1 **VIROLOGY DIVISION NEWS**

2

3 **A taxonomy update for the family *Polyomaviridae***

4

5 **Authors:**

6 **Polyomaviridae Study Group of the International Committee on Taxonomy of Viruses**

7 **Sébastien Calvignac-Spencer[1], Mariet C. W. Feltkamp[2], Matthew D. Daugherty[3], Ugo**

8 **Moens[4], Torbjörn Ramqvist[5], Reimar Johne[6], Bernhard Ehlers[7]**

9

10 [1] Epidemiology of highly pathogenic microorganisms, Robert Koch Institute, 13353, Berlin,

11 Germany; Viral evolution, Robert Koch Institute, 13353, Berlin, Germany, Electronic address:

12 calvignacs@rki.de

13 [2] Department of Medical Microbiology, Leiden University Medical Center, Leiden, The

14 Netherlands. Electronic address: M.C.W.Feltkamp@lumc.nl

15 [3] Division of Basic Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA 98109,

16 USA; Howard Hughes Medical Institute, Fred Hutchinson Cancer Research Center, Seattle,

17 WA 98109, USA. Electronic address: mdaugher@fhcrc.org

18 [4] Faculty of Health Sciences, Department of Medical Biology, University of Tromsø, Norway.

19 Electronic address: ugo.moens@uit.no

20 [5] Department of Oncology-Pathology, Karolinska Institutet, Stockholm, Sweden. Electronic

21 address: Torbjorn.Ramqvist@ki.se

22 [6] Unit Food Hygiene and Virology, Federal Institute for Risk Assessment, Berlin, Germany.

23 Electronic address: Reimar.Johne@bfr.bund.de

24 [7] Division 12 "Measles, mumps, rubella and viruses affecting immunocompromised

25 patients", Robert Koch Institute, 13353, Berlin, Germany. Electronic address: ehlersb@rki.de

26

27 Corresponding author:

28 Bernhard Ehlers

29 Robert Koch-Institut, Division 12

30 Seestr. 10, D-13353 Berlin, Germany

31 Phone 030 18754 2347, fax 0039 30 18754 2598

32 ehlersb@rki.de

Abstract

Many distinct polyomaviruses infecting a variety of vertebrate hosts have recently been discovered and their complete genome sequence could often be determined. To accommodate this fast-growing diversity, the International Committee on Taxonomy of Viruses (ICTV)_*Polyomaviridae*_Study Group designed a host- and sequence-based rationale for an updated taxonomy of the family *Polyomaviridae*. Applying this resulted in numerous recommendations of taxonomical revisions, which were accepted by the Executive Committee of the ICTV in December 2015. New criteria for definition and creation of polyomavirus species were set up that were based on the observed distance between large T antigen coding sequences. Four genera (*Alpha-*, *Beta*, *Gamma-* and *Deltapolyomavirus*) were delineated that altogether include 73 species. Species naming was made as systematic as possible – most species names now consist of the binomial name of the host species followed by *polyomavirus* and a number reflecting the order of discoveries. It is hoped that this important update of the family taxonomy will serve as a stable basis for future taxonomical developments.

Keywords

Polyomavirus; *Polyomaviridae*; virus classification; virus nomenclature; virus taxonomy

Disclaimer

The taxonomic changes summarized here have been submitted as official taxonomic proposal to the International Committee on Taxonomy of Viruses (ICTV) (www.ictvonline.org) and are by now accepted, but not yet ratified. These changes therefore may differ from any new taxonomy that is ultimately approved by the ICTV.

2

Introduction

When it was created, the family *Polyomaviridae* only included a handful of polyomavirus species, whose members had all been discovered by the early 1980s [21]. The situation has now changed dramatically: sequences attributed to relatives of these early polyomaviruses have been published at a much accelerated pace [5, 22], and by September 2015, >1200 fully sequenced polyomavirus genomes representing roughly 100 genetically and biologically distinct polyomaviruses had been deposited in public databases. Nearly all of them were made publicly available in the years 2000-2015, and a number of novel polyomaviruses were published while this report was prepared.

This sudden acceleration found its roots in technological improvements that made polyomavirus discovery much easier, even though still a laborious task (reviewed in [5]). Concomitantly, the first demonstration of the oncogenic potential of a Polyomavirus in humans, the Merkel cell PyV [6], considerably rekindled interest in this viral family. With the ever-growing body of data, new questions will emerge that will likely result in maintaining a firm foot on the discovery throttle. In this respect it is striking to observe that even for the few well-sampled non-human mammalian hosts, e.g. chimpanzees, increasing the sample size often results in identifying new polyomaviruses [4, 9, 13, 16, 19]. Cataloguing the diversity of this family will be a work in progress for many years. Ideally, taxonomy should accompany and help this work.

To enable taxonomic classification, pieces of information have to be identified that are frequently available and that we consider suitable to build a stable and consistent taxonomic system upon. For most novel polyomaviruses, their host and their nucleic acid sequence are the only characters within immediate reach; it is reasonable to anticipate that this will be a long-lasting *status quo*. Therefore, designing a host- and sequence-based taxonomy of the family *Polyomaviridae* seemed to be the best way forward. A first step in this direction had been done by the International Committee on Taxonomy of Viruses (ICTV)_*Polyomaviridae*_Study Group (SG) with the suggestion that entities with >19% whole-genome divergence be considered as members of separate species. In addition, the SG had proposed to create 3 genera within the family (*Avi-, Wuki-, and Orthopolyomavirus*) [11]. However, this approach has not been adopted by the ICTV because it did not account for the observation that some polyomaviruses are recombinants, and the phylogenetic analyses underlying the genus definition were based on different genes. In consideration of the

3

89 committee´s criticisms, the SG developed novel host- and sequence-based criteria for

90 species demarcation and genus delineation. In addition, a standardized scheme for species

91 naming was set up. These taxonomical updates were accepted by the Executive Committee

92 of the ICTV in December 2015 and are described in this article.

93

94 **Criteria for definition of polyomavirus species**

95 Briefly, the 5 delineation criteria aim at ensuring that: i) nucleic acid sequence information is

96 public, verified and unambiguously identifies a polyomavirus (C1-C2), ii) a plausible host is

97 known (C3) and iii) the genetic (and possibly biological) divergence qualifies the new entity

98 as a member of a species distinct from members of all species already recognized (C4-C5).

99 Complying with C1 to C4 is enough to justify the creation of a new species; in cases where C1

100 to C3 are fulfilled but C4 is not, a new species may still be validated by applying C5. The 5

101 delineation criteria are set up as follows:

102 C1.   The complete genome sequence is available in public databases and published in a

103   peer-reviewed journal or an edited journal announcing the availability of sequenced

104   genomes.

105   *Note: As the binomial host species name is part of the polyomavirus species name*

106   *(see below), information on the host of the virus and details regarding how the host*

107   *was determined, are required. Such information is usually included in publications,*

108   *but frequently not available in sequence database entries.*

109 C2.   The genome displays an organization typical for polyomaviruses, i.e., a dsDNA

110   genome with an early region and a late region encoding the T antigens and the

111   structural viral proteins on opposite strands, respectively. Both regions are

112   separated by a noncoding control region.

113   *Note: This criterion was set up to exclude recombinant viruses that associate*

114   *polyomavirus-related coding regions with genomic elements from other viruses, e.g.*

115   *bandicoot papillomatosis viruses [1, 23].*

116 C3.   Sufficient information on the natural host is available.

117   *Note: In cases where the host cannot be firmly identified by host morphology,*

118   *molecular methods should be applied, e.g. mitochondrial cytochrome b typing.*

119 C4.   Observed genetic distance to members of the most closely related species is >15%

120   for large T antigen (LTAg) coding sequence.

4

121   *Note: Under this criterion all publicly available genomes of frequently sequenced*

122   *polyomaviruses fall into their respective species (e.g. BKPyV, HPyV6, HPyV7, JCPyV,*

123   *KIPyV, MCPyV, MWPyV, SV40 and WUPyV genomes). The choice of LTAg as a*

124   *delineating marker was made to keep this criterion in line with the genus delineation*

125   *criteria (see below). Observed genetic distances were chosen after having checked*

126   *that they were very similar to patristic distances (data not shown).*

127   C5.   When two polyomaviruses exhibit <15% observed genetic distance, biological

128   properties (e.g. host specificity, disease association, tissue tropism etc.) can justify

129   the creation of a new species.

130   Example 1:   Two polyomaviruses are regularly detected in the same host, but C4 is not

131   fulfilled (i.e. they exhibit less than 15% divergence). Here, both viruses are

132   assigned to the same species (e.g. BKPyV variants; percentage of identity:

133   93%-100%).

134   Example 2:   Two polyomaviruses are regularly and exclusively detected in separate

135   host, but C4 is not fulfilled (i.e., they exhibit less than 15% divergence). In

136   this case, C5 may result in assigning both viruses to separate species, i.e.,

137   C5 overrides C4. This is exemplified by e.g. the 2 polyomaviruses infecting

138   squirrel monkeys of different species (percentage of identity: 89%; Tab. 1).

139   Example 3:   Two polyomaviruses are regularly detected in the same host and C4 is

140   fulfilled: both polyomaviruses are assigned to separate species (e.g. *Pan*

141   *troglodytes polyomavirus 2* and *3*; percentage of identity: 81%).

142

Naming of polyomavirus species

143

144   As novel polyomaviruses are discovered at a very fast pace, the SG recommended the

145   implementation of standardized species naming, thereby avoiding the nonsystematic

146   inclusion of patient acronyms, geographical and biological designations etc. into the species

147   name. It seems clear that polyomaviruses are host-specific. Despite the use of broad-ranging

148   and flexible detection methods, there are no (or very few) reports about any polyomavirus

149   first discovered in an organism and later detected in another. Exceptions may be SV40 and

150   the lymphotropic polyomavirus but the circulation of these monkey viruses in human

151   populations – or the origin thereof – is still a controversial issue [3, 7, 8, 15, 18]. Therefore,

152   the SG decided to include the host species name into the polyomavirus species name. For

5

this purpose, the binomial host species name was preferred to a common host name, as it is unique at the time of polyomavirus species creation. Naming was achieved by a combination of the Latinized host species name and the term "*polyomavirus*", followed by a number. Numbers are consecutive and follow the chronological order of discovery/publication of the according polyomavirus. Example: the virus known in the literature as bovine polyomavirus (BPyV) belongs to the species *Bos taurus polyomavirus 1*.

Only a few exceptions to this naming scheme were accepted. The ability of the budgerigar fledgling disease polyomavirus (BFDPyV) to infect multiple avian hosts [10] was accounted for by re-naming the respective species *Aves polyomavirus 1*. In addition, all species accommodating human polyomaviruses were named *Human polyomavirus* (instead of *Homo sapiens polyomavirus*), followed by a number. Example: the virus known in the literature as the BK polyomavirus (BKV or BKPyV) belongs to the species *Human polyomavirus 1*.

Definition of novel species, renaming or removal of former species

As of 2015-March-30 (cut-off date for preparation of the current taxonomical update), 68 novel polyomavirus species were defined and named, 8 species were renamed and 5 species were removed from the *Polyomaviridae*, since they do not meet the novel species definition criteria. All in all, 76 species were defined, including 13 polyomavirus species with members infecting humans, 10 ape polyomavirus species (7 chimpanzee, 1 gorilla and 2 orangutan polyomavirus species), 13 monkey polyomavirus species, 21 bat polyomavirus species, 4 rodent polyomavirus species, 7 species with members identified from other mammals, 7 avian polyomavirus species, and 1 fish polyomavirus species. They are listed with their host and accession number in Table 1. Members of 61 species displayed > 15 % divergence to the most closely related polyomavirus of another species. Members of 15 species displayed < 15 % divergence (11 % - 14 %) to the most closely related polyomavirus of another species but originated from different host species (Tab. 1).

Additional mammalian and fish polyomaviruses, including polyomaviruses of 5 previously ICTV-recognized species that are now removed from the *Polyomaviridae* (see above), might give rise to additional species within the *Polyomaviridae* in the near future. They are currently excluded from species definition or removed as species from the family, either because their host species was not reported, because their publication happened

6

185 after the cut-off date, or was not validated by peer-reviewing (GenBank accession numbers:

186 NC_025811, NC_007611, KM496324, NC_025800, NC_004763, AB972942, NC_026766,

187 NC_015123, NC_020065, NC_010107, NC_010817, KJ641707, KJ641705, KJ577598,

188 NC_025259, NC_026244, NC_026012, NC_026015, NC_026942, NC_026944, NC_027531,

189 NC_027532).

190

Creation of genera and assignment of polyomavirus species to genera

191

192 The tremendous diversity of polyomaviruses naturally calls for the identification of some

193 hierarchy within the taxonomical structure of the family, e.g. through the definition of

194 intermediate taxa such as genera. Some years ago, the SG took a first step in this direction

195 and proposed to delineate three genera [11]. The suggestion to create the genus

196 *Avipolyomavirus* aimed at accounting for the distinctive biological properties that avian

197 polyomaviruses display when compared to mammalian ones: broad host range and tissue

198 tropism, no oncogenicity but marked pathogenicity, private genomic features [11]. In line

199 with this, phylogenetic analyses consistently supported the reciprocal monophyly of avian

200 and mammalian polyomaviruses. Most mammalian polyomaviruses are only known from

201 their sequences, which prevented a sound examination and comparison of their biological

202 properties. It was however proposed to create two mammalian genera, respectively coined

203 *Orthopolyomavirus* and *Wukipolyomavirus*, whose existence was essentially backed on

204 sequence divergence of the VP1-encoding gene [11]. The addition of new polyomaviruses

205 uncovered that these genera were unlikely to reflect evolutionary lineages [14] and

206 alternative taxonomical arrangements were proposed, e.g. lumping all polyomaviruses into a

207 single genus [20] or delineating additional genera [5]. The SG also re-examined this question,

208 keeping in mind the important constraint that for most novel polyomaviruses, only the host

209 and nucleic acid sequences are available.

210 There is little evidence for pronounced co-divergence of polyomaviruses with their

211 hosts in family-scale phylogenies [20], but when it comes to the very deep nodes they mostly

212 support the separation of polyomaviruses infecting birds and mammals. Although the lack of

213 observed co-divergence may reflect a mere sampling artifact (and be corrected in the

214 future), at the moment there is no real possibility to use hosts as a major factor (or virus

215 trait) to delineate genera.

7

216   The genomic organization of polyomaviruses is very uniform. Although a number of

217   accessory open reading frames have been described, only a single one (ALTO; [2]) could be

218   ascribed as a landmark characterizing a monophyletic group of polyomaviruses. Altogether it

219   therefore seems that genomic organization could also not generally be used as a driving

220   element for genus-level delineation.

221   The unique option left is to use reconstructed evolutionary relationships for the

222   delineation of genera. Although the SG acknowledges that full-genome analyses would in

223   principle be the ideal tool box [12], the recent realization that recombination events in some

224   instances can significantly reshuffle long-diverged genomes called for precaution [14, 20].

225   The SG therefore recommended using a single of the 3 major coding sequences (LTAg, VP1

226   or VP2) for the delineation of genera. To the best of SG´s knowledge, there was no report

227   thus far of meaningful recombination events within these three coding sequences.

228   The SG proposed that evolutionary relationships derived from analyses of the LTAg

229   amino acid sequences be used for this purpose. Our estimate of amino acid rate variation

230   based on relaxed molecular clock models run with BEAST v1.8.2 was lower for LTAg than for

231   VP1 and VP2 (Fig. 1), which facilitates phylogenetic analysis. In addition, more internal

232   branches appeared as relatively well supported with this same fragment, as notably revealed

233   by overlaying posterior sets of trees generated with BEAST v1.8.2 with DensiTree v2.01 (Fig.

234   2).

235   Fig. 3 represents a chronogram derived from an alignment of conserved amino acid blocks

236   (selected with Gblocks v0.1) reconstructed with BEAST v1.8.2 under the best model of amino

237   acid substitution (LG+F+I+G; as determined with ProtTest v3.2), a relaxed clock (lognormal)

238   and a birth-death model of speciation. Branch thickness is proportional to their posterior

239   probability support (thin branches are less supported). A similar topology was supported by

240   an analysis with PhyML v3 using the BEST tree search algorithm. As far as the SG is aware, it

241   comprises sequences representative of most lineages described to date. Members of species

242   were excluded that displayed an observed amino acid distance in LTAg of less than 5% to a

243   member of one of the species included in the tree, as this tree was constructed to facilitate

244   genus delineation.

245   Based on this, the SG recommended the creation of four genera. These include four

246   relatively large radiations of polyomaviruses that altogether collect 73/76 species created by

247   the SG. To name these genera, the SG decided to follow the example of other SGs which had

8

248 to accommodate a lot of species and to create numerous genera, e.g. *Papillomaviridae*.

249 Genus names will therefore be composed of Greek letters followed by "*polyomavirus*", e.g.

250 *Alphapolyomavirus*. Greek letters will be used consecutively, following the order of

251 description of polyomavirus genera.

252   In brief, virus members of the three genera *Alphapolyomavirus*, *Betapolyomavirus* and

253 *Deltapolyomavirus* are known to infect only mammals; their most recent common ancestors

254 (MRCA) approximately emerged in the same time frame as the MRCA of the genus

255 *Gammapolyomavirus*. This genus (formerly named *Avipoloyomavirus*; [11]) gathers all 7

256 polyomavirus species whose virus members are known to infect birds; its type species is

257 *Aves polyomavirus 1* (Fig. 3; Tab. 1).

258   The type species of the genus *Alphapolyomavirus* is *Mus musculus polyomavirus 1*

259 (member: murine polyomavirus; the first polyomavirus discovered). The genus

260 accommodates 36 species whose members infect primates (humans, apes and monkeys),

261 bats, rodents and other mammals (Fig. 3; Tab. 1). The type species of the genus

262 *Betapolyomavirus* is *Macaca mulatta polyomavirus 1* (member: simian virus 40; the first

263 discovered polyomavirus in this genus). Twenty-six species are included that infect primates

264 (humans and monkeys), bats, rodents and other mammals (Fig. 3; Tab. 1). The type species

265 of the genus *Deltapolyomavirus* is *Human polyomavirus 6* (member: human polyomavirus 6;

266 the first discovered polyomavirus in this genus). The genus is currently only populated by 4

267 human polyomavirus species (Fig. 3; Tab. 1).

268   The 3 polyomavirus species not assigned to any genus are *Bos taurus polyomavirus 1*,

269 *Centropristis striata polyomavirus 1* and *Delphinus delphis polyomavirus 1*. The phylogenetic

270 placement of the polyomaviruses populating the species *Bos taurus polyomavirus 1* and

271 *Delphinus delphis polyomavirus 1* came with some ambiguity which prevented their

272 assignment to the new genera (analyses restricted to mammalian polyomaviruses weakly

273 support their sistership, in disagreement with Fig. 3; data not shown). The virus populating

274 the species *Centropristis striata polyomavirus 1* was at the cut-off date the only published

275 PyV infecting fish. Other fish polyomavirus genomes were available in GenBank but not yet

276 peer-reviewed. The decision was made to wait for their validation before a possible

277 incremental update of the taxonomy focused on non-tetrapod polyomaviruses.

278

279 **Polyomaviruses discovered in the future: Species definition and assignment to genera**

9

280 The assignment of a future polyomavirus to a certain genus will rely on its unambiguous

281 phylogenetic placement within the according clade, as demonstrated by sound phylogenetic

282 analyses of LTAg amino acid sequences. All datasets and methods used to generate the

283 phylogenetic trees that served as the basis for the genus delineation are available as

284 Supplementary Files 1 - 7. The SG suggests that authors willing to accompany future

285 polyomavirus discoveries with taxonomical claims check that their methods are mostly in

286 line with the methods and criteria employed here.

287 Of note, a prerequisite for a correct alignment of LTAg amino acid sequences is the

288 proper identification of the LTAg splice donor and acceptor sites. Ideally, this is done

289 experimentally. However, as is the case for most of the currently known polyomaviruses, it

290 can also rely on *in silico* analysis only. This is usually done by search for canonical splice

291 donor and acceptor sites (http://www.umd.be/HSF3/HSF.html; [17]), followed by a selection

292 of those that are well conserved between the virus in question and the most closely related

293 known polyomaviruses. In addition, the observation might help that the introns of the

294 members of genus *Gammapolyomavirus* are shortest (184 nt-205 nt), followed by those of

295 genus *Betapolyomavirus* (262 nt – 400 nt), genus *Deltapolyomavirus* (346 nt – 406 nt), and

296 genus *Alphapolyomavirus* (353 nt – 565 nt). This is a rough guide predicting which length an

297 LTAg intron should have, once preliminary BLAST and phylogenetic analysis have revealed

298 the genus to which the novel virus may belong. Where help is needed in phylogenetic

299 analysis of novel polyomaviruses, for publication purposes or for proposals of new species

300 and genera to the ICTV, the SG offers to provide appropriate assistance.

301

Conclusions

303 A novel rationale for the taxonomy within the family *Polyomaviridae* was developed. It is

304 mainly based on genomic sequences and host species, information that is available for most

305 of the published polyomaviruses. The novel taxonomical criteria allowed for the assignment

306 of the vast majority of polyomaviruses to species and genera. As after closing the

307 polyomavirus list for preparation of the current taxonomical update (2015-March-30)

308 additional mammalian and fish polyomavirus genomes became publicly available, novel

309 polyomavirus taxa, *i.e.* species and, possibly, genera, can already be seen on the horizon.

310 They will serve as a useful touch-stone for this taxonomy´s robustness.

311

Table 1 Polyomavirus species

No.	PyV species[a,b]	Polyomavirus name (abbreviation)[c]	Isolate/strain	Common host name	Genome length (bp)	GenBank accession number	NCBI Reference Sequence Database accession number	% identity to Py... species (species n...
	Genus Alphapolyomavirus							
1	*Acerodon celebensis polyomavirus 1*	Bat polyomavirus 5b (BatPyV5b)	5b-2	Sulawesi flying fox	5040	AB972940		88 (35)
2	*Artibeus planirostris polyomavirus 2*	Bat polyomavirus 3a (BatPyV3a)	A1055	Flat-faced fruit bat	5019	JQ958886		86 (36)
3	*Artibeus planirostris polyomavirus 3*	Bat polyomavirus 4a (BatPyV4a)	R104	Flat-faced fruit bat	5371	JQ958887		74 (6)
4	*Ateles paniscus polyomavirus 1*	*Ateles paniscus* polyomavirus 1 (ApanPyV1)	1960	Spider monkey	5273	JX159987	NC_019853	69 (13)
5	*Cardioderma cor polyomavirus 1*	*Cardioderma* polyomavirus 1 (*Cardioderma* PyV)	KY336	Heart-nosed bat	5372	JX520659	NC_020067	60 (22)
6	*Carollia perspicillata polyomavirus 1*	Bat polyomavirus 4b (BatPyV4b)	C1109	Seba's short-tailed bat	5352	JQ958889		74 (3)
7	*Chlorocebus pygerythrus polyomavirus 1*	Vervet monkey polyomavirus 1 (VmPyV1)	VMS96	Vervet monkey	5157	AB767298	NC_019844	87 (31)
8	*Chlorocebus pygerythrus polyomavirus 3*	Vervet monkey polyomavirus 3 (VmPyV3)	VMS95/VMV97	Vervet monkey	5055	AB767297	NC_025898	86 (30)
9	*Dobsonia moluccensis polyomavirus 1*	Bat polyomavirus 5a (BatPyV5a)	5a	Moluccan naked-backed fruit bat	5075	AB972945	NC_026768	74 (35)
10	*Eidolon helvum polyomavirus 1*	*Eidolon* polyomavirus 1 (*Eidolon* PyV 1)	KY270	Straw-colored fruit bat	5294	JX520660	NC_020068	57 (12)
11	*Gorilla gorilla polyomavirus 1*	*Gorilla gorilla gorilla* polyomavirus 1 (GgorgPyV1)	5766	Western gorilla	5300	HQ385752	NC_025380	87 (25)
12	*Human polyomavirus 5*	Merkel cell polyomavirus (MCPyV)	R17b	Human	5387	HM011556	NC_010277	82 (25)
13	*Human polyomavirus 8*	*Trichodysplasia spinulosa* polyomavirus (TSPyV)	skin	Human	5232	GU989205	NC_014361	82 (33)
14	*Human polyomavirus 9*	Human polyomavirus 9 (HPyV9)	2540	Human	5026	HQ696595	NC_015150	78 (28)
15	*Human polyomavirus 12*	Human polyomavirus 12 (HPyV12)	1403	Human	5033	JX308829	NC_020890	49 (1)
16	*Human polyomavirus 13*	New Jersey polyomavirus (NJPyV)	NJ-PyV-2013	Human	5108	KF954417	NC_024118	82 (23)
17	*Macaca fascicularis polyomavirus 1*	*Macaca fascicularis* polyomavirus 1 (MfasPyV1)	2085	Crab-eating macaque	5087	JX159986	NC_019851	81 (30)
18	*Mesocricetus auratus polyomavirus 1*	Hamster polyomavirus (HaPV)	Berlin-Buch	Syrian hamster	5372	JX036360	NC_001663	47 (9)
19	*Molossus molossus polyomavirus 1*	Bat polyomavirus 3b (BatPyV3b)	B1130	Velvety free-tailed bat	4903	JQ958893		69 (36)
20	*Mus musculus polyomavirus 1*	Mouse polyomavirus (MPyV)	BG	House mouse	5307	AF442959	NC_001515	46 (18)
21	*Otomops martiensseni polyomavirus 1*	Otomops polyomavirus 1 (Otomops PyV 1)	KY156	Martienssen's free-tailed bat	4914	JX520658	NC_020066	69 (19)
22	*Otomops martiensseni polyomavirus 2*	Otomops polyomavirus 2 (Otomops PyV 2)	KY157	Martienssen's free-tailed bat	5176	JX520664	NC_020071	60 (5)
23	*Pan troglodytes polyomavirus 1*	Chimpanzee polyomavirus (ChPyV)	Bob	Common chimpanzee	5086	FR692334	NC_014743	82 (16)
24	*Pan troglodytes polyomavirus 2*	*Pan troglodytes verus* polyomavirus 1a (PtrovPyV1a)	6444	Common chimpanzee	5303	HQ385746	NC_025368	81 (25)
25	*Pan troglodytes polyomavirus 3*	*Pan troglodytes verus* polyomavirus 2a (PtrovPyV2a)	6512	Common chimpanzee	5309	HQ385748	NC_025370	87 (11)

Pan troglodytes polyomavirus 4	*Pan troglodytes verus polyomavirus 3 (PtrovPyV3)*	3161	Common chimpanzee	5333	JX159980	NC_019855	80 (27)
Pan troglodytes polyomavirus 5	*Pan troglodytes verus polyomavirus 4 (PtrovPyV4)*	3147	Common chimpanzee	5349	JX159981	NC_019856	81 (32)
Pan troglodytes polyomavirus 6	*Pan troglodytes verus polyomavirus 5 (PtrovPyV5)*	5743	Common chimpanzee	4994	JX159982	NC_019857	78 (14)
Pan troglodytes polyomavirus 7	*Pan troglodytes schweinfurthii polyomavirus 2 (PtrosPyV2)*	6350	Common chimpanzee	4970	JX159983	NC_019858	67 (14)
Papio cynocephalus polyomavirus 1	*Yellow baboon polyomavirus 1 (YbPyV1)*	BS20	Yellow baboon	5064	AB767294	NC_025894	86 (8)
Piliocolobus rufomitratus polyomavirus 1	*Piliocolobus rufomitratus polyomavirus 1 (PrufPyV1)*	4601	Red colobus	5140	JX159984	NC_019850	87 (7)
Pongo abelii polyomavirus 1	*Sumatran orang-utan polyomavirus (OraPyV-Sum)*	PI	Sumatran orangutan	5358	FN356901		81 (27)
Pongo pygmaeus polyomavirus 1	*Bornean orang-utan polyomavirus 1 (OraPyV-Bo)*	BO	Bornean orangutan	5168	FN356900	NC_013439	82 (13)
Procyon lotor polyomavirus 1	*Raccoon polyomavirus (RacPyV)*	R45	Raccoon	5016	JQ178241	NC_023845	50 (24)
Pteropus vampyrus polyomavirus 1	*Bat polyomavirus 5b (BatPyV5b)*	5b-1	Large flying fox	5047	AB972944	NC_026767	88 (1)
Sturnira lilium polyomavirus 1	*Bat polyomavirus 3a (BatPyV3a)*	B0454	Little yellow-shouldered bat	5058	JQ958888		86 (2)
Genus Betapolyomavirus							
Aceradon celebensis polyomavirus 2	*Bat polyomavirus 6a (BatPyV6a)*	6a	Sulawesi flying fox	5019	AB972941	NC_026762	66 (43)
Artibeus planirostris polyomavirus 1	*Bat polyomavirus 2c (BatPyV2c)*	A504	Flat-faced fruit bat	5187	JQ958890		74 (58)
Cebus albifrons polyomavirus 1	*Cebus albifrons polyomavirus 1 (CalbPyV1)*	2141	White-fronted capuchin	5013	JX159988	NC_019854	73 (61)
Cercopithecus erythrotis polyomavirus 1	*Cercopithecus erythrotis polyomavirus 1 (CeryPyV1)*	4077	Red-eared guenon	5189	JX159985	NC_025892	86 (57)
Chlorocebus pygerythrus polyomavirus 2	*Vervet monkey polyomavirus 2 (VmPyV2)*	VMK96	Vervet monkey	5167	AB767299	NC_025896	87 (57)
Desmodus rotundus polyomavirus 1	*Bat polyomavirus 2a (BatPyV2a)*	AT7	Vampire bat	5201	JQ958892		66 (59)
Dobsonia moluccensis polyomavirus 2	*Bat polyomavirus 6b (BatPyV6b)*	6b	Moluccan naked-backed fruit bat	5039	AB972947	NC_026770	66 (37)
Dobsonia moluccensis polyomavirus 3	*Bat polyomavirus 6c (BatPyV6c)*	6c	Moluccan naked-backed fruit bat	5046	AB972946	NC_026769	63 (37)
Equus caballus polyomavirus 1	*Equine polyomavirus (EPyV)*	CU03	Horse	4987	JQ412134	NC_017982	59 (57)
Human polyomavirus 1	*BK polyomavirus (BK virus; BKV; BKPyV)*	Dunlop	Human	5153	V01108	NC_001538	82 (41)
Human polyomavirus 2	*JC polyomavirus (JC virus; JCV; JCPyV)*	Mad1	Human	5130	J02226	NC_001699	76 (46)
Human polyomavirus 3	*KI polyomavirus (KIPyV)*	Stockholm 60	Human	5040	EF127906	NC_009238	70 (49)
Human polyomavirus 4	*WU polyomavirus (WU virus; WUPyV)*	B0	Human	5229	EF444549	NC_009539	70 (48)
Loxodonta africana polyomavirus 1	*African elephant polyomavirus 1 (AelPyV1)*	DK-1/2011	African elephant	5722	KF147833	NC_022519	54 (47)
Macaca mulatta polyomavirus 1	*Simian virus 40 (SV40)*		Rhesus monkey	5243	J02400	NC_001669	70 (40)
Mastomys natalensis polyomavirus 1	*Mastomys polyomavirus (MasPyV)*	NR55	Multimammate mouse	4899	AB588640	NC_025895	62 (56)
Meles meles polyomavirus 1	*Meles meles polyomavirus 1 (MmelPyV1)*	French	European badger	5187	KP644238	NC_026473	64 (62)

Species	Virus (abbreviation)	Host	Strain	Size	Accession	RefSeq	%
Miniopterus africanus polyomavirus 1	Miniopterus polyomavirus (Miniopterus PyV)	African long-fingered bat	KY369	5213	JX520661	NC_020069	57 (51)
Mus musculus polyomavirus 2	Mouse pneumotropic virus (MPtV)	House mouse	Kilham	4754	M55904	NC_001505	58 (52)
Myotis lucifugus polyomavirus 1	Myotis polyomavirus (MyPyV)	Little brown bat	VM2008_14	5081	FJ188392	NC_011310	62 (52)
Papio cynocephalus polyomavirus 2	Yellow baboon polyomavirus 2 (YbPyV2)	Yellow baboon	BS94/BC94	5181	AB767295	NC_025897	86 (41)
Pteronotus davyi polyomavirus 1	Pteronotus polyomavirus (Pteronotus PyV)	Naked-backed bat	GTM203	5136	JX520662	NC_020070	78 (59)
Pteronotus parnelli polyomavirus 1	Bat polyomavirus 2b (BatPyV2b)	Mustached bat	R266	5041	JQ958891		78 (58)
Saimiri boliviensis polyomavirus 1	Squirrel monkey polyomavirus (SquiPyV)	Black-capped squirrel monkey	Squi106	5075	AM748741	NC_009951	89 (61)
Saimiri sciureus polyomavirus 1	Saimiri sciureus polyomavirus 1 (SsciPyV1)	Common squirrel monkey	2033	5067	JX159989		89 (60)
Zalophus californianus polyomavirus 1	California sea lion polyomavirus 1 (SLPyV, CSLPyV)	Sea lion	CSL6994	5112	GQ331138	NC_013796	64 (53)
Genus Gammapolyomavirus							
Anser anser polyomavirus 1	Goose hemorrhagic polyomavirus (GHPV)	Goose	Germany 2001	5256	AY140894	NC_004800	60 (66)
Aves polyomavirus 1	Budgerigar fledgling disease virus (BFDV)	Parrots, passerines		4981	AF241168	NC_004764	53 (68)
Corvus monedula polyomavirus 1	Crow polyomavirus (CPyV)	Eurasian jackdaw		5079	DQ192570	NC_007922	67 (66)
Cracticus torquatus polyomavirus 1	Butcherbird polyomavirus (Butcherbird PyV)	Butcherbird	AWH19840	5084	KF360862	NC_023008	67 (65)
Pygoscelis adeliae polyomavirus 1	Adélie penguin polyomavirus (AdPyV)	Adélie penguin	Crozier_2012	4988	KP033140	NC_026141	51 (63)
Pyrrhula pyrrhula polyomavirus 1	Finch polyomavirus (FPyV)	Eurasian bullfinch		5278	DQ192571	NC_007923	53 (64)
Serinus canaria polyomavirus 1	Canary polyomavirus (CaPyV)	Canary	Ha09	5421	GU345044	NC_017085	51 (68)
Genus Deltapolyomavirus							
Human polyomavirus 6	Human polyomavirus 6 (HPyV6)	Human	607a	4926	HM011560	NC_014406	66 (71)
Human polyomavirus 7	Human polyomavirus 7 (HPyV7)	Human	713a	4952	HM011560	NC_014407	66 (70)
Human polyomavirus 10	MW polyomavirus (MWPyV)	Human	MA095	4927	JQ898291	NC_018102	63 (73)
Human polyomavirus 11	STL polyomavirus (STLPyV)	Human	MA138	4776	JX463183	NC_020106	63 (72)
Not assigned to a genus							
Bos taurus polyomavirus 1	Bovine polyomavirus (BPyV)	Cattle	monkey kidney cell	4697	PLYBCG	NC_001442	49 (61)
Centropristis striata polyomavirus 1	Black sea bass-associated polyomavirus 1 (BassPyV1)	Black sea bass	2835	7369	KP071318	NC_025790	37 (23)
Delphinus delphis polyomavirus 1	Dolphin polyomavirus 1 (DPyV-1)	short-beaked common dolphin	Trachea/2010	5159	KC594077	NC_025899	54 (59)

14

15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

[a] cut-off date 2015-March-30; [b] For each species, only one virus is listed. Especially for human polyomaviruses, genomes are available from numerous strains and isolates in public databases that cannot be listed here; [c] Polyomavirus name and abbreviation, as used in the literature [d] Percentage of LTAg coding sequence identity as compared to the closest related species (in parentheses species number as indicated in column 1); all percentage values above 85 % are underlined.

Fig. 1. Bayesian estimates of the coefficient of variation of the amino acid substitution rate (across lineages) in polyomavirus LTAg, VP1, and VP2.

Fig. 2. Superposition of sets of posterior trees. In all cases, 9000 posterior trees were overlaid using DensiTree v2.01. Fuzziness and branch intersections indicate branch length and topological uncertainty.

Fig. 3. LTAg-derived Bayesian chronogram of the family *Polyomaviridae*. The branches supporting the existence of the four genera whose creation is recommended by the SG are highlighted with a red circle. Branch support is reported above branches (SH-aLRT/posterior probability). Detailed methods are described in Supplementary file 1. Tips display the names of species (black), the vernacular names, followed by accession numbers, for viruses not allocated to a polyomavirus species (grey) or, in the case of viruses other than polyomaviruses comprising an LTAg sequence, abbreviations followed by accession numbers (grey). JEECV: Japanese eel endothelial cells-infecting virus. BCPV: bandicoot papillomatosis carcinomatosis virus type 1 and 2 (BPCV1 and 2). Note: as this tree was constructed to enable genus delineation, members of species were excluded that displayed an observed amino acid distance in LTAg of less than 5% to a member of one of the species included in the tree.

Compliance with Ethical Standards

Author A declares that he has no conflict of interest.

Author B declares that she has no conflict of interest.

Author C declares that he has no conflict of interest.

Author D declares that he has no conflict of interest.

Author E declares that he has no conflict of interest.

Author F declares that he has no conflict of interest.

Author G declares that he has no conflict of interest.

Ethical approval: This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Bennett MD, Woolford L, Stevens H, Van Ranst M, Oldfield T, Slaven M, O'Hara AJ, Warren KS, Nicholls PK (2008) Genomic characterization of a novel virus found in papillomatous lesions from a southern brown bandicoot (*Isoodon obesulus*) in Western Australia. Virology 376:173-182

2. Carter JJ, Daugherty MD, Qi X, Bheda-Malge A, Wipf GC, Robinson K, Roman A, Malik HS, Galloway DA (2013) Identification of an overprinting gene in Merkel cell polyomavirus provides evolutionary insight into the birth of viral genes. Proc Natl Acad Sci U S A 110:12744-12749

3. Delbue S, Tremolada S, Branchetti E, Elia F, Gualco E, Marchioni E, Maserati R, Ferrante P (2008) First identification and molecular characterization of lymphotropic polyomavirus in peripheral blood from patients with leukoencephalopathies. J Clin Microbiol 46:2461-2462

4. Deuzing I, Fagrouch Z, Groenewoud MJ, Niphuis H, Kondova I, Bogers W, Verschoor EJ (2010) Detection and characterization of two chimpanzee polyomavirus genotypes from different subspecies. Virol J 7:347

5. Feltkamp MC, Kazem S, van der Meijden E, Lauber C, Gorbalenya AE (2013) From Stockholm to Malawi: recent developments in studying human polyomaviruses. J Gen Virol 94:482-496

6. Feng H, Shuda M, Chang Y, Moore PS (2008) Clonal integration of a polyomavirus in human Merkel cell carcinoma. Science 319:1096-1100

7. Focosi D, Maggi F, Andreoli E, Lanini L, Ceccherini-Nelli L, Petrini M (2009) Polyomaviruses other than JCV are not detected in progressive multifocal leukoencephalopathy. Journal of Clinical Virology 45:161-162

8. Garcea RL, Imperiale MJ (2003) Simian virus 40 infection of humans. J Virol 77:5039-5045

9. Johne R, Enderlein D, Nieper H, Müller H (2005) Novel polyomavirus detected in the feces of a chimpanzee by nested broad-spectrum PCR. J Virol 79:3883-3887

10. Johne R, Muller H (2007) Polyomaviruses of birds: Etiologic agents of inflammatory diseases in a tumor virus family. Journal of Virology 81:11554-11559

11. Johne R, Buck CB, Allander T, Atwood WJ, Garcea RL, Imperiale MJ, Major EO, Ramqvist T, Norkin LC (2011) Taxonomical developments in the family *Polyomaviridae*. Arch Virol 156:1627-1634

12. Lauber C, Gorbalenya AE (2012) Partitioning the genetic diversity of a virus family: approach and evaluation through a case study of picornaviruses. J Virol 86:3890-3904

13. Leendertz FH, Scuda N, Cameron KN, Kidega T, Zuberbuhler K, Leendertz SA, Couacy-Hymann E, Boesch C, Calvignac S, Ehlers B (2011) African great apes are naturally infected with polyomaviruses closely related to Merkel cell polyomavirus. J Virol 85:916-924

14. Lim ES, Reyes A, Antonio M, Saha D, Ikumapayi UN, Adeyemi M, Stine OC, Skelton R, Brennan DC, Mkakosya RS, Manary MJ, Gordon JI, Wang D (2013) Discovery of STL polyomavirus, a polyomavirus of ancestral recombinant origin that encodes a unique T antigen by alternative splicing. Virology 436:295-303

15. Lopez-Rios F, Illei PB, Rusch V, Ladanyi M (2004) Evidence against a role for SV40 infection in human mesotheliomas and high risk of false-positive PCR results owing to presence of SV40 sequences in common laboratory plasmids. Lancet 364:1157-1166

16. Madinda NF, Robbins MM, Boesch C, Leendertz FH, Ehlers B, Calvignac-Spencer S (2015) Genome Sequence of a Central Chimpanzee-Associated Polyomavirus Related

to BK and JC Polyomaviruses, *Pan troglodytes troglodytes* Polyomavirus 1. Genome Announc 3

17. Patel AA, Steitz JA (2003) Splicing double: insights from the second spliceosome. Nat Rev Mol Cell Biol 4:960-970

18. Scuda N, Hofmann J, Calvignac-Spencer S, Ruprecht K, Liman P, Kuhn J, Hengel H, Ehlers B (2011) A novel human polyomavirus closely related to the african green monkey-derived lymphotropic polyomavirus. J Virol 85:4586-4590

19. Scuda N, Madinda NF, Akoua-Koffi C, Adjogoua EV, Wevers D, Hofmann J, Cameron KN, Leendertz SA, Couacy-Hymann E, Robbins M, Boesch C, Jarvis MA, Moens U, Mugisha L, Calvignac-Spencer S, Leendertz FH, Ehlers B (2013) Novel polyomaviruses of nonhuman primates: genetic and serological predictors for the existence of multiple unknown polyomaviruses within the human population. PLoS Pathog 9:e1003429

20. Tao Y, Shi M, Conrardy C, Kuzmin IV, Recuenco S, Agwanda B, Alvarez DA, Ellison JA, Gilbert AT, Moran D, Niezgoda M, Lindblade KA, Holmes EC, Breiman RF, Rupprecht CE, Tong S (2013) Discovery of diverse polyomaviruses in bats and the evolutionary history of the *Polyomaviridae*. J Gen Virol 94:738-748

21. Van Regenmortel MHV, Fauquet CM, Bishop DHL, Carstens EB, Estes MK, Lemon SM, Maniloff J, Mayo MA, McGeoch DJ, Pringle CR, Wickner RB (2000) Virus taxonomy. Seventh report of the International Committee on Taxonomy of Viruses. Academic Press, San Diego

22. White MK, Gordon J, Khalili K (2013) The rapidly expanding family of human polyomaviruses: recent developments in understanding their life cycle and role in human pathology. PLoS Pathog 9:e1003206

23. Woolford L, Rector A, Van Ranst M, Ducki A, Bennett MD, Nicholls PK, Warren KS, Swan RA, Wilcox GE, O'Hara AJ (2007) A novel virus detected in papillomas and carcinomas of the endangered western barred bandicoot (*Perameles bougainville*) exhibits genomic features of both the *Papillomaviridae* and *Polyomaviridae*. J Virol 81:13280-13290
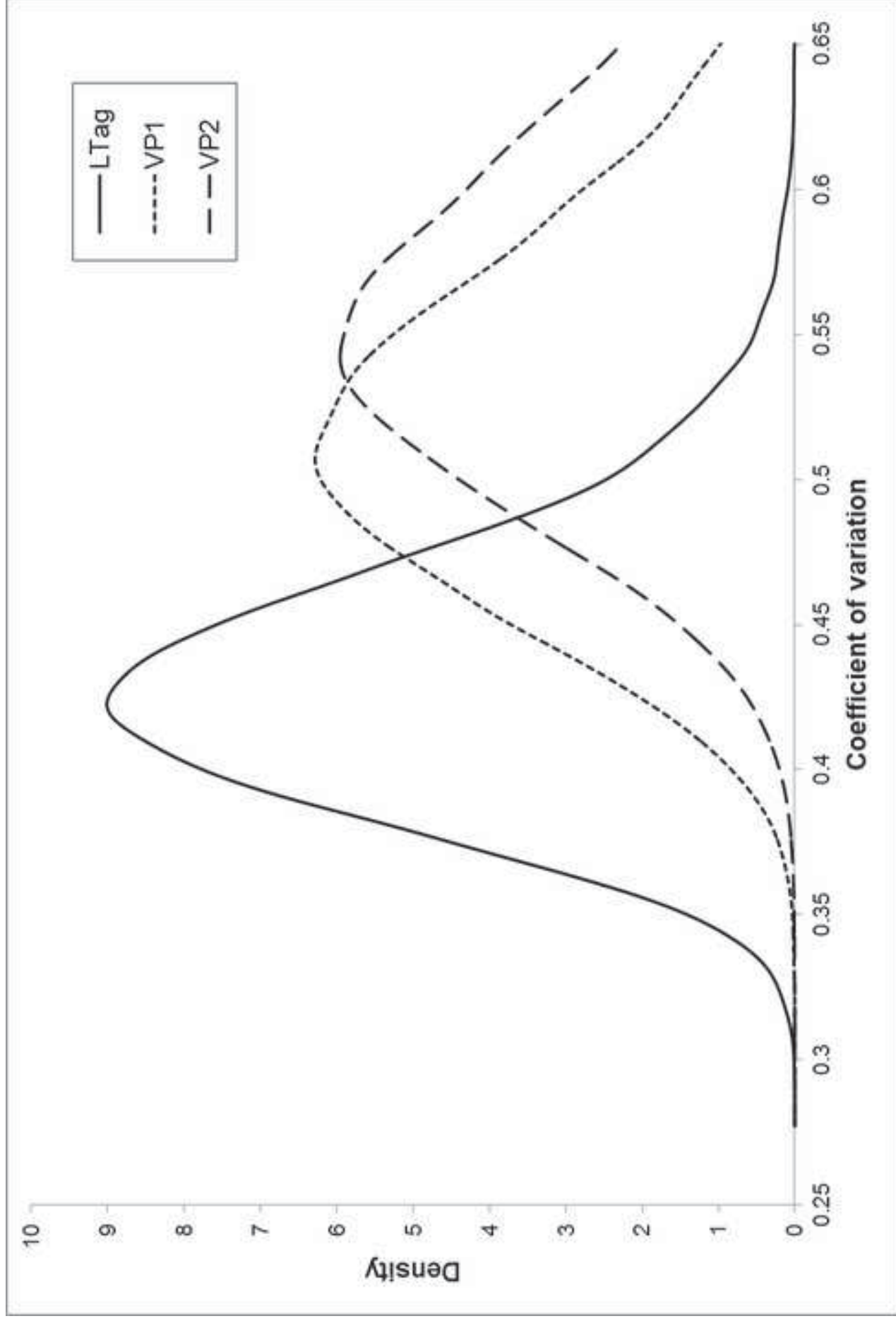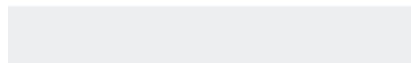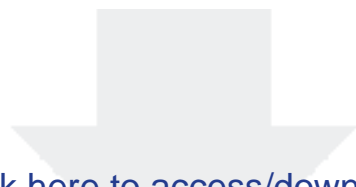
Figure 1

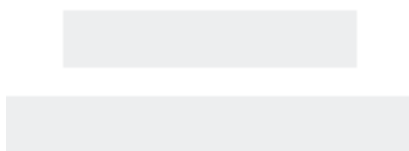Figure 2

Figure 3
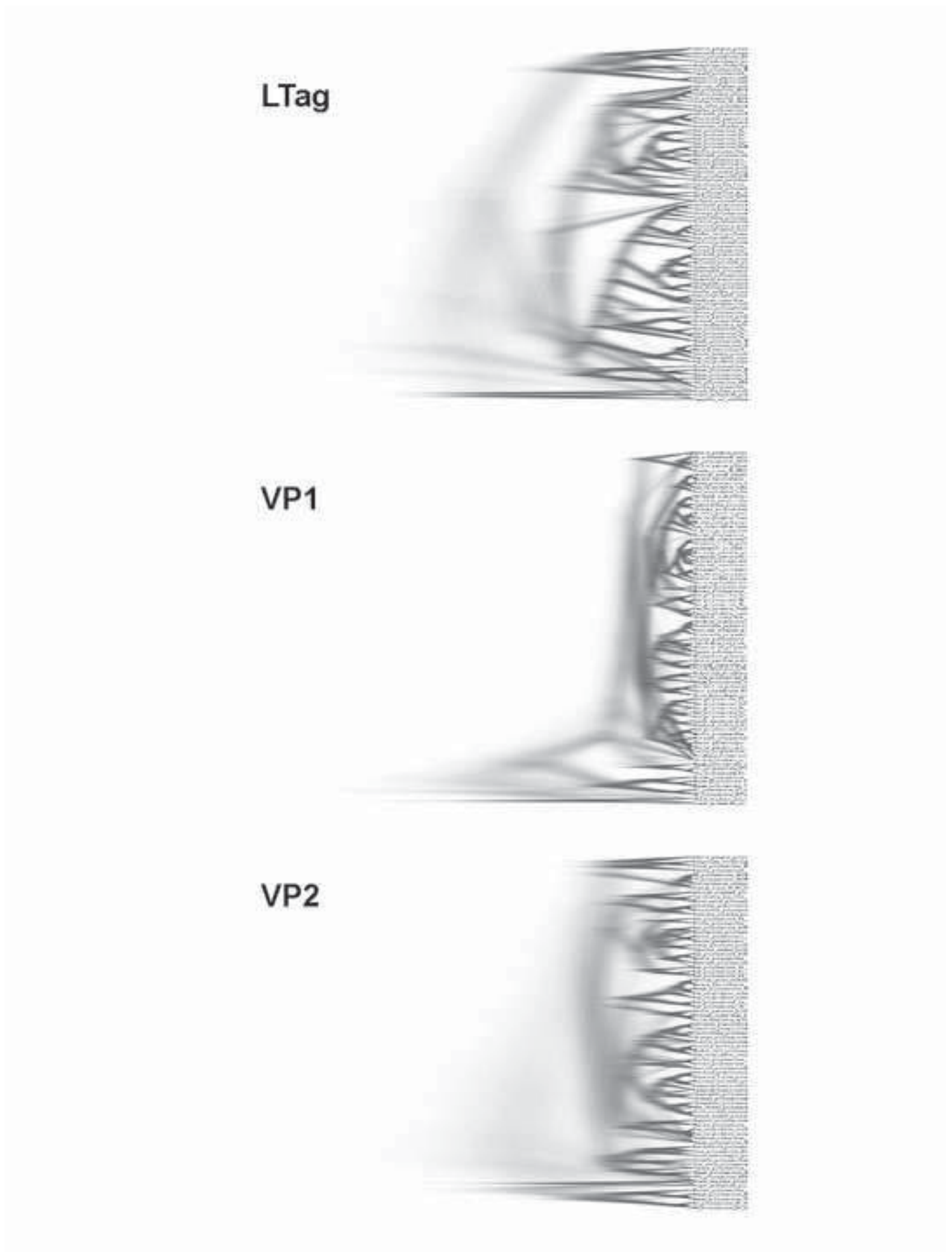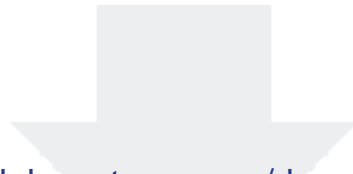
Click here to access/download
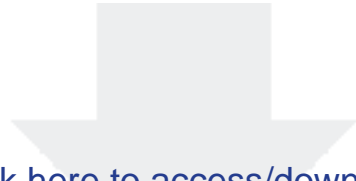**Electronic Supplementary Material**
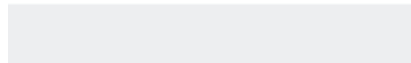S1_Sup_mat_meth_v1-pT.docx

Click here to access/download
Electronic Supplementary Material
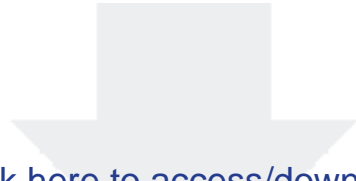S2_LTag-nuc_March-30-2015.txt
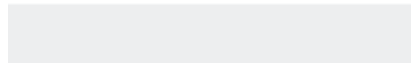
Click here to access/download
Electronic Supplementary Material
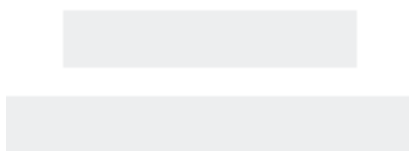S3_LTag-aa_March-30-2015.txt

Click here to access/download
Electronic Supplementary Material
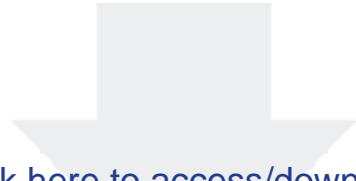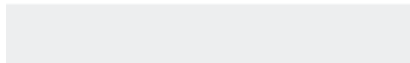S4_VP1-nuc_March-30-2015.txt
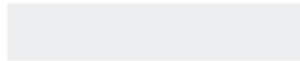
S5

Click here to access/download
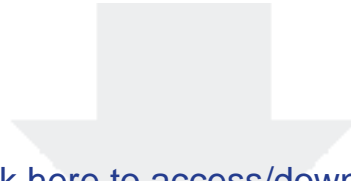**Electronic Supplementary Material**
S5_VP1-aa_March-30-2015.txt

Click here to access/download
Electronic Supplementary Material
S6_VP2-nuc_March-30-2015.txt

Click here to access/download
Electronic Supplementary Material
S7_VP2-aa_March-30-2015.txt