees were constructed with MEGA X, and the visualizations were generated with
829 FigTree. The trees are drawn to scale, with branch lengths measured in the number of
830 substitutions per site.

831

832 **Figure 7. Distribution of GC content across enterovirus species.** GC content distribution for
833 all EV species **(A)**, EV-A species **(B)**, EV-B species **(C)**, and EV-C species **(D)** are plotted. Blue
834 boxes indicate clusters of human derived EV while brown boxes indicate clusters of mainly
835 NHP derived EVs. When possible, three sequences per type were included in the analysis and

836 standard deviations are plotted. Abbreviations: Cult, cultivable; NC, non-cultivable; Resp,
837 respiratory; Hs, *Homo sapiens*; NHP, non-human primate.

838

839 **Figure 8. Dinucleotide ratio analysis of human and non-human primate enterovirus**
840 **genomes. (A)** All dinucleotide ratios are shown. **(B)** The CpG dinucleotide ratio is specifically
841 highlighted for each EV species. Abbreviations: Cult, cultivable; NC, non-cultivable; Resp,
842 respiratory; Hs, *Homo sapiens*; NHP, non-human primate.

843

844 **Figure 9. Codon usage bias analysis of enterovirus species. (A)** Effective number of codons
845 (ENC) values for all NHP EVs and representative human EVs are plotted. When possible, three
846 sequences per type were included in the analysis. **(B)** Relative synonymous codon usage
847 (RSCU) values for all NHP EVs and representative human EVs were summarized in a
848 cladogram. The three main groups separated by the cladogram are labeled A, B and C.
849 Abbreviations: Hs, *Homo sapiens*; NHP, non-human primate.

850

851 **Table 1. Summary of included studies and number of available enterovirus sequences.** For
852 each study, the first author, year of publication, and the number of sequences available are listed.
853 Data were obtained from: (Ao et al., 2016; Han et al., 2021; Handley et al., 2012; Harvala et
854 al., 2011, 2014; Kapusinszky et al., 2017; Li et al., 2020; Mombo et al., 2017, 2021, 2023;
855 Negrey et al., 2022; Nguyen et al., 2014; Sadeuh-Mba et al., 2014; Wang et al., 2016; Zhao et
856 al., 2019).

857

858

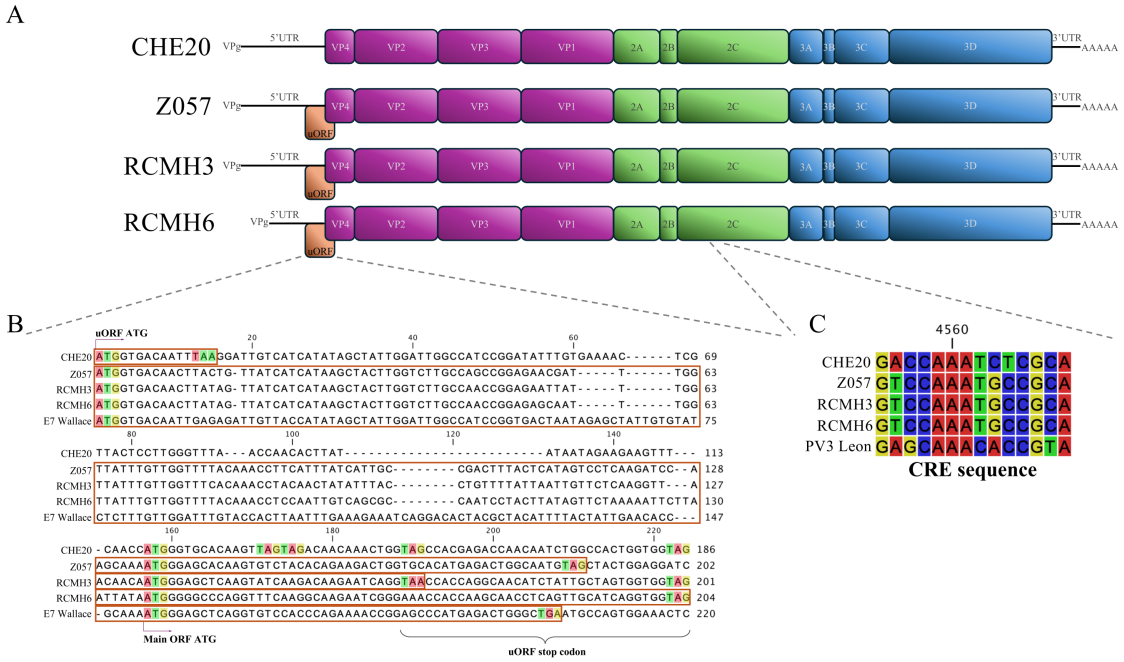


Figure 1. Genome organization of CHE20, Z057, RCHM6, and RCHM3 strains

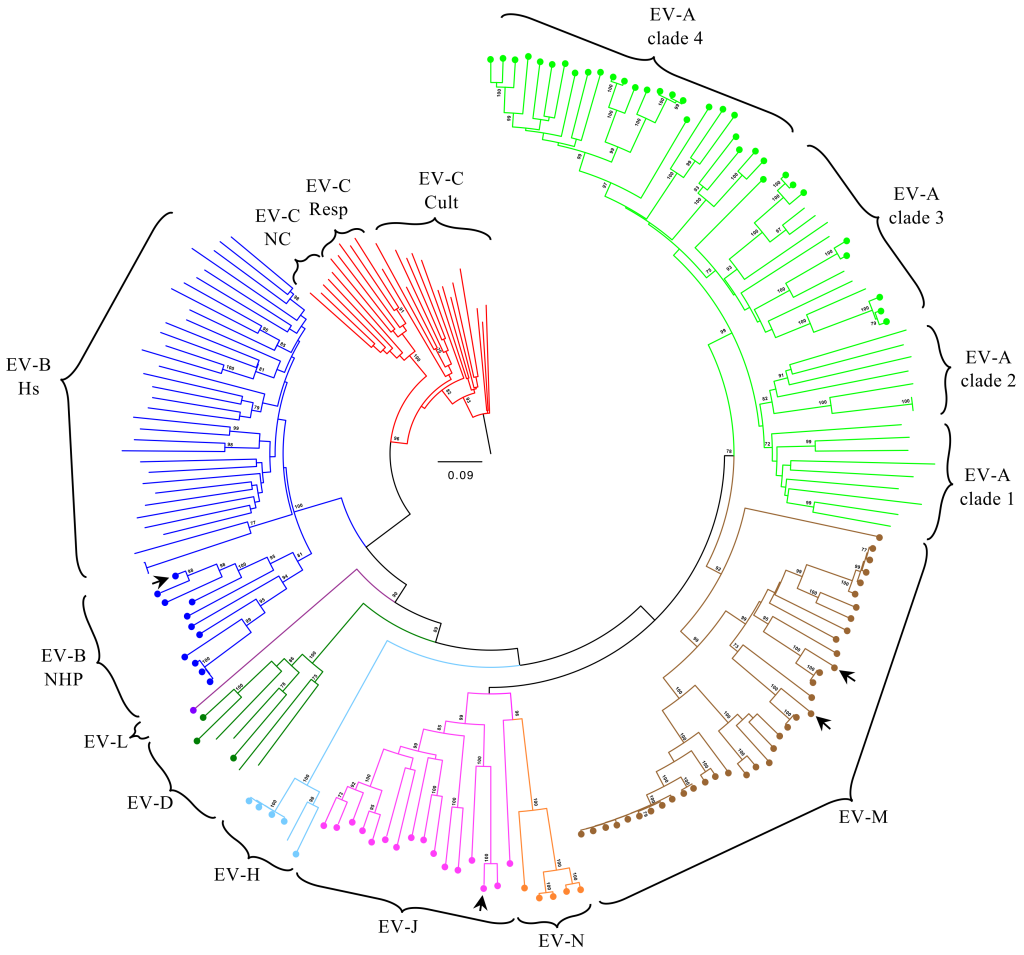


Figure 2. Phylogenetic analysis of complete VP1 sequences from non-human primate and representative human enteroviruses

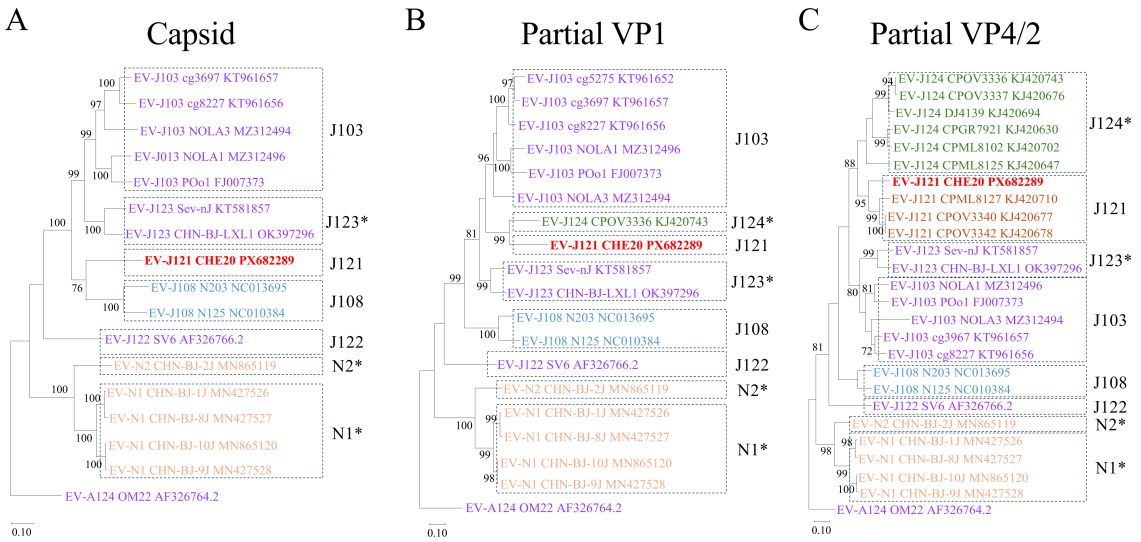


Figure 3. Phylogenetic analysis of enterovirus J species

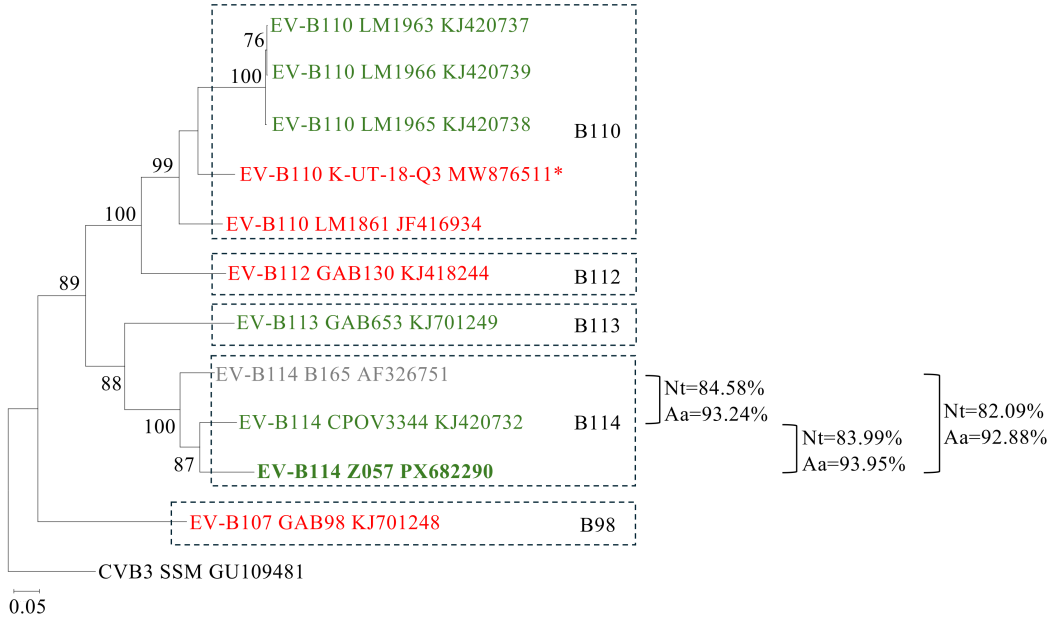


Figure 4. VP1-based phylogenetic tree of enterovirus B species

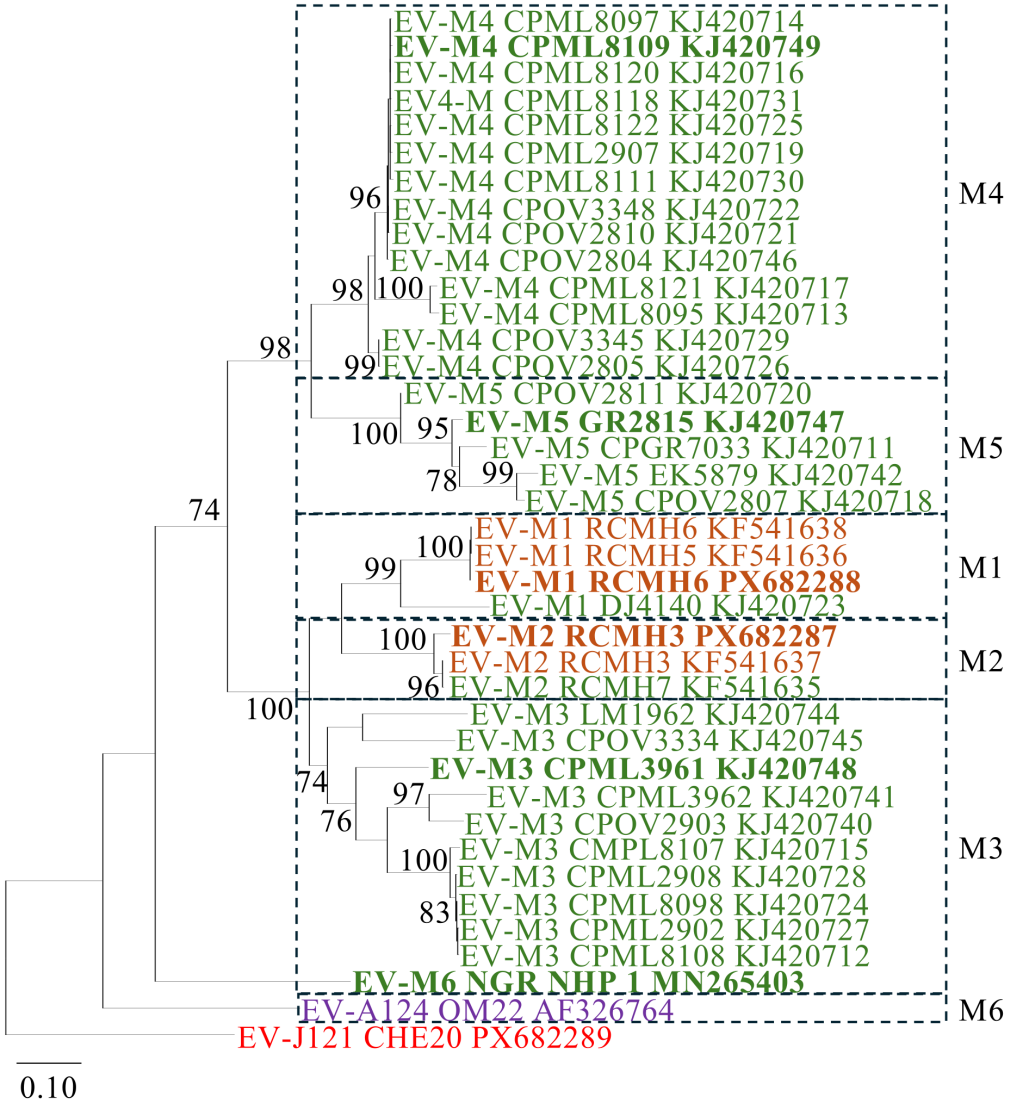


Figure 5. VP1-based phylogenetic tree of the enterovirus M species

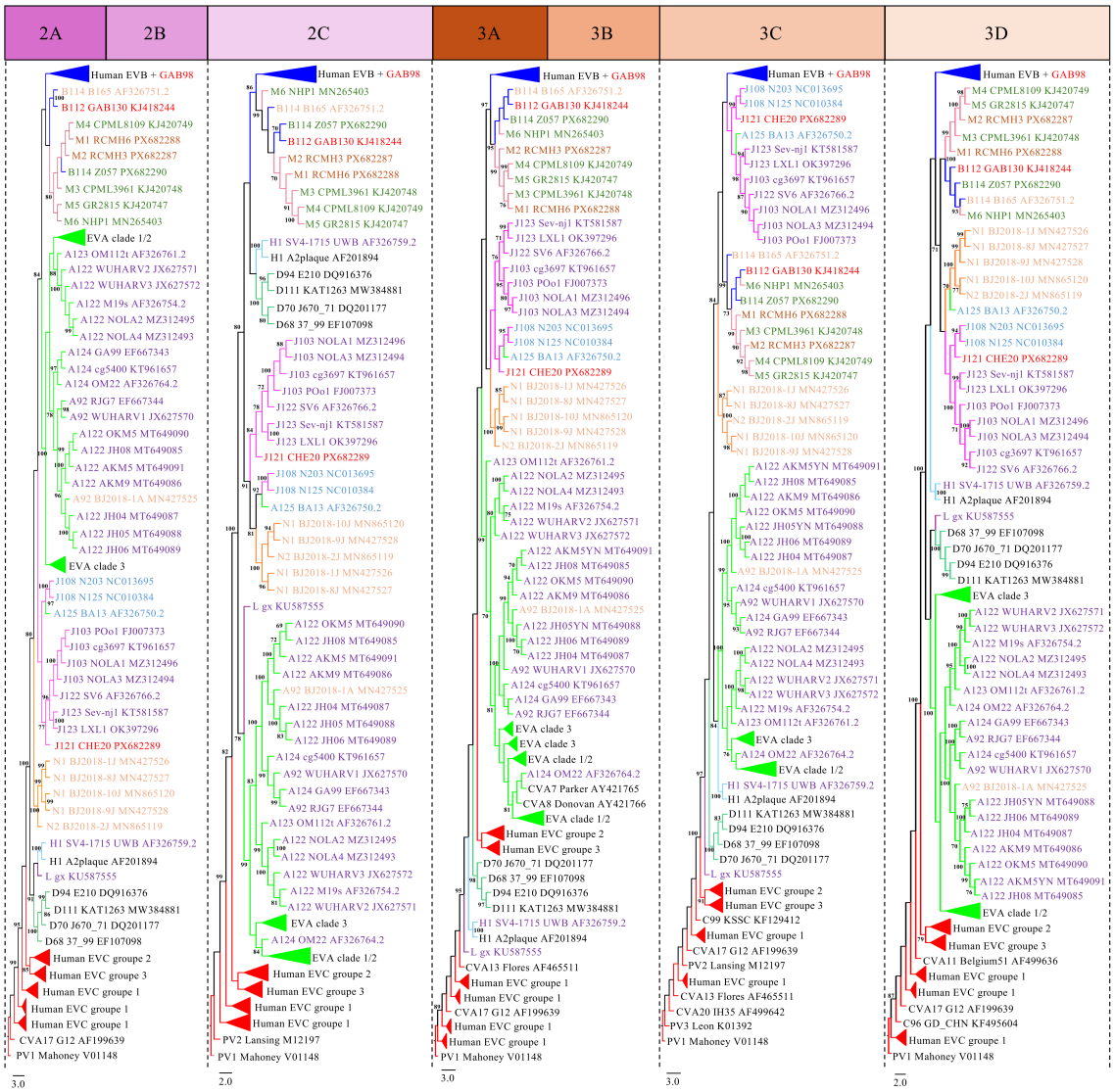


Figure 6. Phylogenetic analysis of non-structural genes derived from complete genomes of NHP and representative human EVs

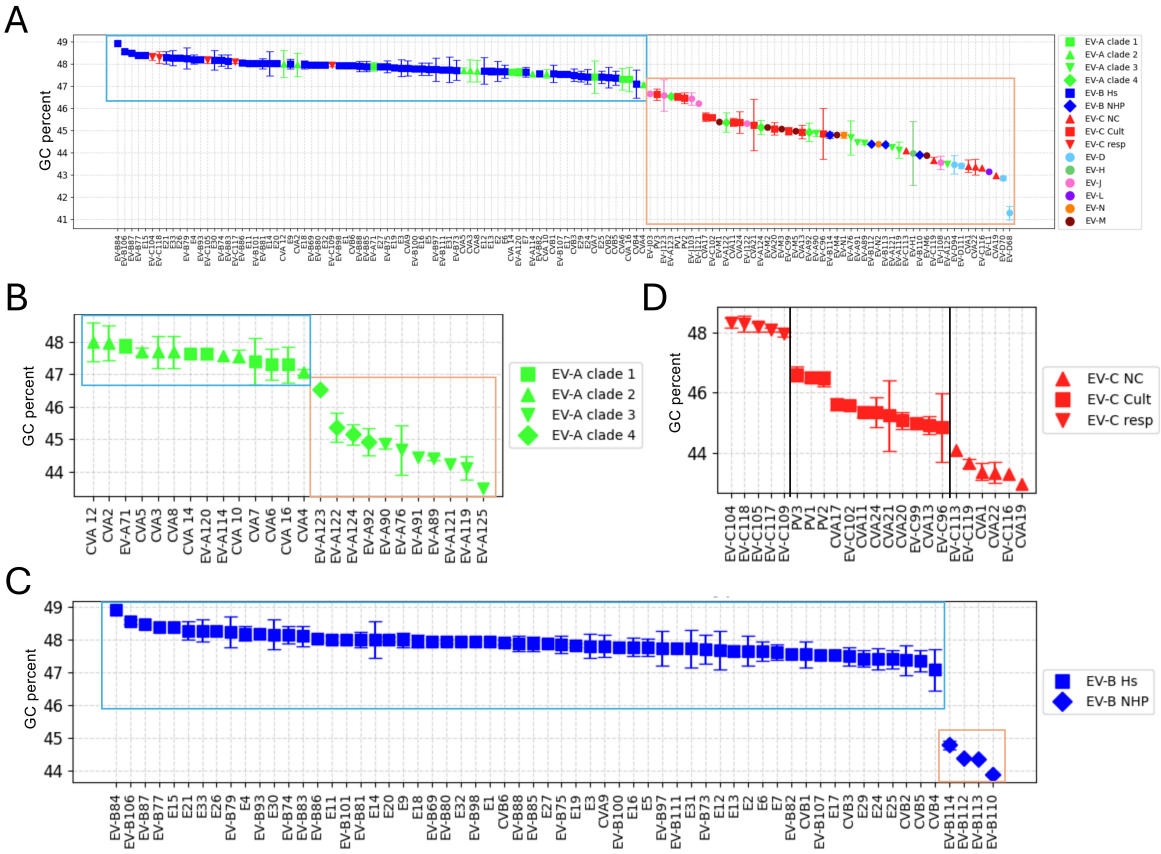


Figure 7. Distribution of GC content across enterovirus species

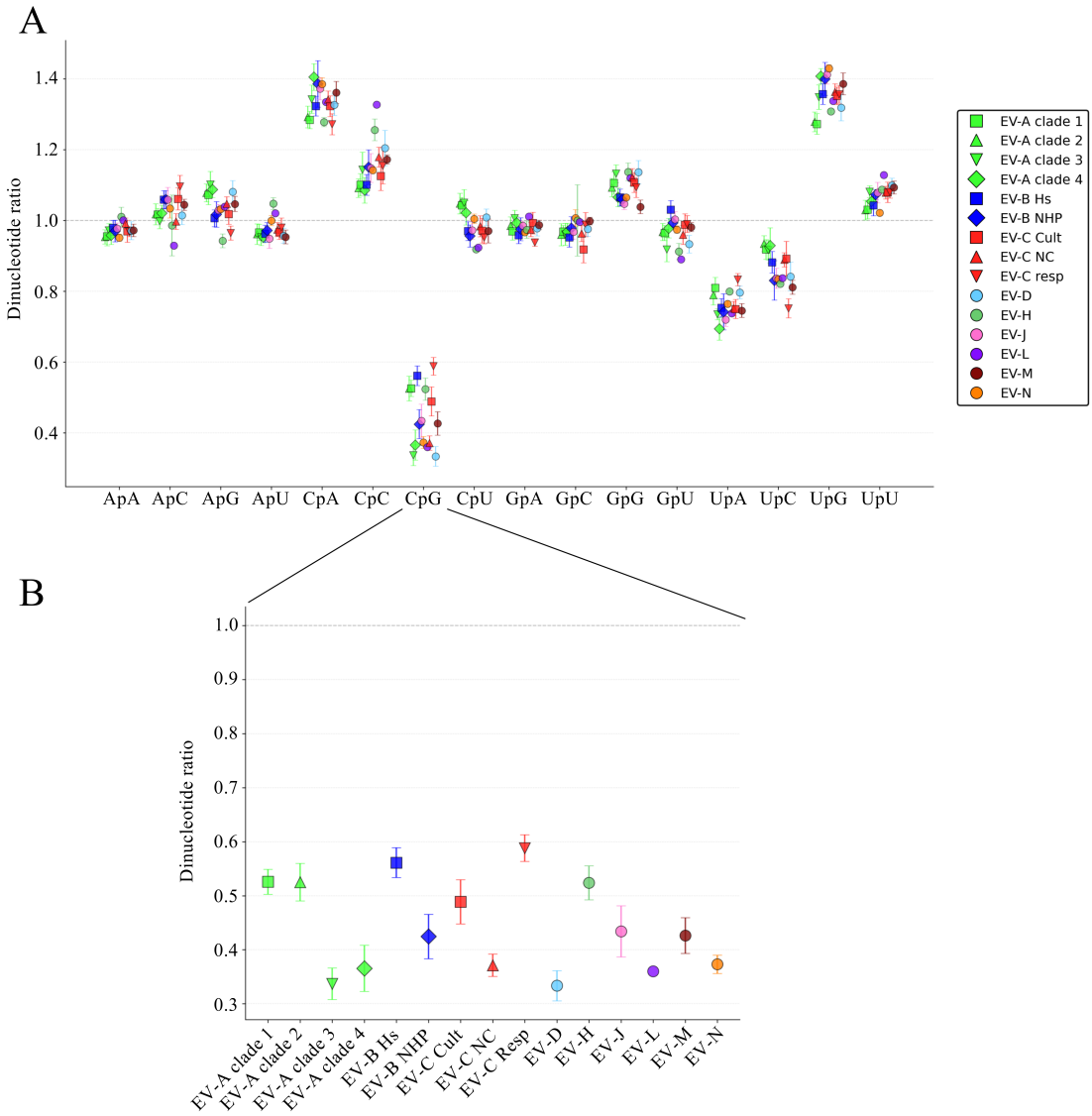


Figure 8. Dinucleotide ratio analysis of human and non-human primate enterovirus genomes

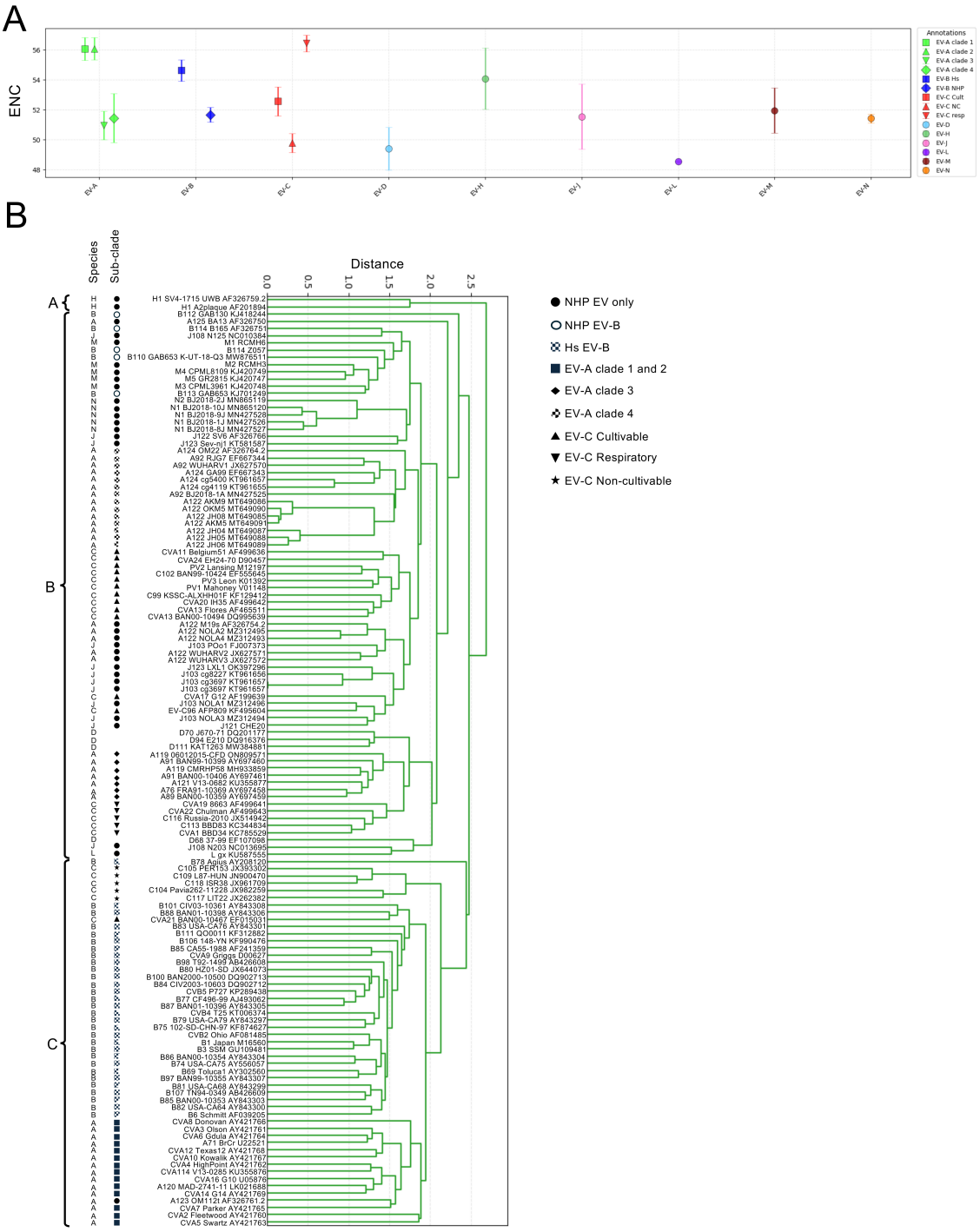


Figure 9. Codon usage bias analysis of enterovirus species

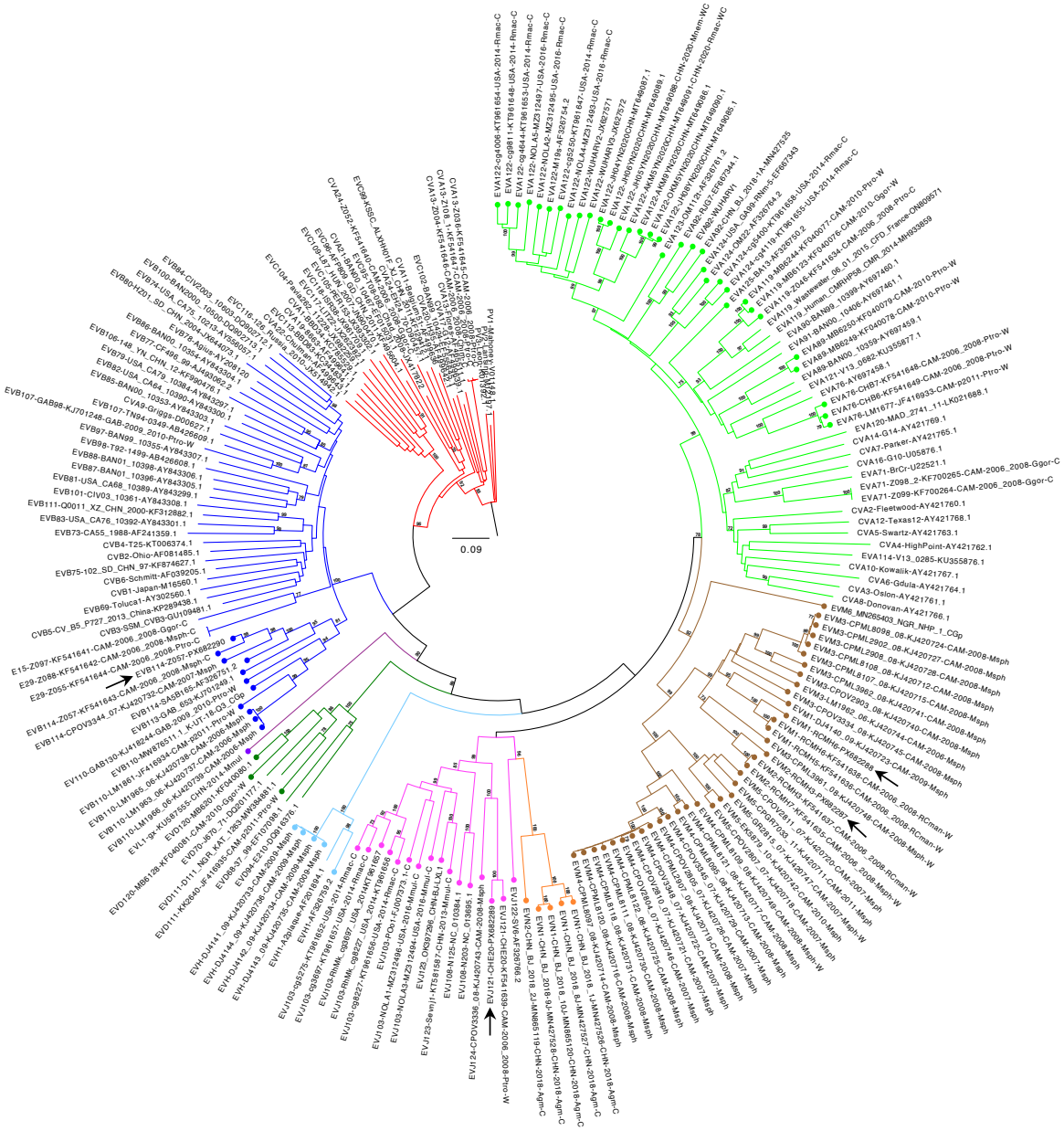
First author	Year	5'UTR	VP4	VP4/2	VP2	VP1	3D	full genome	partial	Total
Harvala	2011	3		3		3	3			12
Handley	2012								3	3
Sadeuh-Mba	2014	36				27	23			86
Nguyen	2014			86		36		3		125
Harvala	2014			14		6				20
Ao	2016							1		1
Wang	2016							1		1
Mombo	2017				19	9		3		31
Kapusinszky	2017								22	22
Zhao	2019							4	1	5
George (unpublished)	2019								1	1
Li	2020							5	1	6
Mombo	2021					10				10
Han	2021							7		7
Negrey	2022								1	1
Mombo	2023				5					5
Total		39	0	103	19	91	26	24	29	336

Table 1. Summary of included studies and number of available enterovirus sequences

877

Supplementary Table 1. Predicted polyprotein cleavage sites of enterovirus Z057, CHE20, RCMH3, and RCHM6. Cleavage site predictions are based on multiple sequence alignment with the annotated poliovirus type 1 Mahoney reference strain (V01148.1). The analysis was performed using CLC Main Workbench v24 software.

Strains	VP4/VP2	VP2/VP3	VP3/VP1	VP1/2A	2A/2B	2B/2C	2C/3A	3A/3B	3B/3C	3C/3D
CHE20	LK/SP	TQ/GL	LQ/SP	TA/GA	EQ/GI	KQ/AD	FQ/GP	FQ/GA	VQ/GP	EQ/GE
Z057	LN/SP	PQ/GL	LQ/SP	TY/GA	EQ/GV	RQ/AD	FQ/GP	FQ/GA	VQ/GP	EQ/GE
RCMH3	LK/SP	TQ/GL	LQ/GD	NT/GA	EQ/GV	RQ/AD	FQ/GP	FQ/GA	VQ/GP	EQ/GE
RCMH6	LK/SP	AQ/GL	LQ/GD	TT/GA	EQ/GV	RQ/AD	FQ/GP	FQ/GA	VQ/GP	EQ/GE



Supplementary Figure 1. Phylogenetic analysis of complete VP1 sequences from non-human primate and representative human enteroviruses. This phylogenetic tree is the same as in the Figure 2, with the name of each EV type indicated. For NHP EV, the following nomenclature is used: Species-Type-Strain name-Accession number-Country-Year-Host species-Condition of life of host. For prototypic EVs, the following nomenclature is used: Species-Type-Strain name-Accession Number. Abbreviations: CAM, Cameroon; CHN, China; USA, United States of America; GAB, Gabon; NGR, Nigeria; Msph, *Mandrillus sphinx*; Mleu, *Mandrillus leucophaeus*; Ctor, *Cercocebus torquatus*; Caet, *Chlorocebus aethiops*; Ptro, *Pan troglodytes*; Ggor, *Gorilla gorilla*; Mmul, *Macaca mulatta*; W, wild; C, captive. The arrows indicate new genomes sequenced in our study (Figure 1). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site.

		1	2	3	4	5	6
EV-H DJ4141 KJ420733	1		100.00	99.88	99.88	85.6	84.09
EV-H DJ4144 KJ420736	2	100.00		99.88	99.88	85.6	84.09
EV-H DJ4143 KJ420735	3	99.65	99.65		100.00	85.71	84.2
EV-H DJ4142 KJ420734	4	99.65	99.65	100.00		85.71	84.2
EV-H1 SV4-1715 UWB AF326759	5	97.56	97.56	97.91	97.91		84.44
EV-H1 A-2 plaque AF201894	6	95.82	95.82	96.17	96.17	96.86	

Percentage of nucleotide identity

Percentage of amino acid identity

Supplementary Figure 2. Type characterization of the enterovirus H species. Nucleotide and amino acid identity matrices of EV-H complete VP1 sequences were calculated using CLC Main Workbench v24.

A

			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
J121	EV-J121 CHE20 KF541639	1		60,89	62,39	61,80	62,27	61,29	60,48	63,33	57,35	55,87	50,45	50,23	50,17	50,63	49,71	49,60
J103	EV-J103 POo1 FJ007373	2	60,89		82,23	75,29	76,81	76,25	71,61	71,71	59,43	57,98	54,98	52,93	50,68	50,68	51,03	51,37
	EV-J103 NOLA1 MZ312496	3	62,39	82,23		75,41	76,31	77,33	71,89	73,89	59,91	58,70	54,57	52,70	51,94	51,71	52,05	52,05
	EV-J103 NOLA3 MZ312494	4	61,80	75,29	75,41		77,82	77,33	72,64	72,55	60,57	60,51	56,00	52,70	52,05	51,71	51,48	51,71
	EV-J103 cg3697 KT961657	5	62,27	76,81	76,31	77,82		87,02	74,48	73,17	61,63	59,84	56,18	53,39	52,85	52,97	53,54	53,77
	EV-J103 cg8227 KT961656	6	61,29	76,25	77,33	77,33	87,02		72,31	73,95	60,30	59,31	55,42	52,82	52,51	52,17	53,88	54,00
J123*	EV-J123 Sev-nJ KT581857	7	60,48	71,61	71,89	72,64	74,48	72,31		82,09	62,53	59,93	57,32	51,78	51,48	51,94	52,17	52,40
	EV-J123 CHN-BJ-LXL1 OK397296	8	63,33	71,71	73,89	72,55	73,17	73,95	82,09		62,87	60,51	56,86	53,86	52,17	52,51	51,71	51,94
J108	EV-J108 N125 NC010384	9	57,35	59,43	59,91	60,57	61,63	60,30	62,53	62,87		75,44	54,84	50,58	50,58	50,35	51,04	51,38
	EV-J108 N203 NC013695	10	55,87	57,98	58,70	60,51	59,84	59,31	59,93	60,51	75,44		54,61	50,12	49,08	49,54	48,96	49,42
J122	EV-J122 SV6 AF326766.2	11	50,45	54,98	54,57	56,00	56,18	55,42	57,32	56,86	54,84	54,61		52,63	53,97	53,29	52,61	52,72
N2*	EV-N2 CHN-BJ-2J MN865119	12	50,23	52,93	52,70	52,70	53,39	52,82	51,78	53,86	50,58	50,12	52,63		72,24	71,78	72,71	72,94
N1*	EV-N1 CHN-BJ-1J MN427526	13	50,17	50,68	51,94	52,05	52,85	52,51	51,48	52,17	50,58	49,08	53,97	72,24		98,12	92,37	92,25
	EV-N1 CHN-BJ-8J MN427527	14	50,63	50,68	51,71	51,71	52,97	52,17	51,94	52,51	50,35	49,54	53,29	71,78	98,12		91,20	91,08
	EV-N1 CHN-BJ-10J MN865120	15	49,71	51,03	52,05	51,48	53,54	53,88	52,17	51,71	51,04	48,96	52,61	72,71	92,37	91,20		99,18
	EV-N1 CHN-BJ-9J MN427528	16	49,60	51,37	52,05	51,71	53,77	54,00	52,40	51,94	51,38	49,42	52,72	72,94	92,25	91,08	99,18	

B

			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
J103	EV-J103 POo1 FJ007373	1		92,31	87,41	89,16	85,31	82,11	83,16	66,10	60,14	62,20	52,90	47,26	46,92	47,60	47,26	46,58
	EV-J103 NOLA1 MZ312496	2	92,31		87,06	88,81	85,66	81,82	83,92	65,87	60,82	62,54	52,22	48,63	48,29	48,97	48,97	47,95
	EV-J103 cg3697 KT961657	3	87,41	87,06		94,76	86,36	84,27	84,62	68,26	61,86	61,86	51,54	48,63	48,63	48,97	48,97	49,32
	EV-J103 cg8227 KT961656	4	89,16	88,81	94,76		87,41	84,97	85,66	66,89	61,51	62,89	52,90	48,97	48,97	49,66	49,66	48,97
	EV-J103 NOLA3 MZ312494	5	85,31	85,66	86,36	87,41		82,17	82,87	65,53	60,82	61,86	50,85	48,63	48,63	47,95	47,95	47,60
J123*	EV-J123 Sev-nJ KT581857	6	82,11	81,82	84,27	84,97	82,17		92,63	67,47	61,86	62,20	52,90	48,63	48,29	48,97	48,97	46,58
	EV-J123 CHN-BJ-LXL1 OK397296	7	83,16	83,92	84,62	85,66	82,87	92,63		68,15	62,89	63,92	54,27	48,97	48,63	48,63	48,63	47,26
J121	EV-J121 CHE20 KF541639	8	66,10	65,87	68,26	66,89	65,53	67,47	68,15		55,07	55,07	47,99	46,28	46,28	46,28	46,28	45,27
J108	EV-J108 N125 NC010384	9	60,14	60,82	61,86	61,51	60,82	61,86	62,89	55,07		89,72	47,78	44,48	44,48	43,79	44,14	44,83
	EV-J108 N203 NC013695	10	62,20	62,54	61,86	62,89	61,86	62,20	63,92	55,07	89,72		51,19	44,83	44,83	44,48	44,48	44,83
J122	EV-J122 SV6 AF326766.2	11	52,90	52,22	51,54	52,90	50,85	52,90	54,27	47,99	47,78	51,19		46,78	46,44	46,44	46,44	46,10
N1*	EV-N1 CHN-BJ-1J MN427526	12	47,26	48,63	48,63	48,97	48,63	48,63	48,97	46,28	44,48	44,83	46,78		99,65	94,37	95,07	79,58
	EV-N1 CHN-BJ-8J MN427527	13	46,92	48,29	48,63	48,97	48,63	48,29	48,63	46,28	44,48	44,83	46,44	99,65		94,01	94,72	79,58
	EV-N1 CHN-BJ-10J MN865120	14	47,60	48,97	48,97	49,66	47,95	48,97	48,63	46,28	43,79	44,48	46,44	94,37	94,01		98,59	79,93
	EV-N1 CHN-BJ-9J MN427528	15	47,26	48,97	48,97	49,66	47,95	48,97	48,63	46,28	44,14	44,48	46,44	95,07	94,72	98,59		80,99
N2*	EV-N2 CHN-BJ-2J MN865119	16	46,58	47,95	49,32	48,97	47,60	46,58	47,26	45,27	44,83	44,83	46,10	79,58	79,58	79,93	80,99	

Supplementary Figure 3. Type characterization of enterovirus J strains based on complete VP1 sequences. Nucleotide **(A)** and amino acid **(B)** identity matrices of VP1 sequences are shown. Translation of nucleotide sequences, multiple sequence alignments, and identity matrices were performed using CLC Main Workbench v24. Putative novel type assignments are indicated by an asterisk (*).

A

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
J103	EV-J103 cg5275 KT961652		91,24	86,66	77,49	77,49	77,36	75,20	72,92	67,25	63,37	59,39	57,92	55,94	53,59	53,86	54,52	54,52	53,19
	EV-J103 cg3697 KT961657	91,24		89,08	79,25	77,63	77,36	75,88	73,05	67,79	63,37	60,45	59,52	56,74	55,19	55,19	56,25	56,52	54,65
	EV-J103 cg8227 KT961656	86,66	89,08		78,03	79,25	78,30	73,85	72,79	68,19	62,57	60,99	58,85	55,14	55,59	55,05	57,05	57,05	54,65
	EV-J103 NOLA3 MZ312494	77,49	79,25	78,03		76,15	76,45	73,78	71,35	65,41	64,43	62,17	60,03	56,21	54,65	54,26	54,26	54,39	53,32
	EV-J103 NOLA1 MZ312496	77,49	77,63	79,25	76,15		83,02	73,05	73,05	64,82	63,37	59,92	58,85	54,07	55,45	55,05	55,98	55,72	53,46
	EV-J103 POo1 FJ007373	77,36	77,36	78,30	76,45	83,02		72,30	71,49	64,46	63,22	59,63	58,43	54,61	53,59	53,59	54,26	54,26	54,26
J123*	EV-J123 Sev-nJ KT581857	75,20	75,88	73,85	73,78	73,05	72,30		81,06	64,91	61,39	61,23	60,96	57,01	53,32	53,99	54,52	54,65	52,79
	EV-J123 CHN-BJ-LXL1 OK397296	72,92	73,05	72,79	71,35	73,05	71,49	81,06		65,17	63,82	62,09	61,82	54,97	53,39	53,79	53,26	53,39	53,52
J124*	EV-J124 CPOV3336 KJ420743	67,25	67,79	68,19	65,41	64,82	64,46	64,91	65,17		65,68	55,21	55,48	52,61	50,20	50,60	50,20	50,33	49,67
J121	EV-J121 CHE20 KF541639	63,37	63,37	62,57	64,43	63,37	63,22	61,39	61,82	65,68		53,66	56,32	51,26	51,20	51,60	50,00	50,00	50,93
J108	EV-J108 N203 NC013695	59,39	60,45	60,99	62,17	59,92	59,63	61,23	62,09	55,21	53,66		79,78	54,04	50,27	50,53	50,13	50,40	53,48
	EV-J108 NJ25 NC010384	57,92	59,52	58,85	60,03	58,85	58,42	60,96	61,82	55,48	56,32	79,78		52,58	51,60	51,47	52,01	51,87	53,88
J122	EV-J122 SV6 AF326766.2	55,94	56,74	55,14	56,21	54,07	54,61	57,01	54,97	52,61	51,26	54,04	52,58		55,01	54,47	54,21	54,34	54,07
N1*	EV-J CHN-BJ-1J MN427526	53,59	55,19	55,59	54,65	55,45	53,59	53,32	53,39	50,20	51,20	50,27	51,60	55,01		98,23	93,18	92,77	73,94
	EV-J CHN-BJ-8J MN427527	53,86	55,19	55,05	54,26	55,05	53,59	53,99	53,79	50,60	51,60	50,53	51,47	54,47	98,23		92,22	91,81	73,67
	EV-J CHN-BJ-10J MN865120	54,52	56,25	57,05	54,26	55,98	54,26	54,52	53,26	50,20	50,00	50,13	52,01	54,21	93,18	92,22		99,32	75,17
	EV-J CHN-BJ-9J MN427528	54,52	56,52	57,05	54,39	55,72	54,26	54,65	53,39	50,33	50,00	50,40	51,87	54,34	92,77	91,81	99,32		75,17
N2*	EV-J CHN-BJ-2J MN865119	53,19	54,65	54,65	53,32	53,46	54,26	52,79	53,52	49,67	50,93	53,48	53,88	54,07	73,94	73,67	75,17	75,17	

B

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
J103	EV-J103 cg5275 KT961652		95,95	93,52	89,07	88,66	85,43	84,21	83,33	73,28	70,00	61,90	61,51	54,84	49,40	49,40	49,80	49,80	49,80
	EV-J103 cg3697 KT961657	95,95		95,14	87,85	88,66	87,04	86,23	83,73	74,09	70,80	63,10	62,70	57,82	50,60	50,60	51,39	51,39	51,00
	EV-J103 cg8227 KT961656	93,52	95,14		90,28	90,28	87,85	85,83	83,73	73,68	68,40	63,49	63,10	53,63	51,00	51,00	52,19	52,19	51,39
	EV-J103 NOLA1 MZ312496	89,07	87,85	90,28		93,12	87,04	82,59	82,54	73,28	67,60	63,10	62,30	52,42	50,20	50,20	50,60	50,60	49,80
	EV-J103 POo1 FJ007373	88,66	88,66	90,28	93,12		86,23	82,52	81,27	74,39	68,27	63,35	62,15	53,23	49,00	49,00	49,40	49,00	48,61
	EV-J103 NOLA3 MZ312494	85,43	87,04	87,85	87,04	86,23		83,81	82,14	73,28	68,00	61,90	61,90	52,02	51,00	51,00	50,60	50,60	49,40
J123*	EV-J123 Sev-nJ KT581857	84,21	86,23	85,83	82,59	82,52	83,81		90,84	73,17	68,67	62,95	63,35	53,63	49,80	49,80	50,60	50,60	48,61
	EV-J123 CHN-BJ-LXL1 OK397296	83,33	83,73	83,73	82,54	81,27	82,14	90,84		71,71	68,50	63,89	63,49	53,75	49,80	49,80	49,80	49,80	49,02
J124*	EV-J CPOV3336 KJ420743	73,28	74,09	73,68	73,28	74,39	73,28	73,17	71,71		75,81	57,66	58,06	49,60	47,77	47,77	48,18	47,77	48,99
J121	EV-J121 CHE20 KF541639	70,00	70,80	68,40	67,60	68,27	68,00	68,67	68,50	75,81		54,72	55,12	50,20	47,83	47,83	47,43	47,04	46,64
J108	EV-J108 N203 NC013695	61,90	63,10	63,49	63,10	63,35	61,90	62,95	63,89	57,66	54,72		90,08	50,20	45,82	45,82	45,82	45,82	45,42
	EV-J108 NJ25 NC010384	61,51	62,70	63,10	62,30	62,15	61,90	63,35	63,49	58,06	55,12	90,08		48,62	45,42	45,42	45,42	45,42	46,22
J122	EV-J122 SV6 AF326766.2	54,84	52,82	53,63	52,42	53,23	52,02	53,63	53,75	49,60	50,20	50,20	48,62		47,01	47,01	47,01	47,01	47,01
N1*	EV-N1 CHN-BJ-1J MN427526	49,40	50,60	51,00	50,20	49,00	51,00	49,80	49,80	47,77	47,83	45,82	45,42	47,01		100,00	95,90	95,90	81,56
	EV-N1 CHN-BJ-8J MN427527	49,40	50,60	51,00	50,20	49,00	51,00	49,80	49,80	47,77	47,83	45,82	45,42	47,01	100,00		95,90	95,90	81,56
	EV-N1 CHN-BJ-10J MN865120	49,80	51,39	52,19	50,60	49,40	50,60	50,60	49,80	48,18	47,43	45,82	45,42	47,01	95,90	95,90		99,18	82,38
	EV-N1 CHN-BJ-9J MN427528	49,80	51,39	52,19	50,60	49,00	50,60	50,60	49,80	47,77	47,04	45,82	45,42	47,01	95,90	95,90	99,18		82,79
N2*	EV-N2 CHN-BJ-2J MN865119	49,80	51,00	51,39	49,80	48,61	49,40	48,61	49,02	48,99	46,64	45,42	46,22	47,01	81,56	81,56	82,38	82,79	

Supplementary Figure 4. Type characterization of enterovirus J strains based on partial VP1 sequences. Nucleotide **(A)** and amino acid **(B)** identity matrices of partial VP1 sequences are shown. Translation of nucleotide sequences, multiple sequence alignments, and identity matrices were performed using CLC Main Workbench 24. Putative novel type assignments are indicated by an asterisk (*).

A

ORF

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
J121	EV-J121 CHE20 KF541639	1		78,44	83,67	82,70	83,03	82,34	82,34	83,71	80,43	80,57	71,82	71,53	71,53	70,85	71,53
J122	EV-J122 SV6 AF326766.2	2	78,44		83,84	83,24	83,84	83,24	83,15	84,48	75,26	74,93	71,93	72,25	72,30	71,84	72,34
J103	EV-J103 cg3697 KT961657	3	83,67	83,84		94,59	95,23	94,13	91,42	92,85	79,62	79,39	72,65	72,51	72,51	72,28	72,74
	EV-J103 POo1 FJ007373	4	82,70	83,24	94,59		96,19	93,21	90,09	91,70	78,88	79,07	71,83	71,83	71,78	71,37	71,87
	EV-J103 NOLA1 MZ312496	5	83,03	83,84	95,23	96,19		94,36	91,15	92,85	79,30	79,48	72,28	72,15	72,05	71,74	72,28
	EV-J103 NOLA3 MZ312494	6	82,34	83,24	94,13	93,21	94,36		90,27	91,38	78,92	78,66	72,15	72,24	72,05	71,51	72,05
J123*	EV-J123 Sev-nJ KT581857	7	82,34	83,15	91,42	90,09	91,15	90,27		95,69	79,06	78,39	71,55	71,83	71,69	71,14	71,74
	EV-J123 CHN-BJ-LXL1 OK397296	8	83,71	84,48	92,85	91,70	92,85	91,38	95,69		79,84	79,43	72,01	72,28	72,15	71,51	72,15
J108	EV-J108 N125 NC010384	9	80,43	75,26	79,62	78,88	79,30	78,92	79,06	79,84		92,73	70,85	70,65	70,52	69,79	70,70
	EV-J108 N203 NC013695	10	80,57	74,93	79,39	79,07	79,48	78,66	78,39	79,43	92,73		70,32	70,14	70,18	69,41	70,05
N2*	EV-N2 CHN-BJ-2J MN865119	11	71,82	71,93	72,65	71,83	72,28	72,15	71,55	72,01	70,85	70,32		92,91	92,72	92,31	93,23
N1*	EV-N1 CHN-BJ-1J MN427526	12	71,53	72,25	72,51	71,83	72,15	72,24	71,83	72,28	70,65	70,14	92,91		99,40	96,45	97,74
	EV-N1 CHN-BJ-8J MN427527	13	71,53	72,30	72,51	71,78	72,05	72,05	71,69	72,15	70,52	70,18	92,72	99,40		96,17	97,46
	EV-N1 CHN-BJ-10J MN865120	14	70,85	71,84	72,28	71,37	71,74	71,51	71,14	71,51	69,79	69,41	92,31	96,45	96,17		98,43
	EV-N1 CHN-BJ-9J MN427528	15	71,53	72,34	72,74	71,87	72,28	72,05	71,74	72,15	70,70	70,05	93,23	97,74	97,46	98,43	

B

2C+3CD

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
J121	EV-J121 CHE20 KF541639	1		88,60	90,04	90,45	90,14	90,14	90,97	91,17	89,85	89,74	84,50	84,19	84,19	84,39	82,85
J123*	EV-J123 Sev-nJ KT581857	2	88,60		93,12	94,15	94,35	94,25	94,46	96,41	84,92	84,62	82,34	82,24	82,24	82,34	81,01
J103	EV-J103 POo1 FJ007373	3	90,04	93,12		96,92	97,43	97,33	94,97	95,28	86,46	86,36	82,65	82,55	82,55	82,65	81,52
	EV-J103 cg3697 KT961657	4	90,45	94,15	96,92		97,95	97,95	96,00	96,30	86,56	86,46	83,37	83,26	83,26	83,37	82,34
	EV-J103 NOLA3 MZ312494	5	90,14	94,35	97,43	97,95		99,28	96,71	96,51	86,67	86,77	83,26	83,16	83,16	83,26	82,03
	EV-J103 NOLA1 MZ312496	6	90,14	94,25	97,33	97,95	99,28		96,51	96,41	86,56	86,67	83,47	83,26	83,26	83,37	82,14
J122	EV-J122 SV6 AF326766.2	7	90,97	94,46	94,97	96,00	96,71	96,51		96,92	86,97	86,46	82,96	82,85	82,85	82,96	81,83
J123*	EV-J123 CHN-BJ-LXL1 OK397296	8	91,17	96,41	95,28	96,30	96,51	96,41	96,92		86,67	86,26	83,16	83,16	83,06	83,16	81,72
J108	EV-J108 N125 NC010384	9	89,85	84,92	86,46	86,56	86,67	86,56	86,97	86,67		94,97	83,18	83,28	83,28	83,49	81,74
	EV-J108 N203 NC013695	10	89,74	84,62	86,36	86,46	86,77	86,67	86,46	86,26	94,97		82,15	81,85	81,85	82,05	80,72
N2*	EV-N2 CHN-BJ-2J MN865119	11	84,50	82,34	82,65	83,37	83,26	83,47	82,96	83,16	83,18	82,15		98,56	98,56	98,77	96,82
N1*	EV-N1 CHN-BJ-1J MN427526	12	84,19	82,24	82,55	83,26	83,16	83,26	82,85	83,16	83,28	81,85	98,56		99,79	99,38	96,71
	EV-N1 CHN-BJ-8J MN427527	13	84,19	82,24	82,55	83,26	83,16	83,26	82,85	83,06	83,28	81,85	98,56	99,79		99,38	96,71
	EV-N1 CHN-BJ-9J MN427528	14	84,39	82,34	82,65	83,37	83,26	83,37	82,96	83,16	83,49	82,05	98,77	99,38	99,38		97,33
	EV-N1 CHN-BJ-10J MN865120	15	82,85	81,01	81,52	82,34	82,03	82,14	81,83	81,72	81,74	80,72	96,82	96,71	96,71	97,33	

C

Capsid

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
J121	EV-J121 CHE20 KF541639	1		74,74	75,06	74,74	73,42	73,92	72,05	68,30	69,66	61,42	55,79	55,97	56,09	55,62	55,74
J123*	EV-J123 Sev-nJ KT581857	2	74,74		95,05	87,97	86,44	87,26	85,38	71,20	71,40	65,65	57,63	58,97	58,86	58,51	58,62
	EV-J123 CHN-BJ-LXL1 OK397296	3	75,06	95,05		88,56	87,60	88,68	85,51	70,85	71,74	65,89	57,63	58,97	58,86	58,39	58,51
J103	EV-J103 cg3697 KT961657	4	74,74	87,97	88,56		90,80	91,39	89,28	70,30	70,96	64,95	58,44	58,74	58,86	58,97	59,09
	EV-J103 POo1 FJ007373	5	73,42	86,44	87,60	90,80		94,46	87,51	69,45	70,58	65,77	57,51	58,39	58,39	58,16	58,28
	EV-J103 NOLA1 MZ312496	6	73,92	87,26	88,68	91,39	94,46		87,40	69,72	70,96	65,06	57,86	58,51	58,39	58,51	58,62
	EV-J103 NOLA3 MZ312494	7	72,05	85,38	85,51	89,28	87,51	87,40		68,99	69,53	63,58	57,97	59,21	58,86	58,39	58,51
J108	EV-J108 N125 NC010384	8	68,30	71,20	70,85	70,30	69,45	69,72	68,99		91,29	59,56	56,29	55,71	55,48	55,36	55,71
	EV-J108 N203 NC013695	9	69,66	71,40	71,74	70,96	70,58	70,96	69,53	91,29		60,21	56,83	56,94	56,94	56,36	56,48
J122	EV-J122 SV6 AF326766.2	10	61,42	65,65	65,89	64,95	65,77	65,06	63,58	59,56	60,21		57,33	59,14	59,25	59,14	59,25
N2*	EV-N2 CHN-BJ-2J MN865119	11	55,79	57,63	57,63	58,44	57,51	57,86	57,97	56,29	56,83	57,33		84,95	84,59	85,07	85,30
N1*	EV-N1 CHN-BJ-1J MN427526	12	55,97	58,97	58,97	58,74	58,39	58,51	59,21	55,71	56,94	59,14	84,95		99,04	95,33	95,57
	EV-N1 CHN-BJ-8J MN427527	13	56,09	58,86	58,86	58,86	58,39	58,39	58,86	55,48	56,94	59,25	84,59	99,04		94,73	94,97
	EV-N1 CHN-BJ-10J MN865120	14	55,62	58,51	58,39	58,97	58,16	58,51	58,39	55,36	56,36	59,14	85,07	95,33	94,73		99,28
	EV-N1 CHN-BJ-9J MN427528	15	55,74	58,62	58,51	59,09	58,28	58,62	58,51	55,71	56,48	59,25	85,30	95,57	94,97	99,28	

Supplementary Figure 5. Species characterization of enterovirus J strains. Amino acid identity matrices of **(A)** the complete ORF, **(B)** the 2C+3CD and **(C)** the capsid sequences are shown. Translation of nucleotide sequences, multiple sequence alignments, and identity matrices were performed using CLC Main Workbench v24. Putative novel type assignments are indicated by an asterisk (*).

A

ORF

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
EV-M3 CPML3961 KJ420748 1		95,02	94,06	89,69	88,50	86,67	77,78	73,85	71,13	69,44	76,19	72,11	71,49	64,97
EV-M1 RCHM6 2	95,02		94,56	89,78	88,36	86,94	78,14	74,16	71,40	70,30	76,46	72,60	71,94	65,20
EV-M2 RCHM3 3	94,06	94,56		90,69	89,18	86,81	78,28	74,12	71,22	69,94	76,73	72,46	71,71	65,25
EV-M4 CPML8109 KJ420749 4	89,69	89,78	90,69		93,17	87,06	78,40	74,23	72,14	70,13	76,80	72,75	71,63	65,20
EV-M5 GR2815 KJ420747 5	88,50	88,36	89,18	93,17		85,69	77,54	73,96	71,05	70,00	76,40	72,60	71,36	64,61
EV-M6 NGR_NHP_1 MN265403 6	86,67	86,94	86,81	87,06	85,69		77,91	74,60	71,03	69,71	76,01	72,47	71,55	63,85
EV-B114 Z057 7	77,78	78,14	78,28	78,40	77,54	77,91		91,48	85,50	81,98	87,92	67,65	67,63	58,76
EV-B114 B165 AF326751 8	73,85	74,16	74,12	74,23	73,96	74,60	91,48		81,20	82,07	87,65	65,52	65,60	57,90
EV-B113 GAB653 KJ701249 9	71,13	71,40	71,22	72,14	71,05	71,03	85,50	81,20		85,24	80,22	61,09	60,81	53,66
EV-B110 K-UT-18-Q3 MW876511* 10	69,44	70,30	69,94	70,13	70,00	69,71	81,98	82,07	85,24		87,21	61,50	61,49	53,48
EV-B112 GAB130 KJ418244 11	76,19	76,46	76,73	76,80	76,40	76,01	87,92	87,65	80,22	87,21		66,74	66,50	57,90
EV-J121 CHE20 12	72,11	72,60	72,46	72,75	72,60	72,47	67,65	65,52	61,09	61,50	66,74		80,64	63,29
EV-J08 N125 NC010384 13	71,49	71,94	71,71	71,63	71,36	71,55	67,63	65,60	60,81	61,49	66,50	80,64		61,42
EV-A122 M19 AF326754 14	64,97	65,20	65,25	65,20	64,61	63,85	58,76	57,90	53,66	53,48	57,90	63,29	61,42	

B

2C+3CD

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
EV-M4 CPML8109 KJ420749 1		99,69	99,59	99,28	98,77	98,46	98,56	97,85	94,15	84,51	81,95	83,50	75,21	65,20
EV-M5 GR2815 KJ420747 2	99,69		99,38	99,18	98,87	98,26	98,36	97,64	93,95	84,31	81,95	83,30	75,02	65,20
EV-M2 RCHM3 3	99,59	99,38		98,97	98,46	98,26	98,36	97,64	94,05	84,31	81,74	83,30	75,02	65,40
EV-M1 RCHM6 4	99,28	99,18	98,97		98,56	98,05	98,26	97,54	94,15	84,41	81,64	83,09	75,12	65,10
EV-M3 CPML3961 KJ420748 5	98,77	98,87	98,46	98,56		97,54	98,15	97,03	93,54	83,79	81,85	82,89	74,75	64,99
EV-B114 Z057 6	98,46	98,26	98,26	98,05	97,54		97,54	97,33	93,54	84,31	81,74	83,20	75,02	64,99
EV-B112 GAB130 KJ418244 7	98,56	98,36	98,36	98,26	98,15	97,54		97,64	93,74	84,38	81,91	82,99	75,02	64,69
EV-M6 NGR_NHP_1 MN265403 8	97,85	97,64	97,64	97,54	97,03	97,33	97,64		94,15	83,28	80,92	83,61	75,58	65,40
EV-B114 B165 AF326751 9	94,15	93,95	94,05	94,15	93,54	93,54	93,74	94,15		80,41	77,44	80,94	73,27	65,40
EV-B110 K-UT-18-Q3 MW876511* 10	84,51	84,31	84,31	84,41	83,79	84,31	84,38	83,28	80,41		94,18	71,00	63,96	55,27
EV-B113 GAB653 KJ701249 11	81,95	81,95	81,74	81,64	81,85	81,74	81,91	80,92	77,44	94,18		67,42	61,01	53,22
EV-J121 CHE20 12	83,50	83,30	83,30	83,09	82,89	83,20	82,99	83,61	80,94	71,00	67,42		80,81	65,51
EV-J08 N125 NC010384 13	75,21	75,02	75,02	75,12	74,75	75,02	75,02	75,58	73,27	63,96	61,01	80,81		57,42
EV-A122 M19 AF326754 14	65,20	65,20	65,40	65,10	64,99	64,99	64,69	65,40	65,40	55,27	53,22	65,51	57,42	

C

Capsid

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
EV-M3 CPML3961 KJ420748 1		89,82	87,49	75,93	74,18	71,84	63,51	57,21	56,40	47,56	47,79	48,01	47,73	47,84
EV-M1 RCHM6 2	89,82		87,94	75,35	73,60	71,96	64,51	58,78	57,32	48,07	48,30	49,20	48,92	48,35
EV-M2 RCHM3 3	87,49	87,94		77,34	75,35	71,61	64,05	57,97	56,86	47,84	47,84	48,30	48,58	49,03
EV-M4 CPML8109 KJ420749 4	75,93	75,35	77,34		85,14	71,94	63,93	58,56	56,66	48,63	48,97	50,68	48,80	48,69
EV-M5 GR2815 KJ420747 5	74,18	73,60	75,35	85,14		69,27	62,77	59,21	56,89	48,23	48,57	49,71	49,31	48,05
EV-M6 NGR_NHP_1 MN265403 6	71,84	71,96	71,61	71,94	69,27		60,07	56,89	55,79	49,72	49,72	50,28	49,37	48,92
EV-A122 M19 AF326754 7	63,51	64,51	64,05	63,93	62,77	60,07		59,44	56,78	47,61	47,84	48,41	47,55	47,78
EV-J121 CHE20 8	57,21	58,78	57,97	58,56	59,21	56,89	59,44		69,04	47,43	47,09	47,89	48,28	47,60
EV-J08 N125 NC010384 9	56,40	57,32	56,86	56,66	56,89	55,79	56,78	69,04		47,61	47,49	47,38	47,43	47,21
EV-B114 Z057 10	47,56	48,07	47,84	48,63	48,23	49,72	47,61	47,43	47,61		94,35	85,90	79,55	79,55
EV-B114 B165 AF326751 11	47,79	48,30	47,84	48,97	48,57	49,72	47,84	47,09	47,49	94,35		86,02	80,14	78,85
EV-B113 GAB653 KJ701249 12	48,01	49,20	48,30	50,68	49,71	50,28	48,41	47,89	47,38	85,90	86,02		78,26	77,32
EV-B112 GAB130 KJ418244 13	47,73	48,92	48,58	48,80	49,31	49,37	47,55	48,28	47,43	79,55	80,14	78,26		88,92
EV-B110 K-UT-18-Q3 MW876511* 14	47,84	48,35	49,03	48,69	48,05	48,92	47,78	47,60	47,21	79,55	78,85	77,32	88,92	

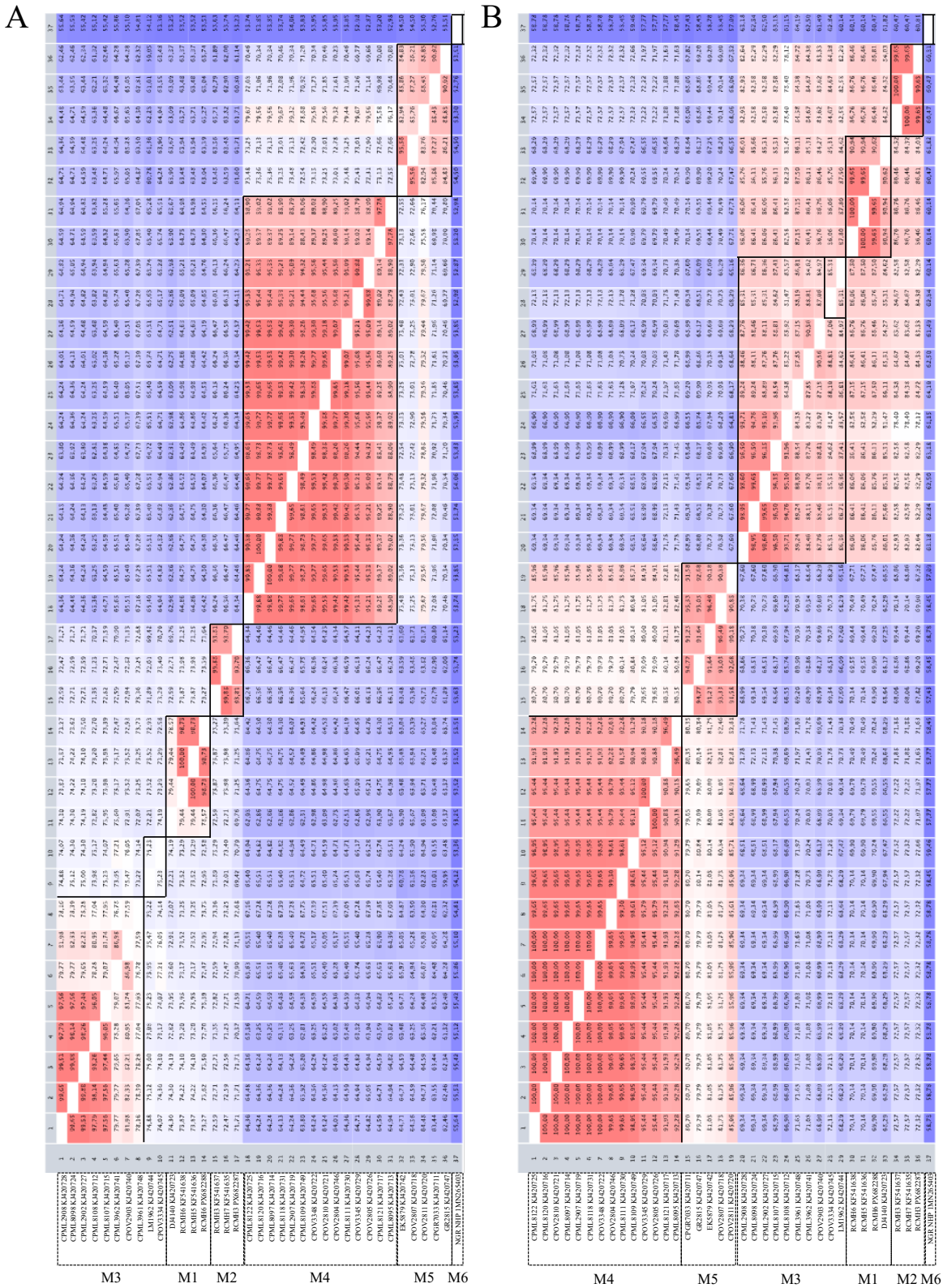
Supplementary Figure 6. Species characterization of enterovirus M strains. Amino acid identity matrices of **(A)** the complete ORF, **(B)** the 2C+3CD and **(C)** the capsid sequences are shown. Translation of nucleotide sequences, multiple sequence alignments, and identity matrices were performed using CLC Main Workbench v24. The asterisk (*) indicates the sequence where the last 399 nt are missing in the 3D gene.

		1	2	3	4	5	6
EV-M2 RCMH3	1		72.59	67.59	68.4	65.63	59.01
EV-M1 RCMH6	2	86.81		64.59	64.13	64.13	57.77
EV-M3 CPML3961 KJ420748	3	84.72	87.5		74.45	73.36	58.22
EV-M4 CPML8109 KJ420749	4	72.66	70.24	71.97		73.82	62.5
EV-M5 GR2815 KJ420747	5	69.2	69.9	69.9	80.84		60.02
EV-M6 NGR_NHP_1 MN265403	6	60.81	60.47	64.19	59.46	58.45	

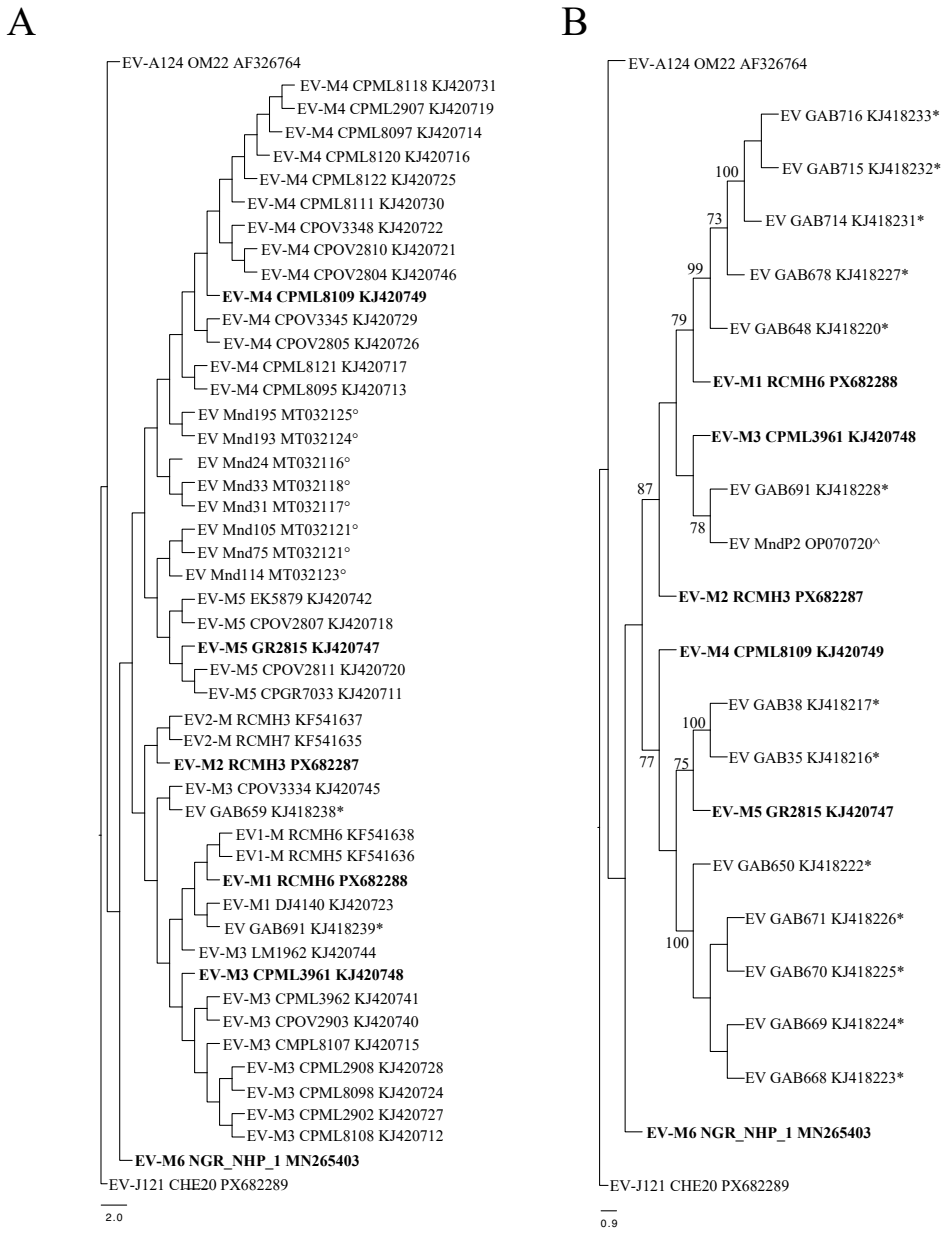
Percentage of nucleotide identity

Percentage of amino acid identity

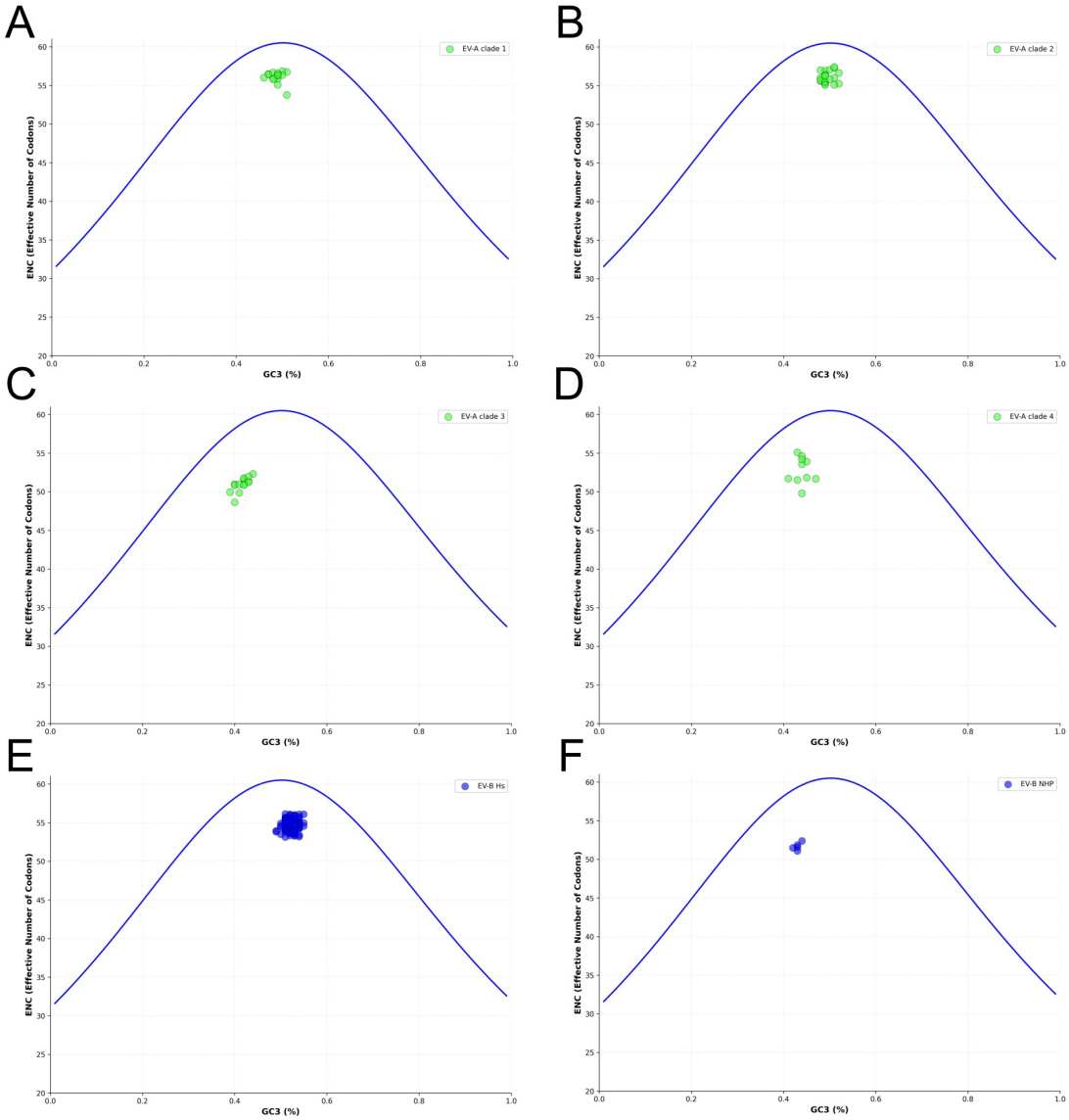
Supplementary Figure 7. Type characterization of enterovirus M strains based on complete VP1 sequences. Nucleotide and amino acid identity matrices of EV-M VP1 sequences derived from complete genomes were calculated using CLC Main Workbench v24.



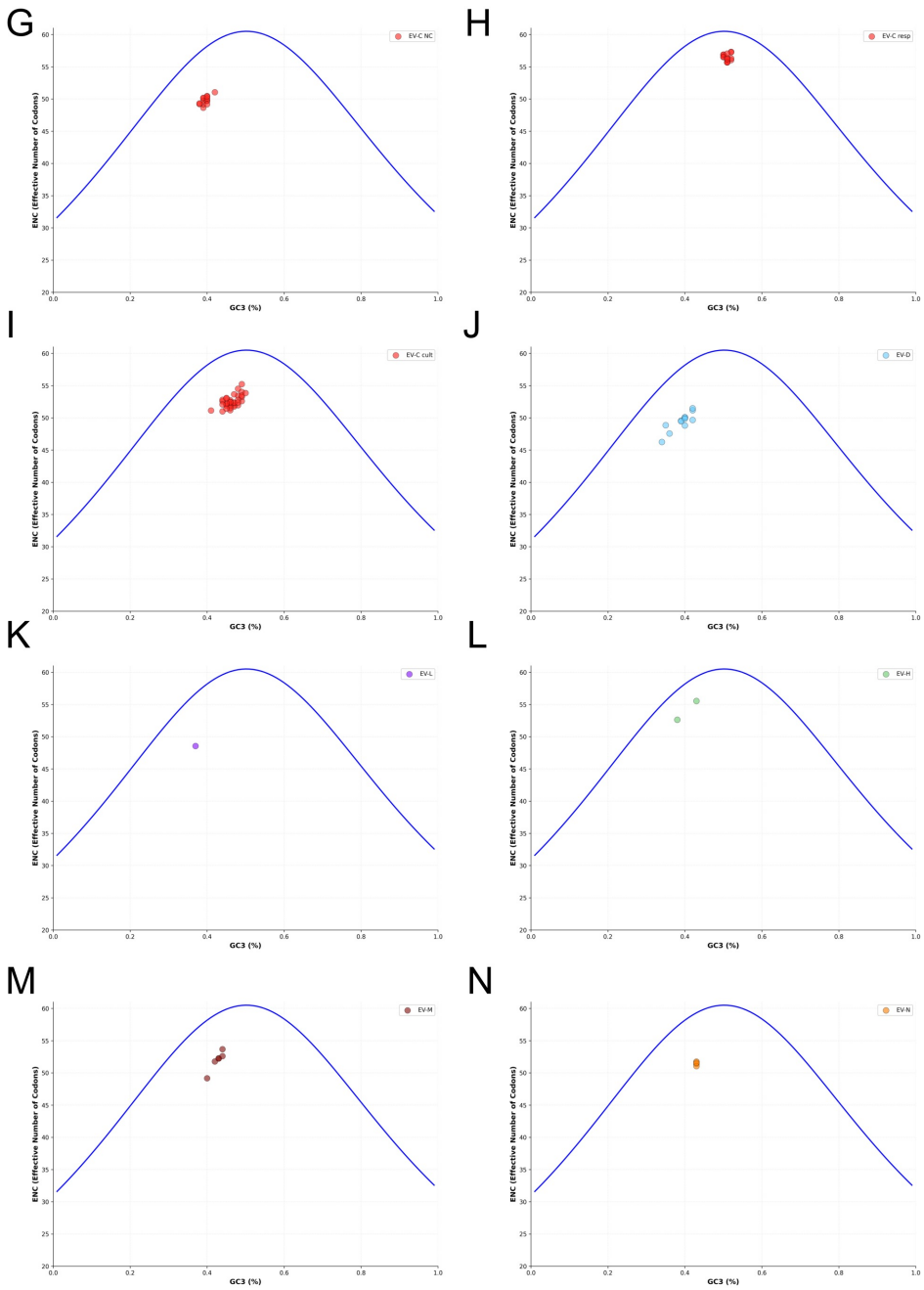
Supplementary Figure 8. Type characterization of enterovirus M strains based on all available VP1 sequences. Nucleotide **(A)** and amino acid **(B)** identity matrices of VP1 sequences are shown. Translation of nucleotide sequences, multiple sequence alignments, and identity matrices were performed using CLC Main Workbench v24.



Supplementary Figure 9. Phylogenetic analysis of enterovirus M strains based on partial VP1 and VP2 sequences. Maximum likelihood phylogenetic trees based on **(A)** VP1 (average length: 354 nucleotides) and **(B)** VP2 (average length: 320 nucleotides) partial sequences are shown. Sequence alignments were performed using CLC Main Workbench v24 and phylogenetic trees were constructed with MEGA X (version 10.2.6). The trees are drawn to scale, with branch lengths measured in the number of substitutions per site. Complete genomes are in bold and symbols indicate references of published sequences: ^, Mombo et al. 2023; *, Mombo et al. 2017; ° , Mombo et al. 2021.



Supplementary Figure 10. ENC–GC3 plots to determine bias mediated by GC3 position. The blue curve indicates the expected ENC values if GC3 composition alone drives codon usage bias. **(A)** EV-A clade 1 species. **(B)** EV-A clade 2 species. **(C)** EV-A clade 3 species. **(D)** EV-A clade 4 species. **(E)** Human EV-B. **(F)** NHP EV-B.



Supplementary Figure 10. ENC–GC3 plots to determine bias mediated by GC3 position. The blue curve indicates the expected ENC values if GC3 composition alone drives codon usage bias. **(G)** Non-cultivable EV-C. **(H)** Respiratory EV-C. **(I)** Cultivable EV-C. **(J)** EV-D. **(K)** EV-L. **(L)** EV-H. **(M)** EV-M. **(N)** EV-N. Resp, Respiratory; NC, Non-cultivable; Cult: cultivable.