

1 **The landscape of persistent human DNA viruses in femoral bone**

2 Mari Toppinen¹, Diogo Pratas^{1,2,3}, Elina Väisänen¹, Maria Söderlund-Venermo¹, Klaus Hedman^{1,4},
3 Maria F. Perdomo^{1*} and Antti Sajantila^{5, 6 *}

4

5 ¹ Department of Virology, University of Helsinki, Finland

6 ² Department of Electronics, Telecommunications and Informatics, University of Aveiro, Portugal

7 ³ Institute of Electronics and Informatics Engineering of Aveiro, University of Aveiro, Portugal

8 ⁴ HUSLAB, Helsinki University Hospital, Finland

9 ⁵ Department of Forensic Medicine, University of Helsinki, Finland

10 ⁶ Forensic Medicine Unit, Finnish Institute of Health and Welfare, Finland

11

12 * Shared last authorship

13

14 **Keywords:** DNA viruses, femoral bone, NGS, human provenance, parvovirus B19 genotype 2

15

16 corresponding authors: Maria F. Perdomo maria.perdomo@helsinki.fi

17 Antti Sajantila antti.sajantila@helsinki.fi

18

19 **Abstract**

20 The imprints left by persistent DNA viruses in the tissues can testify to the changes driving virus
21 evolution as well as provide clues on the provenance of modern and ancient humans. However, the
22 history hidden in skeletal remains is practically unknown, as only parvovirus B19 and hepatitis B
23 virus DNA have been detected in hard tissues so far. Here, we investigated the DNA prevalences of
24 38 viruses in femoral bone of recently deceased individuals. To this end, we used quantitative PCRs
25 and a custom viral targeted enrichment followed by next-generation sequencing. The data was
26 analyzed with a tailor-made bioinformatics pipeline. Our findings revealed bone to be a much richer
27 source of persistent DNA viruses than earlier perceived, discovering ten additional ones, including
28 several members of the herpes- and polyomavirus families, as well as human papillomavirus 31 and
29 torque teno virus. Remarkably, many of the viruses found have oncogenic potential and/or are likely
30 to reactivate in the elderly and immunosuppressed individuals. Thus, their persistence warrants
31 careful evaluation of their clinical significance and impact on bone biology. Our findings open new
32 frontiers for the study of virus evolution from ancient relics as well as provide new tools for the
33 investigation of human skeletal remains in forensic and archeological contexts.

34 **Introduction**

35 DNA viruses commonly establish persisting infections in humans, remaining thus, their genetic
36 material imprinted in the tissues throughout life. These DNAs exhibit phylogeographies resembling
37 those of humans[1], pointing to shared evolutionary and dispersal paths that can, together, add to
38 the understanding of human population history and migrations[2–8]. In addition, the specific
39 geographical distributions of DNA viruses, in both global[2,9] and local[10–12] scales, may provide
40 new insights into the origins of unidentified individuals.

41 In this regard, the most widely studied virus is JC polyomavirus (JCPyV), a highly prevalent pathogen
42 with three main phylogenetic clusters spread across Africa, Asia, and Europe. JCPyV's genotype-
43 specific global spread has been proposed as an indicator of the provenances of both modern[12]
44 and ancient humans[13–15]. Although its timescale of evolution has been debated[16–18], recent
45 work by Forni et al.,[8] based on ~1100 worldwide strains, supports co-dispersal of this virus with
46 major human migratory routes as well as its co-divergence with human mitochondrial and nuclear
47 markers.

48 In addition to the spatial and temporal dimensions, viruses also display intriguing age-dependent
49 distributions[19–21]. This is the case with genotype 2 of parvovirus B19, the traces of which went
50 missing in Europe around 1970s. Thus, contemporary DNA findings of this virus variant in tissues
51 are confined to elderly individuals or historic human remains[22].

52 Despite these encouraging data, the utility of the viral DNAs has not been fully addressed in forensic
53 and anthropological settings. Reasons for this include the lack of a comprehensive picture of the
54 overall distribution of viral DNAs in various organs, limiting the use of some human tissues in versatile
55 scenarios, and most importantly, the fact that their prevalences in bones and teeth are almost entirely
56 unknown.

57 Of the latter, we were the first to detect parvovirus B19 DNA in bones from soldiers of World War
58 II[23], followed by findings of this same virus[24] and of hepatitis B[25,26] virus in archaeological
59 remains dated from the Neolithic to medieval times. These singular discoveries confirm the presence
60 and preservation of viral DNA in hard tissues and call for the investigation of the full landscape of
61 viral DNAs that here persist.

62 In this study, we searched for 38 highly prevalent DNA viruses in the femoral bones of recently
63 deceased individuals. To this end, we used in-house quantitative PCRs (qPCRs) and virus-targeted
64 enrichment coupled with next-generation sequencing (NGS).

65 Our findings significantly expand the present knowledge on the virome of human bone, opening new
66 frontiers for the study of virus evolution as well new tools for the investigation of forensic and
67 anthropological cases.

68 **Materials and methods**

69 *Study subjects*

70 The study cohort consisted of 27 deceased individuals of Finnish origin, aged 36 to 85 years (mean
71 67.6) with a male:female ratio of 19:8. The postmortem interval ranged from 4 to 30 days (mean 8).
72 The manners of death were disease (n=19), occupational disease (n=1), injury (n=5), and suicide
73 (n=2). In none of the cases the medical records revealed the use of immunosuppressants or pre-
74 conditions related to infectious-diseases, except for one with a history of herpes-zoster a few weeks
75 before death. The mode and cause of death are presented in Figure 2. The study protocol was
76 reviewed by the Ethics Committee of Helsinki and Uusimaa Hospital District (approval
77 164/13/03/00/2014).

78 *Specimen collection and preparation*

79 A 5-10 cm piece of the femoral diaphysis was collected and cut longitudinally with an oscillating tip
80 saw. The bone was cleaned using a toothbrush and washed sequentially in distilled water, 0.1 %
81 sodium hypochlorite, and 96.1 w/w ethanol. The bones were let to dry for 5 to 7 days at room
82 temperature under flow in a laminar hood. From each individual, the external and internal surfaces
83 of femoral bone were sampled using a dentist drill (Schick Qube) except for three cases, in which
84 the bone was cryomilled (Spex 6775 Freezer/mill; Spex). The sample preparations and drillings were
85 performed in enclosed dedicated facilities, using full-body protection suits.

86 *DNA extraction*

87 Approximately 0.1-0.4g of bone powder were lysed and extracted as specified before[23] and eluted
88 in a final volume of 50-100 µl. The DNA extractions were performed in enclosed facilities, dedicated
89 to the extraction of human DNA. No viral work is performed in these rooms.

90 The total DNA was quantitated with Quant-iT™ PicoGreen™ dsDNA Assay Kit (Thermo Fisher
91 Scientific).

92 The quality of the total DNA was evaluated with the Genomic DNA Reagent Kit in a LabChip GX
93 Instrument (Perkin Elmer).

94 *Quantitative PCRs*

95 The quantifications of parvovirus B19, Merkel cell, JC and BK polyomaviruses, and nine human
96 herpesviruses as well as of the human single-copy gene RNase P were performed with in-house
97 quantitative PCRs (qPCRs) as described[27–30] [Pyöriä *et al.* in press]. The quantification of
98 hepatitis B virus and human papillomavirus type 31 DNAs were performed with commercial kits
99 (Hepatitis B Virus PCR Kit, GeneProof; Genesig Human papillomavirus 31 Standard kit,
100 PrimerDesign) according to the manufacturer instructions.

101 For the quantification of torque teno virus, a qPCR was optimized[31,32] to amplify and detect the
102 conserved untranslated region (UTR) of the virus using the following primers: AMTS fwd (5'-
103 GTGCCGNAGGTGAGTTTA-'3), AMTAS rev (5'-AGCCCGGCCAGTCC-'3), AMTASgr4 rev (5'-
104 AGCCCGGCCAGACC-'3) and AMTPTU probe (5'-FAM-TCAAGGGGCAATTCGGGCT-BHQ1-'3).
105 The qPCR reaction consisted of 1x Maxima probe qPCR Master mix (Thermo Fisher Scientific) with
106 0.03 µM of ROX passive reference dye, 0.5 µM of each of the primers, 0.4 µM of the probe, 5 µl of
107 the template, and nuclease-free water to a final volume of 25 µl. After initial denaturation at +95°C
108 for 10 min, the qPCR cycles were +95 °C for 15 s and +55 °C for 1 min for 45 cycles.

109 The qPCR amplicons were 63 to 154 nucleotides in length.

110 Plasmid dilution series were used in all the qPCR runs as positive controls and to create standard
111 curves for quantification. The plasmids of parvovirus B19, human herpesvirus 1-8, Merkel cell, JC
112 and BK polyomaviruses, and RNaseP are described elsewhere[33][Pyöriä *et al.* in press]. For torque
113 teno virus, a plasmid, named 10B, containing 1184 nucleotides of the virus was cloned from a healthy
114 blood donor's plasma (GenBank MT448658).

115 The virus amplifications were completed with AriaMx Real-Time PCR System except for those of
116 torque teno virus, RNaseP, and human papillomavirus 31 that were analyzed with Stratagene
117 Mx3005P qPCR System (both Agilent).

118 The qPCR mixes, sample handling, plasmid dilutions, and amplifications, were performed each in
119 completely separate rooms, following strict protocols and work-flows to prevent contamination.
120 Negative controls (PCR-grade water) were included in every step starting from the DNA extraction.

121 *PCR inhibition tests*

122 PCR efficiency due to carryover of inhibitors following DNA extraction was controlled using DNA
123 extracts from bone together with or in parallel to a pre-quantified RNaseP plasmid[33].

124 To evaluate the performance of different polymerases and the quantification accuracy, four different
125 commercial master mixes were tested. From this, Maxima probe qPCR master mix was selected
126 and used throughout the study.

127 The impact of residual EDTA on the qPCR performance was examined by testing varying EDTA
128 concentrations in the lysis buffer (0.5 to 500 nM) and by adding excess EDTA (0.05 to 50 nM) to the
129 qPCR reaction. In addition, pre-quantified plasmid dilutions were extracted following the bone
130 extraction protocol (i.e. lysis buffer containing 500 nM EDTA).

131 The impact of residual Ca²⁺-ions from bone in the extracts was investigated by addition of MgCl₂ to
132 the qPCR reactions (in final concentrations of 4 to 7 mM).

133

134 *Library preparation, viral enrichment, and sequencing*

135 The sequencing libraries were prepared on 10 to 1000 ng of total DNA using the KAPA HyperPlus
136 library preparation kit (Roche), following the manufacturer protocol with two modifications: 1)
137 mechanical fragmentation with a Covaris E220 of the DNA with target fragments of 200 nt and 2) the
138 use of xGen Dual Index UMI Adapters (Integrated DNA Technologies).

139 After sonication, the fragment length distributions were analyzed with the DNA High Sensitivity
140 Reagent Kit in a LabChip GX Instrument (Perkin Elmer).

141 The viral enrichment was performed using two consecutive rounds of hybridization on individual
142 samples following recommendations for low input DNA (MyBaits v4 kit; Arbor Biosciences). For each
143 library, 200 ng per round of biotinylated RNA-baits were used. The baits were 100 nt in length and
144 designed with 2X tiling (Supplementary Table 1 for a list of viruses). xGen Universal Blockers-TS
145 Mix (Integrated DNA Technologies) were used to block unspecific binding to the adapters during
146 hybridization.

147 During library preparation and viral enrichment, the libraries were amplified 3x13-25 cycles. The
148 clean-up steps were performed with either KAPA Pure Beads (Roche) or MinElute PCR Purification
149 Kit (Qiagen).

150 The enriched libraries were quantified with KAPA Library Quantification Kit (Roche) using Stratagene
151 3005P qPCR System (Agilent) and subsequently pooled for sequencing on NovaSeq 6000 (SP
152 PE151 reagent kit; Illumina).

153 *NGS data analysis*

154 The viral genomic sequences were reconstructed after removal of PCR duplicates, using a
155 customized bioinformatics pipeline (TRACESPipe; [Pratas *et al.* in revision; available for download
156 at <https://github.com/viromelab/tracespipe>).

157 The consensus, as well as single sequences (when in low coverage), were confirmed by BLAST
158 (NCBI). The highest similarity was used to classify the virus genotype. For parvovirus B19, the
159 sequences covering >70 % of the viral genome (n=7) were aligned with previously published full or
160 near-full length sequences in EMBL-EBI Clustal Omega and analyzed with Bioedit v.7.2.5 (Ibis
161 Biosciences).

162 For the following viruses, the consensus sequences will be available in GenBank with respective
163 accession numbers: parvovirus B19 (7 sequences; MT410184-MT410190); human papillomavirus
164 type 31 (MT410191); hepatitis B virus (MT410192); Merkel cell polyomavirus (MT410193).

165

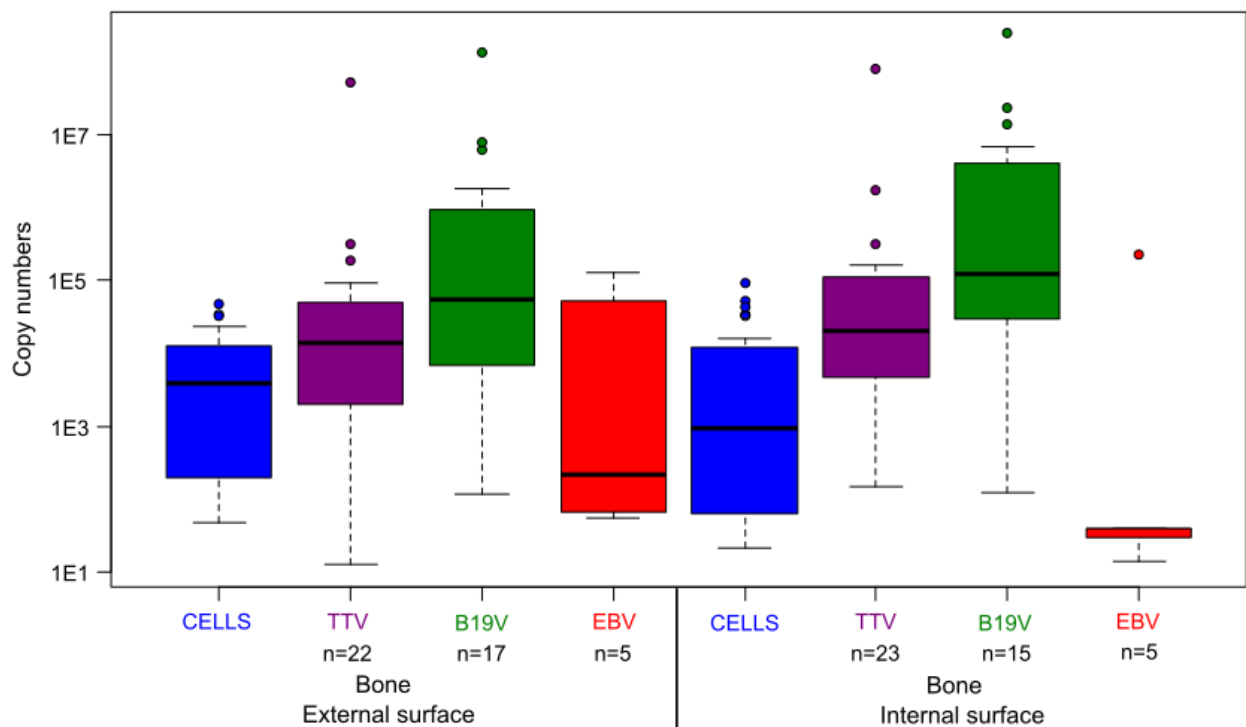
166 *Statistical analysis*

167 The differences in viral or cell copy numbers in the external and internal surfaces of femoral bone
168 were calculated with Student's t-test in RStudio (version 1.0.153).

169 **Results**

170 *Viral DNA prevalences in the external and internal surfaces of the femoral bone.*

171 To determine the most optimal site for sampling, we first investigated the differences in prevalence
172 and quantity of persisting viral DNAs in the external and internal surfaces of 27 femoral bones (Fig
173 1). To this end, we examined three ubiquitous viruses (torque teno, parvovirus B19, and Epstein-
174 Barr virus) known to infect >70% of the global population and to persist in several soft tissues in the
175 body. To control for bias by uneven amounts of DNA in the extracts, we performed in parallel a qPCR
176 for the human single-copy gene RNase P. We found no significant differences in the viral DNA
177 prevalences ($p > 0.1$) nor copy numbers (p values > 0.5) between these two surfaces. Hence,
178 subsequent analyses were performed only on samples taken from the external surface.



179

180 **Fig 1. Viral loads in the external and internal surfaces of the femoral bone.** The numbers of cells as well
181 as the DNA copies of torque teno virus (TTV), parvovirus B19 (B19V), and Epstein-Barr virus (EBV) from the
182 external and internal surfaces of femoral bone were determined. No significant differences between these two
183 locations were observed ($p > 0.5$). The cell counts are expressed per 1 μ l of DNA extract and the viral DNA
184 copy numbers per 1E6 cells.

185 *Viruses are highly prevalent in human femoral bone*

186 We then investigated the prevalences of altogether 38 persistent virus genomes using targeted
187 enrichment and confirmatory qPCRs. Overall, the 27 study subjects harbored on average 2.6 virus-

188 types in their femoral bones, with a maximum of seven in one individual. Altogether, we detected 12
 189 different virus-types in 92.6 % of the bones, with only two individuals (> 60 years of age) being
 190 completely negative for all viruses tested. The viral findings are presented in Table 1 and Figure 2.

191 The viruses most prevalent in bone were torque teno virus and parvovirus B19 with genoprevalences
 192 of 81.5% and 74.1%, respectively (Table 1). The third most prevalent was Merkel cell polyomavirus
 193 at 33.3%, followed by Epstein-Barr virus (25.9%), human papillomavirus (22.2%), human
 194 herpesvirus 7 (18.5%) and JC polyomavirus (14.8%). Other viral sequences detected were of herpes
 195 simplex 1, varicella-zoster, cytomegalovirus, human herpesvirus 6B, and hepatitis B virus.

196 The median viral copy numbers per one million cells were 1.9E4 for torque teno virus, 2.4E3 for
 197 parvovirus B19, 8.7E3 for Merkel cell polyomavirus, and 2.1E1 for Epstein-Barr virus (Table 1).

198 The most common co-occurrences were of parvovirus B19 and torque teno virus, found in 70.3% of
 199 the samples. Interestingly, we detected in the bone of one individual three cancer-associated viruses:
 200 hepatitis B virus, human papillomavirus type 31, and Merkel cell polyomavirus.

201 **Table 1. Viral DNA prevalence in human femoral bone**

Family	Virus	Genome size (kb)	Geno-prevalence	NGS+ cases	qPCR+ cases	Breadth coverage (range or value of single sample)	Viral load (copies/million cells)
Parvoviridae	B19V	5.6	74.1 %	17	16	4.5-100.0 %	2.4E3
Herpesviridae	HSV1	152	3.7 %	1	1	3.4 %	4.9E1
	VZV	125	3.7 %	1	1	13.6 %	2.0E0
	EBV	170	25.9 %	6	4	1.0-3.3 %	2.1E1
	CMV	236	3.7 %	0	1	-	9.8E1
	HHV6B	162	11.1 %	3	0	2.0-3.1 %	-
	HHV7	150	18.5 %	5	0	4.8-12.0 %	-
Polyomaviridae	JCPyV	5.1	14.8 %	3	2	3.7-14.8 %	2.7E1
	MCPyV	5.4	33.3 %	9	4	1.4-58.6 %	8.7E3
Papillomaviridae	HPV	8	22.2 %	6	1*	1.8-89.6 %	8.7E3
Hepadnaviridae	HBV	3.2	7.4 %	2	1	15.6-44.5 %	1.7E3
Anelloviridae	TTV	3.8	81.5 %	9	22	2.1-56.3 %	1.9E4

202 * Only HPV type 31 qPCR was performed. B19V: parvovirus B19; HSV1: herpes simplex virus-1; VZV: varicella-zoster virus; EBV:
 203 Epstein-Barr virus; CMV: cytomegalovirus; HHV6B: human herpesvirus 6B; HHV7: human herpesvirus 7; JCPyV: JC polyomavirus;
 204 MCPyV: Merkel cell polyomavirus; HPV: human papillomavirus; HBV: hepatitis B virus; TTV: torque teno virus

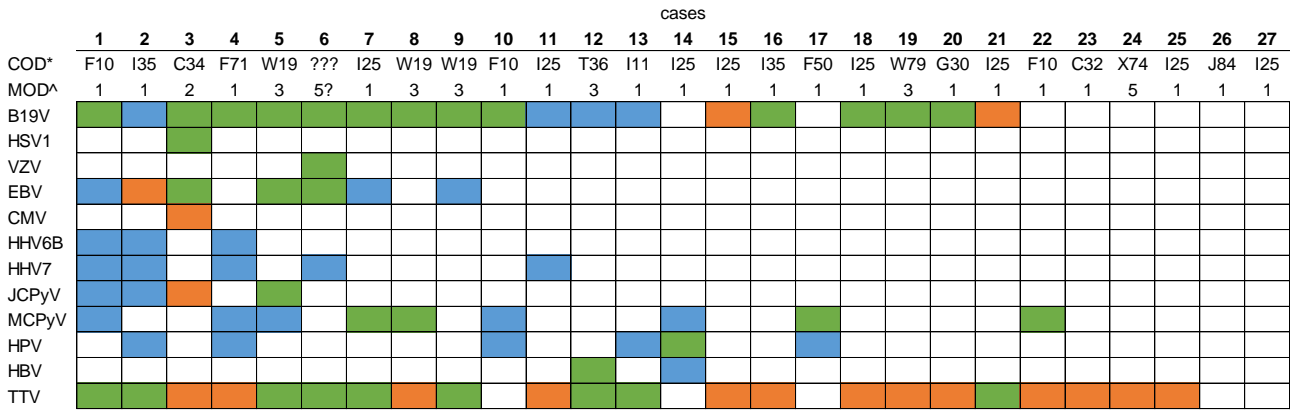
205 We verified the accuracy of the qPCR results by examining the patterns of DNA fragmentation in the
 206 extracts and by evaluating the impact of potential inhibitors carried over during lysis and extraction.
 207 The quality of the total DNA in each sample was analyzed with a LabChip GX Instrument, which
 208 revealed 100 to 500 nt fragments in addition to intact genomic DNA (≤ 40 kb; Fig 3). We found no
 209 PCR inhibition accountable to excess EDTA or Ca^{2+} in the extracts.

210

211

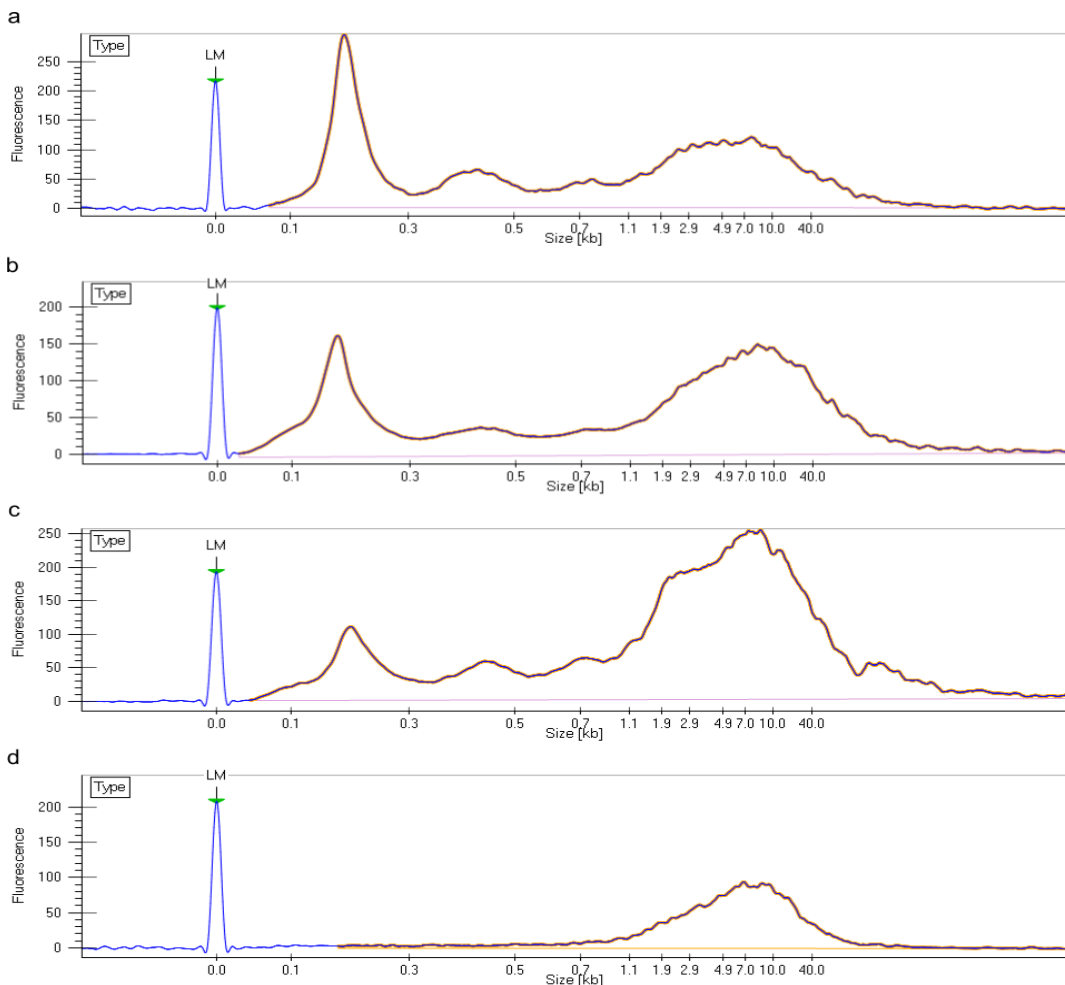
212

213



214 **Fig 2. Viral DNA occurrence by NGS and qPCR in bone per study subject.** The viral findings are
 215 presented as green (NGS+, qPCR+), yellow (NGS+, qPCR-), orange (NGS-, qPCR+), and white
 216 (NGS-, qPCR-). *The cause of death (COD) is given according to the WHO ICD10 classification.
 217 ^The manner of death (MOD) is presented as 1=disease, 2=occupational disease, 3=injury,
 218 5=suicide. The results are presented from left to right according to the highest and lowest virus
 219 prevalences in bone. B19V= parvovirus B19; HSV1=herpes simplex virus-1; VZV= varicella-zoster
 220 virus; EBV= Epstein-Barr virus; CMV= cytomegalovirus; HHV6B= human herpesvirus 6B; HHV7=
 221 human herpesvirus 7; JCPyV= JC polyomavirus; MCPyV= Merkel cell polyomavirus; HPV= human
 222 papillomavirus; HBV= hepatitis B virus; TTV= torque teno virus.

223

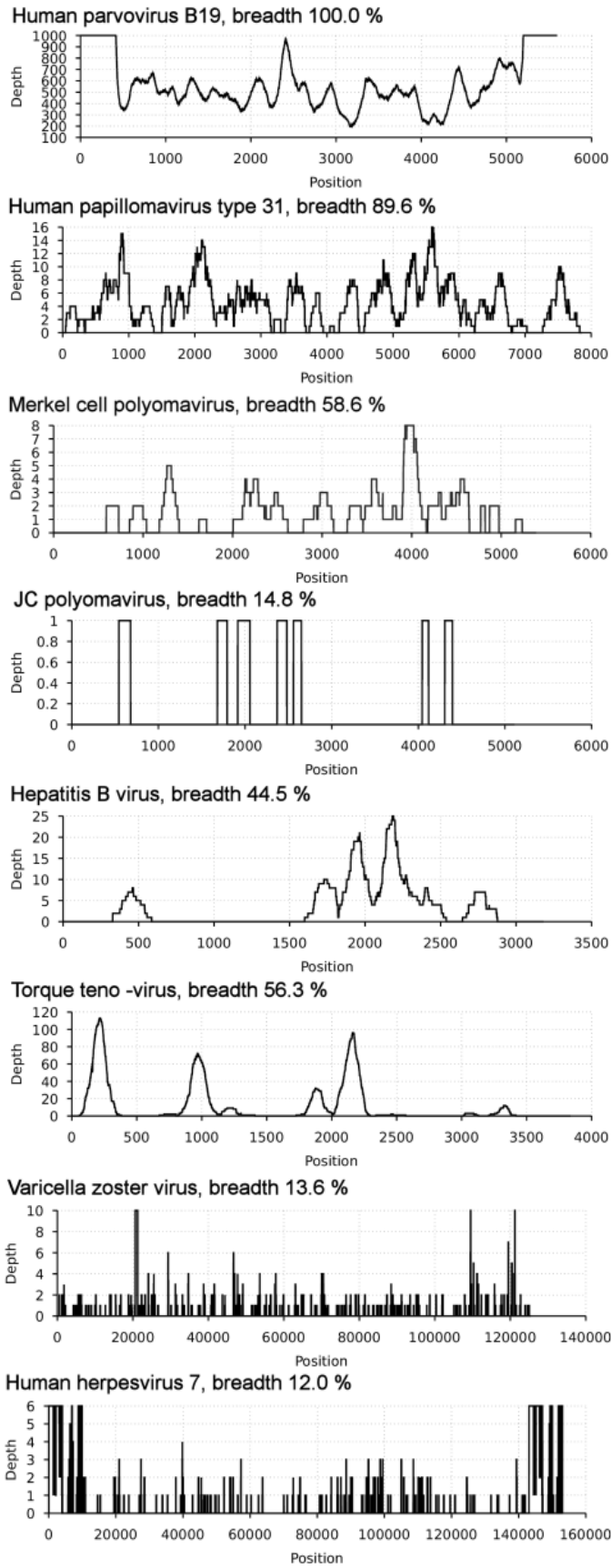


224 **Fig 3. DNA fragment analysis of total DNA.** Fragment length distribution of patterns of genomic
 225 DNA for four representative samples as established with LabChip GX.

226 *Viral capture and sequence analysis*

227 As established from the qPCR results, the persisting viral DNA quantities were far below 1% of total
228 DNA present in a sample. Thus, to enrich this fraction, we performed in-solution capture with
229 biotinylated RNA oligonucleotides prior to sequencing in Novaseq 6000. Subsequently, we analyzed
230 the NGS data with a custom pipeline, TRACESPipe (Pratas *et al.* in revision), which reconstructs the
231 viral sequences using both reference-based alignment and *de-novo* assembly.

232 We reconstructed a total of 15 viral genomic sequences, with a minimum of 15% breadth coverage.
233 The highest qualities were attained for parvovirus B19 (n=7; average breadth coverage 86.1%),
234 human papillomavirus (n=1; breadth coverage 89.6%), Merkel cell polyomavirus (n=1; breath
235 coverage 58.6%), JC polyomavirus (n=1; breadth coverage 14.8%), hepatitis B virus (n=2; breadth
236 coverages 44.5% and 15.6%), and torque teno virus (n=3; breadth coverages 56.3%, 30.6% and
237 24.9%). The breadth and depth coverages of representative viruses are presented in Figure 4.



238

239 **Fig 4. Coverage profiles of reconstructed viral DNA sequences.** NGS coverage profiles (breadth and
 240 depth) of representative viruses.

241 By NCBI BLAST, we found close relation of the reconstructed genomes to previously published
242 sequences. We recovered seven full/near-full B19V genomic sequences, three of which were of
243 genotype 1 (95.16% to 99.61% similarity to AY504945.1) and four of genotype 2 (90.1% to 98.6%,
244 similarity to AB550331.1). The latter genotype represents an extinct form of this virus and was found
245 in the present study exclusively in individuals older than 68 years. Of this genotype, only a single
246 full-length sequence has ever been published (AB550331.1). We found that this genotype's hairpins
247 present similar flip and flop configurations to those of genotype 1.

248 The human papillomavirus sequence showed 99.1% similarity to type 31 (KU298889.1); Merkel cell
249 polyomavirus 98.7% to the European/Caucasian type (KF266963.1) and JC polyomavirus 100.0%
250 to genotype 1 (MF662198.1). The two hepatitis B virus sequences matched 99.4% and 99.6% to
251 genotypes D (JX898691.1) and A (MN507849.1), respectively; and the torque teno virus sequences
252 showed 90.2%, 99.0% and 96.2% similarity to strains AF122920.1, KT163880.1 and FR751497.1,
253 respectively.

254 Moreover, we confirmed by BLAST unique sequences (mean length 100 nt) mapping to the following
255 viruses: herpes simplex 1 (1 case, 28 reads), varicella-zoster (1 case, 362 reads), Epstein Barr (6
256 cases, altogether 270 reads), human herpesvirus 6B (3 cases, total of 66 reads), human herpesvirus
257 7 (5 cases, total of 274 reads), JC polyomavirus (3 additional cases, total of 13 reads), Merkel cell
258 polyomavirus (8 additional cases, total of 15 reads), human papillomavirus (5 additional cases, total
259 of 14 reads), torque teno virus (6 additional cases, total of 26 reads), and parvovirus B19 (10
260 additional cases, total of 673 reads).

261 **Discussion**

262 The analysis of viral DNAs shows great potential as complementary markers for human identification
263 as well as for estimation of provenance and migration. However, one limitation of their use in these
264 contexts is insufficient knowledge of the landscape of viruses persisting in the host, in particular in
265 human bone, with only parvovirus B19[23,24] and hepatitis B[25] virus having been detected in this
266 tissue so far.

267 To this end, we systematically explored the prevalences of 38 ubiquitous viruses in human femoral
268 bone. We discovered an unprecedented number of viral DNAs, detecting up to seven per individual.
269 Besides the already known, we report here on ten new viruses including several members of the
270 herpesvirus family (herpes simplex-1, varicella-zoster, Epstein-Barr, cytomegalovirus, human
271 herpesviruses 6B and 7), JC- and Merkel cell polyomaviruses, human papillomavirus 31, and torque
272 teno virus. Intriguingly, a common feature shared by these viruses (except for HHV-6B) is their
273 persistence in soft tissues in episomal form[34]. Although the methods used in this study prevent us
274 from confirming whether this is also the case in bone, the low copy numbers detected are indicative
275 of a quiescent infection. Moreover, some of them (e.g. herpes-, polyoma-, and papillomaviruses) are

276 known to establish latency as densely packed nucleosomes, as mechanism to regulate gene
277 expression[35,36]. Thus, the circular form and the tight histone packaging may confer DNA viruses
278 superior preservation in relation to the host DNA (nuclear and mitochondrial)[37–39]. In this regard,
279 the analysis of viral DNA could be of utmost value for forensic investigations or studies of ancient
280 human remains dealing with highly compromised samples.

281 One challenging factor in the study of the human virome continues to be its extremely low proportion
282 in relation to other sources of DNA. With this in mind, we carried out our screening using two
283 approaches, qPCR and NGS. We found that the viral DNAs were exceeded on average a 1000-fold
284 by the nuclear DNA alone, in line with the loads reported for persisting viruses in soft tissues[40–42].
285 By targeted enrichment and NGS we identified a higher number of viruses than by qPCR, an
286 observation that can in part be explained by moderate DNA fragmentation in the samples. The only
287 exception was torque teno virus (TTV), for which we found a significantly higher number of positive
288 samples by qPCR. This circular ssDNA virus exhibits substantial heterogeneity[43], which was
289 under-represented by the three reference sequences upon which our baits were designed. Hence,
290 a better characterization of it could require a broader bait coverage of its five genogroups together
291 with analysis at the amino acid level, as certain motifs are more likely to be shared by different TTVs.
292 The qPCR on the other hand, targeted the conserved UTR region, to recognize a wide a repertoire
293 of TT-viruses.

294 Unsurprisingly, larger genome coverages were recovered from the viruses with the highest copy
295 numbers ($\sim\log_3$ copies/million cells). However, many of the viral loads fell below this threshold,
296 whereby, follow-up singleplex enrichments may be beneficial to increase the analytical resolution of
297 the data. Importantly, even for the low coverage genomes, we confirmed the sequences retrieved to
298 be virus-specific and hence, genuine findings of the femoral bone.

299 In all, we unveil that the human femoral bone is a much richer source of persistent DNA viruses than
300 earlier known. We propose that a “Human Virome Panel” could be built as an efficient tool for
301 assessment of human provenance in conjunction with the standard human DNA profiling. Indeed,
302 the phylogeographical distributions of JC[13,14,44], BK[2], varicella-zoster[7], and Epstein Barr[6]
303 viruses have been examined in this context and shown to add power to forensic cases and
304 anthropological studies. Undoubtedly, larger cohorts, both global and local, are required to validate
305 the benefits and boundaries of such panel in multiple taphonomic conditions. This work could also
306 help to identify patterns in the age distributions of certain viruses and their genotypes[19,20], to
307 support biometric estimations.

308 Moreover, our data raise intriguing questions on the clinical significance of the long-term presence
309 of these viruses – or their genomes – in the human skeleton. Indeed, many of the viruses we found
310 can reactivate, and several have oncogenic potential. Among the latter, we detected Merkel cell

311 polyomavirus and human papillomavirus 31, two oncoviruses which, remarkably, are
312 mucocutaneous. Certainly, these unexpected findings warrant further studies on the transcriptional,
313 translational and reactivation potential of these pathogens as bone residents, particularly among the
314 elderly and immune suppressed.

315 **Conclusions**

316 Our work substantially expands the current knowledge on the spectrum of DNA viruses persisting in
317 human bone and opens new perspectives on their applicability in the investigation of human skeletal
318 remains. It also supports the search for viruses from ancient relics, which can foreseeably remodel
319 our understanding of virus evolution.

320 **Acknowledgments**

321 We would like to thank Teemu Smura, Anna-Maija Sulonen and the Sequencing unit of the Institute
322 for Molecular Medicine Finland for their support with NGS, and Anna-Liina Mustaniemi for instruction
323 on bone handling. This study was funded by grants from the Finnish Medical Society (MFP), the
324 Jane and Aatos Erkko Foundation (KH), the Finnish Cultural Foundation (MT, MFP), the Juhani Aho
325 Foundation for Medical Research (MT), the Sigrid Jusélius Foundation (MS-V, KH), the Life and
326 Health Medical Foundation (MS-V, AS), the Finnish Society of Sciences and Letters (KH, AS), the
327 Kone Foundation (AS), and Finska Läkaresällskapet (KH, MFP). Fundação para a Ciência e a
328 Tecnologia, I.P., under the Scientific Employment Stimulus - Institutional Call - CI-CTTI-94-
329 ARH/2019 (DP).

330

331 **Conflicts of Interest**

332 The authors declare no conflict of interest.

333 **Contributions**

334 M.T., K.H., M.F.P and A.S designed the study. M.T. executed the experiments. MT, D.P. and
335 M.F.P analyzed the data, E.V. and M.S.V contributed with a plasmid and qPCR optimization, M.T.,
336 M.F.P. and A.S. drafted the manuscript. All authors commented, edited and approved the final
337 version of the manuscript.

338 **References**

- 339 [1] E.C. Holmes, Evolutionary History and Phylogeography of Human Viruses, *Annu. Rev.*
340 *Microbiol.* 62 (2008) 307–328. <https://doi.org/10.1146/annurev.micro.62.081307.162912>.
- 341 [2] S. Zhong, P.S. Randhawa, H. Ikegaya, Q. Chen, H.-Y. Zheng, M. Suzuki, T. Takeuchi, A.
342 Shibuya, T. Kitamura, Y. Yogo, Distribution patterns of BK polyomavirus (BKV) subtypes
343 and subgroups in American, European and Asian populations suggest co-migration of BKV
344 and the human race., *J. Gen. Virol.* 90 (2009) 144–52. <https://doi.org/10.1099/vir.0.83611-0>.
- 345 [3] H. Ikegaya, H. Motani, P. Saukko, K. Sato, T. Akutsu, K. Sakurada, BK virus genotype
346 distribution offers information of tracing the geographical origins of unidentified cadaver,

- 347 Forensic Sci. Int. 173 (2007) 41–46. <https://doi.org/10.1016/j.forsciint.2007.01.022>.
- 348 [4] D. Miyamori, N. Ishikawa, N. Idota, Y. Kakiuchi, S. Mclean, T. Kitamura, H. Ikegaya, Tracing
349 Jomon and Yayoi ancestries in Japan using ALDH2 and JC virus genotype distributions,
350 (2015). <https://doi.org/10.1186/s13323-015-0031-1>.
- 351 [5] H. Ikegaya, H. Iwase, Trial for the geographical identification using JC viral genotyping in
352 Japan, Forensic Sci. Int. 139 (2004) 169–172.
353 <https://doi.org/10.1016/J.FORSCIINT.2003.10.019>.
- 354 [6] H. Ikegaya, H. Motani, K. Sakurada, K. Sato, T. Akutsu, M. Yoshino, Forensic application of
355 Epstein-Barr virus genotype: Correlation between viral genotype and geographical area, J.
356 Virol. Methods. 147 (2008) 78–85. <https://doi.org/10.1016/j.jviromet.2007.08.010>.
- 357 [7] H. Inoue, H. Motani-Saitoh, K. Sakurada, H. Ikegaya, D. Yajima, M. Hayakawa, Y. Sato, K.
358 Otsuka, K. Kobayashi, S. Nagasawa, H. Iwase, Determination of the geographical origin of
359 unidentified cadavers based on geographical differences in genotype of varicella-zoster
360 virus, J. Med. Virol. 82 (2010) 903–908. <https://doi.org/10.1002/jmv.21666>.
- 361 [8] D. Forni, R. Cagliani, M. Clerici, U. Pozzoli, M. Sironi, You Will Never Walk Alone:
362 Codispersal of JC Polyomavirus with Human Populations, Mol. Biol. Evol. (2019).
363 <https://doi.org/10.1093/molbev/msz227>.
- 364 [9] A.W. Kolb, C. Ané, C.R. Brandt, Using HSV-1 Genome Phylogenetics to Track Past Human
365 Migrations, PLoS One. 8 (2013). <https://doi.org/10.1371/journal.pone.0076267>.
- 366 [10] H. Ikegaya, H.Y. Zheng, P.J. Saukko, L. Varesmaa-Korhonen, T. Hovi, T. Vesikari, H.
367 Suganami, T. Takasaka, C. Sugimoto, Y. Ohasi, T. Kitamura, Y. Yogo, Genetic diversity of
368 JC virus in the Saami and the Finns: Implications for their population history, Am. J. Phys.
369 Anthropol. (2005). <https://doi.org/10.1002/ajpa.20189>.
- 370 [11] C.D. Bowen, H. Paavilainen, D.W. Renner, J. Palomäki, J. Lehtinen, T. Vuorinen, P.
371 Norberg, V. Hukkanen, M. Szpara, HSV-1 strains circulating in Finland demonstrate
372 uncoupling of geographic and phenotypic variation, BioRxiv. (2018) 1–46.
373 <http://biorxiv.org/content/early/2018/10/12/424408.abstract>.
- 374 [12] H. Ikegaya, H. Iwase, Trial for the geographical identification using JC viral genotyping in
375 Japan, Forensic Sci. Int. 139 (2004) 169–172.
376 <https://doi.org/10.1016/J.FORSCIINT.2003.10.019>.
- 377 [13] H.T. Agostini, R. Yanagihara, V. Davis, C.F. Ryschkewitsch, G.L. Stoner, Asian genotypes
378 of JC virus in Native Americans and in a Pacific Island population: Markers of viral evolution
379 and human migration, Proc. Natl. Acad. Sci. U. S. A. 94 (1997) 14542–14546.

- 380 <https://doi.org/10.1073/pnas.94.26.14542>.
- 381 [14] C. Sugimoto, T. Kitamura, J. Guo, M.N. Al-Ahdal, S.N. Shchelkunov, B. Otova, P. Ondrejka,
382 J.Y. Chollet, S. El-Safi, M. Ettayebi, G. Grésenguët, T. Kocagöz, S. Chaiyarasamee, K.Z.
383 Thant, S. Thein, K. Moe, N. Kobayashi, F. Taguchi, Y. Yogo, Typing of urinary JC virus DNA
384 offers a novel means of tracing human migrations., *Proc. Natl. Acad. Sci. U. S. A.* 94 (1997)
385 9191–6. <http://www.ncbi.nlm.nih.gov/pubmed/9256458> (accessed March 20, 2018).
- 386 [15] C. Sugimoto, M. Hasegawa, H.Y. Zheng, V. Demenev, Y. Sekino, K. Kojima, T. Honjo, H.
387 Kida, T. Hovi, T. Vesikari, J.A. Schalken, K. Tomita, Y. Mitsunobu, H. Ikegaya, N.
388 Kobayashi, T. Kitamura, Y. Yogo, JC virus strains indigenous to northeastern Siberians and
389 Canadian Inuits are unique but evolutionally related to those distributed throughout Europe
390 and Mediterranean areas, *J. Mol. Evol.* 55 (2002) 322–335. [https://doi.org/10.1007/s00239-](https://doi.org/10.1007/s00239-001-2329-2)
391 [001-2329-2](https://doi.org/10.1007/s00239-001-2329-2).
- 392 [16] L.A. Shackelton, A. Rambaut, O.G. Pybus, E.C. Holmes, JC Virus Evolution and Its
393 Association with Human Populations, *J. Virol.* 80 (2006) 9928–9933.
394 <https://doi.org/10.1128/jvi.00441-06>.
- 395 [17] P.M. Sharp, P. Simmonds, Evaluating the evidence for virus/host co-evolution, *Curr. Opin.*
396 *Virol.* 1 (2011) 436–441. <https://doi.org/10.1016/j.coviro.2011.10.018>.
- 397 [18] A. Kitchen, M.M. Miyamoto, C.J. Mulligan, Utility of DNA viruses for studying human host
398 history: Case study of JC virus, *Mol. Phylogenet. Evol.* (2008).
399 <https://doi.org/10.1016/j.ympev.2007.09.005>.
- 400 [19] S. Zhong, H.-Y. Zheng, M. Suzuki, Q. Chen, H. Ikegaya, N. Aoki, S. Usuku, N. Kobayashi,
401 S. Nukuzuma, Y. Yasuda, N. Kuniyoshi, Y. Yogo, T. Kitamura, Age-related urinary excretion
402 of BK polyomavirus by nonimmunocompromised individuals., *J. Clin. Microbiol.* 45 (2007)
403 193–8. <https://doi.org/10.1128/JCM.01645-06>.
- 404 [20] J.S. Smith, N.J. Robinson, Age-Specific Prevalence of Infection with Herpes Simplex Virus
405 Types 2 and 1: A Global Review, 2002.
- 406 [21] P. Norja, K. Hokynar, L.-M. Aaltonen, R. Chen, A. Ranki, E.K. Partio, O. Kiviluoto, I.
407 Davidkin, T. Leivo, A.M. Eis-Hübinger, B. Schneider, H.-P. Fischer, R. Tolba, O. Vapalahti,
408 A. Vaheri, M. Söderlund-Venermo, K. Hedman, Bioportfolio: Lifelong persistence of variant
409 and prototypic erythrovirus DNA genomes in human tissue, *Proc. Natl. Acad. Sci. U. S. A.*
410 103 (2006) 7450–7453. <http://www.pnas.org/content/pnas/103/19/7450.full.pdf> (accessed
411 June 8, 2018).
- 412 [22] M. Toppinen, M.F. Perdomo, J.U. Palo, P. Simmonds, S.J. Lycett, M. Söderlund-Venermo,

- 413 A. Sajantila, & K. Hedman, Bones hold the key to DNA virus history and epidemiology, *Sci.*
414 *Rep.* 5 (2015). <https://doi.org/10.1038/srep17226>.
- 415 [23] M. Toppinen, M.F. Perdomo, J.U. Palo, P. Simmonds, S.J. Lycett, M. Söderlund-Venermo,
416 A. Sajantila, & K. Hedman, Bones hold the key to DNA virus history and epidemiology, *Sci.*
417 *Rep.* 5 (2015). <https://doi.org/10.1038/srep17226>.
- 418 [24] B. Mühlemann, A. Margaryan, P. de B. Damgaard, M.E. Allentoft, L. Vinner, A.J. Hansen, A.
419 Weber, V.I. Bazaliiskii, M. Molak, J. Arneborg, W. Bogdanowicz, C. Falys, M. Sablin, V.
420 Smrčka, S. Sten, K. Tashbaeva, N. Lynnerup, M. Sikora, D.J. Smith, R.A.M. Fouchier, C.
421 Drosten, K.-G. Sjögren, K. Kristiansen, E. Willerslev, T.C. Jones, Ancient human parvovirus
422 B19 in Eurasia reveals its long-term association with humans., *Proc. Natl. Acad. Sci. U. S.*
423 *A.* 115 (2018) 7557–7562. <https://doi.org/10.1073/pnas.1804921115>.
- 424 [25] B. Mühlemann, T.C. Jones, P. De Barros Damgaard, M.E. Allentoft, I. Shevnina, A. Logvin,
425 E. Usmanova, I.P. Panyushkina, B. Boldgiv, T. Bazartseren, K. Tashbaeva, V. Merz, N. Lau,
426 V. Smrčka, D. Voyakin, E. Kitov, A. Epimakhov, D. Pokutta, M. Vicze, T.D. Price, V.
427 Moiseyev, A.J. Hansen, L. Orlando, S. Rasmussen, M. Sikora, L. Vinner, A.D.M.E.
428 Osterhaus, D.J. Smith, D. Glebe, R.A.M. Fouchier, C. Drosten, K.G. Sjögren, K. Kristiansen,
429 E. Willerslev, Ancient hepatitis B viruses from the Bronze Age to the Medieval period,
430 *Nature.* 557 (2018) 418–423. <https://doi.org/10.1038/s41586-018-0097-z>.
- 431 [26] B. Krause-Kyora, J. Susat, F.M. Key, D. Kühnert, E. Bosse, A. Immel, C. Rinne, S.-C.
432 Kornell, D. Yepes, S. Franzenburg, H.O. Heyne, T. Meier, S. Lösch, H. Meller, S. Friederich,
433 N. Nicklisch, K.W. Alt, S. Schreiber, A. Tholey, A. Herbig, A. Nebel, J. Krause, Neolithic and
434 medieval virus genomes reveal complex evolution of hepatitis B, *Elife.* 7 (2018).
435 <https://doi.org/10.7554/eLife.36666>.
- 436 [27] S. Goh, C. Lindau, A. Tiveljung-Lindell, T. Allander, Merkel cell polyomavirus in respiratory
437 tract secretions., *Emerg. Infectious Dis.* 15 (2009) 489–91.
438 <https://doi.org/10.3201/eid1503.081206>.
- 439 [28] A. Dumoulin, H.H. Hirsch, Reevaluating and optimizing polyomavirus BK and JC real-time
440 PCR assays to detect rare sequence polymorphisms., *J. Clin. Microbiol.* 49 (2011) 1382–8.
441 <https://doi.org/10.1128/JCM.02008-10>.
- 442 [29] N.G. Hoffman, L. Cook, E.E. Atienza, A.P. Limaye, K.R. Jerome, Marked variability of BK
443 virus load measurement using quantitative real-time PCR among commonly used assays.,
444 *J. Clin. Microbiol.* 46 (2008) 2671–80. <https://doi.org/10.1128/JCM.00258-08>.
- 445 [30] M. Toppinen, P. Norja, L.-M. Aaltonen, S. Wessberg, L. Hedman, M. Söderlund-Venermo,

- 446 K. Hedman, A new quantitative PCR for human parvovirus B19 genotypes, *J. Virol.*
447 *Methods.* 218 (2015) 40–45. <https://doi.org/10.1016/j.jviromet.2015.03.006>.
- 448 [31] M. Pistello, A. Morrica, F. Maggi, M.L. Vatteroni, G. Freer, C. Fornai, F. Casula, S. Marchi,
449 P. Ciccorossi, P. Rovero, M. Bendinelli, TT virus levels in the plasma of infected individuals
450 with different hepatic and extrahepatic pathology., *J. Med. Virol.* 63 (2001) 189–95.
451 <http://www.ncbi.nlm.nih.gov/pubmed/11170056> (accessed February 23, 2018).
- 452 [32] F. Maggi, C. Fornai, M.L. Vatteroni, G. Siciliano, F. Menichetti, C. Tascini, S. Specter, M.
453 Pistello, M. Bendinelli, Low prevalence of TT virus in the cerebrospinal fluid of viremic
454 patients with central nervous system disorders., *J. Med. Virol.* 65 (2001) 418–22.
- 455 [33] A.L. McNees, Z.S. White, P. Zanwar, R.A. Vilchez, J.S. Butel, Specific and quantitative
456 detection of human polyomaviruses BKV, JCV, and SV40 by real time PCR, *J. Clin. Virol.* 34
457 (2005) 52–62. [https://doi.org/S1386-6532\(05\)00047-8](https://doi.org/S1386-6532(05)00047-8) [pii].
- 458 [34] P.M. Lieberman, Epigenetics and Genetics of Viral Latency, *Cell Host Microbe.* 19 (2016)
459 619–628. <https://doi.org/10.1016/j.chom.2016.04.008>.
- 460 [35] S.B. Kutluay, S.J. Triezenberg, Role of chromatin during herpesvirus infections, *Biochim.*
461 *Biophys. Acta - Gen. Subj.* 1790 (2009) 456–466.
462 <https://doi.org/10.1016/j.bbagen.2009.03.019>.
- 463 [36] B.I. Milavetz, L. Balakrishnan, Viral Epigenetics, (n.d.). [https://doi.org/10.1007/978-1-4939-](https://doi.org/10.1007/978-1-4939-1804-1_30)
464 [1804-1_30](https://doi.org/10.1007/978-1-4939-1804-1_30).
- 465 [37] D. Higgins, A.B. Rohrlach, J. Kaidonis, G. Townsend, J.J. Austin, Differential nuclear and
466 mitochondrial DNA preservation in post-mortem teeth with implications for forensic and
467 ancient DNA studies, *PLoS One.* 10 (2015) 1–17.
468 <https://doi.org/10.1371/journal.pone.0126935>.
- 469 [38] J.S. Pedersen, E. Valen, A.M. Velazquez, B.J. Parker, M. Rasmussen, S. Lindgreen, B.
470 Lilje, D.J. Tobin, T.K. Kelly, S. Vang, R. Andersson, P.A. Jones, C.A. Hoover, A. Tikhonov,
471 E. Prokhortchouk, E.M. Rubin, A. Sandelin, M.T. Gilbert, A. Krogh, E. Willerslev, L. Orlando,
472 Genome-wide nucleosome map and cytosine methylation levels of an ancient human
473 genome, *Genome Res.* 24 (2014) 454–466. <https://doi.org/10.1101/gr.163592.113>.
- 474 [39] L. Orlando, M.T.P. Gilbert, E. Willerslev, Reconstructing ancient genomes and epigenomes,
475 *Nat Rev Genet.* 16 (2015) 395–408. <https://doi.org/10.1038/nrg3935>.
- 476 [40] L. Pyöriä, M. Toppinen, E. Mäntylä, L. Hedman, L.-M. Aaltonen, M. Vihinen-Ranta, T.
477 Ilmarinen, M. Söderlund-Venermo, K. Hedman, M.F. Perdomo, Extinct type of human
478 parvovirus B19 persists in tonsillar B cells, *Nat. Commun.* 8 (2017) 14930.

- 479 <https://doi.org/10.1038/ncomms14930>.
- 480 [41] C. Berger, M. Hug, C. Gysin, L. Molinari, M. Frei, W. Bossart, D. Nadal, Distribution patterns
481 of β - and γ -herpesviruses within Waldeyer's ring organs, *J. Med. Virol.* 79 (2007) 1147–
482 1152. <https://doi.org/10.1002/jmv.20899>.
- 483 [42] P. Randhawa, R. Shapiro, A. Vats, Quantitation of DNA of Polyomaviruses BK and JC in
484 Human Kidneys, *J. Infect. Dis.* 192 (2005) 504–509. <https://doi.org/10.1086/431522>.
- 485 [43] S. Hino, H. Miyata, Torque teno virus (TTV): current status, *Rev. Med. Virol.* 17 (2007) 45–
486 57. <https://doi.org/10.1002/rmv.524>.
- 487 [44] P. Randhawa, R. Shapiro, A. Vats, Quantitation of DNA of Polyomaviruses BK and JC in
488 Human Kidneys, *J. Infect. Dis.* 192 (2005) 504–509. <https://doi.org/10.1086/431522>.
- 489