



HAL
open science

Why HPV16? Why, now, HPV42? How the discovery of HPV42 in rare cancers provides an opportunity to challenge our understanding about the transition between health and disease for common members of the healthy microbiota

Ignacio Bravo, Sophia Belkhir, Philippe Paget-Bailly

► To cite this version:

Ignacio Bravo, Sophia Belkhir, Philippe Paget-Bailly. Why HPV16? Why, now, HPV42? How the discovery of HPV42 in rare cancers provides an opportunity to challenge our understanding about the transition between health and disease for common members of the healthy microbiota. *FEMS Microbiology Reviews*, 2024, 10.1093/femsre/fuae029 . hal-04797356

HAL Id: hal-04797356

<https://hal.science/hal-04797356v1>

Submitted on 22 Nov 2024

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

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

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32

Why HPV16? Why, now, HPV42?

**How the discovery of HPV42 in rare cancers provides an opportunity
to challenge our understanding about the transition between health and disease
for common members of the healthy microbiota**

Ignacio G. Bravo, Sophia Belkhir and Philippe Paget-Bailly

Laboratory MIVEGEC (Univ Montpellier, CNRS, IRD)
French National Center for Scientific Research (CNRS), Montpellier, France.

Corresponding author contact information:

Ignacio G. Bravo
National Center for Scientific Research (CNRS)
Laboratory MIVEGEC, Team VIROSTYLE
911 Avenue Agropolis
Montpellier, France
Tlf: +33 467 41 5123
E-mail: ignacio.bravo@cnr.fr

Key words: Papillomavirus; diversity; ecology; evolution; virome; microbiote; infection and cancer; cell type; phenotype; virus-host interactions; genotype-to-phenotype; phenotype-to-environment

33 **Abstract:**

34 In 2022, a bioinformatic, agnostic approach identified HPV42 as causative agent of a rare
35 cancer, later confirmed experimentally. This unexpected association offers an opportunity
36 to reconsider our understanding about papillomavirus infections and cancers.

37 We have expanded our knowledge about the diversity of papillomaviruses and the
38 diseases they cause. Yet we still lack answers to fundamental questions, such as what
39 makes HPV16 different from the closely related HPV31 or HPV33; or why the very
40 divergent HPV13 and HPV32 cause focal epithelial hyperplasia, while HPV6 or HPV42 do
41 not, despite their evolutionary relatedness.

42 Certain members of the healthy skin microbiota are associated to rare clinical conditions.
43 We propose that a focus on cellular phenotypes, most often transient and influenced by
44 intrinsic and extrinsic factors, may help understand the continuum between health and
45 disease. A conceptual switch is required towards an interpretation of biology as a diversity
46 of states connected by transition probabilities, rather than quasi-deterministic programs.
47 Under this perspective, papillomaviruses may only trigger malignant transformation
48 when specific viral genotypes interact with precise cellular states.

49 Drawing on Canguilhem's concepts of normal and pathological, we suggest that
50 understanding the transition between fluid cellular states can illuminate how commensal-
51 like infections transition from benign to malignant.

52

53

54 FROM THE VIRAL THEORY OF CANCER TO THE MOST ONCOGENIC BIOLOGICAL AGENT FOR
55 HUMANS.

56 Research in the first half of the 20th century solidly established that chronic infections by
57 certain viruses could cause tumours, as pioneered by Rous (Rous 1911) and by Fujinami
58 (Fujinami and Inamoto 1914). Indeed, before the paradigm shift towards a genetic
59 component of cancer susceptibility, the viral theory of cancer “*aspired to explain most or all*
60 *types of the disease as being induced by viruses*” (Duran-Reynals, F. 1953). The viral aetiology
61 of human warts had been established through classic filtration and inoculation
62 experiments, moving the authors to propose that “*if a filtrable virus can produce marked*
63 *epithelial hypertrophy, may not this have a bearing on the growth of epithelium in other tumors, for*
64 *example, malignant tumors?*” (Wile and Kingery 1919). The ability of rabbit
65 papillomaviruses to induce warts in wild rabbits and tumours in domestic rabbits was
66 consistently demonstrated later (Shope and Hurst 1933). The consensus in the 1960s
67 classified papillomaviruses of rabbits and humans, polyomaviruses of mice, and
68 vacuolating viruses of monkeys, together into a papova virus group, as they all shared
69 morphological and genomic similarities and all of them “*produce latent and chronic infections*
70 *in their natural hosts, and all are tumorigenic in their natural or other host species, or both*”
71 (Melnick 1962).

72 In this intellectually favourable context, Harald zur Hausen pinpointed that cervical
73 cancer presented an epidemiological pattern similar to that of genital warts and
74 condyloma acuminata, suggesting that the same virus could be responsible for both
75 clinical presentations (zur Hausen 1976). This was at the time a risky hypothesis, because
76 even if a viral aetiology had been proposed for genital warts ((Ciuffo 1907), cited by
77 (Karamanou *et al.* 2010)), no papillomavirus DNA had been detected in genital warts,
78 including attempts by the group of zur Hausen himself (zur Hausen *et al.* 1974a; DeLap,
79 Friedman Kjen and Rush 1976). The origin of the inconsistency was the unexpected
80 diversity among human papillomaviruses, as at the time the “*human papova (wart) virus*”
81 was thought to be an individual, single entity (Rowson and Mahy 1967). However, the
82 fever for viral isolation and characterisation rapidly revealed the large genetic
83 heterogeneity among human papillomaviruses (HPVs) (Gissmann and zur Hausen 1976;
84 Gissmann, Pfister and zur Hausen 1977): i) the genome of HPV1 was characterised in 1977

85 (Orth, Favre and Croissant 1977); DNA from a novel virus, HPV6, was eventually
86 identified from a condyloma acuminata (Gissmann and zur Hausen 1980); and the genome
87 of HPV16, which was not responsible for condyloma acuminata but turned to be the most
88 oncogenic biological agent to humans, was identified (Dürst *et al.* 1983, 1987) and
89 sequenced (Seedorf *et al.* 1985) shortly thereafter. It is now well established that a handful
90 of HPVs are oncogenic for humans (Walboomers *et al.* 1999; Arbyn *et al.* 2014), and that
91 they are globally responsible for around 5% of all new cancer cases every year (de Martel
92 *et al.* 2020). The current diversity estimates, based on targeted as well as on broad
93 metagenomic surveys, suggest that there exist around 800 distinct HPV genotypes (Arroyo
94 Mühr, Eklund and Dillner 2021), extremely diversified, and that do not share a single,
95 exclusive recent common ancestor (Willemsen and Bravo 2019).

96

97 THE TWISTED ROAD BETWEEN GENOTYPE AND PHENOTYPE IN PAPILOMAVIRUSES

98 PVs display a conserved, syntenic core set of early and late ORFs. The *l1* ORF, encoding
99 for the main capsid protein, displays the lowest ratio in amino acid substitutions over
100 nucleotide substitutions (Mengual-Chuliá *et al.* 2016), and pairwise genetic distances
101 between *l1* sequences display a multimodal distribution that can be used to define
102 taxonomic categories (de Villiers *et al.* 2004). The finest-grain category for PV taxonomy
103 recognised by the International Committee for the Taxonomy of Viruses is the viral
104 species, which comprises PV genomic sequences sharing between 71% and 89% nucleotide
105 identity within the complete *l1* ORF
106 (<https://ictv.global/report/chapter/papillomaviridae/papillomaviridae>). However, for
107 the PV research community the clinically relevant taxonomy level is rather the type, a
108 category that comprises viral genomic sequences sharing between 90% and 98% nucleotide
109 identity within the complete *l1* ORF (de Villiers *et al.* 2004; Van Doorslaer 2022). Indeed,
110 types within a same PV viral species can display large biological differences. A first,
111 clinically important example of large quantitative differences among types within a PV
112 species is AlphaPV-species-9, which includes seven HPVs, with a close evolutionary
113 relationship, similar gene composition and with similar tropism (Willemsen and Bravo
114 2019) but that span a large variation in oncogenic potential, the extremes being HPV16, the

115 most oncogenic biological agent to humans, twenty times more prevalent in cervical
116 cancers than in normal cervical samples, and HPV67, closely related to it and recognised
117 as possibly carcinogenic, but with similarly very low prevalence in cervical cancer and in
118 normal cervical samples (HPV information centre 2024). A second cogent example of the
119 large phenotypic diversity for diseases associated to closely related HPVs is AlphaPV-
120 species-10, which includes seven HPVs and two non-human primates PVs. Members of
121 this species comprise HPV6 (Gissmann and zur Hausen 1980) and HPV11 (Dartmann *et al.*
122 1986), sister viruses associated to genital warts and to respiratory recurrent papillomatosis,
123 but also HPV13 (Pfister *et al.* 1983), PpanPV1 infecting bonobos (Van Ranst *et al.* 1991) and
124 PtroPV1 (Scinicariello *et al.* 1997) infecting chimpanzees, the latter three associated to focal
125 epithelial hyperplasia in their respective host species. Finally, a third example of distantly
126 related HPVs associated to a similar clinical presentation is focal epithelial hyperplasia.
127 This proliferative disease displays a particular epidemiology in terms of human genetic
128 background and geographical distribution (Archard, Heck and Stanley 1965), and seems to
129 be associated to two, only distantly related viruses: HPV32, member of AlphaPV-species-1
130 (Beaudenon *et al.* 1987b) and HPV13, member of Alpha-PV-species-10 (Pfister *et al.* 1983).

131 But for HPVs, the differences between evolutionary and genomic relatedness on the one
132 hand and clinical presentation of the infection on the other hand are also noticeable at the
133 shallower taxonomic level of viral variants. For PVs, viral variants comprise viral genome
134 sequences sharing above 98% identity, usually evaluated across the complete genome
135 (Burk, Harari and Chen 2013), although the distribution of polymorphisms is not
136 homogeneous among ORFs (Mengual-Chuliá *et al.* 2016; Mirabello *et al.* 2017). Despite this
137 very close genetic relationship, variants within the same HPV type present differences in
138 geographical distribution and prevalence (Eschle *et al.* 1992; Xi *et al.* 1997; Yamada *et al.*
139 1997; Clifford *et al.* 2019), epidemiology (Zehbe *et al.* 1998; Villa *et al.* 2000; Sichero *et al.*
140 2007; Mirabello *et al.* 2016, 2017) or function (Stoppler *et al.* 1996; Hu *et al.* 2002; Sichero,
141 Franco and Villa 2005; Asadurian *et al.* 2007; Lace *et al.* 2009). A fraction of these biological
142 differences between variants can be accounted for by a *virus genotype*host genotype*
143 interaction, in which the parallel evolution of viruses and hosts would have resulted in
144 differences of variant prevalence and of association with disease (Ong *et al.* 1993; Xi *et al.*
145 2006). Further, episodes of recombination, isolation and transmission between related

146 hosts could have disrupted virus-host parallel evolution and lead to differential cancer
147 risk (Gottschling *et al.* 2011), as proposed for the transmission of specific HPV16 variants
148 from Neandertals to anatomically modern humans (Pimenoff, De Oliveira and Bravo
149 2017). Another fraction of the biological differences between variants is probably related to
150 specific *virus genotype*cell (pheno)type* interactions, including specific internal cues (*e.g.*
151 biochemical status or cell cycle), changes across cell ontogeny (*e.g.* keratinocyte
152 differentiation) and/or external cues (*e.g.* stress signals, tissue homeostasis or paracrine
153 regulation). Overall, even closely related variant genotypes may be expressed differently
154 in different cellular geno/phenotypes, resulting in differences in the natural history and
155 eventually in the clinical presentation of the infection. Such interactions have been
156 extensively documented for the differential prevalence of HPV6 and HPV11 variants in
157 genital warts (Flores-Díaz *et al.* 2017a, 2017b) and in respiratory recurrent papillomatosis
158 (Combrinck *et al.* 2012; Godinez *et al.* 2013; Sichero *et al.* 2021); for the differential
159 prevalence of HPV16 variants in anogenital cancers as a function of the anatomical
160 location (Nicolás-Párraga *et al.* 2016); or for the differential prevalence of HPV16 variants
161 in squamous and in adenocarcinomas (Nicolás-Párraga *et al.* 2017; Clifford *et al.* 2019).

162

163 THE UNEXPECTED DESCRIPTION OF HPV42 AS AN ONCOGENIC VIRUS

164 Historical efforts to identify viruses as etiological drivers of cancers had typically focused
165 on detecting targeted suspects through a guided search of genetic material, as in *e.g.* (zur
166 Hausen *et al.* 1974a, 1974b; Wolf *et al.* 1975). However, high throughput sequencing
167 techniques applied to cancer samples in clinical settings have provided a plethora of off-
168 target, metagenomic sequences that can potentially contain traces of viral genetic material
169 present in cancer tissue. Starting from an agnostic search for viral sequences in cancer
170 cases, the PV and pathology communities have generated in a very short period of time a
171 sound corpus of knowledge that suggests an etiological role of HPV42 in Digital Papillary
172 Adenocarcinoma (DPA), a rare tumour of the sweat glands. This involvement was first
173 proposed by Vanderbilt and coworkers based on a systematic survey of shotgun
174 sequencing data aiming at unveiling unknown virus-cancer associations. These authors
175 searched for the presence of viral DNA in almost 60,000 tumour samples, reporting novel

176 potential associations, especially between human herpes virus 6 and neuroblastoma,
177 between human herpes virus 7 and esophagogastric cancer, and, of especial interest for the
178 HPV community, between HPV42 and DPA, based on four DPA cases (Vanderbilt *et al.*
179 2022b). Later, this same group confirmed the findings with the detection of HPV42 by *in-*
180 *situ* hybridisation in ten DPA cases and its absence in 20 cases of other adnexal tumours
181 (Vanderbilt *et al.* 2022a). Independently, Leindecker and coworkers provided strong
182 evidence about the presence of HPV42 DNA and RNA in 45 out of a large series of 47 DPA
183 cases, as well as about the possible mechanisms underlying cellular transformation
184 (Leindecker *et al.* 2023). Only few months after the first descriptions, a number of
185 confirmation studies have further underpinned the role of HPV42 in DPA, in large scale
186 surveys of DPA cases in acral (Kervarrec *et al.* 2023b) and in non-acral sites (Kervarrec *et al.*
187 2023a), as well as in individual case reports (Cascardo, Cornell and Moesch 2023). The
188 specificity of the HPV42-DPA association is remarkable because: i) the acral anatomical
189 location *per se* is not related to a high prevalence or aggressivity of HPV42, as a
190 communication about digital squamous cell carcinomas reports a notable association with
191 different HPVs, chiefly HPV33, but not with HPV42 (Gormley *et al.* 2011); and ii) HPV42 is
192 not detected in other cancers ontogenically and/or phenotypically related to DPA, such as
193 hidradenomas or hidradenocarcinomas (Vanderbilt *et al.* 2022a; Kervarrec *et al.* 2023b).

194 HPV42 had never been identified as a possible/probable oncogenic agent for humans.
195 HPV42 and HPV32 are the only known members of AlphaPV-species-1. HPV42 was
196 isolated (Beaudenon *et al.* 1987a) and characterised (Philipp *et al.* 1992) from flat vulvar
197 papillomas, while HPV32 was isolated from a focal epithelial hyperplasia of the oral
198 mucosa (Beaudenon *et al.* 1987b). Despite their close evolutionary relationship, these two
199 viruses display divergent epidemiological patterns and are associated with largely
200 divergent clinical presentations and natural histories of the infection.

201 HPV42 is often present in different anatomical locations, without any obvious sign of
202 proliferative lesions. The world prevalence of HPV42 in normal cervical cytology samples
203 amounts to 0.6% [95% CI 0.5-0.6%], compared with 2.8% [95% CI 2.8-2.9%] for HPV16, the
204 most prevalent viral genotype in this kind of samples. Very interestingly, the prevalence of
205 HPV42 in normal cytology samples varies largely among geographical regions, being
206 highest in Africa (1.2% [95% CI 1.0-1.5%], compared with 2.4% [95% CI 2.2-2.6%] for

207 HPV16) and being lowest in Asia (0.3% [95% CI 0.2-0.3%], compared with 2.3% [95% CI
208 2.2-2.4%] for HPV16 (HPV information centre 2024). Substantially higher HPV42
209 prevalence values have been reported in particular populations and anatomical locations,
210 reaching 8% in cervical screening in Ghana (Donkoh *et al.* 2022), 17.6% in cervical self-
211 sampling in Malagasy women (Vassilakos *et al.* 2016), or 15% in anal samples of men
212 having sex with men and living with HIV (Damay *et al.* 2010). Further, HPV42 is the fourth
213 most prevalent HPV found in the semen from infertile men, reaching 8.8% prevalence,
214 with HPV16 being the most frequent type, reaching 22% prevalence (Boeri *et al.* 2019).
215 Finally, HPV42 has been reported to be very often present in oral samples in a Polish
216 population, with a high oro-genital concordance in the same individual, as well as with a
217 good correspondence of positivity status within couples (Kiwerska *et al.* 2019).

218 The HPV42 prevalence in anogenital cancers is low and varies depending on the
219 anatomical location, reaching 0.2% [95% CI 0.2-0.3%] for cervical cancers, 0.2% [95% CI 0.0-
220 1.1%] for anal cancers or 0.1% [95% CI 0.0-0.6%] for penile cancers (HPV information
221 centre 2024). The true contribution of HPV42 to anogenital cancers may actually be lower,
222 after having accounted for the detection of multiple viral genotypes in cancers (Pimenoff *et*
223 *al.* 2019). In a series of over 15,000 HPV-positive anogenital cancers, only three cases (one
224 cervical, one vulvar and one penile lesion, all of them squamous cell carcinomas)
225 contained HPV42 as sole viral genotype and possible etiological agent (Guimera *et al.*
226 2013). Other HPV42-related squamous lesions have been described, such as a case of
227 vulvar intraepithelial neoplasia and intravulvar condyloma acuminata (Kacerovská *et al.*
228 2014), a case of anal dysplasia (Kreuter *et al.* 2018) and a case of squamous cell carcinoma
229 in a non-anogenital location (Da *et al.* 2018). HPV42 has also been sporadically associated
230 to particular proliferative lesions, such as a periungual bowenoid papulosis (Gómez
231 Vázquez and Navarra Amayuelas 2013), but also to seborrheic keratosis-like lesions of the
232 cervix and the vagina (Hennell *et al.* 2012; Talia and McCluggage 2017; Pujari *et al.* 2021),
233 as well as in seborrheic keratosis-like lesions of the cervix with high-grade dysplasia (Talia
234 *et al.* 2022).

235

237 EVOLUTIONARY CORRELATES OF DIFFERENTIAL ONCOGENICITY AMONG HUMAN ALPHAPV

238 The case for the oncogenicity of HPV42 in DPA offers an opportunity to challenge and
239 update our classical view of HPV-related carcinogenesis. An evolutionary analysis
240 suggests that the oncogenic potential for HPVs is not biunivocally associated to any single
241 evolutionary event or genetic trait. AlphaPVs are the sister group to DyoomikronPVs,
242 isolated from cervicovaginal samples from squirrel monkeys (Chen *et al.* 2018). The
243 evolutionary relationships among AlphaPV species differ depending on the region of the
244 genome being considered, so that regarding the early ORFs all possibly/probably
245 oncogenic AlphaHPVs are closely related, while regarding the late ORFs (notably the *l1*
246 ORF, used for taxonomy assignments) they are only distantly related and appear
247 dispersed among non-oncogenic HPVs (Bravo and Alonso 2004; Garcia-Vallve, Alonso
248 and Bravo 2005; Angulo and Carvajal-Rodriguez 2007; Gottschling *et al.* 2011). When using
249 a genomic approach, AlphaPVs have diverged into four evolutionary clades, all of them
250 spanning viruses that infect catarrhines, from hominoids, (*e.g.* humans and chimpanzees)
251 to cercopithecoids (*e.g.* macaques, colobuses or baboons) (Figure 1). There is a rough
252 correspondence between the large, four basal AlphaPV clades and the main clinical
253 manifestations of their infections in humans, roughly skin warts, genital warts and
254 anogenital lesions (Bravo and Alonso 2004; Schiffman *et al.* 2005; Bravo and Felez-Sanchez
255 2015). At this genomic level, the strongest evolutionary signal clusters all
256 possibly/probably oncogenic HPVs into a monophyletic group together with PVs
257 infecting cercopithecoids, all of them often associated to anogenital infections, but not all
258 of them identified as oncogenic in their respective hosts (Kloster *et al.* 1988; Ostrow *et al.*
259 1993; Bravo and Alonso 2004; Chen *et al.* 2009, 2019; Gottschling *et al.* 2011; Bergin *et al.*
260 2013; Long *et al.* 2022).

261 There are few genomic differences among the four AlphaPV clades. The first one is the
262 presence in the genome of AlphaPVs of a nucleotide stretch between the *e2* and the *l2*
263 ORFs, and the nature of the ORFs therein encoded. At the level of the *Papillomaviridae*
264 family, there are four main lineages that present a long DNA segment between the early
265 and the late gene expression cassettes (Garcia-Vallve, Alonso and Bravo 2005), and these

266 four segments are not evolutionary related (Willemsen, Félez-Sánchez and Bravo 2019).
267 Unfortunately, all ORFs encoded in this *inter_e2_l2* region have historically been named as
268 *e5*, even if they do not share a common ancestor, which can generate confusion when
269 trying to identify common mechanisms between the cellular effects of distantly related
270 proteins bearing the same name (e.g. BPV1_E5 and HPV16_E5 (Dimaio and Petti 2013)).
271 Conversely, evolutionary analyses support monophyly for the *inter_e2_l2* region in
272 AlphaPVs at the DNA level, but not at the protein level (Willemsen, Félez-Sánchez and
273 Bravo 2019). This means that a single event resulted in the emergence of this *ca.* 400 bp
274 *inter_e2_l2* DNA fragment in AlphaPVs, but that the different E5 proteins therein encoded
275 probably have an independent origin (Figure 2). Indeed, the phylogeny and the chemistry
276 of the E5 proteins in AlphaHPVs suggests multiple origins and divergent functionality,
277 and match well the clinical presentations of the infection (Bravo and Alonso 2004), with
278 E5_alpha being encoded in the genome of (possibly/probably) oncogenic AlphaHPVs,
279 E5_beta being encoded in AlphaHPVs associated to skin warts (e.g. HPV2 or HPV3), and
280 E5_gamma and E5_delta being encoded in AlphaHPVs associated to genital warts (e.g.
281 HPV6).

282 The individual histories of evolution and conservation of the ORF encoded in the
283 *inter_e2_l2* region of oncogenic AlphaPVs highlight the connection between evolution and
284 oncogenicity (Bravo and Alonso 2004; Van Doorslaer et al. 2017; Willemsen et al. 2019).
285 Members of the AlphaPVs-species-12 are potentially oncogenic for their cercopithecinae
286 host species (Wood *et al.* 2004, 2007; Chen *et al.* 2009; Bergin *et al.* 2013), and encode for two
287 proteins in tandem, E5_epsilon and E5_delta, which are distinct in chemistry and
288 sequence from the E5_alpha encoded by the sister taxa of possibly/probably oncogenic
289 HPVs. Further, AlphaPVs-species-5 and species-6 (respectively represented by HPV26 and
290 HPV30 (Delius and Hofmann 1994)), display a conserved but pseudogenised *e5_alpha* ORF
291 that lacks a starting ATG codon (Figure S1), and precisely these two AlphaPV-species are
292 associated to a lower oncogenic potential than other probably/possibly oncogenic HPVs.
293 We confirm thus that the nature of the E5 protein encoded in the *inter_e2_l2* segment of
294 AlphaPVs displays a strong correspondence with the clinical presentation and with the
295 oncogenic potential of the corresponding viruses (Figure 1). This connection between
296 epidemiology and evolution for E5 proteins (Bravo and Alonso 2004; Schiffman *et al.* 2005)

297 contrasts with the relatively little, indirect evidence of translation of the E5 ORFs, as they
298 are located in the 3'-end of the early viral transcripts (Dimairo and Petti 2013; Trammel *et*
299 *al.* 2024), as well as with the virtual absence of functional studies focusing on E5 proteins
300 other than E5_alpha (Cartin and Alonso 2003).

301 The second main qualitative difference between oncogenic and non-oncogenic AlphaPVs
302 are the transcription patterns of the *e6* and *e7* ORFs. In oncogenic AlphaHPVs the *e6_e7*
303 expression cassette is transcribed from a single transcription start site, and generates a
304 bicistronic mRNA that can be spliced in the *e6* ORF to encode for a shorter E6*I protein
305 (Schneider-Gädicke and Schwarz 1986; Smotkin and Wettstein 1986; Inagaki *et al.* 1988).
306 Even if the precise splice donor and acceptor sites are not strictly conserved among closely
307 related viruses, the fact of producing alternative mRNAs by splicing of the *e6* ORF is a
308 signature shared by all oncogenic AlphaHPVs (Cornelissen *et al.* 1990; Nakagawa *et al.*
309 2000; Sotlar *et al.* 2004; Halec *et al.* 2013) (Figure 1). In contrast, non-oncogenic AlphaHPVs
310 present a second transcription start site located within the *e7* ORF, that allows for the
311 expression of two different mRNA molecules, a bicistronic mRNA that can conceptually
312 encode for E6 and E7, and a second one that encodes only for E7 (Inagaki *et al.* 1988;
313 Smotkin, Prokoph and Wettstein 1989). Further, the *e6* and *e7* ORFs in the bicistronic *e6_e7*
314 mRNAs do not overlap in oncogenic AlphaHPVs while they do overlap in non-oncogenic
315 HPVs, so that in the latter the *e6* stop codon is located downstream the *e7* AUG codon
316 (Auslander *et al.* 2019). The common knowledge in the PV field is that the E7 protein is
317 essentially translated from the *e6*I_e7* mRNA in oncogenic HPVs, and from the *e7* mRNA
318 in non-oncogenic HPVs (Schwartz 2013). However, viruses have evolved a plethora of
319 mechanisms that hijack the eukaryotic ribosome for non-canonical translation (Sorokin *et*
320 *al.* 2021). Translation of both E6 and E7 could be conceived from the *e6_e7* bicistronic
321 mRNA by means of leaky scanning in both oncogenic and non-oncogenic AlphaHPVs
322 (Stacey *et al.* 2000), and/or by means of reinitiation upon termination in oncogenic
323 AlphaHPVs only. The lack of overlap between the *e6* and *e7* ORFs is an evolutionary
324 novelty of the monophyletic group of oncogenic primate AlphaPVs (Tang *et al.* 2006)
325 (Figure 2). It can be thus hypothesised that the equilibrium between E6 and E7 levels and
326 the downstream associated effects could be under (post)transcriptional control, through
327 modulation of splicing activity or under translational control, through modulation of non-

328 canonical translation events. Such differences in the E6:E7 ratio could thus vary as a
329 function of the cell type, cell differentiation stage, cell cycle, or local or systemic immune
330 status, ultimately modulating the probability of the different infection outcomes. Should
331 this be the case, the combination of the ability of generating *e6*I_e7* transcripts together
332 with the lack of overlap between *e6* and *e7* could have provided the clade of
333 possibly/probably oncogenic AlphaPVs with an additional mechanism to regulate the
334 E6:E7 ratio, and thus allow novel ways of exploiting the host cell and eventually lead to
335 particular clinical presentations of the infection.

336 The third main qualitative difference between oncogenic and non-oncogenic AlphaHPVs
337 is the ability of the E6 protein in oncogenic HPVs to induce degradation of the cellular
338 protein p53, as described first for HPV16 and HPV18 (Scheffner *et al.* 1990; Werness,
339 Levine and Howley 1990), *via* an interaction with the LxxLL motif of the cellular E3
340 ubiquitin ligase E6-associated protein (E6AP) (Huibregtse, Scheffner and Howley 1991;
341 Zanier *et al.* 2013; Martinez-Zapien *et al.* 2016). The E6-E6AP interaction is an evolutionary
342 ancient signature shared by members of the Alpha-OmikronPVs, while PVs in other
343 distantly related crown groups such as such as Beta-Xi or Delta-Zeta-PVs display rather an
344 E6-MAML1 interaction (Brimer, Drews and Vande Pol 2017). In its turn, the ability of this
345 E6-E6AP complex to promote p53 degradation is an evolutionary recent signature,
346 displayed by a well-defined clade of AlphaPVs (Hiller *et al.* 2006; Fu *et al.* 2010; Mesplède
347 *et al.* 2012) that are monophyletic when considering the early genes (Bravo and Alonso
348 2004). The evolutionary explanation proposed to account for this pattern is that in the most
349 recent common ancestor of AlphaPVs-species-5, 6 ,7, 9, 11 and 12 the *e6* ORF acquired
350 changes that allowed for the interaction E6-E6AP-p53, ultimately resulting in the
351 degradation of cellular p53 (Bravo and Felez-Sanchez 2015; Willemsen and Bravo 2019)
352 (Figure 2). Indeed, a virus-host co-divergence pattern has been experimentally
353 demonstrated for species-specific interactions between E6 and p53 (Long *et al.* 2022), and a
354 fraction of the differences in oncogenic potential among oncogenic HPVs may actually be
355 related to small differences in the E6-E6AP-p53 interaction, that eventually modulate the
356 intensity of p53 degradation (Conrady *et al.* 2020). Beyond p53 degradation, E6 proteins
357 from oncogenic AlphaHPVs display a vast interaction repertoire with cellular proteins
358 mediated by a small PDZ-binding motif present in the E6 C-terminal sequence (Nomine *et*

359 *al.* 2003; Thomas *et al.* 2016; Webb Strickland *et al.* 2018). Within AlphaPVs, only the E6
360 proteins of possibly/probably HPVs, *i.e.* species 5, 6, 7, 9 and 11, display such a PDZ-
361 binding motif in their C-termini (Figure 1). Closely related E6 sequences show qualitative
362 and quantitative differences in their protein-protein interaction repertoire, associated to
363 minor disparities in the PDZ-binding motif display (Gogl *et al.* 2022). Many viral proteins
364 display indeed a PDZ-binding motif in their C-termini, and the protein-protein interaction
365 networks provided through this small sequence stretch may be an essential element of the
366 viral parasitic lifestyle (Davey, Travé and Gibson 2011; Javier and Rice 2011). The strong
367 concordance in oncogenic HPVs between the lack of overlapping between *e6* and *e7*,
368 possibly related to a small nucleotide insertion (Auslander *et al.* 2019), and the presence of
369 a PDZ-binding domain, makes it tempting to hypothesise that both traits are related to a
370 single evolutionary event.

371 A final qualitative difference between the main clades within AlphaPVs are the codon
372 usage preferences of early and late genes (Félez-Sánchez *et al.* 2015). Among the very few
373 things that all viruses have in common, a fundamental one is that all viruses are parasites
374 of the translation machinery of the infected cell. Codon usage preferences and the
375 associated frequency of CpG and TpA dinucleotides are thus central to the virus-host
376 interplay during gene expression, because they condition the expression ability of a gene
377 (Plotkin and Kudla 2011; Brule and Grayhack 2017; Callens *et al.* 2021) and also because
378 they modulate the immune response against the viral genes (Li *et al.* 2012; Mordstein *et al.*
379 2020). Indeed, codon recoding strongly modifies expression of viral genes (Burns *et al.*
380 2006; Lauring *et al.* 2012), to the extent that a full novel generation of vaccines, including
381 RNA-based ones, applies synonymous recoding (Martínez *et al.* 2016; Zhang *et al.* 2023).
382 Papillomaviruses genomes are A+T rich and their codon usage preferences differ from the
383 average ones of their hosts (Zhao, Liu and Frazer 2003; Bravo and Müller 2005). However,
384 different tissues express different repertoires of tRNAs and codon usage preferences may
385 be tissue-dependent (Dittmar, Goodenbour and Pan 2006; Kames *et al.* 2020; Akins *et al.*
386 2023), and it has been proposed that viruses display different codon usage preferences
387 depending on the tissues they infect (Hernandez-Alias *et al.* 2021). In the case of PVs,
388 codon usage preferences of capsid genes are limited by the tRNA availability in the upper
389 skin layers (Zhou *et al.* 1999; Gu *et al.* 2004; Zhao *et al.* 2005). These findings are consistent

390 with the changes in viral gene expression changes across the differentiation stages of the
391 infected keratinocyte. Further, codon usage preferences largely differ between AlphaHPVs
392 and Beta- and GammaHPVs (Cladel, Bertotto and Christensen 2010; Félez-Sánchez *et al.*
393 2015), suggesting that viral life style strongly conditions codon usage in viruses that infect
394 the same human host. Finally, for AlphaHPVs, codon usage patterns cluster together
395 viruses in three large groups, corresponding to evolutionary close viruses associated to
396 similar natural histories of the infection (Félez-Sánchez *et al.* 2015), corresponding to a
397 differential match to the average human codon usage preferences (Figure 3) (Bourret,
398 Alizon and Bravo 2019). Thus, oncogenic HPVs display codon usage preferences
399 qualitatively distinct from those of other AlphaHPVs, albeit no quantitative differences
400 can be established among oncogenic HPVs connecting codon usage preferences and
401 differential oncogenicity.

402

403

404 WHY HPV16? WHY, NOW HPV42?

405 After several decades of research about the oncogenicity of certain HPVs, nothing of what
406 we knew about HPV42 let us think that it could be related to any cancer. The intellectual
407 challenge is thus to reconcile the experimental evidence strongly supporting an etiological
408 role of HPV42 infections in DPA, a rare cancer of sweat glands, with the absence of any of
409 the classical signatures of oncogenic HPVs in the genome of HPV42: i) it lacks an ORF in
410 the *inter_e2_l2*, region and does not display an eroded coding signal, while the closely
411 related HPV32 encodes for an E5_delta protein; ii) the *e6_e7* expression cassette
412 corresponds to a non-oncogenic pattern, the *e7* AUG start codon being located 25
413 nucleotides upstream the *e6* UAA stop codon; iii) no splicing signals for *e6*le7* transcripts
414 are predicted (Van Doorslaer *et al.* 2017), and no spliced transcripts have been detected
415 (Leiendecker *et al.* 2023); iv) HPV42_E6 can bind human E6AP (Leiendecker *et al.* 2023),
416 but the expression of HPV42_E6 in human cells does not induce degradation of p53 (Fu *et al.*
417 2010; Mesplède *et al.* 2012; Leiendecker *et al.* 2023); finally, v) HPV42_E6 lacks the
418 terminal PDZ-binding domain present in the C-terminus of E6 proteins in oncogenic
419 HPVs. The only element that could point to a carcinogenic potential for HPV42 is that

420 HPV42_E7 presents an aspartate residue upstream the LxCxE motif (Guimera *et al.* 2013),
421 which enhances E7-pRb binding (Heck *et al.* 1992), and indeed, experimental data show
422 that overexpression of HPV42_E7 results in cellular transformation and *in vivo* tumour
423 growth (Leiendecker *et al.* 2023).

424 We propose here that a perspective of the virus-host interactions centered on cell
425 phenotypes could help understand the riddle of rare clinical and phenotypic presentations
426 associated to common, largely asymptomatic viral infections. Under this perspective, the
427 differential probabilities for the repertoire of natural histories of the infection depend on
428 the particular interaction of a specific viral genotype with a specific cell phenotype. This
429 model does not stem from incomplete knowledge, but integrates instead the inherent non-
430 determinism of biological processes, where gene expression and cellular states fluctuate
431 unpredictably (Raser and O’Shea 2005), including a multitude of “non canonical”, poorly
432 understood mechanisms that are most likely related to biological noise (Pickrell *et al.* 2010;
433 Sibley, Blazquez and Ule 2016; Liu, François and Capp 2019; Yu and Kim 2020; Sorokin *et*
434 *al.* 2021; Porter 2023). Our common understanding of viral infections partly integrates a
435 static perspective in this regard, when referring to viral tropism or to infection-permissive
436 cells or opportunistic infections. But we want to stress here that the cellular phenotype that
437 is relevant for viral infections is a dynamic one: it can be stable, (*e.g.* a definite cell type, a
438 particular epigenetic state) or transient, (*e.g.* a stage along a cell differentiation process, or
439 over cell cycle, or a transient epigenetic state); it may depend on the context and vary
440 across space (*e.g.* for similar cell types in different tissues or organs), or across time (*e.g.*
441 with life history traits such as puberty, menstruation, reproduction, menopause or
442 senescence); and it can further depend on the immune status of the individual (*e.g.* local or
443 systemic immunodepression of idiopathic origin, or linked to infections such as HIV, or
444 associated to medical treatments).

445 We understand here cellular phenotypes as a quasi-discontinuous repertoire of states
446 available to a given cell, that are connected by probabilities of transition from one to
447 another. Such transitions can be stochastic, or favored by intrinsic or extrinsic factors. A
448 particular cellular state leads to particular replication, transcription and translation viral
449 patterns, allowing for different infection profiles, even at the very fine scale of cellular
450 cycling timing (Schulte and Andino 2014; Petkidis *et al.* 2024). This cellular context will be

451 modified by infection itself, so that the actual phenotype of the infected virocell (Forterre
452 2011), and thus the individual outcome of the infection, may be subject to complex
453 feedback cycles. This approach for interpreting the diversity of natural histories associated
454 to a parasite-host interaction becomes especially useful when applied to resident members
455 of the microbiota.

456 Small DNA viruses, notably PVs and polyomaviruses (PyVs), are fundamental
457 components of the healthy human microbiota (Foulongne *et al.* 2012; Ma *et al.* 2014;
458 Rascovan, Duraisamy and Desnues 2016). The prevalence of each individual viral
459 genotype varies among individuals, but also differs among anatomical sites and changes
460 over lifetime (de Koning *et al.* 2007; Hannigan *et al.* 2015; Amorrortu *et al.* 2022). As it is the
461 case for most individual PV and polyomavirus genotypes, HPV42 is part and parcel of this
462 healthy skin human microbiota (Ma *et al.* 2014) but most HPV42 infections are not
463 associated to any clinical sign. In our perspective of interpreting the outcomes of the
464 infection under the light of specific virus genotype-cell phenotype interactions, we
465 propose that the vast majority of HPV42 infections on the human skin are asymptomatic,
466 but that in rare occasions HPV42 may gain access to a rare cell (pheno)type present in the
467 sweat glands, with an increased probability of resulting in transformation and cancer.
468 Specific cells in skin appendages have actually been proposed to be “infection reservoirs”
469 for certain HPVs, mostly specific cells in hair follicles (Schmitt *et al.* 1996; Boxman *et al.*
470 1997) but also in sweat glands (Egawa 2005). The skin is indeed a highly heterogeneous,
471 complex cellular environment that we start only to be able to engineer and reproduce (Lee
472 *et al.* 2020). The variety of skin cell (pheno)types allows for a large variety of virus-host
473 interactions, as the precise viral transcription and translation patterns will depend on the
474 biochemical and cellular contexts provided by the individual infected cell, (Butz and
475 Hoppe Seyler 1993; Cid *et al.* 1993; Schmitt *et al.* 1996; Schenkel *et al.* 1999). This overall
476 perspective on cellular phenotypes can provide a general framework for the hit-and-run
477 mechanism invoked for the long-lasting puzzle of the role of BetaHPVs infections in non-
478 melanoma skin cancer (Rollison *et al.* 2019), reconciling this view with the conflicting
479 proposal of a cancer-protective role of cutaneous PV infections as modulators of the T-cell
480 immunity (Strickley *et al.* 2019), and helping to frame the interesting scientific controversy
481 it has generated (Lambert *et al.* 2020; Strickley *et al.* 2020). Recent transcriptomic data

482 confirm this view, with specific viral transcripts being detected in specific cellular
483 subpopulations across the keratinocyte differentiation gradient (Bedard *et al.* 2023), as well
484 as recent proteomic data providing evidence of spatial translation heterogeneity in
485 cutaneous HPVs infections (Schäfer *et al.* 2023).

486 By focusing on the interaction between viral genotypes and the transient, context-
487 dependent nature of cell phenotypes, we will gain a deeper understanding of the how
488 common asymptomatic infections may be related to rare cancers like DPA.

489

490 HOW A VIRUS GENOTYPE - CELL PHENOTYPE PERSPECTIVE CAN HELP GUIDE OUR
491 UNDERSTANDING OF THE DIVERSITY OF NATURAL HISTORIES OF THE INFECTION

492 Traditional pathogen-host paradigms, such as the Koch-Pasteur model, rely on a *one-to-one*
493 connection (or a simplified version of it) between pathogen and disease. But this paradigm
494 does not help account for the diversity of natural histories of viral infections, particularly
495 those from common microbiota members. In many cases we observe one virus causing a
496 large spectrum of clinical presentations of the infection (*one-to-many*), or distinct viruses
497 causing the same disease (*many-to-one*). A cell phenotype-centered perspective provides a
498 probabilistic framework to understand these diverse relationships, and it generates key
499 questions that will need to be answered in order to properly understand the underlying
500 causal mechanisms.

501 **One-to-one relationships: HPV42 and DPA.** In the case of HPV42 and DPA the
502 evolutionary question is, given the close relationship between HPV32 and HPV42, why
503 there are no HPV32-driven DPAs?; from the ontogeny standpoint, are HPV32 and/or
504 HPV42 associated to malignant proliferations of cell (pheno)types close to those associated
505 to DPA? In analogy, the causative role of Merkel cell polyomavirus in a substantial
506 fraction of Merkel cell carcinomas (Feng *et al.* 2008), raises the question of whether the
507 closely related *Pan troglodytes verus* polyomavirus 2a and *Gorilla gorilla gorilla*
508 polyomavirus 1 (Polyomaviridae Study Group of the International Committee on
509 Taxonomy of Viruses *et al.* 2016) might be involved in malignant proliferations analogous
510 to Merkel cell carcinomas in their respective host species.

511 One-to-many relationships:

512 *i) the extreme oncogenicity of HPV16.* Possibly, the most enduring puzzle after fifty
513 years of research on HPVs and cancers is not how certain HPVs cause certain
514 cancers, but why HPV16 is uniquely oncogenic compared to its close relatives, *e.g.*
515 HPV31 and HPV35. Indeed, HPV16 is an outlier compared to any other AlphaHPV
516 in terms of epidemiology of anogenital asymptomatic infections, considering
517 incidence, prevalence and duration of asymptomatic infections, but also regarding
518 the individual contribution to HPV-driven cancers (Figure 3) (Goodman *et al.* 2008;
519 Trottier *et al.* 2008; Schmeink *et al.* 2013; Ramanakumar *et al.* 2016; Malagón *et al.*
520 2023; HPV information centre 2024). Genomic and functional hallmarks allow to
521 pinpoint (possibly/probably) oncogenic HPVs from other AlphaHPVs (Bravo and
522 Alonso 2004; Schiffman *et al.* 2005), but they do not serve to establish qualitative nor
523 quantitative differences in oncogenic potential among them, as noted previously
524 (Hiller *et al.* 2006; Muench *et al.* 2009). The viral genotype - cell phenotype
525 perspective suggests to focus on specific cell precursors related to particular
526 anogenital transformations, as it has been shown that a discrete population of cells in
527 the cervical transition zone is enriched in HPVs-driven cervical cancer (Herfs *et al.*
528 2012). Analogous specific keratinocyte precursors are more prone to develop
529 cutaneous squamous cell carcinomas upon MmuPV1 infection (Moreno, Buehler and
530 Lambert 2022). It could thus be hypothesized that the higher relative contribution of
531 HPV16 to anal or to oropharyngeal cancers compared to cervical cancers could be
532 related to the differential frequency of a particular cell phenotype in these
533 anatomical locations. Identifying such particular cell phenotypes would be of
534 interest because the microanatomy differs between cervical and anal
535 squamocolumnar junctions (Yang *et al.* 2015), but also because we still lack a proper
536 understanding of the natural history of HPV-driven oropharyngeal cancers in terms
537 of precursor lesions and progression pace and stages (Roberts *et al.* 2019).

538 *ii) HPV18 and adenocarcinomas.* The relative contribution of HPV18 is substantially
539 higher in adenocarcinomas compared to squamous cell carcinomas (Reynders *et al.*
540 2023), which could point towards a particular interaction between AlphaPVs-
541 species-7 and certain mucosal-related cell phenotypes.

542 *iii) HPV33 and multiphenotypic sinonasal carcinomas.* HPV-related multiphenotypic
543 sinonasal carcinomas present features of adenoid cystic carcinomas (Bishop *et al.*
544 2013) and are most often associated to HPV33 and other AlphaHPVs-species-9, with
545 only a minor contribution of HPV16 to this specific disease (Bishop *et al.* 2017). HPV-
546 related multiphenotypic sinonasal carcinomas share typical diagnostic features with
547 other HPV-driven cancers, such as p16 overexpression (Bishop *et al.* 2017). Given the
548 substantial albeit low contribution of HPV33 to anogenital cancers (ten to twenty
549 times lower than that of HPV16) and given that little qualitative differences have
550 been pinpointed so far at the mechanistic level between HPV16 and HPV33 (Muench
551 *et al.* 2009; Thomas *et al.* 2016; Gogl *et al.* 2022), we interpret that the oncogenic
552 potential of closely related HPVs may sharply differ depending on the specific
553 phenotype of the infected cell.

554 *Many-to-one relationships.* Multifocal epithelial hyperplasia is a benign proliferative
555 condition of the oral mucosa (Estrada, L. 1956; Archard, Heck and Stanley 1965) that
556 presents increased prevalence in small human populations, mostly of American-native
557 ancestry, although cases have been reported in individuals of very different genetic
558 background (Sethi *et al.* 2022). This disease exemplifies a many-to-one relationship, as it is
559 associated to infections by two distant genotypes, HPV13 and HPV32, (Beaudenon *et al.*
560 1987b; Henke *et al.* 1987, 1989). The sole case-control study available confirms HPV13 as
561 causative agent in children of the Embera-Chami community in Colombia (Cuberos *et al.*
562 2006). A clinical presentation similar to focal epithelial hyperplasia has been
563 communicated in chimpanzees, reaching high prevalence in a small zoo community, 15
564 out of 44 animals showing symptoms (Hollander and Van Noord 1972). The viral
565 causative agents of focal epithelial hyperplasia in bonobos and in chimpanzees are close
566 relatives of HPV13 (Van Ranst *et al.* 1991; Scinicariello *et al.* 1997), all of them belonging
567 together in a monophyletic clade within AlphaPVs-species-10. The epidemiological and
568 evolutionary relationships among these viruses raise the question of whether a specific
569 oral mucosal cell phenotype is present in the oral mucosa of primates that, when subject to
570 viral-induced benign proliferation, results in the clinical phenotype of focal epithelial
571 hyperplasia, independently of the precise genotype of the virus that triggered the disease.
572 Indeed, viruses closely related to HPV13, such as HPV44 or HPV74, members of the same

573 species, have not been reported in focal epithelial hyperplasia, and the same holds true for
574 HPV42, close relative of HPV32. To further complicate the picture, PV-driven focal
575 epithelial hyperplasia has also been reported in brown howler monkeys (Pulecio-Santos,
576 de Souza and Sá 2024), and the associated PV belongs into DyoomikronPVs, the sister
577 group of AlphaPVs (Silvestre *et al.* 2016). Thus, on the one hand three closely related
578 viruses are associated to a same, distinct clinical presentation of the infection in sister host
579 species, which is useful as it provides a good yardstick for dating PV evolution (Willemsen
580 and Bravo 2019), but on the other hand the same clinical presentation in humans is
581 associated to two genetically very distant HPVs, while the respectively closely related
582 HPVs are not. Finally, a distantly related virus is associated to a similar disease in a
583 distantly related host species. The mechanisms for which these PVs induce proliferation in
584 this putative cell phenotype would not necessarily need to be similar, if the clinical
585 presentation is mostly determined by the proliferating cell phenotype, which would allow
586 for the large genetic gap between the etiological agents of the disease. The key
587 determinant of the clinical presentation of the infection would thus be the cell phenotype
588 and not the viral genotype itself.

589 Focusing on viral genotype-cell phenotype interactions may thus provide a unifying
590 framework to understand the repertoire of natural histories of the infection. Identifying
591 the specific cellular phenotypes involved in proliferative diseases could help explain why
592 certain viruses are more or less prone to cause cancers in certain anatomical locations and
593 under certain conditions, while closely related viruses do not. This perspective allows to
594 integrate long-lasting genetic and environmental factors with more fluid and transient
595 cellular states, acknowledging the role of stochasticity and biological noise as a central
596 factor at determining the outcome of viral infections.

597

598 COROLLARY

599 Every surface in our bodies is a small universe, homeplace to millions of viruses, bacteria
600 and fungi. DNA viruses, particularly small DNA viruses, are central components of the
601 human skin virome. The vast majority of infections by HPVs and by human
602 polyomaviruses are asymptomatic, but a few, evolutionary close HPVs cause both benign

603 and malignant proliferative diseases. Given the high prevalence of PVs asymptomatic
604 infections in the human population, symptomatic infections, and especially cancers, play a
605 minimal role in the population and evolutionary dynamics of these viruses. We interpret
606 that the actual eco-evolutionary pressures driving PVs evolution are mostly related to the
607 asymptomatic, commensal-like virus lifestyle as their main strategy to exploit the host.
608 The different clinical manifestations are in most cases unfortunate spill-overs of this
609 commensal-like nature.

610 HPVs have not evolved “to cause” or “because they cause” common cancers, such as
611 cervical cancer, or rare cancers, such as DPA. The oncogenicity of a handful of HPVs may
612 play only a minor role on virus’ evolution. We propose that a perspective centered on the
613 differential probabilistic interactions between viral genotypes and host cell phenotypes
614 offers a powerful framework to understand how common viruses can cause rare
615 proliferative diseases.

616 Host cell phenotypes are inherently dynamic. They can be stable, such as differentiated
617 cell types, but are most often transient, fluctuating over time and space, across
618 developmental and life history changes, across tissues and hormonal cycles. Cell
619 phenotypes vary also with local and systemic changes in the immune status such as in
620 infection-mediated or iatrogenic immune suppression. Finally, cell phenotypes will vary
621 depending on the interaction with the local microbiota, so that the presence of a given
622 assembly of intracellular or extracellular viruses/bacteria/fungi may modify the cellular
623 phenotype and thus the response to a same individual cell to future microbial challenges.

624 The main challenge lies in moving beyond the deterministic models of biology focusing on
625 genetic (be they on the side of the pathogen or the host) and environmental factors,
626 towards an understanding that integrates transient, fluid interactions related to
627 probabilistic transitions between different biological states. Most often uncertainty in
628 biology is conceived as lack of knowledge, so that if we could have more data, finer in
629 space and in time, addressing multiple integration levels, we would be able to understand
630 complex biological systems and to predict biological interactions. Unfortunately, this
631 Newtonian-like conception of biology collides with the probabilistic nature of all
632 information transfer processes in biology, with heritable non-genetic heterogeneity, and

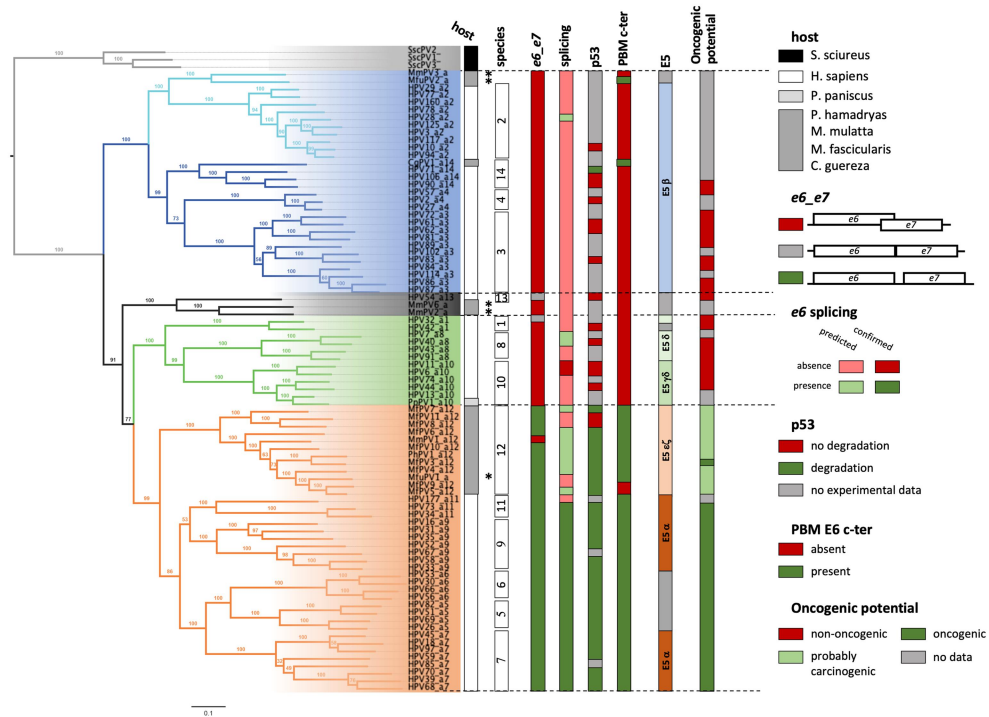
633 with the growing evidence about the importance of intrinsic and of extrinsic noise in
634 biology.

635 We require a conceptual transition towards a conception of biological reality as an
636 ensemble of quasi-discontinuous states, connected by transition probabilities. For common
637 members of the microbiota that can trigger disease under very specific cellular contexts,
638 crucially for HPVs and the cancers they cause, we urgently need to understand the
639 continuum between the normal and the pathological. As Georges Canguilhem suggested,
640 *“the continuity of a transition between one state and another can certainly be compatible with the*
641 *heterogeneity of these states”* (Canguilhem, Georges 1966). We hope that our proposal to
642 focus on transient cellular phenotypes can contribute to this understanding.

643

644

ORIGINAL UNEDITED MANUSCRIPT



646

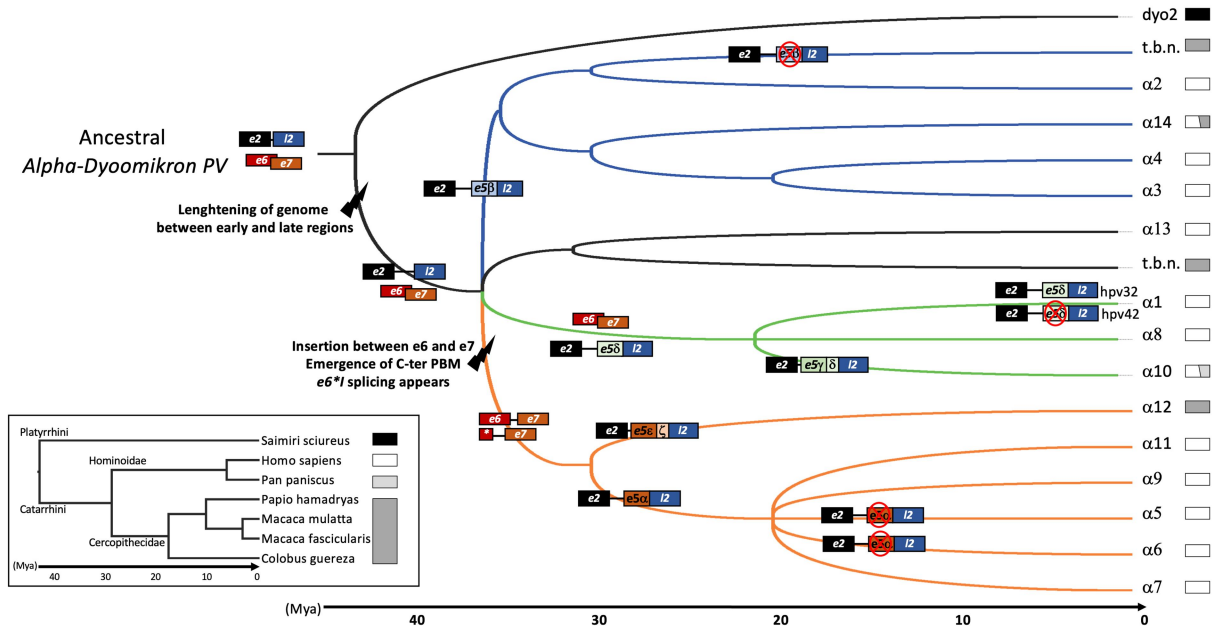
647 **Figure 1.- Phylogeny, genomic and phenotypic traits of Alpha- and DyoomikronPVs.**

648 Phylogenetic relationships inferred using maximum likelihood (RAxML) based on the
 649 concatenated *e1e1l2l1* ORFs. Numbers in the branches present support after 5,000
 650 bootstrap cycles. The tree is rooted using DyoomikronPVs as outgroup. The four main
 651 clades within AlphaPVs are identified in blue, grey, green and orange. Scalebar presents
 652 number of substitutions per site. **SPECIES:** taxonomy species for AlphaPVs. Genotypes
 653 labelled by an asterisk are not classified yet. **HOST:** taxonomy of the host for each
 654 genotype. **E6_E7:** organization of the e6e7 ORFs; red, e7 start precedes e6 stop with a range
 655 of 3 to 39 nucleotides; gray, means e7 start and e6 stop are merged in a TAATG sequence;
 656 green, e6 stop precedes e7 start with a range of 2 to 17 nucleotides. **SPLICING:** presence
 657 (green) or absence (red) of a splicing event in the e6 ORF either predicted (light colour)
 658 (Van Doorslaer *et al.* 2017) or experimentally confirmed (dark colour) (Cornelissen *et al.*
 659 1990; Nakagawa *et al.* 2000; Sotlar *et al.* 2004; Halec *et al.* 2013). **p53:** empirical evidence of
 660 E6 protein ability to induce degradation of p53 (Fu *et al.* 2010; Mesplède *et al.* 2012; Long *et al.*
 661 *al.* 2022). **PBM:** presence (green) or absence (red) of a C-terminal PDZ binding Motif (PBM)

662 in the E6 protein. **E5**: type of E5 ORF found in each AlphaPV species (Bravo and Alonso
663 2004). **ONCOGENIC POTENTIAL**: IARC/ICO classification for HPVs (HPV information
664 centre 2024), and literature data for AlphaPVs-species-12 (Kloster *et al.* 1988; Ostrow *et al.*
665 1993; Chen *et al.* 2009, 2019; Bergin *et al.* 2013; Long *et al.* 2022).

666

ORIGINAL UNEDITED MANUSCRIPT



668

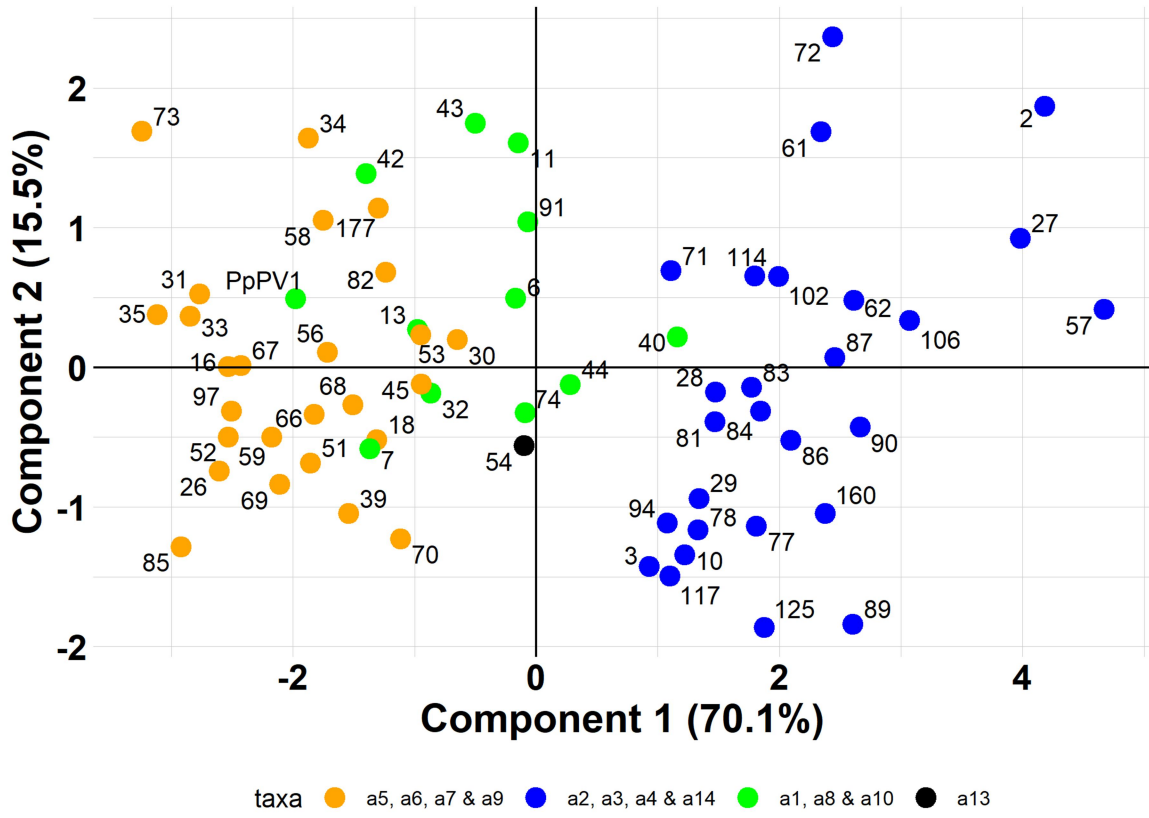
669 **Figure 2.- Global scenario of AlphaPV evolution.**

670 Cartoon depicting the evolutionary relationships among the four large clades of AlphaPV
 671 species, with branches in blue, black, green and orange. DyoomikronPVs are used as an
 672 outgroup. Branches comprising viruses not classified yet are labelled as t.b.n. (to be
 673 assigned). Blue and black boxes indicate the genomic organization of the *e2I2* interface
 674 between early and late ORF cassettes, as well as the type of E5 protein encoded, if
 675 applicable. Red and orange boxes indicate the organization of the *e6* and *e7* ORFs,
 676 overlapping or not. The inset presents the evolutionary history of the host species, with
 677 the scale in millions of years before present. For each viral branch, coloured box matches
 678 the code for the corresponding host species. A time scale is included as a suggestion for
 679 the timing of the different evolutionary events in the history of AlphaPVs, assuming that a
 680 large co-phylogeny between viruses and hosts. The genome of the AlphaDyoomikronPV
 681 ancestor, predating the split between Platyrrhini and Catarrhini, did not contain any
 682 region between early and late ORF cassettes and the *e6* and *e7* ORFs overlapped. In the
 683 genome of the AlphaPV ancestor, an integration event introduced a sequence between the
 684 early and the late gene cassettes. This individual event is possibly linked to the basal
 685 radiation in AlphaPVs that generated the extant four main clades, all of them spanning

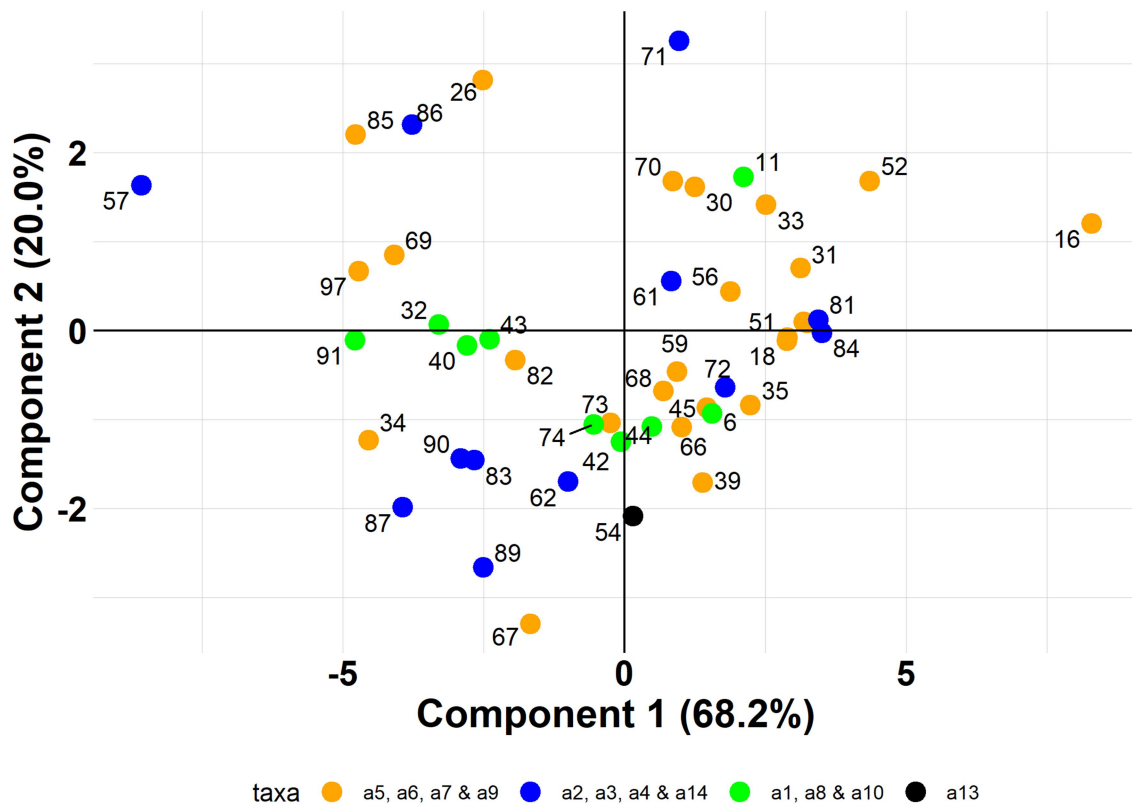
686 PVs infecting hosts in the Hominoidea and in the Cercopithecoidea primate families.
687 Different types of E5 ORFs emerged separately in each of these large clades, and
688 eventually disappeared in specific taxa, such as E5delta in HPV42, or were recently
689 pseudogenised, such as E5_alpha in AlphaPVs species 5 and 6. In the genome of the
690 ancestor of AlphaPVs species 5, 6, 7, 9, 11 and 12, an insertion occurred between the *e6* and
691 *e7* ORFs, possibly linked to the emergence of the C-terminal PDZ binding Motif in the E6
692 protein, to the gain of function of E6 to induce degradation of p53 and to the appearance of
693 the splicing event in the *e6* ORF. This viral clade, labelled in orange, contains all
694 (probably | possibly) oncogenic HPVs, classified after their oncogenic risk for anogenital
695 and oropharyngeal cancers. HPV42 causative agent of Digital Papillary Adenocarcinoma,
696 a rare tumour of the sweat glands, is a member of AlphaPVs-species 1.

697

ORIGINAL UNEDITED MANUSCRIPT



ORIGINAL UNEDITED



700

701

702 **Figure 3.- Clustering of AlphaPVs after epidemiological characteristics and codon usage**
 703 **preferences.**

704 Principal component analyses for epidemiological variables (A) and for codon usage
 705 preferences (B) of AlphaHPVs. Values in the axes reflect the percentage of the total
 706 variance captured by the corresponding component. Viral species are stratified into the
 707 main four evolutionary clades. Numbers correspond to the HPV genotype.
 708 Epidemiological variables used: genotype-specific prevalence in asymptomatic cervical
 709 samples and in cervical, vaginal, vulvar, anal and penile cancers (HPV information centre
 710 2024), genotype-specific median duration of the cervical infection (Goodman *et al.* 2008;
 711 Trottier *et al.* 2008; Schmeink *et al.* 2013; Ramanakumar *et al.* 2016), and genotype-specific
 712 one-year and five-years cumulative first detections of cervical infection (Malagón *et al.*
 713 2023). Codon usage was evaluated using the codon usage similarity index (COUSIN) for
 714 each of the *e6*, *e7*, *e1*, *e2*, *l2* and *l1* ORFs against the average human codon usage
 715 preferences (Bourret, Alizon and Bravo 2019).

716

717 **Bibliography**

718

719 Akins RB, Ostberg K, Cherlin T *et al.* The Typical tRNA Co-Expresses Multiple 5' tRNA Halves Whose
720 Sequences and Abundances Depend on Isodecoder and Isoacceptor and Change with
721 Tissue Type, Cell Type, and Disease. *Noncoding RNA* 2023;**9**:69.

722 Amorrortu RP, Zhao Y, Fenske NA *et al.* Natural History of Incident and Persistent Cutaneous
723 Human Papillomavirus and Human Polyomavirus Infections. *J Infect Dis* 2022;**226**:1162–74.

724 Angulo M, Carvajal-Rodriguez A. Evidence of recombination within human alpha-papillomavirus.
725 *Viol J* 2007;**4**:33.

726 Arbyn M, Tommasino M, Depuydt C *et al.* Are 20 human papillomavirus types causing cervical
727 cancer? *J Pathol* 2014;**234**:431–5.

728 Archard HO, Heck JW, Stanley HR. FOCAL EPITHELIAL HYPERPLASIA: AN UNUSUAL ORAL MUCOSAL
729 LESION FOUND IN INDIAN CHILDREN. *Oral Surg Oral Med Oral Pathol* 1965;**20**:201–12.

730 Arroyo Mühr LS, Eklund C, Dillner J. Misclassifications in human papillomavirus databases. *Virology*
731 2021;**558**:57–66.

732 Asadurian Y, Kurilin H, Lichtig H *et al.* Activities of human papillomavirus 16 E6 natural variants in
733 human keratinocytes. *J Med Virol* 2007;**79**:1751–60.

734 Auslander N, Wolf YI, Shabalina SA *et al.* A unique insert in the genomes of high-risk human
735 papillomaviruses with a predicted dual role in conferring oncogenic risk. *F1000Res*
736 2019;**8**:1000.

737 Beaudenon S, Kremsdorf D, Obalek S *et al.* Plurality of genital human papillomaviruses:
738 characterization of two new types with distinct biological properties. *Virology*
739 1987a;**161**:374–84.

740 Beaudenon S, Praetorius F, Kremsdorf D *et al.* A new type of human papillomavirus associated
741 with oral focal epithelial hyperplasia. *J Invest Dermatol* 1987b;**88**:130–5.

742 Bedard MC, Chihanga T, Carlile A *et al.* Single cell transcriptomic analysis of HPV16-infected
743 epithelium identifies a keratinocyte subpopulation implicated in cancer. *Nat Commun*
744 2023;**14**:1975.

745 Bergin IL, Bell JD, Chen Z *et al.* Novel genital alphapapillomaviruses in baboons (*Papio hamadryas*
746 *anubis*) with cervical dysplasia. *Vet Pathol* 2013;**50**:200–8.

747 Bishop JA, Andreasen S, Hang J-F *et al.* HPV-related Multiphenotypic Sinonasal Carcinoma: An
748 Expanded Series of 49 Cases of the Tumor Formerly Known as HPV-related Carcinoma
749 With Adenoid Cystic Carcinoma-like Features. *American Journal of Surgical Pathology*
750 2017;**41**:1690–701.

- 751 Bishop JA, Guo TW, Smith DF *et al.* Human papillomavirus-related carcinomas of the sinonasal
752 tract. *Am J Surg Pathol* 2013;**37**:185–92.
- 753 Boeri L, Capogrosso P, Ventimiglia E *et al.* High-risk human papillomavirus in semen is associated
754 with poor sperm progressive motility and a high sperm DNA fragmentation index in
755 infertile men. *Hum Reprod* 2019;**34**:209–17.
- 756 Bourret J, Alizon S, Bravo IG. COUSIN (COdon Usage Similarity INdex): A Normalized Measure of
757 Codon Usage Preferences. *Genome Biology and Evolution* 2019;**11**:3523–8.
- 758 Boxman ILA, Berkhout RJM, Mulder LHC *et al.* Detection of human papillomavirus DNA in plucked
759 hairs from renal transplant recipients and healthy volunteers. *J Invest Dermatol*
760 1997;**108**:712–5.
- 761 Bravo IG, Alonso A. Mucosal human papillomaviruses encode four different E5 proteins whose
762 chemistry and phylogeny correlate with malignant or benign growth. *J Virol*
763 2004;**78**:13613–26.
- 764 Bravo IG, Felez-Sanchez M. Papillomaviruses: viral evolution, cancer and evolutionary medicine.
765 *Evol Med Public Health* 2015, DOI: eov003 [pii] 10.1093/emph/eov003.
- 766 Bravo IG, Müller M. Codon usage in papillomavirus genes: practical and functional aspects.
767 *Papillomavirus report* 2005;**16**:1–9.
- 768 Brimer N, Drews CM, Vande Pol SB. Association of papillomavirus E6 proteins with either MAML1
769 or E6AP clusters E6 proteins by structure, function, and evolutionary relatedness. Lambert
770 PF (ed.). *PLoS Pathog* 2017;**13**:e1006781.
- 771 Brule CE, Grayhack EJ. Synonymous Codons: Choose Wisely for Expression. *Trends in genetics: TIG*
772 2017;**33**:283–97.
- 773 Burk RD, Harari A, Chen Z. Human papillomavirus genome variants. *Virology* 2013;**445**:232–43.
- 774 Burns CC, Shaw J, Campagnoli R *et al.* Modulation of poliovirus replicative fitness in HeLa cells by
775 deoptimization of synonymous codon usage in the capsid region. *J Virol* 2006;**80**:3259–72.
- 776 Butz K, Hoppe Seyler F. Transcriptional control of human papillomavirus (HPV) oncogene
777 expression: composition of the HPV type 18 upstream regulatory region. *J Virol*
778 1993;**67**:6476–86.
- 779 Callens M, Pradier L, Finnegan M *et al.* Read between the Lines: Diversity of Nontranslational
780 Selection Pressures on Local Codon Usage. *Genome Biology and Evolution* 2021;**13**, DOI:
781 <https://doi.org/10.1093/gbe/evab097>.
- 782 Canguilhem, Georges. *Le Normal et Le Pathologique*. 2nd edition, 1972. Paris: Presses
783 universitaires de France, 1966.
- 784 Cartin W, Alonso A. The human papillomavirus HPV2a E5 protein localizes to the Golgi apparatus
785 and modulates signal transduction. *Virology* 2003;**314**:572–9.

- 786 Cascardo CA, Cornell G, Moesch JR. Human papillomavirus 42-associated digital papillary
787 adenocarcinoma. *JAAD Case Rep* 2023;**32**:52–4.
- 788 Chen Z, van Doorslaer K, DeSalle R *et al.* Genomic diversity and interspecies host infection of
789 alpha12 Macaca fascicularis papillomaviruses (MfPVs). *Virology* 2009;**393**:304–10.
- 790 Chen Z, Long T, Wong PY *et al.* Non-human Primate Papillomaviruses Share Similar Evolutionary
791 Histories and Niche Adaptation as the Human Counterparts. *Front Microbiol* 2019;**10**:2093.
- 792 Chen Z, Wood CE, Abee CR *et al.* Complete Genome Sequences of Three Novel Saimiri sciureus
793 Papillomavirus Types Isolated from the Cervicovaginal Region of Squirrel Monkeys.
794 *Genome Announc* 2018;**6**:e01400-17.
- 795 Cid A, Auewarakul P, Garcia Carranca A *et al.* Cell-type-specific activity of the human
796 papillomavirus type 18 upstream regulatory region in transgenic mice and its modulation
797 by tetradecanoyl phorbol acetate and glucocorticoids. *J Virol* 1993;**67**:6742–52.
- 798 Ciuffo G. Innesto positivo con filtrato di verruca volgare. *Giorn Ital Mal Venereol* 1907;**12**:7.
- 799 Cladel NM, Bertotto A, Christensen ND. Human alpha and beta papillomaviruses use different
800 synonymous codon profiles. *Virus Genes* 2010;**40**:329–40.
- 801 Clifford GM, Tenet V, Georges D *et al.* Human papillomavirus 16 sub-lineage dispersal and cervical
802 cancer risk worldwide: Whole viral genome sequences from 7116 HPV16-positive women.
803 *Papillomavirus Res* 2019;**7**:67–74.
- 804 Combrinck CE, Seedat RY, Randall C *et al.* Novel HPV-6 variants of human papillomavirus causing
805 recurrent respiratory papillomatosis in southern Africa. *Epidemiol Infect* 2012;**140**:1095–
806 101.
- 807 Conrady MC, Suarez I, Gogl G *et al.* Structure of High-Risk Papillomavirus 31 E6 Oncogenic Protein
808 and Characterization of E6/E6AP/p53 Complex Formation. *J Virol* 2020;**95**:e00730-20.
- 809 Cornelissen MT, Smits HL, Briet MA *et al.* Uniformity of the splicing pattern of the E6/E7 transcripts
810 in human papillomavirus type 16-transformed human fibroblasts, human cervical
811 premalignant lesions and carcinomas. *J Gen Virol* 1990;**71**:1243–6.
- 812 Cuberos V, Perez J, Lopez CJ *et al.* Molecular and serological evidence of the epidemiological
813 association of HPV 13 with focal epithelial hyperplasia: A case-control study. *Journal of*
814 *Clinical Virology* 2006;**37**:21–6.
- 815 Da MH, Dong ZB, Yan Q *et al.* Detection of human papillomavirus 42 in a squamous cell carcinoma.
816 *Australas J Dermatol* 2018;**59**:e283–4.
- 817 Damay A, Fabre J, Costes V *et al.* Human papillomavirus (HPV) prevalence and type distribution,
818 and HPV-associated cytological abnormalities in anal specimens from men infected with
819 HIV who have sex with men. *J Med Virol* 2010;**82**:592–6.

- 820 Dartmann K, Schwarz E, Gissmann L *et al.* The nucleotide sequence and genome organization of
821 human papilloma virus type 11. *Virology* 1986;**151**:124–30.
- 822 Davey NE, Travé G, Gibson TJ. How viruses hijack cell regulation. *Trends Biochem Sci* 2011;**36**:159–
823 69.
- 824 DeLap RD, Friedman Kien AE, Rush MG. The absence of human papilloma viral DNA sequences in
825 condylomata acuminata. *Virology* 1976;**74**:268–72.
- 826 Delius H, Hofmann B. Primer-directed sequencing of human papillomavirus types.
827 *CurrTopMicrobiolImmunol* 1994;**186**:13–31.
- 828 Dimaio D, Petti LM. The E5 proteins. *Virology* 2013;**445**:99–114.
- 829 Dittmar KA, Goodenbour JM, Pan T. Tissue-specific differences in human transfer RNA expression.
830 *PLoS Genet* 2006;**2**:e221.
- 831 Donkoh ET, Asmah RH, Agyemang-Yeboah F *et al.* Prevalence and Distribution of Vaccine-
832 Preventable Genital Human Papillomavirus(HPV) Genotypes in Ghanaian Women
833 Presenting for Screening. *Cancer Control* 2022;**29**:10732748221094721.
- 834 Duran-Reynals, F. Virus-Induced Tumors and the Virus Theory of Cancer. In: Homburger, F. F WH
835 (ed.). *The Physiopathology of Cancer : A Treatise for Investigators, Physicians and Students.*
836 London, UK: Cassell / Co. Ltd., 1953.
- 837 Dürst M, Croce CM, Gissmann L *et al.* Papillomavirus sequences integrate near cellular oncogenes
838 in some cervical carcinomas. *ProcNatlAcadSciUSA* 1987;**84**:1070–4.
- 839 Dürst M, Gissmann L, Ikenberg H *et al.* A papillomavirus DNA from a cervical carcinoma and its
840 prevalence in cancer biopsy samples from different geographic regions.
841 *ProcNatlAcadSciUSA* 1983;**80**:3812–5.
- 842 Egawa K. Eccrine-centred distribution of human papillomavirus 63 infection in the epidermis of the
843 plantar skin. *BrJ Dermatol* 2005;**152**:993–6.
- 844 Eschle D, Dürst M, ter Meulen J *et al.* Geographical dependence of sequence variation in the E7
845 gene of human papillomavirus type 16. *JGenViro* 1992;**73**:1829–32.
- 846 Estrada, L. Nuevo aporte al estudio odontológico de los Indios Katios. *Heraldo Dental* 1956;**2**:59–
847 73.
- 848 Féllez-Sánchez M, Trösemeier J-H, Bedhomme S *et al.* Cancer, Warts, or Asymptomatic Infections:
849 Clinical Presentation Matches Codon Usage Preferences in Human Papillomaviruses.
850 *Genome Biology and Evolution* 2015;**7**:2117–35.
- 851 Feng H, Shuda M, Chang Y *et al.* Clonal integration of a polyomavirus in human Merkel cell
852 carcinoma. *Science* 2008;**319**:1096–100.
- 853 Flores-Díaz E, Sereday KA, Ferreira S *et al.* HPV-6 Molecular Variants Association With the
854 Development of Genital Warts in Men: The HIM Study. *J Infect Dis* 2017a;**215**:559–65.

- 855 Flores-Díaz E, Sereday KA, Ferreira S *et al.* HPV-11 variability, persistence and progression to
856 genital warts in men: the HIM study. *J Gen Virol* 2017b;**98**:2339–42.
- 857 Forterre P. Manipulation of cellular syntheses and the nature of viruses: The virocell concept.
858 *Comptes Rendus Chimie* 2011;**14**:392–9.
- 859 Foulongne V, Sauvage V, Hebert C *et al.* Human skin microbiota: high diversity of DNA viruses
860 identified on the human skin by high throughput sequencing. *PLoS One* 2012;**7**:e38499.
- 861 Fu L, Van Doorslaer K, Chen Z *et al.* Degradation of p53 by human Alphapapillomavirus E6 proteins
862 shows a stronger correlation with phylogeny than oncogenicity. *PLoS One* 2010;**5**, DOI:
863 10.1371/journal.pone.0012816.
- 864 Fujinami A, Inamoto K. Ueber Geschwülste bei japanischen Haushühnern, insbesondere über
865 einen transplantablen Tumor. *Z Krebs-forsch* 1914;**14**:94–119.
- 866 Garcia-Vallve S, Alonso A, Bravo IG. Papillomaviruses: different genes have different histories.
867 *Trends Microbiol* 2005;**13**:514–21.
- 868 Gissmann L, zur Hausen H. Human papilloma virus DNA: physical mapping and genetic
869 heterogeneity. *Proc Natl Acad Sci USA* 1976;**73**:1310–3.
- 870 Gissmann L, zur Hausen H. Partial characterization of viral DNA from human genital warts
871 (Condylomata acuminata). *Int J Cancer* 1980;**25**:605–9.
- 872 Gissmann L, Pfister H, zur Hausen H. Human papilloma viruses (HPV): characterization of four
873 different isolates. *Virology* 1977;**76**:569–80.
- 874 Godinez JM, Nicolas-Parraga S, Pimenoff VN *et al.* Phylogenetically related, clinically different:
875 Human Papillomaviruses 6 and 11 variants distribution in genital warts and in laryngeal
876 papillomatosis. *Clin Microbiol Infect* 2013, DOI: 10.1111/1469-0691.12420.
- 877 Gogl G, Zambo B, Kostmann C *et al.* Quantitative fragmentomics allow affinity mapping of
878 interactomes. *Nat Commun* 2022;**13**:5472.
- 879 Gómez Vázquez M, Navarra Amayuelas R. Periungual bowenoid papulosis due to human
880 papillomavirus type 42. *Actas Dermosifiliogr* 2013;**104**:932–4.
- 881 Goodman MT, Shvetsov YB, McDuffie K *et al.* Prevalence, acquisition, and clearance of cervical
882 human papillomavirus infection among women with normal cytology: Hawaii Human
883 Papillomavirus Cohort Study. *Cancer Res* 2008;**68**:8813–24.
- 884 Gormley RH, Groft CM, Miller CJ *et al.* Digital squamous cell carcinoma and association with
885 diverse high-risk human papillomavirus types. *J Am Acad Dermatol* 2011;**64**:981–5.
- 886 Gottschling M, Goker M, Stamatakis A *et al.* Quantifying the phylodynamic forces driving
887 papillomavirus evolution. *Mol Biol Evol* 2011;**28**:2101–13.
- 888 Gu W, Li M, Zhao WM *et al.* tRNASer(CGA) differentially regulates expression of wild-type and
889 codon-modified papillomavirus L1 genes. *Nucleic Acids Res* 2004;**32**:4448–61.

- 890 Guimera N, Lloveras B, Lindeman J *et al.* The occasional role of low-risk human papillomaviruses 6,
891 11, 42, 44, and 70 in anogenital carcinoma defined by laser capture microdissection/PCR
892 methodology: results from a global study. *Am J Surg Pathol* 2013;**37**:1299–310.
- 893 Halec G, Schmitt M, Dondog B *et al.* Biological activity of probable/possible high-risk human
894 papillomavirus types in cervical cancer. *Int J Cancer* 2013;**132**:63–71.
- 895 Hannigan GD, Meisel JS, Tyldsley AS *et al.* The human skin double-stranded DNA virome:
896 topographical and temporal diversity, genetic enrichment, and dynamic associations with
897 the host microbiome. *mBio* 2015;**6**:e01578-01515.
- 898 zur Hausen H. Condylomata acuminata and human genital cancer. *Cancer Res* 1976;**36**:794.
- 899 zur Hausen H, Meinhof W, Scheiber W *et al.* Attempts to detect virus-specific DNA in human
900 tumors. I. Nucleic acid hybridizations with complementary RNA of human wart virus. *Int J*
901 *Cancer* 1974a;**13**:650–6.
- 902 zur Hausen H, Schulte-Holthausen H, Wolf H *et al.* Attempts to detect virus-specific DNA in human
903 tumors. II. Nucleic acid hybridizations with complementary RNA of human herpes group
904 viruses. *Int J Cancer* 1974b;**13**:657–64.
- 905 Heck DV, Yee CL, Howley PM *et al.* Efficiency of binding the retinoblastoma protein correlates with
906 the transforming capacity of the E7 oncoproteins of the human papillomaviruses.
907 *Proc Natl Acad Sci USA* 1992;**89**:4442–6.
- 908 Henke RP, Guerin Reverchon I, Milde Langosch K *et al.* In situ detection of human papillomavirus
909 types 13 and 32 in focal epithelial hyperplasia of the oral mucosa. *J Oral Pathol Med*
910 1989;**18**:419–21.
- 911 Henke RP, Milde Langosch K, Loning T *et al.* Human papillomavirus type 13 and focal epithelial
912 hyperplasia of the oral mucosa: DNA hybridization on paraffin-embedded specimens.
913 *Virchows Arch A Pathol Anat Histopathol* 1987;**411**:193–8.
- 914 Hennell C, Jamison J, Wells M *et al.* Inverted papilloma of the cervix and vagina: report of 2 cases
915 of a rare lesion associated with human papillomavirus 42. *Hum Pathol* 2012;**43**:435–9.
- 916 Herfs M, Yamamoto Y, Laury A *et al.* A discrete population of squamocolumnar junction cells
917 implicated in the pathogenesis of cervical cancer. *Proc Natl Acad Sci U S A*
918 2012;**109**:10516–21.
- 919 Hernandez-Alias X, Benisty H, Schaefer MH *et al.* Translational adaptation of human viruses to the
920 tissues they infect. *Cell Rep* 2021;**34**:108872.
- 921 Hiller T, Poppelreuther S, Stubenrauch F *et al.* Comparative analysis of 19 genital human
922 papillomavirus types with regard to p53 degradation, immortalization, phylogeny, and
923 epidemiologic risk classification. *Cancer Epidemiol Biomarkers Prev* 2006;**15**:1262–7.
- 924 Hollander CF, Van Noord MJ. Focal epithelial hyperplasia: A virus-induced oral mucosal lesion in
925 the chimpanzee. *Oral Surgery, Oral Medicine, Oral Pathology* 1972;**33**:220–6.

- 926 HPV information centre. ICO/IARC Information Centre on HPV and Cancer (<https://hpcvcentre.net>).
927 2024.
- 928 Hu J, Cladel NM, Pickel MD *et al.* Amino acid residues in the carboxy-terminal region of cottontail
929 rabbit papillomavirus E6 influence spontaneous regression of cutaneous papillomas. *J Virol*
930 2002;**76**:11801–8.
- 931 Huibregtse JM, Scheffner M, Howley PM. A cellular protein mediates association of p53 with the
932 E6 oncoprotein of human papillomavirus types 16 or 18. *EMBO J* 1991;**10**:4129–35.
- 933 Inagaki Y, Tsunokawa Y, Takebe N *et al.* Nucleotide sequences of cDNAs for human papillomavirus
934 type 18 transcripts in HeLa cells. *J Virol* 1988;**62**:1640–6.
- 935 Javier RT, Rice AP. Emerging theme: cellular PDZ proteins as common targets of pathogenic
936 viruses. *J Virol* 2011;**85**:11544–56.
- 937 Kacerovská D, Michal M, Kašpírková J *et al.* Epidermolytic hyperkeratosis of the vulva associated
938 with basal cell carcinoma in a patient with vaginal condyloma acuminatum and vaginal
939 intraepithelial neoplasia harboring HPV, type 42. *Cesk Patol* 2014;**50**:92–4.
- 940 Kames J, Alexaki A, Holcomb DD *et al.* TissueCoCoPUTs: Novel Human Tissue-Specific Codon and
941 Codon-Pair Usage Tables Based on Differential Tissue Gene Expression. *Journal of*
942 *Molecular Biology* 2020;**432**:3369–78.
- 943 Karamanou M, Agapitos E, Kousoulis A *et al.* From the humble wart to HPV: a fascinating story
944 throughout centuries. *Oncol Rev* 2010;**4**:133–5.
- 945 Kervarrec T, Imbeaud S, Veyer D *et al.* Digital Papillary Adenocarcinoma in Nonacral Skin:
946 Clinicopathologic and Genetic Characterization of 5 Cases. *Am J Surg Pathol*
947 2023a;**47**:1077–84.
- 948 Kervarrec T, Tallet A, Macagno N *et al.* Sweat Gland Tumors Arising on Acral Sites: A Molecular
949 Survey. *Am J Surg Pathol* 2023b;**47**:1096–107.
- 950 Kiwerska K, Jozefiak A, Markowska J *et al.* Oral-genital human papillomavirus infection in Polish
951 couples: frequent detection of HPV 42. *BMC Infect Dis* 2019;**19**:122.
- 952 Kloster BE, Manias DA, Ostrow RS *et al.* Molecular cloning and characterization of the DNA of two
953 papillomaviruses from monkeys. *Virology* 1988;**166**:30–40.
- 954 de Koning MN, Struijk L, Bavinck JN *et al.* Betapapillomaviruses frequently persist in the skin of
955 healthy individuals. *J Gen Virol* 2007;**88**:1489–95.
- 956 Kreuter A, Hufbauer M, Silling S *et al.* Simultaneous Induction of Benign Condyloma and High-
957 grade Anal Dysplasia Induced by Low-risk Human Papillomavirus Type 42. *Acta Derm*
958 *Venerol* 2018;**98**:616–7.

- 959 Lace MJ, Isacson C, Anson JR *et al.* Upstream regulatory region alterations found in human
960 papillomavirus type 16 (HPV-16) isolates from cervical carcinomas increase transcription,
961 ori function, and HPV immortalization capacity in culture. *J Virol* 2009;**83**:7457–66.
- 962 Lambert PF, Münger K, Rösl F *et al.* Beta human papillomaviruses and skin cancer. *Nature*
963 2020;**588**:E20–1.
- 964 Lauring AS, Acevedo A, Cooper SB *et al.* Codon usage determines the mutational robustness,
965 evolutionary capacity, and virulence of an RNA virus. *Cell Host Microbe* 2012;**12**:623–32.
- 966 Lee J, Rabbani CC, Gao H *et al.* Hair-bearing human skin generated entirely from pluripotent stem
967 cells. *Nature* 2020;**582**:399–404.
- 968 Leiendecker L, Neumann T, Jung PS *et al.* Human Papillomavirus 42 Drives Digital Papillary
969 Adenocarcinoma and Elicits a Germ Cell-like Program Conserved in HPV-Positive Cancers.
970 *Cancer Discov* 2023;**13**:70–84.
- 971 Li M, Kao E, Gao X *et al.* Codon-usage-based inhibition of HIV protein synthesis by human schlafen
972 11. *Nature* 2012;**491**:125–8.
- 973 Liu J, François J-M, Capp J-P. Gene Expression Noise Produces Cell-to-Cell Heterogeneity in
974 Eukaryotic Homologous Recombination Rate. *Front Genet* 2019;**10**:475.
- 975 Long T, Burk RD, Chan PKS *et al.* Non-human primate papillomavirus E6-mediated p53 degradation
976 reveals ancient evolutionary adaptation of carcinogenic phenotype to host niche. *PLoS*
977 *Pathog* 2022;**18**:e1010444.
- 978 Ma Y, Madupu R, Karaoz U *et al.* Human papillomavirus community in healthy persons, defined by
979 metagenomics analysis of human microbiome project shotgun sequencing data sets. *J*
980 *Virology* 2014;**88**:4786–97.
- 981 Malagón T, Trottier H, El-Zein M *et al.* Human Papillomavirus Intermittence and Risk Factors
982 Associated With First Detections and Redetections in the Ludwig-McGill Cohort Study of
983 Adult Women. *The Journal of Infectious Diseases* 2023;**228**:402–11.
- 984 de Martel C, Georges D, Bray F *et al.* Global burden of cancer attributable to infections in 2018: a
985 worldwide incidence analysis. *The Lancet Global Health* 2020;**8**:e180–90.
- 986 Martínez MA, Jordan-Paiz A, Franco S *et al.* Synonymous Virus Genome Recoding as a Tool to
987 Impact Viral Fitness. *Trends Microbiol* 2016;**24**:134–47.
- 988 Martínez-Zapien D, Ruiz FX, Poirson J *et al.* Structure of the E6/E6AP/p53 complex required for
989 HPV-mediated degradation of p53. *Nature* 2016;**529**:541–5.
- 990 Melnick JL. Papova virus group. *Science* 1962;**135**:1128–30.
- 991 Mengual-Chuliá B, Bedhomme S, Lafforgue G *et al.* Assessing parallel gene histories in viral
992 genomes. *BMC Evolutionary Biology* 2016;**16**:32.

- 993 Mesplède T, Gagnon D, Bergeron-Labrecque F *et al.* p53 degradation activity, expression, and
994 subcellular localization of E6 proteins from 29 human papillomavirus genotypes. *J Virol*
995 2012;**86**:94–107.
- 996 Mirabello L, Yeager M, Cullen M *et al.* HPV16 Sublineage Associations With Histology-Specific
997 Cancer Risk Using HPV Whole-Genome Sequences in 3200 Women. *J Natl Cancer Inst*
998 2016;**108**:d1w100.
- 999 Mirabello L, Yeager M, Yu K *et al.* HPV16 E7 Genetic Conservation Is Critical to Carcinogenesis. *Cell*
1000 2017;**170**:1164–1174.e6.
- 1001 Mordstein C, Savisaar R, Young RS *et al.* Codon Usage and Splicing Jointly Influence mRNA
1002 Localization. *Cell Systems* 2020;**10**:351–362.e8.
- 1003 Moreno R, Buehler D, Lambert PF. MmuPV1-Induced Cutaneous Squamous Cell Carcinoma Arises
1004 Preferentially from Lgr5+ Epithelial Progenitor Cells. *Viruses* 2022;**14**:1751.
- 1005 Muench P, Hiller T, Probst S *et al.* Binding of PDZ proteins to HPV E6 proteins does neither
1006 correlate with epidemiological risk classification nor with the immortalization of foreskin
1007 keratinocytes. *Virology* 2009;**387**:380–7.
- 1008 Nakagawa S, Yoshikawa H, Yasugi T *et al.* Ubiquitous presence of E6 and E7 transcripts in human
1009 papillomavirus-positive cervical carcinomas regardless of its type. *J Med Virol*
1010 2000;**62**:251–8.
- 1011 Nicolás-Párraga S, Alemany L, de Sanjosé S *et al.* Differential HPV16 variant distribution in
1012 squamous cell carcinoma, adenocarcinoma and adenosquamous cell carcinoma. *Int J*
1013 *Cancer* 2017;**140**:2092–100.
- 1014 Nicolás-Párraga S, Gandini C, Pimenoff VN *et al.* HPV16 variants distribution in invasive cancers of
1015 the cervix, vulva, vagina, penis, and anus. *Cancer Med* 2016;**5**:2909–19.
- 1016 Nomine Y, Charbonnier S, Ristriani T *et al.* Domain substructure of HPV E6 oncoprotein:
1017 biophysical characterization of the E6 C-terminal DNA-binding domain. *Biochemistry*
1018 2003;**42**:4909–17.
- 1019 Ong C-K, Chan S-Y, Campo MS *et al.* Evolution of human papillomavirus type 18: an ancient
1020 phylogenetic root in Africa and intratype diversity reflect coevolution with human ethnic
1021 groups. *J Virol* 1993;**67**:6424–31.
- 1022 Orth G, Favre M, Croissant O. Characterization of a new type of human papillomavirus that causes
1023 skin warts. *J Virol* 1977;**24**:108–20.
- 1024 Ostrow RS, Liu Z, Schneider JF *et al.* The products of the E5, E6, or E7 open reading frames of RhPV
1025 1 can individually transform NIH 3T3 cells or in cotransfections with activated ras can
1026 transform primary rodent epithelial cells. *Virology* 1993;**196**:861–7.

- 1027 Petkidis A, Suomalainen M, Andriasyan V *et al.* Preexisting cell state rather than stochastic noise
1028 confers high or low infection susceptibility of human lung epithelial cells to adenovirus.
1029 *mSphere* 2024:e0045424.
- 1030 Pfister H, Hettich I, Runne U *et al.* Characterization of human papillomavirus type 13 from focal
1031 epithelial hyperplasia Heck lesions. *J Virol* 1983;**47**:363–6.
- 1032 Philipp W, Honore N, Sapp M *et al.* Human papillomavirus type 42: new sequences, conserved
1033 genome organization. *Virology* 1992;**186**:331–4.
- 1034 Pickrell JK, Pai AA, Gilad Y *et al.* Noisy splicing drives mRNA isoform diversity in human cells. *PLoS*
1035 *Genet* 2010;**6**:e1001236.
- 1036 Pimenoff VN, De Oliveira CM, Bravo IG. Transmission between archaic and modern human
1037 ancestors during the evolution of the oncogenic human papillomavirus 16. *Molecular*
1038 *Biology and Evolution* 2017;**34**:4–19.
- 1039 Pimenoff VN, Tous S, Benavente Y *et al.* Distinct geographic clustering of oncogenic human
1040 papillomaviruses multiple infections in cervical cancers: Results from a worldwide cross-
1041 sectional study. *Int J Cancer* 2019;**144**:2478–88.
- 1042 Plotkin JB, Kudla G. Synonymous but not the same: the causes and consequences of codon bias.
1043 *Nature Reviews Genetics* 2011;**12**:32–42.
- 1044 Polyomaviridae Study Group of the International Committee on Taxonomy of Viruses, Calvignac-
1045 Spencer S, Feltkamp MCW *et al.* A taxonomy update for the family Polyomaviridae. *Arch*
1046 *Virol* 2016;**161**:1739–50.
- 1047 Porter LL. Fluid protein fold space and its implications. *BioEssays* 2023;**45**:2300057.
- 1048 Pujari R, Newman MR, Talia KL *et al.* Seborrhic Keratosis-Like Lesion of the Cervix: First Report of
1049 the Cytological Features of a Low-Risk HPV 42-Associated Lesion. *Acta Cytol* 2021;**65**:448–
1050 52.
- 1051 Pulecio-Santos SL, de Souza AJS, Sá LRM de. Epidemiological characterization of oral focal
1052 epithelial hyperplasia in brown howler monkeys (*Alouatta guariba clamitans*). *J Med*
1053 *Primatol* 2024;**53**:e12728.
- 1054 Ramanakumar AV, Naud P, Roteli-Martins CM *et al.* Incidence and duration of type-specific human
1055 papillomavirus infection in high-risk HPV-naïve women: results from the control arm of a
1056 phase II HPV-16/18 vaccine trial. *BMJ Open* 2016;**6**:e011371.
- 1057 Rascovan N, Duraisamy R, Desnues C. Metagenomics and the Human Virome in Asymptomatic
1058 Individuals. *Annu Rev Microbiol* 2016;**70**:125–41.
- 1059 Raser JM, O’Shea EK. Noise in gene expression: origins, consequences, and control. *Science*
1060 2005;**309**:2010–3.

- 1061 Reynders C, Lerho T, Goebel EA *et al.* Prevalence and genotype distribution of human
1062 papillomavirus in cervical adenocarcinoma (usual type and variants): A systematic review
1063 and meta-analysis. *J Med Virol* 2023;**95**:e29190.
- 1064 Roberts S, Evans D, Mehanna H *et al.* Modelling human papillomavirus biology in oropharyngeal
1065 keratinocytes. *Phil Trans R Soc B* 2019;**374**:20180289.
- 1066 Rollison DE, Viarasio D, Amorrortu RP *et al.* An Emerging Issue in Oncogenic Virology: the Role of
1067 Beta Human Papillomavirus Types in the Development of Cutaneous Squamous Cell
1068 Carcinoma. Sullivan CS (ed.). *J Virol* 2019;**93**:e01003-18.
- 1069 Rous P. A SARCOMA OF THE FOWL TRANSMISSIBLE BY AN AGENT SEPARABLE FROM THE TUMOR
1070 CELLS. *J Exp Med* 1911;**13**:397–411.
- 1071 Rowson KE, Mahy BW. Human papova (wart) virus. *Bacteriol Rev* 1967;**31**:110–31.
- 1072 Schäfer M, Schneider M, Müller T *et al.* Spatial tissue proteomics reveals distinct landscapes of
1073 heterogeneity in cutaneous papillomavirus-induced keratinocyte carcinomas. *J Med Virol*
1074 2023;**95**:e28850.
- 1075 Scheffner M, Werness BA, Huibregtse JM *et al.* The E6 oncoprotein encoded by human
1076 papillomavirus types 16 and 18 promotes the degradation of p53. *Cell* 1990;**63**:1129–36.
- 1077 Schenkel J, Gaissert H, Protopapa EE *et al.* The human papillomavirus type 11 upstream regulatory
1078 region triggers hair-follicle-specific gene expression in transgenic mice. *J Invest Dermatol*
1079 1999;**112**:893–8.
- 1080 Schiffman M, Herrero R, Desalle R *et al.* The carcinogenicity of human papillomavirus types reflects
1081 viral evolution. *Virology* 2005;**337**:76–84.
- 1082 Schmeink CE, Massuger LF a. G, Lenselink CH *et al.* Prospective follow-up of 2,065 young
1083 unscreened women to study human papillomavirus incidence and clearance. *Int J Cancer*
1084 2013;**133**:172–81.
- 1085 Schmitt A, Rochat A, Zeltner R *et al.* The primary target cells of the high-risk cottontail rabbit
1086 papillomavirus colocalize with hair follicle stem cells. *J Virol* 1996;**70**:1912–22.
- 1087 Schneider-Gädicke A, Schwarz E. Different human cervical carcinoma cell lines show similar
1088 transcription patterns of human papillomavirus type 18 early genes. *EMBO J* 1986;**5**:2285–
1089 92.
- 1090 Schulte MB, Andino R. Single-cell analysis uncovers extensive biological noise in poliovirus
1091 replication. *J Virol* 2014;**88**:6205–12.
- 1092 Schwartz S. Papillomavirus transcripts and posttranscriptional regulation. *Virology* 2013;**445**:187–
1093 96.

- 1094 Scinicariello F, Soza I, Brasky KM *et al.* DNA Sequence of a Novel Papillomavirus (CCPV1) from an
1095 Outbreak of Focal Epithelial Hyperplasia-like Disease in a Common Chimpanzee (*Pan*
1096 *Troglodytes*) Colony., 1997.
- 1097 Seedorf K, Krämer G, Dürst M *et al.* Human papillomavirus type 16 DNA sequence. *Virology*
1098 1985;**145**:181–5.
- 1099 Sethi S, Ali A, Ju X *et al.* An update on Heck's disease-a systematic review. *J Public Health (Oxf)*
1100 2022;**44**:269–85.
- 1101 Shope RE, Hurst EW. INFECTIOUS PAPILOMATOSIS OF RABBITS : WITH A NOTE ON THE
1102 HISTOPATHOLOGY. *J Exp Med* 1933;**58**:607–24.
- 1103 Sibley CR, Blazquez L, Ule J. Lessons from non-canonical splicing. *Nat Rev Genet* 2016;**17**:407–21.
- 1104 Sichero L, Ferreira S, López RVM *et al.* Prevalence of human papillomavirus 6 and 11 variants in
1105 recurrent respiratory papillomatosis. *J Med Virol* 2021;**93**:3835–40.
- 1106 Sichero L, Ferreira S, Trottier H *et al.* High grade cervical lesions are caused preferentially by non-
1107 European variants of HPVs 16 and 18. *Int J Cancer* 2007;**120**:1763–8.
- 1108 Sichero L, Franco EL, Villa LL. Different P105 promoter activities among natural variants of human
1109 papillomavirus type 18. *J Infect Dis* 2005;**191**:739–42.
- 1110 Silvestre RVD, de Souza AJS, Júnior ECS *et al.* First New World Primate Papillomavirus Identification
1111 in the Atlantic Forest, Brazil: \textlessi\textgreaterAlouatta guariba papillomavirus
1112 1\textless/i\textgreater. *Genome Announcements* 2016;**4**:e00725-16.
- 1113 Smotkin D, Prokoph H, Wettstein FO. Oncogenic and nononcogenic human genital
1114 papillomaviruses generate the E7 mRNA by different mechanisms. *J Virol* 1989;**63**:1441–7.
- 1115 Smotkin D, Wettstein FO. Transcription of human papillomavirus type 16 early genes in a cervical
1116 cancer and a cancer-derived cell line and identification of the E7 protein.
1117 *Proc Natl Acad Sci USA* 1986;**83**:4680–4.
- 1118 Sorokin II, Vassilenko KS, Terenin IM *et al.* Non-Canonical Translation Initiation Mechanisms
1119 Employed by Eukaryotic Viral mRNAs. *Biochemistry Moscow* 2021;**86**:1060–94.
- 1120 Sotlar K, Stubner A, Diemer D *et al.* Detection of high-risk human papillomavirus E6 and E7
1121 oncogene transcripts in cervical scrapes by nested RT-polymerase chain reaction. *J Med*
1122 *Virol* 2004;**74**:107–16.
- 1123 Stacey SN, Jordan D, Williamson AJ *et al.* Leaky scanning is the predominant mechanism for
1124 translation of human papillomavirus type 16 E7 oncoprotein from E6/E7 bicistronic mRNA.
1125 *J Virol* 2000;**74**:7284–97.
- 1126 Stoppler MC, Ching K, Stoppler H *et al.* Natural variants of the human papillomavirus type 16 E6
1127 protein differ in their abilities to alter keratinocyte differentiation and to induce P53
1128 degradation. *J Virol* 1996;**70**:6987–93.

- 1129 Strickley JD, Messerschmidt JL, Awad ME *et al.* Immunity to commensal papillomaviruses protects
1130 against skin cancer. *Nature* 2019;**575**:519–22.
- 1131 Strickley JD, Messerschmidt JL, Jenson AB *et al.* Reply to: Beta human papillomaviruses and skin
1132 cancer. *Nature* 2020;**588**:E22–3.
- 1133 Talia KL, McCluggage WG. Seborrhoeic Keratosis-like Lesions of the Cervix and Vagina: Report of a
1134 New Entity Possibly Related to Low-risk Human Papillomavirus Infection. *Am J Surg Pathol*
1135 2017;**41**:517–24.
- 1136 Talia KL, Rahimi S, Hawkes D *et al.* HPV42-associated Seborrhoeic Keratosis-like Lesion of the
1137 Cervix: First Reported Case With High-grade Morphology. *Int J Gynecol Pathol*
1138 2022;**41**:649–54.
- 1139 Tang S, Tao M, McCoy JP *et al.* The E7 oncoprotein is translated from spliced E6*I transcripts in
1140 high-risk human papillomavirus type 16- or type 18-positive cervical cancer cell lines via
1141 translation reinitiation. *J Virol* 2006;**80**:4249–63.
- 1142 Thomas M, Myers MP, Massimi P *et al.* Analysis of Multiple HPV E6 PDZ Interactions Defines Type-
1143 Specific PDZ Fingerprints That Predict Oncogenic Potential. *PLoS Pathog*
1144 2016;**12**:e1005766.
- 1145 Trammel J, Amusan O, Hultgren A *et al.* Epidermal growth factor receptor-dependent stimulation
1146 of differentiation by human papillomavirus type 16 E5. *Virology* 2024;**590**:109952.
- 1147 Trottier H, Mahmud S, Prado JCM *et al.* Type-specific duration of human papillomavirus infection:
1148 implications for human papillomavirus screening and vaccination. *J Infect Dis*
1149 2008;**197**:1436–47.
- 1150 Van Doorslaer K. Revisiting Papillomavirus Taxonomy: A Proposal for Updating the Current
1151 Classification in Line with Evolutionary Evidence. *Viruses* 2022;**14**:2308.
- 1152 Van Ranst M, Fuse A, Sobis H *et al.* A papillomavirus related to HPV type 13 in oral focal epithelial
1153 hyperplasia in the pygmy chimpanzee. *J Oral Pathol Med* 1991;**20**:325–31.
- 1154 Vanderbilt C, Brenn T, Moy AP *et al.* Association of HPV42 with digital papillary adenocarcinoma
1155 and the use of in situ hybridization for its distinction from acral hidradenoma and
1156 diagnosis at non-acral sites. *Mod Pathol* 2022a;**35**:1405–10.
- 1157 Vanderbilt CM, Bowman AS, Middha S *et al.* Defining Novel DNA Virus-Tumor Associations and
1158 Genomic Correlates Using Prospective Clinical Tumor/Normal Matched Sequencing Data. *J*
1159 *Mol Diagn* 2022b;**24**:515–28.
- 1160 Van Doorslaer K, Li Z, Xirasagar S *et al.* The Papillomavirus Episteme: a major update to the
1161 papillomavirus sequence database. *Nucleic Acids Res* 2017;**45**:D499–506.
- 1162 Vassilakos P, Catarino R, Bougel S *et al.* Use of swabs for dry collection of self-samples to detect
1163 human papillomavirus among Malagasy women. *Infect Agent Cancer* 2016;**11**:13.

- 1164 Villa LL, Sichero L, Rahal P *et al.* Molecular variants of human papillomavirus types 16 and 18
1165 preferentially associated with cervical neoplasia. *J Gen Virol* 2000;**81**:2959–68.
- 1166 de Villiers E-M, Fauquet C, Broker TR *et al.* Classification of papillomaviruses. *Virology*
1167 2004;**324**:17–27.
- 1168 Walboomers JM, Jacobs MV, Manos MM *et al.* Human papillomavirus is a necessary cause of
1169 invasive cervical cancer worldwide. *J Pathol* 1999;**189**:12–9.
- 1170 Webb Strickland S, Brimer N, Lyons C *et al.* Human Papillomavirus E6 interaction with cellular PDZ
1171 domain proteins modulates YAP nuclear localization. *Virology* 2018;**516**:127–38.
- 1172 Werness BA, Levine AJ, Howley PM. Association of Human Papillomavirus Types 16 and 18 E6
1173 Proteins with p53. *Science* 1990;**248**:76–9.
- 1174 Wile UJ, Kingery LB. THE ETIOLOGY OF COMMON WARTS: PRELIMINARY REPORT OF AN
1175 EXPERIMENTAL STUDY. *JAMA* 1919;**73**:970.
- 1176 Willemsen A, Bravo IG. Origin and evolution of papillomavirus (onco)genes and genomes.
1177 *Philosophical Transactions of the Royal Society B: Biological Sciences* 2019;**374**:20180303.
- 1178 Willemsen A, Félez-Sánchez M, Bravo IG. Genome Plasticity in Papillomaviruses and De Novo
1179 Emergence of E5 Oncogenes. Baptiste E (ed.). *Genome Biology and Evolution*
1180 2019;**11**:1602–17.
- 1181 Wolf H, Zur Hausen H, Klein G *et al.* Attempts to detect virus-specific DNA sequences in human
1182 tumors. III. Epstein-Barr viral DNA in non-lymphoid nasopharyngeal carcinoma cells. *Med*
1183 *Microbiol Immunol* 1975;**161**:15–21.
- 1184 Wood CE, Borgerink H, Register TC *et al.* Cervical and Vaginal Epithelial Neoplasms in Cynomolgus
1185 Monkeys. *Vet Pathol* 2004;**41**:108–15.
- 1186 Wood CE, Chen Z, Cline JM *et al.* Characterization and Experimental Transmission of an Oncogenic
1187 Papillomavirus in Female Macaques. *J Virol* 2007;**81**:6339–45.
- 1188 Xi LF, Kiviat NB, Hildesheim A *et al.* Human Papillomavirus Type 16 and 18 Variants: Race-Related
1189 Distribution and Persistence. *JNCI: Journal of the National Cancer Institute* 2006;**98**:1045–
1190 52.
- 1191 Xi LF, Koutsky LA, Galloway DA *et al.* Genomic variation of human papillomavirus type 16 and risk
1192 for high grade cervical intraepithelial neoplasia. *J Natl Cancer Inst* 1997;**89**:796–802.
- 1193 Yamada T, Manos MM, Peto J *et al.* Human papillomavirus type 16 sequence variation in cervical
1194 cancers: a worldwide perspective. *J Virol* 1997;**71**:2463–72.
- 1195 Yang EJ, Quick MC, Hanamornroongruang S *et al.* Microanatomy of the cervical and anorectal
1196 squamocolumnar junctions: a proposed model for anatomical differences in HPV-related
1197 cancer risk. *Mod Pathol* 2015;**28**:994–1000.

- 1198 Yu S, Kim VN. A tale of non-canonical tails: gene regulation by post-transcriptional RNA tailing. *Nat*
1199 *Rev Mol Cell Biol* 2020;**21**:542–56.
- 1200 Zanier K, Charbonnier S, Sidi AOMO *et al.* Structural basis for hijacking of cellular LxxLL motifs by
1201 papillomavirus E6 oncoproteins. *Science* 2013;**339**:694–8.
- 1202 Zehbe I, Wilander E, Delius H *et al.* Human papillomavirus 16 E6 variants are more prevalent in
1203 invasive cervical carcinoma than the prototype. *Cancer Res* 1998;**58**:829–33.
- 1204 Zhang H, Zhang L, Lin A *et al.* Algorithm for optimized mRNA design improves stability and
1205 immunogenicity. *Nature* 2023;**621**:396–403.
- 1206 Zhao K-N, Gu W, Fang NX *et al.* Gene codon composition determines differentiation-dependent
1207 expression of a viral capsid gene in keratinocytes in vitro and in vivo. *Mol Cell Biol*
1208 2005;**25**:8643–55.
- 1209 Zhao K-N, Liu WJ, Frazer IH. Codon usage bias and A+T content variation in human papillomavirus
1210 genomes. *Virus Res* 2003;**98**:95–104.
- 1211 Zhou J, Liu WJ, Peng SW *et al.* Papillomavirus capsid protein expression level depends on the
1212 match between codon usage and tRNA availability. *J Virol* 1999;**73**:4972–82.
- 1213

ORIGINAL UNEDITED MANUSCRIPT