

Global Genomic Diversity of Human Papillomavirus 11 Based on 433 Isolates and 78 Complete Genome Sequences

Mateja M. Jelen,^a Zigui Chen,^b Boštjan J. Kocjan,^c Lea Hošnjak,^c Felicity J. Burt,^d Paul K. S. Chan,^b Diego Chouhy,^e Catharina E. Combrinck,^d Christine Estrade,^f Alison Fiander,^g Suzanne M. Garland,^{h,i,j} Adriana A. Giri,^e Joaquín Víctor González,^k Arndt Gröning,^l Sam Hibbitts,^g Tommy N. M. Luk,^{m,n} Karina Marinic,^o Toshihiko Matsukura,^p Anna Neumann,^l Anja Oštrbenk,^c Maria Alejandra Picconi,^k Martin Sagadin,^c Roland Sahli,^f Riaz Y. Seedat,^d Katja Seme,^c Alberto Severini,^r Jessica L. Sinchi,^o Jana Smahelova,^s Sepehr N. Tabrizi,^{h,i,j} Ruth Tachezy,^s Sarah Tohme Faybush,^r Virgilijus Uloza,^t Ingrida Uloziene,^t Yong Wee Wong,^u Snježana Židovec Lepej,^v Robert D. Burk,^w Mario Poljak^c

Institute of Pathology, Faculty of Medicine, University of Ljubljana, Ljubljana, Slovenia^a; Department of Microbiology, Faculty of Medicine, The Chinese University of Hong Kong, Prince of Wales Hospital, Hong Kong Special Administrative Region, China^b; Institute of Microbiology and Immunology, Faculty of Medicine, University of Ljubljana, Ljubljana, Slovenia^c; Department of Medical Microbiology and Virology, Faculty of Health Sciences, University of the Free State, Bloemfontein, South Africa^d; Human Virology Group, National Council of Scientific and Technical Research, Institute of Molecular and Cell Biology of Rosario, Rosario, Argentina^e; Institute of Microbiology, Lausanne University Hospital (CHUV), and University of Lausanne, Lausanne, Switzerland^f; Department of Obstetrics and Gynaecology, Cardiff University School of Medicine, Institute of Cancer and Genetics, Cardiff, United Kingdom^g; Department of Microbiology and Infectious Diseases, The Royal Women's Hospital, Parkville, Victoria, Australia^h; Department of Obstetrics and Gynaecology, University of Melbourne, Melbourne, Victoria, Australiaⁱ; Murdoch Childrens Research Institute, Parkville, Victoria, Australia^j; Oncogenic Viruses Service, National Institute of Infectious Diseases-ANLIS Dr. Carlos G. Malbrán, Buenos Aires, Argentina^k; Medizinisches Versorgungszentrum wagnerstibbe für Laboratoriumsmedizin und Pathologie GmbH, Hannover, Germany^l; Dermatology Research Centre, The Chinese University of Hong Kong, Prince of Wales Hospital, Hong Kong Special Administrative Region, China^m; Department of Health, Social Hygiene Service, Centre for Health Protection, Hong Kong Special Administrative Region, Chinaⁿ; Histocompatibility and Molecular Genetics Laboratory, Dr. Julio C. Perrando Hospital, Resistencia, Chaco, Argentina^o; Momate Memorial Laboratory, Tokyo, Japan^p; Department of Otorhinolaryngology, Faculty of Health Sciences, University of the Free State, Bloemfontein, South Africa^q; Viral Exanthemata and STD Section, National Microbiology Laboratory, Public Health Agency of Canada, Winnipeg, Manitoba, Canada^r; Department of Experimental Virology, National Reference Laboratory for Papillomaviruses, Institute of Hematology and Blood Transfusion, Prague, Czech Republic^s; Department of Otolaryngology, Lithuanian University of Health Sciences, Medical Academy, Kaunas, Lithuania^t; DNA Laboratories Sdn. Bhd., UKM-MTDC Technology Centre, Universiti Kebangsaan Malaysia, Bangi, Malaysia^u; Department of Molecular Diagnostics, University Hospital for Infectious Diseases Dr. Fran Mihaljević, Zagreb, Croatia^v; Departments of Pediatrics, Microbiology and Immunology, Epidemiology and Population Health, Obstetrics and Gynecology, and Women's Health, Albert Einstein College of Medicine, New York, New York, USA^w

ABSTRACT

Human papillomavirus 11 (HPV11) is an etiological agent of anogenital warts and laryngeal papillomas and is included in the 4-valent and 9-valent prophylactic HPV vaccines. We established the largest collection of globally circulating HPV11 isolates to date and examined the genomic diversity of 433 isolates and 78 complete genomes (CGs) from six continents. The genomic variation within the 2,800-bp E5a-E5b-L1-upstream regulatory region was initially studied in 181/207 (87.4%) HPV11 isolates collected for this study. Of these, the CGs of 30 HPV11 variants containing unique single nucleotide polymorphisms (SNPs), indels (insertions or deletions), or amino acid changes were fully sequenced. A maximum likelihood tree based on the global alignment of 78 HPV11 CGs (30 CGs from our study and 48 CGs from GenBank) revealed two HPV11 lineages (lineages A and B) and four sublineages (sublineages A1, A2, A3, and A4). HPV11 (sub)lineage-specific SNPs within the CG were identified, as well as the 208-bp representative region for CG-based phylogenetic clustering within the partial E2 open reading frame and noncoding region 2. Globally, sublineage A2 was the most prevalent, followed by sublineages A1, A3, and A4 and lineage B.

IMPORTANCE

This collaborative international study defined the global heterogeneity of HPV11 and established the largest collection of globally circulating HPV11 genomic variants to date. Thirty novel complete HPV11 genomes were determined and submitted to the available sequence repositories. Global phylogenetic analysis revealed two HPV11 variant lineages and four sublineages. The HPV11 (sub)lineage-specific SNPs and the representative region identified within the partial genomic region E2/noncoding region 2 (NCR2) will enable the simpler identification and comparison of HPV11 variants worldwide. This study provides an important knowledge base for HPV11 for future studies in HPV epidemiology, evolution, pathogenicity, prevention, and molecular assay development.

Human papillomavirus 11 (HPV11), which belongs to species 10 of the *Alphapapillomavirus* genus (*Alpha-PV*), is etiologically associated with approximately 20% of anogenital warts and 30 to 40% of laryngeal papillomas (1–11). HPV11 is generally considered a low-risk HPV type due to its rare presence in HPV-related cancers in humans, especially cervical cancer (9). HPV11 is present in 0.5% of samples from HPV-positive women with a normal cytology worldwide and causes 2.3% of cervical low-grade squamous cell intraepithelial lesions (12, 13). However, rare case

reports of HPV11-positive cases of cervical and anal squamous cell carcinomas, malignantly transformed laryngeal papillomas, and sinonasal inverted papillomas associated with squamous cell carcinoma can be found in the literature (14–21). Due to its clinical significance, HPV11 has been included in the current quadrivalent and nonavalent prophylactic HPV vaccines (22, 23).

The genetic diversity of HPV11 was studied for the first time in 1995, when Heinzel et al. sequenced the noncoding upstream regulatory region (URR) of 40 HPV11 isolates from the head and

neck and anogenital regions from patients on five continents (Europe, Asia, North America, South America, and Africa) (24). In 2011, a Slovenian research group performed the first extensive genetic analysis of HPV11 by sequencing approximately half of the viral genome of 63 HPV11 isolates from various anatomical locations of infection. The 10 most divergent HPV11 isolates were also fully sequenced, and phylogenetic analysis confirmed the existence of two HPV11 variant groups, as initially suspected from a comparison of the concatenated L1, URR, E6, E5a, and E5b sequences of all studied HPV11 isolates (25). In 2011, the classification and nomenclature system for *Alpha*-PV full-length genomic variants was established (26). On the basis of the analysis of 32 complete HPV11 genomes originating mainly from Slovenia, HPV11 genomic variants were classified into two sublineages: sublineage A1 (previously termed the “prototypic variant”) and sublineage A2 (previously termed the “nonprototypic variant”). Three additional HPV11 genomic diversity studies that focused on samples from the head and neck and the anogenital regions and that described the variability of the HPV11 genome from ethnogeographically closed cohorts and/or focused on individual complete or partial HPV11 genomic regions have recently been published (7, 27, 28). Several HPV11 genomic variants with highly similar nucleotide profiles were identified in these studies. Apart from the observation that isolates from sublineage A2 prevail globally (7, 28), no other ethnogeographical or clinical associations with HPV11 variants have been described to date (7, 24, 25, 27, 28).

This international collaborative study established the largest database of globally circulating HPV11 isolates to date and analyzed 433 HPV11 isolates and 78 complete genome (CG) sequences from six continents in order to determine the global distribution of HPV11 genomic variants.

MATERIALS AND METHODS

Study samples. In order to explore the intratype genomic diversity of HPV11, 207 HPV11 DNA-positive samples were obtained from 14 countries covering six continents: Argentina, Australia, Canada, Croatia, the Czech Republic, Germany, Hong Kong, Japan, Lithuania, Malaysia, South Africa, Switzerland, the United Kingdom, and the United States. HPV11 DNA-positive samples were collected from the anogenital region (105 samples; 50.7%) and the head and neck region (56 samples; 27.1%); the anatomical location was unavailable for 46 samples (22.2%). The majority of samples were anogenital wart tissue samples (76 samples; 36.7%), followed by laryngeal papilloma tissue samples (52 samples; 25.1%) and cervical swab specimens (22 samples; 10.6%) (Table 1).

In order to identify ethnogeographical associations among HPV11

TABLE 1 Characteristics of 207 HPV11-positive study samples

Anatomical location and sample type	No. (%) of samples
Anogenital region	105 (50.7)
Cervical swab ^a	22 (20.9)
Anogenital wart tissue	76 (72.4)
Unspecified sample type	7 (6.7)
Head and neck region ^b	56 (27.1)
Laryngeal papilloma tissue	52 (92.8)
Unspecified sample type	3 (5.4)
Unspecified anatomical location	46 (22.2)

^a Cervical swab samples included 2 (9.2%) samples with atypical squamous cells of undetermined significance, 1 (4.5%) sample with a low-grade squamous intraepithelial lesion, and 1 (4.5%) sample with cervical intraepithelial neoplasia grade 1. Data were not available for 18 (81.8%) samples.

^b Samples obtained from the head and neck region included 1 (1.8%) sample of tracheal papilloma tissue.

lineages and sublineages, 500 HPV11 sequences were selected, as follows: (i) 30 complete genomes and 177 partial sequences ranging from 453 to 2,765 bp generated in this study and (ii) 48 complete genomes and 245 partial sequences ranging from 114 to 3,218 bp obtained from the GenBank NCBI database and the PapillomaVirus Episteme database (PaVE) (data collection was closed in December 2014). HPV11 sequences with assigned (sub)lineages and information on geographical origin (433/500 sequences; 86.6%) were eligible for further statistical analysis.

Whole-genome amplification and sequencing. A 4,851-bp-long fragment including the 3' end of the E2 open reading frame (ORF), the complete E5a, E5b, L2, and L1 ORFs, the complete URR, and the 5' end of the E6 ORF was amplified with primer pair HPV-11-S2F (5'-TTACAACAA GCACCAAAGAAG-3', nucleotides [nt] 3529 to 3549) and HPV-11-S2R (5'-TTCTATTTACACAACGGCT-3', nt 427 to 446). The entire genomes of HPV11 genomic variants with unique single nucleotide polymorphisms (SNPs), amino acid changes, or insertions or deletions (indels) within E5a, E5b, L1, and/or URR were sequenced. To obtain the complete viral genome, an additional 4,521-bp-long overlapping fragment encompassing the 3' end of the L1 ORF, the complete URR, E6, E7, and E1 ORFs, and a greater part of the E2 ORF was generated using primer pair HPV-11-S1MS-F (5'-GGATATGAGTTTTTGGGAGGT-3', nt 7084 to 7104) and HPV-11-S1MS-R (5'-ATGCCACGTTGAAGATGCTA-3', nt 3660 to 3679).

All primers used in this study were designed according to the corrected HPV11 reference sequence (GenBank accession no. M14119; PaVE accession no. HPV11REF) using CLC Main Workbench software, version 6.5 (CLC Bio, Denmark), and synthesized at TIB Molbiol (Berlin, Germany).

For both overlapping PCRs, a Platinum *Taq* DNA polymerase high-fidelity kit (Invitrogen, Carlsbad, CA) was utilized according to the manufacturer's instructions. Each reaction mixture contained 100 ng of template DNA, 2.5 μ l of 10 \times High-Fidelity PCR buffer, 0.1 μ l of each primer (50 μ M), 0.5 μ l of the deoxynucleoside triphosphates (10 mM), 1 μ l of MgSO₄ (50 mM), 0.1 μ l of enzyme (1 U/50 μ l), and water up to 25 μ l. The cycling conditions were carried out as follows: 94°C for 2 min, followed by 50 cycles of 94°C for 30 s, 53°C or 57°C (depending on the primer pair) for 30 s, and 68°C for 5 min. The final extension step was carried out at 68°C for 7 min.

PCR products were verified by gel electrophoresis, purified using a QIAquick PCR purification kit (Qiagen, Hilden, Germany), and submitted to MacroGen Europe (Amsterdam, Netherlands) or to the in-house sequencing facility of the Institute of Microbiology and Immunology, Faculty of Medicine, University of Ljubljana, for Sanger sequencing of both strands using previously published primers (25) and a previously described sequencing protocol (29).

The acquired nucleotide sequences were assembled and analyzed

Received 18 December 2015 Accepted 19 March 2016

Accepted manuscript posted online 30 March 2016

Citation Jelen MM, Chen Z, Kocjan BJ, Hošnjak L, Burt FJ, Chan PKS, Chouhy D, Combrinck CE, Estrade C, Fiander A, Garland SM, Giri AA, González JV, Gröning A, Hibbitts S, Luk TNM, Marinic K, Matsukura T, Neumann A, Oštrbenk A, Picconi MA, Sagadin M, Sahli R, Seedat RY, Seme K, Severini A, Sinchi JL, Smahelova J, Tabrizi SN, Tachezy R, Tohme Faybush S, Uloza V, Uloziene I, Wong YW, Židovec Lepej S, Burk RD, Poljak M. 2016. Global genomic diversity of human papillomavirus 11 based on 433 isolates and 78 complete genome sequences. *J Virol* 90:5503–5513. doi:10.1128/JVI.03149-15.

Editor: L. Banks, International Centre for Genetic Engineering and Biotechnology
Address correspondence to Mario Poljak, mario.poljak@mf.uni-lj.si.

Copyright © 2016, American Society for Microbiology. All Rights Reserved.

with CLC Main Workbench software, version 6.5 (CLC Bio, Denmark), using the corrected HPV11 reference sequence (PaVE accession no. HPV11REF) as a standard for nucleotide position comparison and numbering. All samples with wobble nucleotides or unique nucleotide profiles compared to the reference sequence were selected for repeated amplification and sequencing.

Phylogenetic tree construction. The global phylogenetic tree was generated using the 30 complete HPV11 genomes determined in this study (GenBank accession no. LN833161 to LN833190) and 48 complete genomes available in GenBank (GenBank accession no. EU918768 [30]; FR872717 [31]; FN870021, FN870022, and FN907957 to FN907964 [25]; JQ773408 to JQ773412 [32]; HE574701 to HE574705 [19]; JN644141 [33]; and HE611258 to HE611274, HE962023 to HE962025, and HE962365 to HE962368 [34]), including the sequence with PaVE accession no. HPV11REF. Among the 48 complete genomes from GenBank, 34 (70.1%) originated from two Slovenian studies (25, 34). All 78 complete HPV11 genomes were linearized at position 101 within the URR of the sequence with PaVE accession no. HPV11REF and globally aligned using MAFFT software (version 6.864b) (35). Maximum likelihood trees were constructed using RAXML HPC2 software (version 8.1.11), employing 1,000 bootstrap values and a general time-reversible model with a gamma-distributed rate variation and a proportion of invariable sites (GTR+G+I) (36). Bayesian trees were constructed using MrBayes software (version 3.2.3) (37) with 10,000,000 cycles for Markov chain Monte Carlo (MCMC) runs. In order to identify the most informative region(s) for whole-genome-based phylogenetic clustering, maximum likelihood trees were constructed using RAXML HPC2 (version 8.1.11) and MEGA (version 5) software (38), employing individual HPV11 ORFs, the URR, and multiple genomic regions selected according to the nucleotide positions of (sub)lineage-specific SNPs.

Identification of variant lineages/sublineages and lineage/sublineage-specific SNPs. Identification and naming of HPV11 variant lineages and sublineages were based on the 78 globally aligned HPV11 genomes, using the classification and nomenclature system proposed previously (26): nucleotide differences of >1% to <10% and >0.5% to <1% of the complete genomes were used to define variant lineages and sublineages, respectively. The pairwise identities (*p* distances) calculated in MEGA software (version 5) (38) were used for heat map construction, with color gradients indicating maximum (100%, blue) and minimum (98.6%, red) pairwise identities.

HPV11 (sub)lineage-specific SNPs were identified on the basis of the nucleotide sequence alignment of all 78 complete genomes using MacClade software (version 4.08) (39).

Statistical methods. The ethnogeographical associations for the HPV11 variants were analyzed for a total of 433 HPV11 samples obtained from 23 countries covering six continents. Univariate logistic regression comparing sublineage A2 and sublineage A1 was performed. Sublineage A1 was absent in Australia and Africa; therefore, the distribution of sublineages A2 and A1 was compared only for Europe, North America, South America, and Asia. However, because sublineage A2 was more common than sublineage A1 on all continents and the proportion of sublineage A1 isolates was higher in Asia than in the other five continents, we additionally combined sublineage A1 isolates from Europe, North and South America, Australia, and Africa into a single category. Isolates of lineage B and sublineages A3 and A4 were excluded from the analysis due to the small number of samples containing isolates of this lineage and these sublineages (less than 10). Odds ratios (ORs) were calculated, and the significance level was set at an α value of 0.05. All tests were two-sided. Additionally, to evaluate the independence of associations, a multiple logistic regression was performed using geographical location, which included all continents combined except Asia. All analyses were performed using SPSS software (version 22.0; IBM Corp.).

Nucleotide sequence accession numbers. The 30 complete HPV11 genome sequences determined in this study were deposited in the Euro-

pean Nucleotide Archive (ENA) of the EMBL-EBI database under accession numbers LN833161 to LN833190.

RESULTS

A complete 4,851-bp-long fragment covering the E5a, E5b, L1, and URR genomic regions was successfully amplified and sequenced from 181/207 (87.4%) HPV11 DNA-positive samples collected in this study. At least one SNP, amino acid change, and/or indel in the E5a, E5b, L1, and/or URR genomic region was found in 27/181 (14.9%) HPV11 isolates, whose genomes were subsequently fully sequenced. Three complete HPV11 genomes with concatenated E5a-E5b-L1-URR nucleotide sequence profiles highly similar to those of 27 genomes previously completely sequenced were additionally sequenced. Thus, overall, 30 complete HPV11 genome sequences originating from six continents were generated in this study.

The minimum and maximum lengths of 78 complete HPV11 genomes (30 from this study and 48 retrieved from GenBank) were 7,930 and 7,949 bp, respectively. A total of 239 (3.0%) variable nucleotide positions and 76 (3.0%) variable amino acid positions were identified. The nucleotide and amino acid sequence variabilities among the 78 complete HPV11 genomes are summarized in Table 2. In brief, the rates of variability in the nucleotide sequences of coding regions varied from 1.4% (4/297) in the E7 ORF to 5.8% (13/225) in the E5b ORF, and the rates of variability in the amino acid sequences ranged from 1.5% (10/650) in the E1 ORF to 11.0% (12/109) in the E4 ORF. In total, 10 indels of from 1 up to 15 nucleotides were identified in the URR, and one in-frame insertion of 9 nucleotides was detected in the L2 ORF. Among the 30 completely sequenced genomes, only 1 contained a single double peak that was marked as a wobble nucleotide in its final genome sequence. The maximum pairwise nucleotide difference between two genomes was 1.3%.

Phylogenetic analysis and variant lineage/sublineage identification. The global phylogenetic tree constructed from the 78 complete HPV11 genomes (30 from our study and 48 from GenBank) revealed that the tree topology was dichotomous, dividing HPV genomic variants into two variant lineages: lineages A and B (Fig. 1). Lineage A was further divided into four sublineages: sublineages A1, A2, A3, and A4. The prototype genome sequence, the sequence with PaVE accession no. HPV11REF, was designated a representative of the HPV11 variants clustering within sublineage A1. Sublineage A2 included HPV11 variants with sequences closely related to the sequence of the CAC86 genomic variant (GenBank accession no. FN907962). Sublineage A2 prevailed with 68/78 (87.2%) HPV11 isolates, followed by sublineage A1 with 6/78 (7.7%) isolates and the newly discovered sublineages A3 and A4 and lineage B, which were represented by 2/78 (2.6%), 1/78 (1.3%), and 1/78 (1.3%) HPV11 isolates, respectively (Fig. 2).

The topology of the phylogenetic tree based on concatenated E5a-E5b-L1-URR sequences of approximately 2,800 bp completely reproduced the topology of a global phylogenetic tree constructed from 78 complete HPV11 genomes. The topologies of all of the nine individually used HPV11 genomic coding regions and the noncoding region URR reproduced the dichotomous tree topology, allowing identification of variant lineages A and B; however, the clustering of individual HPV11 genomic regions (nine ORFs and the URR) did not reproduce the clustering of HPV11 isolates into the four sublineages, sublineages A1 to A4. Among all

TABLE 2 Comparison of nucleotide and amino acid sequence variabilities within 78 HPV11 CGs

ORF or noncoding region ^a	Maximum % nucleotide pairwise difference ^b	No. of nucleotides	No. (%) of variable nucleotide positions ^c	No. of variable nucleotides at each codon position			Maximum % amino acid pairwise difference ^d	No. of amino acids	No. (%) of variable amino acid positions
				1st	2nd	3rd			
E6	1.99	453	15 (3.31)	4	2	9	4.08	151	7 (4.64)
E7	1.01	297	4 (1.35)	4	0	0	2.06	99	3 (3.03)
NCR1	0.00	5	0 (0)						
E1	0.87	1,950	36 (1.85)	10	4	22	0.62	650	10 (1.54)
E2	1.72	1,104	38 (3.44)	9	8	21	2.20	368	17 (4.62)
E4	2.45	327	14 (4.28)	7	5	2	6.70	109	12 (11.0)
NCR2	4.55	44	3 (6.82)						
E5a	2.17	276	11 (3.99)	5	3	3	5.65	92	8 (8.70)
E5b	2.67	225	13 (5.78)	3	2	8	7.00	75	5 (6.67)
NCR3	4.35	46	2 (4.35)						
L2	1.46	1,368	38 (2.78)	7	3	28	1.77	456	11 (2.41)
L1	0.86	1,506	37 (2.46)	6	8	23	1.00	502	13 (2.59)
URR	3.08	802	42 (5.24)						
CG ^e	1.28	7,933	239 (3.01)					2,502	76 (3.03)

^a ORF, open reading frame (E6, E7, E1, E2, E4, E5a, E5b, L2, and L1); NCR1, noncoding region 1 (between the E7 and E1 ORFs); NCR2, noncoding region 2 (between the E2 and E5a ORFs); NCR3, noncoding region 3 (between the E5b and L2 ORFs); URR, upstream regulatory region between L1 and E6; CG, complete genome.

^b The maximum pairwise difference was calculated from the global nucleotide sequence alignment.

^c Nucleotide variations included SNPs. Indels and wobble nucleotides were excluded from the count.

^d The maximum pairwise difference was calculated from the global amino acid sequence alignment.

^e A single genome size was inferred from the nucleotide sequence alignment. The total number and percentage of variable nucleotides included SNPs; indels and wobble nucleotides were excluded from the count. The cumulative number of amino acids was taken from all 9 ORFs (E6, E7, E1, E2, E4, E5a, E5b, L2, and L1).

HPV11 genomic regions, only URR, E2, and L2 generated phylogenetic trees with the power to optimally discriminate between sublineages A3 and A4, but the trees did not effectively discriminate between sublineages A1 and A2. In addition, the tree obtained with a 208-bp-long fragment from nucleotide positions 3626 to 3833 of the sequence with PaVE accession no. HPV11REF reproduced the whole-genome phylogenetic tree, and this fragment could be used as the shortest representative region for optimal (sub)lineage assignment (Fig. 3). The 208-bp region consists of the 3' end of the E2 ORF and the 5' end of noncoding region 2 (NCR2) and contains at least one (sub)lineage-specific SNP (Fig. 4).

Lineage- and sublineage-specific SNPs. Lineage-specific SNPs were equally dispersed throughout the genome, allowing reliable discrimination of lineages A and B within all genomic regions of approximately 1,000 bp in length. In contrast, sublineage-specific SNPs were not evenly distributed throughout the HPV11 genome. For example, the first sublineage A4-specific SNP was detected after position 1300 of the sequence with PaVE accession no. HPV11REF. Lineage- and sublineage-specific SNPs were more frequently observed in the genomic regions from the E2 to L2 ORFs (nucleotide positions 3000 to 5000 of the sequence with PaVE accession no. HPV11REF) and in the URR (nucleotide positions 7000 to 7933 of the sequence with PaVE accession no. HPV11REF) (Fig. 4).

Interestingly, some SNPs or SNP patterns independent of the HPV11 variant lineage/sublineage may be associated with the ethnicity of the host. For example, among the 78 complete genomes, the combination of the C2358T and A3391G nucleotide changes showed significantly different predominances in isolates of the A2 sublineage from Europe and Asia (76% versus 53%, $P < 0.001$).

Ethnogeographical associations of HPV11 (sub)lineages. The worldwide distribution of HPV11 genomic variant (sub)lineages is presented in Fig. 5. Sublineage A2 was prevalent globally

(373/433; 86.1%), followed by sublineage A1 (54/433; 12.5%), sublineage A3 (3/433; 0.7%), sublineage A4 (2/433; 0.5%), and lineage B (1/433; 0.2%) (Fig. 5A). Sublineage A2 was also found in each continent, with its prevalence ranging from 73.9% in Asia (Fig. 5C) to 98.7% in Australia (Fig. 5F). Sublineage A1 was found only in Europe, Asia, and North and South America, and its prevalence ranged from 9.8% in North America (Fig. 5D) to 26.1% in Asia (Fig. 5C). Sublineage A3 was present in Europe, North America, and Australia, with a maximal prevalence of 1.6% in North America (Fig. 5D). Sublineage A4 was found only in Europe, with a maximal prevalence of 1.7% (Fig. 5B), and lineage B was found only in Africa, with a maximal prevalence of 2.9% (Fig. 5G).

In agreement with the observation that sublineage A2 was present in a higher proportion than sublineage A1 in each continent except Asia (Fig. 5), univariate analysis confirmed the statistically significant higher odds for the presence of sublineage A2 than sublineage A1 in Europe (OR, 2.6; 95% confidence interval [CI], 1.3 to 5.2; $P = 0.007$) and North America (OR, 3.2; 95% CI, 1.2 to 8.2; $P = 0.016$) than Asia. Furthermore, the analysis showed a higher odds for the presence of sublineage A2 than sublineage A1 in all continents other than Asia (OR, 4.1; 95% CI, 2.3 to 7.4; $P < 0.001$) (Table 3). However, multiple regression analysis did not confirm the statistically significant difference in the geographical distribution of sublineages A2 and A1.

DISCUSSION

We present the first study of the global genomic diversity of HPV11, an etiological agent of the two most frequent benign tumors in the anogenital region and upper respiratory tract: anogenital warts and laryngeal papillomas, respectively (2, 5, 6). For the purpose of this study, a total of 207 HPV11 isolates were collected from 14 countries covering six continents, and four

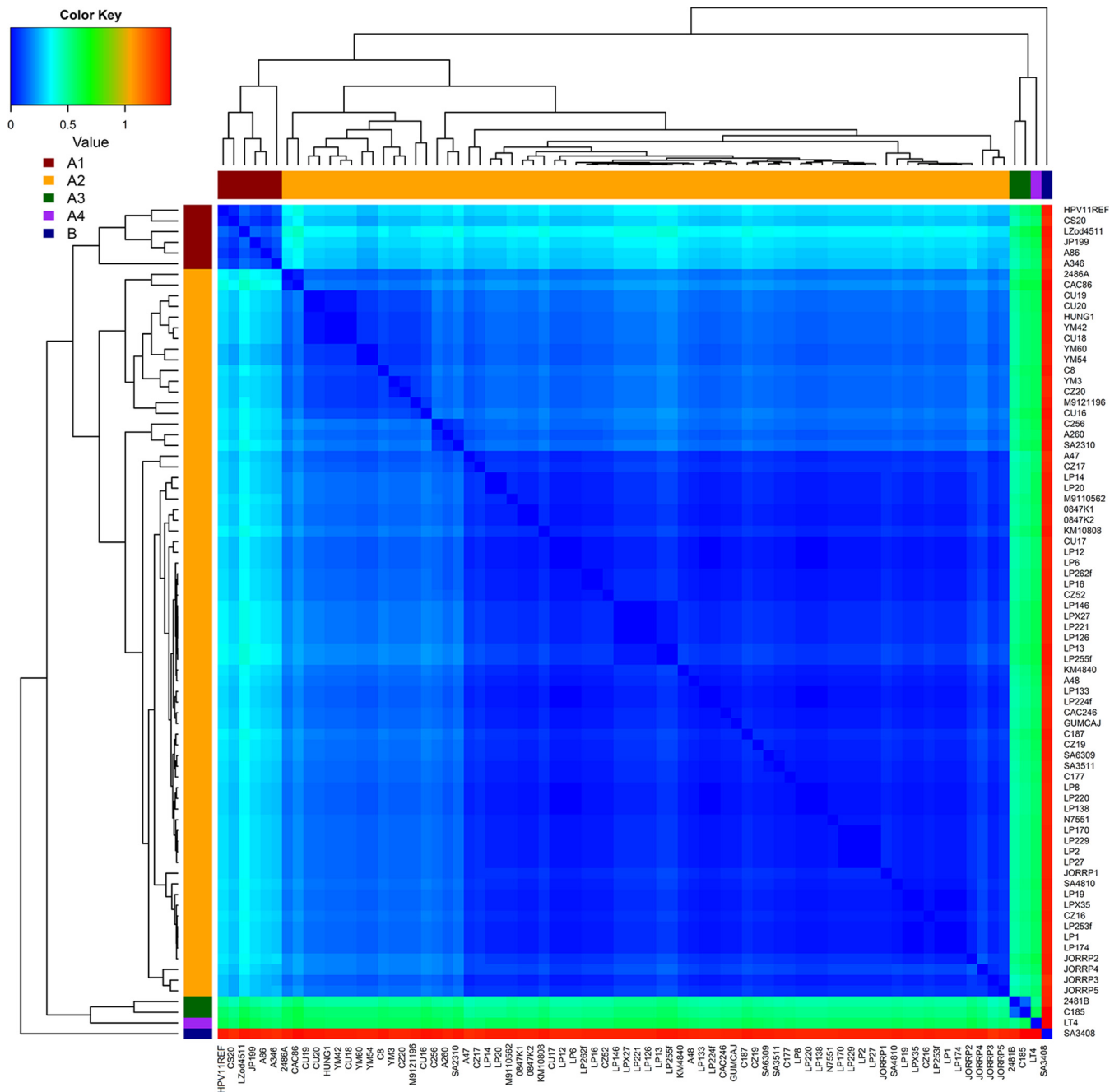


FIG 1 HPV11 heat map and maximum likelihood trees inferred from 78 complete HPV11 genomes. All 78 complete genomes were aligned, and the pairwise differences were calculated. The minimum sequence differences (0% pairwise difference, dark blue) and the maximum sequence differences (1.3% pairwise difference, red) are displayed. To demonstrate how the percent nucleotide sequence differences relate to the phylogenetic tree topology, maximum likelihood trees are shown alongside the heat map. Lineages and sublineages are labeled and have 90 to 100% bootstrap support.

genomic regions, including the E5a, E5b, and L1 ORFs and the noncoding region URR were sequenced. These particular regions were selected for initial sequencing on the basis of findings from previous studies: (i) E5a, L1, and URR contain sequence information sufficient for unambiguous differentiation between HPV11 sublineages A1 and A2 (25, 26, 28); (ii) the E5a and E5b genes evolve faster and mutate more frequently than other HPV genes (40, 41); and (iii) our recent genomic diversity study of HPV6 (the closest relative of HPV11) showed that the concatenated E5a-E5b-

L1-URR sequence can be reliably used as a surrogate for the whole-genome phylogenetic clustering of HPV6 variants (29), and thus, sequencing of the same genomic regions would additionally allow comparison of the genetic diversity of a large number of HPV6 and HPV11 isolates. The genomes of a total of 30 HPV11 isolates showing the highest diversity in the E5a, E5b, L1, and/or URR genomic region were fully sequenced and phylogenetically evaluated together with 48 complete HPV11 genome sequences from GenBank and PaVE, of which 70.1% (34/48) mainly originated from two pre-

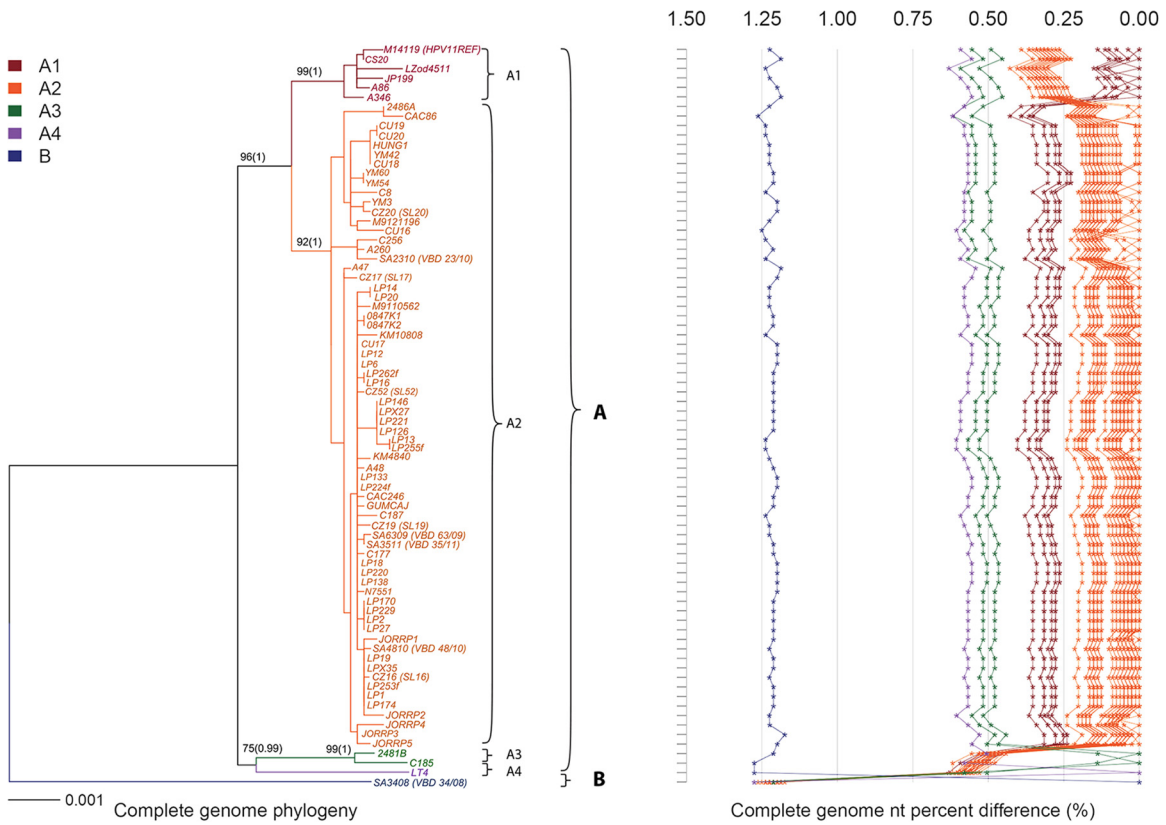


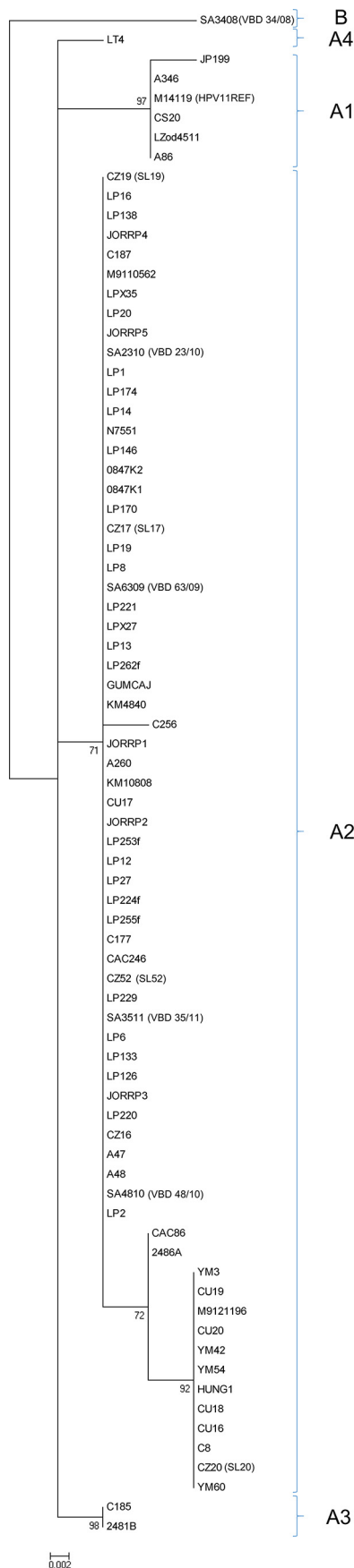
FIG 2 Tree topology and pairwise comparisons of 78 complete HPV11 genomes. RAxML and Bayesian trees were inferred from the global alignment of 78 complete HPV11 genomes (the RAxML tree is shown on the left). Support scores alongside the branches indicate bootstrap percentages obtained from the RAxML HPC2 software; additionally, Bayesian credibility values obtained with MrBayes software are indicated in parentheses. The pairwise nucleotide sequence differences were calculated for each isolate and are shown on the right, with the scale shown at the top of the panel. Values for each comparison for a given isolate are connected by lines, and the comparison to self is indicated as 0.00%.

vious Slovenian studies (25, 34). Overall, 64/78 (82.1%) of the complete HPV11 genome sequences determined to date were identified by members of our international HPV11 consortium.

Nucleotide sequence differences across the complete HPV genome of 1.0% to 10.0% and 0.5% to 1.0% define distinct HPV variant lineages and sublineages, respectively (26, 42). In this study, the calculated maximal pairwise difference between nucleotide sequences among the 78 complete HPV11 genomes analyzed was 1.3%, indicating the existence of two HPV11 variant lineages. The maximum likelihood tree obtained with 78 complete HPV11 genome sequences further confirmed two deeply separated variant clusters, named variant lineages A and B. In addition, lineage A consisted of four sublineages, A1, A2, A3, and A4, with sublineages A3 and A4 being identified for the first time in this study. Due to the identification of a novel lineage B, originating from a single African isolate, the calculated maximum pairwise dissimilarity of the complete HPV11 genome variants (e.g., isolate LT4 versus isolate SA3408) reached 1.3%, exceeding the previously reported rate of diversity by approximately 4-fold (25, 26, 28). Hence, the genomic diversity of HPV11 of 1.3% is highly similar to that of its closest relative, HPV6 (1.6%) (29). It is distinctly lower than that of the two most important high-risk HPV types, HPV16 (2.3%) and HPV18 (2.1%), and similar to that of some of their closest relatives, HPV31 (1.4%), HPV45 (1.5%), and HPV70 (1.6%) (43–45). Although this study consisted of an uneven initial sample set of HPV11

variants from specific continents/countries, it has shown that HPV11 acquired nucleotide sequence changes that over time led to the fixation of the lineages. However, the newly identified lineage B consisted of only one African isolate, and so additional studies will be needed to further investigate the increased heterogeneity of HPV11 observed in this study.

In accordance with our findings from an HPV6 global genomic diversity study (29), the topology of the HPV11 phylogenetic tree constructed from concatenated E5a-E5b-L1-URR sequences resembled the topology of the tree constructed from the complete HPV11 genome. On the other hand, the topologies of the trees constructed with each of the nine individual ORFs and the URR were dichotomous, clearly identifying lineages A and B but poorly resolving sublineages A1 to A4. These observations further support the established concept of complete genome sequencing for the basis of HPV variant classification (42, 46). Because nucleotide changes specific to genetic (sub)lineages are not evenly dispersed throughout the HPV genome, the generation of phylogenetic trees with sequences from individual genomic regions may result in trees with different topologies. Therefore, only stable (sub)lineage-specific SNPs could serve as diagnostic sites in studies in which complete genome sequences cannot be obtained (26, 42, 46). In this study, following the identification of several (sub)lineage-specific SNPs, the most suitable genomic region (208 bp) for HPV11 whole-genome tree reconstruction and the identification of all currently known vari-



ant (sub)lineages was found at the 3' end of the E2 ORF and the 5' end of NCR2.

In order to examine ethnogeographical associations among HPV11 (sub)lineages, a total of 433 HPV11 sequences were evaluated. The analysis showed that sublineage A2 prevailed globally and was detected significantly more frequently in all continents except Asia. Sublineage A1 was more common in Asia than the other continents but was not present in samples from Australia and Africa. A limited number of samples from Africa, all originating from a single African country, is an important disadvantage of our study, since the genetic variation of high-risk HPV types has been shown to be the greatest in Africa. Thus, in order to clarify the potential imbalance in the distribution of sublineage A1 among continents observed in our study, further studies with greater numbers of samples from Australia and several African countries are warranted. The newly identified sublineages A3 and A4 and lineage B were all found in low numbers. Given the low mutation rate of papillomaviruses (PVs) (47), the majority of HPV variant lineages/sublineages diverged from their common ancestors within the last million years (48), a time period when archaic hominids began to diverge and spread across the globe. Thus, the widespread HPV11 sublineages A2 and A1 probably adapted to the host well before the global spread of modern humans out of Africa, whereas some (sub)lineages that were identified only in specific continents/populations nearly became extinct in archaic hominids due to the unsuccessful transmission of the virus to modern human ancestors.

This study showed no correlation between HPV11 intratype diversification and the worldwide migration patterns of the human host, a phenomenon that has been well established for the high-risk HPV types HPV16 and HPV18 (49, 50). As with HPV6, we strongly believe that the absence of a strong geographical clustering of HPV11 genomic variants is most probably due to the lower replication rate and delayed fixation of (sub)lineage-specific SNPs throughout the HPV11 genome (24, 26, 29, 43, 51). Therefore, the major HPV11 lineages and sublineages may have evolved well before humans spread to all the continents. Similar (phylo)geographical associations have previously been reported for some of the less frequently encountered high-risk HPV types, such as HPV31, HPV35, HPV52, and HPV58 (51). The underlying force driving the natural variation in viruses is still unclear. During over millions of years of PV evolution, random mutations could have caused a significant portion of functionally neutral genetic changes (with no changes to the phenotype), while some variations or their patterns may have caused differences in viral fitness, as a result of niche adaptation or virus-host codivergence. Newly introduced prophylactic HPV vaccines represent an unforeseen potential driving force of PV evolution. The generation of HPV11 vaccine escape mutants that could impact the efficiency of quadrivalent and nonavalent prophylactic HPV vaccines is of potential concern. However, on the basis of preliminary results of our ongoing study, it seems that, at least for HPV6 and HPV11, this is highly unlikely but warrants further investigation.

When interpreting our data, two important limitations of our study should be considered. First, due to the uneven sample set of

FIG 3 Maximum likelihood tree based on a 208-bp region at the 3' end of the E2 ORF and part of the NCR2 sequence for 78 HPV11 sequences.

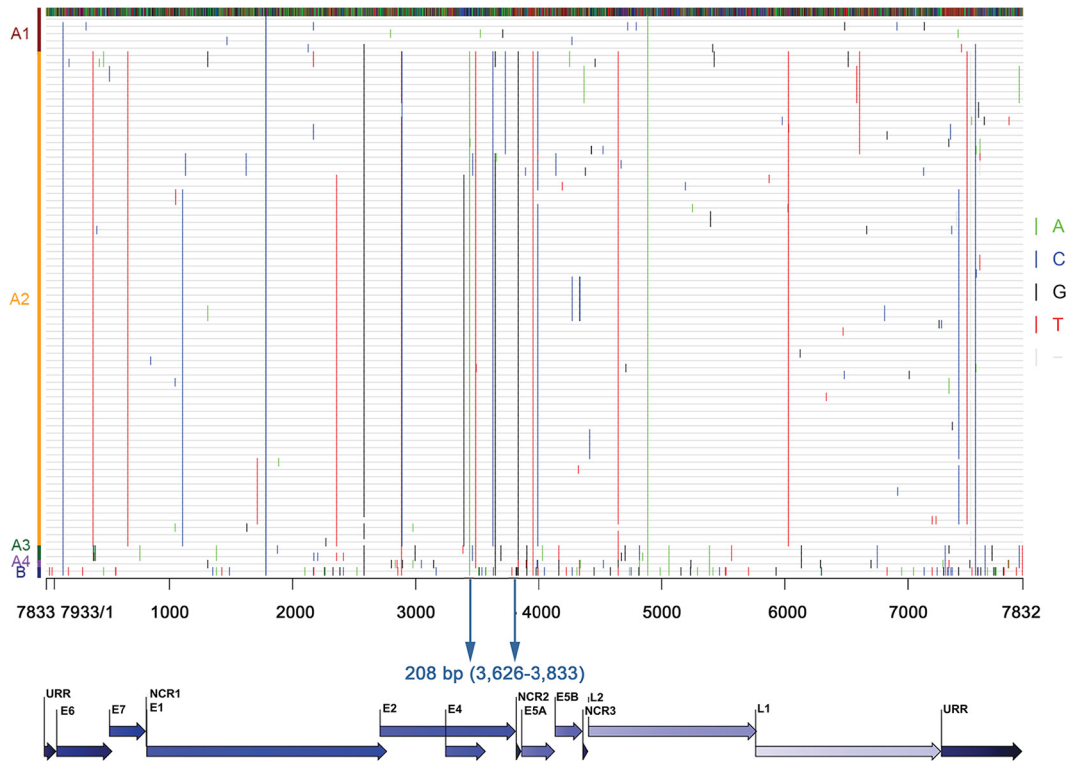


FIG 4 Diagnostic lineage/sublineage-specific SNPs identified from comparisons of the 78 complete HPV11 genomes. SNP positions relative to the genome of the prototype HPV11 isolate (PaVE accession no. HPV11REF) are displayed across lineages (lineages A and B) and sublineages (sublineages A1, A2, A3, and A4). The representative region (208 bp) used for whole-genome-based phylogenetic clustering is marked with vertical arrows.

HPV11 isolates from specific geographical regions, especially South America, North America, Australia, and Africa, we could not exclude the possibility that (sub)lineages detected in low numbers in this study could also exist in isolated, unsampled populations. Second, the

use of the Sanger sequencing method, which has a significantly lower sensitivity for genomic variant detection than next-generation sequencing techniques, could have prevented us from detecting genomic variants present at low copy numbers, especially in rare cases

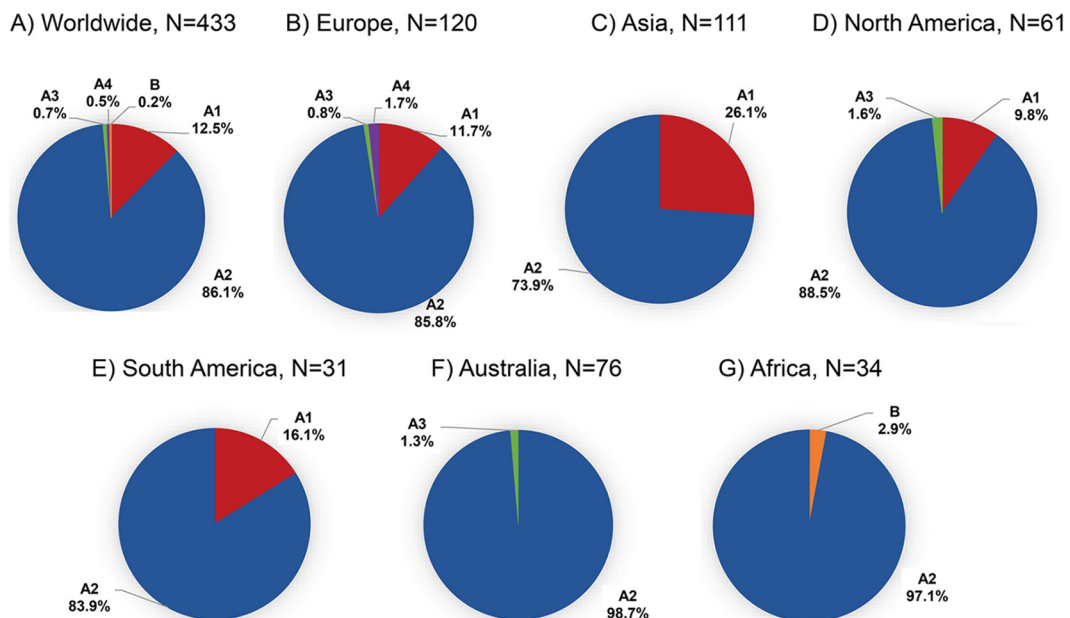


FIG 5 Worldwide and continental distributions of HPV11 variant lineage B and sublineages A1, A2, A3, and A4. (A) Worldwide; (B) Europe; (C) Asia; (D) North America; (E) South America; (F) Australia; (G) Africa.

TABLE 3 Univariate and multiple logistic regression analyses of associations between HPV11 sublineages A1 and A2 and geographical location (continent)

Continent ^a	Univariate regression analysis for A2 vs A1		Multiple regression analysis for A2 vs A1	
	OR (95% CI) ^b	P value	OR (95% CI) ^b	P value
Europe	2.6 (1.3–5.2)	0.007		
North America	3.2 (1.2–8.2)	0.016		
South America	1.8 (0.6–5.2)	0.254		
All continents ^c	4.1 (2.3–7.4)	<0.001	1.6 (0.8–3.4)	0.207

^a Asia was used as the reference category.

^b OR, odds ratio; CI, confidence interval.

^c Europe, North America, South America, Australia, and Africa combined.

of the concurrent presence of multiple HPV11 variants in a single sample (42, 43). Therefore, a more in-depth sampling across all continents, coupled with next-generation sequencing, would probably reveal an even higher level of heterogeneity of HPV11.

In summary, this study is the largest international, global study of the genomic diversity of HPV11 performed to date, analyzing the complete genomes of globally circulating HPV11 variants. Thirty new complete HPV11 genomes were determined and submitted to the available sequence repositories. The global phylogenetic analysis of HPV11 revealed the existence of two variant lineages, lineages A and B, and four sublineages, sublineages A1, A2, A3, and A4. Sublineage A2 prevailed globally, followed by sublineage A1, which was detected in the highest proportion in Asia. Several HPV11 (sub)lineage-specific SNPs and a 208-bp representative region within the E2 ORF and NCR2 for whole-genome-based phylogenetic clustering were identified. This study provides the most comprehensive data on the global genomic diversity of HPV11 to date and will, it is hoped, significantly contribute to future evolutionary, epidemiological, vaccination, and molecular assay development studies, as well as to studies on the pathogenesis of HPV11 and other *Alpha*-PV types.

ACKNOWLEDGMENTS

We thank Ariana Harari for technical advice and critical discussions and Vanja Erčulj from RHO Sigma, Ljubljana, Slovenia, for support and help with statistical analysis. We thank Richard Gibson from ENA, EMBL-EBI, Cambridge, United Kingdom, for technical assistance.

FUNDING INFORMATION

This work, including the efforts of Mateja M. Jelen, was funded by Slovenian Research Agency (1000-10-310163). This work, including the efforts of Lea Hošnjak, Anja Oštrbenk, and Mario Poljak, was funded by European Commission FP7 Program (603019). This work, including the efforts of Mario Poljak, was funded by Merck Sharp and Dohme (research grant). This work, including the efforts of Robert D. Burk, was funded by HHS | NIH | National Cancer Institute (NCI) (CA78527 and P30CA013330). This work, including the efforts of Robert D. Burk, was funded by HHS | National Institutes of Health (NIH) (AI-51519). This work, including the efforts of Mateja M. Jelen, was funded by European Society of Clinical Microbiology and Infectious Diseases (ESCMID) (research grant 2014).

REFERENCES

- Wiatrak BJ, Wiatrak DW, Broker TR, Lewis L. 2004. Recurrent respiratory papillomatosis: a longitudinal study comparing severity associated with human papilloma viral types 6 and 11 and other risk factors in a large pediatric population. *Laryngoscope* 114(Suppl 104):1–23. <http://dx.doi.org/10.1097/01.mlg.000148224.83491.0f>.
- Aubin F, Pr  tet JL, Jacquard AC, Saunier M, Carcopino X, Jaroud F, Pradat P, Soubeyrand B, Leocmach Y, Moug  n C, Riethmuller D. 2008. Human papillomavirus genotype distribution in external acuminata condylomata: a large French national study (EDiTH IV). *Clin Infect Dis* 47: 610–615. <http://dx.doi.org/10.1086/590560>.
- Garland SM, Steben M, Sings HL, James M, Lu S, Railkar R, Barr E, Haupt RM, Joura EA. 2009. Natural history of genital warts: analysis of the placebo arm of 2 randomized phase III trials of a quadrivalent HPV (types 6, 11, 16, 18) vaccine. *J Infect Dis* 199:805–814. <http://dx.doi.org/10.1086/597071>.
- Bernard HU, Burk RD, Chen Z, van Doorslaer K, zur Hausen H, de Villiers EM. 2010. Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology* 401:70–79. <http://dx.doi.org/10.1016/j.virol.2010.02.002>.
- Gillison M, Alemany L, Snijders PJF, Chaturvedi A, Steinberg BM, Schwartz S, Castellsague X. 2012. Human papillomavirus and diseases of the upper airway: head and neck cancer and respiratory papillomatosis. *Vaccine* 30(Suppl 5):F34–F54. <http://dx.doi.org/10.1016/j.vaccine.2012.05.070>.
- Komlo   KF, Kocjan BJ, Ko  orok P, Luzar B, Megli   L, Poto  nik M, Ho  cvar-Bolte  zar I, Gale N, Seme K, Poljak M. 2012. Tumor-specific and gender-specific pre-vaccination distribution of human papillomavirus types 6 and 11 in anogenital warts and laryngeal papillomas: a study on 574 tissue specimens. *J Med Virol* 84:1233–1241. <http://dx.doi.org/10.1002/jmv.23318>.
- Danielewski JA, Garland SM, McCloskey J, Hillman RJ, Tabrizi SN. 2013. Human papillomavirus type 6 and 11 genetic variants found in 71 oral and anogenital epithelial samples from Australia. *PLoS One* 8:e63892. <http://dx.doi.org/10.1371/journal.pone.0063892>.
- Buchinsky FJ, Donfack J, Derkay CS, Choi SS, Conley SF, Myer CM, III, McClay JE, Campisi P, Wiatrak BJ, Sobol SE, Schweinfurth JM, Tsuji DH, Hu FZ, Rockette HE, Ehrlich GD, Post JC. 2008. Age of child, more than HPV type, is associated with clinical course in recurrent respiratory papillomatosis. *PLoS One* 3:e2263. <http://dx.doi.org/10.1371/journal.pone.0002263>.
- Bouvard V, Baan R, Straif K, Grosse Y, Secretan B, El Ghissassi F, Benbrahim-Tallaa L, Guha N, Freeman C, Galichet L, Cogliano V. 2009. A review of human carcinogens. Part B. *Biological agents*. *Lancet Oncol* 10:321–322. [http://dx.doi.org/10.1016/S1470-2045\(09\)70096-8](http://dx.doi.org/10.1016/S1470-2045(09)70096-8).
- Seedat RY, Thukane M, Jansen AC, Rossouw I, Goedhals D, Burt FJ. 2010. HPV types causing juvenile recurrent laryngeal papillomatosis in South Africa. *Int J Pediatr Otorhinolaryngol* 74:255–259. <http://dx.doi.org/10.1016/j.ijporl.2009.11.016>.
- Tjon Pian Gi RE, San Giorgi MR, Slagter-Menkema L, van Hemel BM, van der Laan BF, van den Heuvel ER, Dijkers FG, Schuurings EM. 2015. Clinical course of recurrent respiratory papillomatosis: Comparison between aggressiveness of human papillomavirus-6 and human papillomavirus-11. *Head Neck* 37:1625–1632. <http://dx.doi.org/10.1002/hed.23808>.
- Pr  tet J-C, Jacquard A-C, Saunier M, Clavel C, Dachez R, Gondry J, Pradat P, Soubeyrand B, Leocmach Y, Moug  n C, Riethmuller D. 2008. Human papillomavirus genotype distribution in low-grade squamous intraepithelial lesions in France and comparison with CIN2/3 and invasive cervical cancer—the EDiTH III study. *Gynecol Oncol* 110:179–184. <http://dx.doi.org/10.1016/j.ygyno.2008.04.012>.
- Bzhalava D, Guan P, Franceschi S, Dillner J, Clifford G. 2013. A systematic review of the prevalence of mucosal and cutaneous human papillomavirus types. *Virology* 445:224–231. <http://dx.doi.org/10.1016/j.virol.2013.07.015>.
- Gillou L, Sahli R, Chaubert P, Monnier P, Cuttat JF, Costa J. 1991. Squamous cell carcinoma of the lung in a nonsmoking, nonirradiated patient with juvenile laryngotracheal papillomatosis. Evidence of human papillomavirus-11 DNA in both carcinoma and papillomas. *Am J Surg Pathol* 15:891–898.
- But-Had  zi   J, Jenko K, Poljak M, Kocjan BJ, Gale N, Strojan P. 2011. Sinonasal inverted papilloma associated with squamous cell carcinoma. *Radiol Oncol* 45:267–272. <http://dx.doi.org/10.2478/v10019-011-0033-4>.
- Katsenos S, Becker HD. 2011. Recurrent respiratory papillomatosis: a rare chronic disease, difficult to treat, with potential to lung cancer transformation: apropos of two cases and a brief literature review. *Case Rep Oncol* 4:162–171. <http://dx.doi.org/10.1159/000327094>.
- Yuan H, Myers S, Wang J, Zhou D, Woo JA, Kallakury B, Ju A,

- Bazylewicz M, Carter YM, Albanese C, Grant N, Shad A, Dritschilo A, Liu X, Schlegel R. 2012. Use of reprogrammed cells to identify therapy for respiratory papillomatosis. *N Engl J Med* 367:1220–1227. <http://dx.doi.org/10.1056/NEJMoa1203055>.
18. Cornall AM, Roberts JM, Garland SM, Hillman RJ, Grulich AE, Tabrizi SN. 2013. Anal and perianal squamous carcinomas and high-grade intraepithelial lesions exclusively associated with “low-risk” HPV genotypes 6 and 11. *Int J Cancer* 133:2253–2258. <http://dx.doi.org/10.1002/ijc.28228>.
19. Gáll T, Kis A, Tatár TZ, Kardos G, Gergely L, Szarka K. 2013. Genomic differences in the background of different severity in juvenile-onset respiratory papillomatosis associated with human papillomavirus type 11. *Med Microbiol Immunol* 202:353–363. <http://dx.doi.org/10.1007/s00430-013-0297-y>.
20. Guimerà N, Lloveras B, Alemany L, Iljazovic E, Shin HR, Jung-II S, de Sanjose S, Jenkins D, Bosch FX, Quint W. 2013. Laser capture microdissection shows HPV11 as both a causal and a coincidental infection in cervical cancer specimens with multiple HPV types. *Histopathology* 63:287–292. <http://dx.doi.org/10.1111/his.12137>.
21. Mauz PS, Zago M, Kurth R, Pawlita M, Holderried M, Thiericke J, Iftner A, Stubenrauch F, Sotlar K, Iftner T. 2014. A case of recurrent respiratory papillomatosis with malignant transformation, HPV11 DNAemia, high L1 antibody titre and a fatal papillary endocardial lesion. *Virology* 11:114. <http://dx.doi.org/10.1186/1743-422X-11-114>.
22. Chow EPF, Read TRH, Wigan R, Donovan B, Chen MY, Bradshaw CS, Fairley CK. 2015. Ongoing decline in genital warts among young heterosexuals 7 years after the Australian human papillomavirus (HPV) vaccination programme. *Sex Transm Infect* 91:214–219. <http://dx.doi.org/10.1136/sextrans-2014-051813>.
23. Joura EA, Giuliano AR, Iversen OE, Bouchard C, Mao C, Mehlsen J, Moreira ED, Jr, Ngan Y, Petersen LK, Lazcano-Ponce E, Pitisuttithum P, Restrepo JA, Stuart G, Woelber L, Yang YC, Cuzick J, Garland SM, Huh W, Kjaer SK, Bautista OM, Chan IS, Chen J, Gesser R, Moeller E, Ritter M, Vuocolo S, Luxembourg A, Broad Spectrum HPV Vaccine Study. 2015. A 9-valent HPV vaccine against infection and intraepithelial neoplasia in women. *N Engl J Med* 372:711–723. <http://dx.doi.org/10.1056/NEJMoa1405044>.
24. Heinzel PA, Chan SY, Ho L, O'Connor M, Balaram P, Campo MS, Fujinaga K, Kiviati N, Kuypers J, Pfister H, Steinberg BM, Tay S-K, Villa LL, Bernard H-U. 1995. Variation of human papillomavirus type 6 (HPV-6) and HPV-11 genomes sampled throughout the world. *J Clin Microbiol* 33:1746–1754.
25. Maver PJ, Kocjan BJ, Seme K, Potočnik M, Gale N, Poljak M. 2011. Prevaccination genomic diversity of human papillomavirus genotype 11: a study on 63 clinical isolates and 10 full-length genome sequences. *J Med Virol* 83:461–470. <http://dx.doi.org/10.1002/jmv.21994>.
26. Burk RD, Chen Z, Harari A, Smith BC, Kocjan BJ, Maver PJ, Poljak M. 2011. Classification and nomenclature system for human alphapapillomavirus variants: general features, nucleotide landmarks and assignment of HPV6 and HPV11 isolates to variant lineages. *Acta Dermatovenereol Alp Panonica Adriat* 20:113–123.
27. de Matos RP, Sichero L, Mansur IM, do Bonfim CM, Bittar C, Nogueira RL, Küpper DS, Valera FC, Nogueira ML, Villa LL, Calmon MF, Rahal P. 2013. Nucleotide and phylogenetic analysis of human papillomavirus types 6 and 11 isolated from recurrent respiratory papillomatosis in Brazil. *Infect Genet Evol* 16:282–289. <http://dx.doi.org/10.1016/j.meegid.2012.12.033>.
28. Godínez JM, Nicolás-Párraga S, Pimenoff VN, Mengual-Chuliá B, Muñoz N, Bosch FX, Sánchez GI, McCloskey J, Bravo IG. 2014. Phylogenetically related, clinically different: human papillomaviruses 6 and 11 variants distribution in genital warts and in laryngeal papillomatosis. *Clin Microbiol Infect* 20:O406–O413. <http://dx.doi.org/10.1111/1469-0691.12420>.
29. Jelen MM, Chen Z, Kocjan BJ, Burt FJ, Chan PK, Chouhy D, Combrinck CE, Coutlée F, Estrade C, Ferenczy A, Fiander A, Franco EL, Garland SM, Giri AA, González JV, Gröning A, Heidrich K, Hibbitts S, Hošnjak L, Luk TN, Marinic K, Matsukura T, Neumann A, Oštrbenk A, Piccini MA, Richardson H, Sagadin M, Sahli R, Seedat RY, Seme K, Severini A, Sinchi JL, Smahelova J, Tabrizi SN, Tachezy R, Tohme S, Uloza V, Vitkauskienė A, Wong YW, Zidovec Lepej S, Burk RD, Poljak M. 2014. Global genomic diversity of human papillomavirus 6 based on 724 isolates and 190 complete genome sequences. *J Virol* 88:7307–7316. <http://dx.doi.org/10.1128/JVI.00621-14>.
30. Wu X, Zhang C, Feng S, Liu C, Li Y, Yang Y, Gao J, Li H, Meng S, Li L, Zhang Y, Hu X, Wu X, Lin L, Li X, Wang Y. 2009. Detection of HPV types and neutralizing antibodies in Gansu Province, China. *J Med Virol* 81:693–702. <http://dx.doi.org/10.1002/jmv.21435>.
31. Gáll T, Kis A, Fehér E, Gergely L, Szarka K. 2011. Virological failure of intralesional cidofovir therapy in recurrent respiratory papillomatosis is not associated with genetic or epigenetic changes of HPV11: complete genome comparison of sequential isolates. *Antiviral Res* 92:356–358. <http://dx.doi.org/10.1016/j.antiviral.2011.09.007>.
32. Chansaenroj J, Theamboonlers A, Junyangdikul P, Supiyaphan P, Poovorawan Y. 2012. Whole genome analysis of human papillomavirus genotype 11 from cervix, larynx and lung. *Asian Pac J Cancer Prev* 13:2619–2623. <http://dx.doi.org/10.7314/APJCP.2012.13.6.2619>.
33. Yuan H, Zhou D, Wang J, Schlegel R. 2013. Divergent human papillomavirus associated with recurrent respiratory papillomatosis with lung involvement. *Genome Announc* 1(4):e00474-13. <http://dx.doi.org/10.1128/genomeA.00474-13>.
34. Kocjan BJ, Gale N, Hočevar Boltežar I, Seme K, Fujs Komloš K, Hošnjak L, Maver PJ, Jelen MM, Zupanič Pajnič I, Balazič J, Poljak M. 2013. Identical human papillomavirus (HPV) genomic variants persist in recurrent respiratory papillomatosis for up to 22 years. *J Infect Dis* 207:583–587. <http://dx.doi.org/10.1093/infdis/jis733>.
35. Katoh K, Toh H. 2010. Parallelization of the MAFFT multiple sequence alignment program. *Bioinformatics* 26:1899–1900. <http://dx.doi.org/10.1093/bioinformatics/btq224>.
36. Stamatakis A. 2014. RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313. <http://dx.doi.org/10.1093/bioinformatics/btu033>.
37. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* 61:539–542. <http://dx.doi.org/10.1093/sysbio/sys029>.
38. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28:2731–2739. <http://dx.doi.org/10.1093/molbev/msr121>.
39. Maddison DR, Maddison WP. 2005. MacClade 4: analysis of phylogeny and character evolution, version 4.08a. <http://macclade.org/index.html>. Accessed 2 April 2016.
40. Bravo IG, Alonso A. 2004. Mucosal human papillomaviruses encode four different E5 proteins whose chemistry and phylogeny correlate with malignant or benign growth. *J Virol* 78:13613–13626. <http://dx.doi.org/10.1128/JVI.78.24.13613-13626.2004>.
41. Van Doorslaer K. 2013. Evolution of the Papillomaviridae. *Virology* 445:11–20. <http://dx.doi.org/10.1016/j.virol.2013.05.012>.
42. Burk RD, Harari A, Chen Z. 2013. Human papillomavirus genome variants. *Virology* 445:232–243. <http://dx.doi.org/10.1016/j.virol.2013.07.018>.
43. Chen Z, Schiffman M, Herrero R, Desalle R, Anastos K, Segondy M, Sahasrabudde VV, Gravitt PE, Hsing AW, Burk RD. 2011. Evolution and taxonomic classification of human papillomavirus 16 (HPV16)-related variant genomes: HPV31, HPV33, HPV35, HPV52, HPV58 and HPV67. *PLoS One* 6:e20183. <http://dx.doi.org/10.1371/journal.pone.0020183>.
44. Smith B, Chen Z, Reimers L, Van Doorslaer K, Schiffman M, Desalle R, Herrero R, Yu K, Wacholder S, Wang T, Burk RD. 2011. Sequence imputation of HPV16 genomes for genetic association studies. *PLoS One* 6:e21375. <http://dx.doi.org/10.1371/journal.pone.0021375>.
45. Chen Z, Schiffman M, Herrero R, DeSalle R, Anastos K, Segondy M, Sahasrabudde VV, Gravitt PE, Hsing AW, Burk RD. 2013. Evolution and taxonomic classification of alphapapillomavirus 7 complete genomes: HPV18, HPV39, HPV45, HPV59, HPV68 and HPV70. *PLoS One* 8:e72565. <http://dx.doi.org/10.1371/journal.pone.0072565>.
46. Harari A, Chen Z, Burk RD. 2014. Human papillomavirus genomics: past, present and future. *Curr Probl Dermatol* 45:1–18. <http://dx.doi.org/10.1159/000355952>.
47. Rector A, Lemey P, Tachezy R, Mostmans S, Ghim SJ, Van Doorslaer K, Roelke M, Bush M, Montali RJ, Joslin J, Burk RD, Jenson AB, Sundberg JP, Shapiro B, Van Ranst M. 2007. Ancient papillomavirus-host co-speciation in Felidae. *Genome Biol* 8:R57. <http://dx.doi.org/10.1186/gb-2007-8-4-r57>.

48. Chen Z, DeSalle R, Schiffman M, Herrero R, Burk RD. 2009. Evolutionary dynamics of variant genomes of human papillomavirus types 18, 45, and 97. *J Virol* 83:1443–1455. <http://dx.doi.org/10.1128/JVI.02068-08>.
49. Ho L, Chan SY, Burk RD, Das BC, Fujinaga K, Icenogle JP, Kahn T, Kiviat N, Lancaster W, Mavromara-Nazos P. 1993. The genetic drift of human papillomavirus type 16 is a means of reconstructing prehistoric viral spread and the movement of ancient human populations. *J Virol* 67:6413–6423.
50. Ong CK, Chan SY, Campo MS, Fujinaga K, Mavromara-Nazos P, Labropoulou V, Pfister H, Tay SK, ter Meulen J, Villa LL. 1993. Evolution of human papillomavirus type 18: an ancient phylogenetic root in Africa and intratype diversity reflect coevolution with human ethnic groups. *J Virol* 67:6424–6431.
51. Calleja-Macias IE, Villa LL, Prado JC, Kalantari M, Allan B, Williamson A-L, Chung L-P, Collins RJ, Zuna RE, Dunn ST, Chu T-Y, Cubie HA, Cuschieri K, von Knebel-Doeberitz M, Martins CR, Sanchez GI, Bosch FX, Munoz N, Bernard H-U. 2005. Worldwide genomic diversity of the high-risk human papillomavirus types 31, 35, 52, and 58, four close relatives of human papillomavirus type 16. *J Virol* 79:13630–13640. <http://dx.doi.org/10.1128/JVI.79.21.13630-13640.2005>.