Infection, Genetics and Evolution 26 (2014) 8–13

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Human papillomavirus type 16 long control region and E6 variants stratified by cervical disease stage

Luigi Marongiu, Anna Godi, John V. Parry, Simon Beddows *

Virus Reference Department, Public Health England, London, UK

ARTICLE INFO

Article history:
Received 20 March 2014
Received in revised form 30 April 2014
Accepted 5 May 2014
Available online 10 May 2014

Keywords:
Human papillomavirus
Cervical cancer
HPV16
Variants
LCR
E6

ABSTRACT

Objective: Certain intra-type variants of HPV16 have been shown to be associated with an increased risk of developing high grade cervical disease, but their potential association is confounded by apparent geographic and phylogenetic lineage dependency. The objective of this study was to evaluate the relationship between HPV16 sequence variants and cervical disease stage in monospecific infection samples from a single lineage (European, EUR) in England.

Methods: One hundred and twelve women singly infected with HPV16 and displaying normal and abnormal cytology grades were selected. An 1187 bp fragment encompassing the entire LCR and a portion of the E6 open reading frame was sequenced to identify intra-type variants. Intra-type diversity was estimated using Shannon entropy.

Results: Almost all samples (110/112; 98%) were assigned to the EUR lineage, one sample was classified as European-Asian (EAS) and another African (Afr1a). The mean pairwise distance of the EUR sequences in this study was low (0.29%; 95%CI 0.13–0.45%) but there were nevertheless several sites in the LCR ($n = 5$) and E6 ($n = 2$) that exhibited a high degree of entropy. None of these sites, however, including the T350G non-synonymous (L83V) substitution in E6, alone or in combination, were found to be associated with cervical disease stage.

Conclusions: Despite using single infection samples and samples from a single variant lineage, intra-type variants of HPV16 were not differentially associated with cervical disease. Monitoring intra-lineage, site-specific variants, such as T350G, is unlikely to be of diagnostic value.

1. Introduction

Persistent infection with oncogenic genotypes of genital Human Papillomavirus (HPV) is associated with the development of cervical cancer, a significant cause of morbidity and mortality of women worldwide (Schiffman et al., 2007). There are about a dozen HPV genotypes associated with the development of cervical disease (Bouvard et al., 2009), differing in their relative contributions to the overall prevalence of disease (Guan et al., 2012), with HPV16 displaying the highest prevalence worldwide and the strongest disease association. The underlying mechanisms that drive a minority of HPV16 infections to persist, leading to the development of precancerous lesions and ultimately cervical cancer are poorly understood. Principal viral cofactors include the oncoproteins, E6 and E7, that drive cellular deregulation under the control of the upstream Long Control Region (LCR) (Bernard, 2013; Moody and Laimins, 2010).

The DNA-based HPV genome is efficiently replicated by cell polymerases with an error rate of *ca.* 1×10^{-8} base substitutions, per site, per year (Chen et al., 2009) which is considerably lower than for the HIV-1 RNA genome (*ca.* 1×10^{-3}) for example (Korber et al., 2000). Nevertheless, over millennia, the HPV genome has diverged into genotypes and intra-genotype variants (Burk et al., 2013), compounded by diversifying selection pressure(s) (Chen et al., 2005). The intra-genotype variation of HPV16 appears to be closely aligned with human migration patterns (Chan et al., 1992), leading to the designation of pseudo-geographic lineages based upon LCR diversity: European (EUR), European-Asian (EAS), African 1 (Afr1a, Afr1b), African 2 (Afr2a, Afr2b), Asian-American (AA1 and AA2) and North-American (NA), although the actual geographical distribution of these lineage variants is somewhat more complex (Cornet et al., 2012; Yamada et al., 1997).

It has been observed that Non-European variants of HPV16, particularly of the AA lineage, have a higher propensity for persistence (Schiffman et al., 2010; Villa et al., 2000; Zuna et al., 2009), and

* Corresponding author. Address: Virus Reference Department, Public Health England, 61 Colindale Avenue, London NW9 5EQ, UK. Tel.: +44 (0)20 8327 6169.
E-mail address: simon.beddows@phe.gov.uk (S. Beddows).

http://dx.doi.org/10.1016/j.meegid.2014.05.009
1567-1348/

perhaps because of this have a stronger association with high grade disease (CIN2/3) (Xi et al., 1997, 2007). A recent meta-analysis of worldwide HPV16 lineage distribution data confirmed the association of certain lineages with an increased risk of cervical disease but also noted some geographical dependency to these associations (Cornet et al., 2013b). Within the EUR variant lineage, a T350G substitution in the E6 gene leads to an altered amino acid residue (L83V) and has been associated with HPV16 persistence (Gheit et al., 2011) and cervical disease (Andersson et al., 2000), though not in all cases (Chan et al., 2002; Cornet et al., 2013a; Nindl et al., 1999; Zuna et al., 2009). Two recent meta-analyses demonstrate that the E350 codon is associated with cervical disease but is likely to be geographically dependent (Cornet et al., 2013b; Tornesello et al., 2011). The LCR is traditionally used to ascertain the variant lineage of a sequence as this region contains a number of lineage-dependent diagnostic markers (Cornet et al., 2012), but data on the potential for an association of individual sites within the LCR and cervical disease are limited. Some studies have found an association with individual sites in the LCR (Chang et al., 2013; Schmidt et al., 2001), while others have not (Kammer et al., 2002), although the significance of these observations have been confounded by the use of relatively small panels of samples containing mixed lineages.

In addition, mixed infections are common throughout the course of cervical disease (Guan et al., 2012). Few of these studies have explicitly used, or separately analyzed, samples harboring a single HPV type wherein the association between the HPV type under evaluation and cervical disease can be made with some confidence.

The biological basis for the differential disease outcomes following infection by HPV16 lineage variants is unclear. The HPV16 AA lineage variant and the EUR T350G (L83V) variant exhibit an increased propensity to transform primary human foreskin keratinocytes or induce apoptosis in organotypic raft cultures *in vitro* compared to the HPV16 prototype (Niccoli et al., 2012; Richard et al., 2010; Zehbe et al., 2011). HPV16 LCR variant lineages have been shown to differentially drive transcriptional activity of the p97 promoter (Kammer et al., 2000), as have other combinations of site-specific variations within and outside the principal transcriptional motifs in the LCR (Bernard, 2013; Lace et al., 2009), although their impact individually is less certain (Lace et al., 2009; Veress et al., 1999).

In this study, we evaluated LCR–E6 sequence variation within the EUR lineage of HPV16 in single infection samples to understand the distribution of such variation and its potential role in cervical disease.

2. Materials and methods

2.1. Samples

The present study made use of extracted DNA (archived at −25 °C) from individuals singly infected with HPV16 ($n = 112$) from a cohort of 4719 women attending cervical screening in England wherein HPV genotyping was conducted using the LINEAR ARRAY® HPV Genotyping Test (Roche Molecular Diagnostics) (Howell-Jones et al., 2010). Cytology samples were graded using Dysplasia nomenclature (Schiffman et al., 2007; Sherman, 2003). Under this system, cytology grades of moderate and severe dysplasia are equivalent to a designation of High grade Squamous Intraepithelial Lesion (HSIL) using Bethesda nomenclature. Similarly, grades of borderline (Atypical Squamous Cells of Undetermined Significance; ASCUS) and mild (Low grade Squamous Intraepithelial Lesion; LSIL) dysplasia are equivalent to the indicated grades using the Bethesda system. For the purposes of analysis, cytology grades of normal, borderline and mild dyskaryosis were categorized as low grade (controls) and cytology grades of moderate and severe dyskaryosis as high grade (cases). Accompanying histological data were only available for *ca.* 15% of the cytology samples in the total study cohort. Amongst these, CIN2+ was diagnosed in 18% of borderline or mild dyskaryosis and in 79% of moderate or severe dyskaryosis. The median age of the individuals for which samples were used in the present study was 36 (inter-quartile range, IQR, 29–46). There was no significant difference between the ages of the individuals in the low grade (controls) group (median 34; IQR 27–46) and those in the high grade (cases) group (median 39; IQR 32–47) ($p = 0.348$). The testing of residual, anonymized DNA extracts for the purposes of improved understanding of cervical disease was approved by the Harrow Research Ethics Committee, UK (08/H0719/17).

The full genome HPV16 plasmid was kindly provided by the German Cancer Research Centre, Heidelberg, Germany (E.M. de Villiers). Site-specific numbering was made according to the HPV16 reference sequence, K02718 (http://pave.niaid.nih.gov).

2.2. LCR–E6 PCR and sequencing

HPV16 samples were amplified in a single 1329 bp fragment with primers 16-F101 (5′-ACCCACCACCTCATCTACC-3′) and 16-R20 (5′-TGCAACAAGACATACATCGACC-3′), annealing at positions 7100–7118 and 524–503, respectively. Template amplification was performed in a 25 μL reaction mix containing Kapa HiFi HotStart ReadyMix (Kapa Biosystems), 5 pmol of each primer and 5 μL of template DNA sample under the indicated cycling conditions (95 °C for 5 min followed by 40 cycles of 98 °C for 20 s, 63 °C for 15 s, 72 °C for 2 min, and final elongation at 72 °C for 5 min) on a PTC-200 thermal cycler (MS research). The resulting PCR product was evaluated for its molecular weight using a DNA mass ladder (Invitrogen) and GelAnalyzer software (www.gelanalyzer.com) and sequenced using 5 pmol of each indicated sequencing primer 16-F101, 16-R20, 16-F2 (5′-GAAACCGGTTAGTATAAAA GCAGAC-3′, nucleotides 53–77), 16-F480 (5′-AATGTGTTTTTTTAAATAG-3′, nucleotides 7481–7499), 16-R28 (5′-AGTTGTTTGCAGCTCTG TGA-3′, nucleotides 169–149), and 16-R38 (5′-GTAAG GCGTTGGCGCATAGTG-3′, nucleotides 7686–7666). All primers were designed using Oligo Explorer software (www.genelink.com). The HPV16 plasmid was re-sequenced during the study ($n = 10$) with 100% identity. In addition, 22 clinical samples were re-amplified and re-sequenced to ensure quality of the procedure and integrity of the data returning 100% identity to their initial sequences. The resulting sequence fragment corresponds to an 1178 bp stretch covering the entire LCR and the first 427 bp of the E6 ORF. The accession numbers for sequences generated in this study are KJ754940–KJ755051. They comprise samples graded as cytologically Normal (No_1 to No_6; KJ755002–KJ755007), and those exhibiting Borderline (Bo_1 to Bo_9; KJ754940–KJ754948), Mild (mi_1 to mi_19; KJ754949–KJ754967), Moderate (Mo_1 to Mo_34; KJ754968–KJ755001) or Severe (Se_1 to Se_44; KJ755 008–KJ755051) dysplasia.

2.3. Phylogenetic and statistical analyses

EUR ($n = 145$) and non-EUR ($n = 208$) LCR–E6 HPV16 background sequences (Cornet et al., 2012) were downloaded from the National Center for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov/). Sequence alignments and Neighbor-Joining (NJ) tree construction (500 bootstrapped replicates) were made using MEGA v4.1 (Tamura et al., 2007). The average number of base substitutions per site was based on the pairwise analysis of resulting sequence alignments including an estimate of standard error (500 bootstrapped replicates) using the Maximum Composite Likelihood model (MEGA).

Site-specific Shannon entropy was estimated using the *Entropy One* program (http://www.hiv.lanl.gov/content/sequence/ENTROPY/entropy_one.html) (Korber et al., 1994). For comparison purposes, a level of 5% site variation equates to an entropy score of *ca.* 0.2.

The Fisher's exact test was used to test for differences in proportions of variants between cases and controls with crude odds ratios (95%CI) also given. Based on an equal distribution of prototype (E350T) and variant (E350G) sequences overall (Cornet et al., 2013b; Tornesello et al., 2011) we ascertained that 50 samples from each of the cases and controls would be sufficient to observe a 50% difference in the distribution of the prototype and variant sequences between the cases and controls ($p < 0.05$). The non-parametric Mann–Whitney *U* test was used to examine the age distribution within the cases and control cytology groups. Tests were 2-tailed where appropriate and all tests were carried out using Stata 12.1 (StataCorp, USA).

3. Results and discussion

LCR–E6 sequences were generated from 112 liquid-based cytology samples singly infected with HPV16. To ascertain the variant lineage of these sequences, we initially aligned the LCR fragment with 353 LCR sequences from a recent study on the global distribution of HPV16 variants (Cornet et al., 2012) (Fig. 1A). Almost all (110/112, 98%) sequences from this present study were of the European (EUR) variant lineage apart from one sequence that aligned with the EAS variant lineage and one sequence that aligned with the Afr1a variant lineage. These two sequences were removed from further analysis. The remaining 110 EUR LCR–E6 sequences were then aligned with 145 global EUR LCR–E6 sequences (Cornet et al., 2012), to identify any potential sub-lineages (Fig. 1B). Sub-clustering within the EUR lineage was only weakly supported, with bootstrap values ranging between 50% and 70%.

We next used Shannon entropy to estimate site-specific variation (Fig. 2) which highlighted seven sites of significant variation: five in the LCR (G7193T, A7316C, T7449C, T7495C and G7520) and two in the E6 gene (T109C and T350G). The distribution of these variant sites was not clearly associated with any potential sub-clustering of the EUR lineage (Fig. 1B), with few exceptions. For example, T7449C was overrepresented in South Asian sequences (65% of sequences; 95%CI 45–81%) compared to the average (18%; 95%CI 14–23%; $n = 255$) ($p < 0.001$) and the T350G was underrepresented in East Asian sequences (8%; 95%CI 2–38%) compared to the average (44%; 95%CI 38–51%) ($p = 0.014$). These differences aside, EUR LCR–E6 variant sequences from England could be found distributed throughout these potential EUR variant sub-clusters (Fig. 1B). There were many other variant sites within the LCR region ($n = 48$ sites; 0.9–3.6% of total sequences) and E6 gene ($n = 8$; 0.9–2.7%), as exemplified by entropy scores of <0.2 (Fig. 2), but with each represented by only a handful of sequences within the panel. The mean pairwise distance of EUR LCR–E6 sequences in this study based on a single European country (0.29%; 95%CI 0.13–0.45%; $n = 110$) was similar to the mean pairwise distance of the geographically dispersed EUR sequences

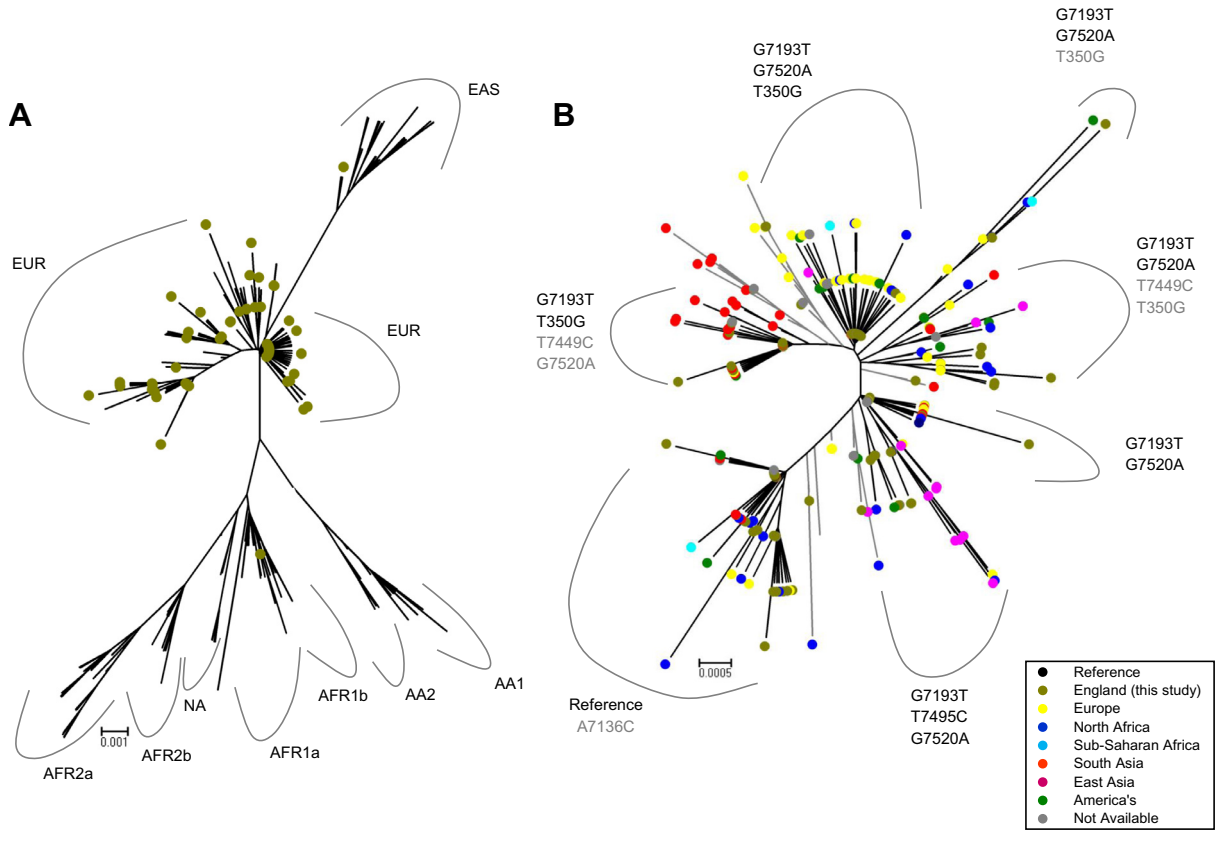

Fig. 1. Phylogenetic distribution of LCR–E6 variants. Radial neighbor-joining phylogenetic trees derived using the HPV16 LCR–E6 EUR variants from this study ($n = 112$) and (A) worldwide ($n = 353$) or (B) only background EUR variants ($n = 145$) (Cornet et al., 2012). The geographical source of each sequence is represented by colored dot, as indicated by the key. Variant lineages are indicated as follows: European (EUR), European-Asian (EAS), African 1 (Afr1a, Afr1b), African 2 (Afr2a, Afr2b), Asian-American (AA1 and AA2) and North-American (NA). EUR branches indicated in gray had less than 50% bootstrap support (500 iterations). Variant sites that form the majority of the sequences within each potential sub-cluster within the EUR variant lineage are indicated, with minority site-specific variants indicated in gray text.

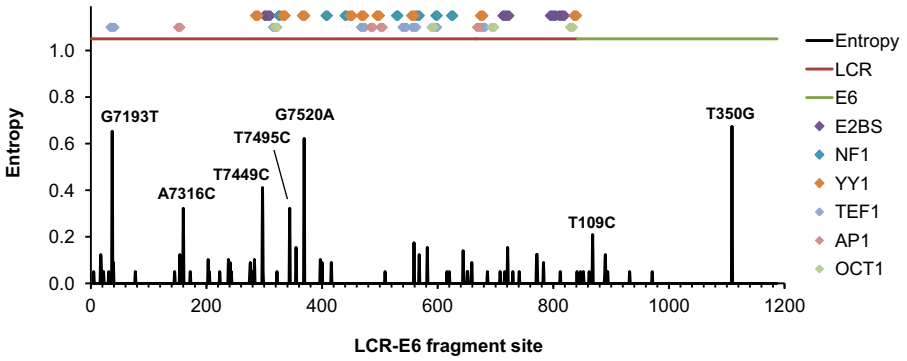

Fig. 2. Site-specific intra-type sequence diversity. Variation within the 1178 bp fragment is estimated using Shannon entropy, wherein a value of zero reflects site-specific conservation and higher values indicate increasing degrees of site-specific variation. A level of 5% variation equates to an entropy score of *ca.* 0.2. Site-specific variations above this level are indicated by convention, numbered according to the reference HPV16 sequence, K02718. The boundaries of the LCR (red) and E6 (green) are indicated as are major elements within the LCR, including the E2 binding site (E2BS), and transcriptional elements NF1, YY1, TEF1, AP1 and OCT1.

(0.35%; 0.21–0.49%; $n = 145$; Fig. 1B) (Cornet et al., 2012). Both of these were considerably less diverse than the panel of background HPV16 sequences regardless of variant lineage (1.34%; 0.99–1.69%; $n = 353$) (Cornet et al., 2012).

The proportion of individual variant sites and sequence variants displaying combinations of these major variant sites were examined in relation to cervical disease stage (Table 1). Of the five LCR variant positions, G7193T and G7520A were present in around 65% and 69% of sequences, respectively, while A7316C, T7449C and T7495C were present in around 10–15% of sequences, mostly in combination with other LCR or E6 variant sites. G7193T was found in 61% (95%CI 42–77%) of control sequences and 66% (55–77%) of cases giving a crude OR of 1.28 (0.50–3.20; $p = 0.665$). Similarly, G7520A was found in 64% (45–80%) of control sequences and 71% (60–81%) of cases giving a crude OR of 1.43 (0.54–3.66; $p = 0.501$). G7193T (found in 65% of sequences) is positioned within a TEF1 binding site and G7520A is within a YY1 binding site. Variation within YY1 binding sites, including G7520A, can affect individual site-specific binding of YY1 (Schmidt et al., 2001) but multiple YY1 binding site disruptions, akin to those found in the major AA and Afr variant lineages, are required to affect E6 promoter (p97) activity significantly (Kammer et al., 2000; Lace et al., 2009). In addition, variants bearing both G7193T and G7520A substitutions did not alter the ability of p97 to drive luciferase reporter expression in one study (Veress et al., 1999).

The T109C substitution in E6 (found in 5% of sequences) is a silent change that does not affect the Phe residue at codon 2. Overall, the T350G (L83V) substitution was present in 41% (95%CI 32–51%) of sequences in this study which was lower than the 54% (51–57%) reported across Europe in a recent meta-analysis ($p = 0.009$) (Tornesello et al., 2011). In this latter study, the T350G variant was over-represented in cases compared to controls (Tornesello et al., 2011), while in another analysis the T350G variant was under-represented in cases from Europe/Central Asia while being over-represented in cases from South/Central America (Cornet et al., 2013b). In the present study, the T350G substitution was found in 45% (28–64%) of control sequences and 39% (28–51%) of cases giving a crude OR of 0.77 (0.31–1.91; $p = 0.534$).

The data presented in this study represent the only available information on HPV16 sequence variation and cervical disease in the UK. While the use of single infection samples should have reduced potential confounding factors in relation to the assignment of disease status, the data overall suggest that no individual or combination of LCR–E6 site variants, including T350G, displayed any apparent association with cervical disease.

We identified samples singly infected with HPV16 using a generic PCR and genotyping test (Howell-Jones et al., 2010). This is a common approach and one of the more robust genotyping tests (Eklund et al., 2013), but such generic genotyping tests cannot exclude the potential for masking within mixed infections (van Alewijk et al., 2013). The selection of samples exhibiting an appar-

Table 1
Distribution of LCR–E6 variants in cases and controls.

Variant[a]	Controls ($n = 33$)			Cases ($n = 77$)			OR	95%CI	p value
	n	%	95%CI (%)	n	%	95%CI (%)			
G7193T	20	61	42–77	51	66	55–77	1.28	0.50–3.20	0.665
A7316C	4	12	3–28	7	9	4–18	0.73	0.17–3.65	0.731
T7449C	6	18	7–35	10	13	6–23	0.67	0.20–2.49	0.557
T7495C	2	6	1–20	9	12	5–21	2.05	0.39–20.50	0.500
G7520A	21	64	45–80	55	71	60–81	1.43	0.54–3.66	0.501
T109C	2	6	1–20	4	5	1–13	0.85	0.11–9.86	1.000
T350G	15	45	28–64	30	39	28–51	0.77	0.31–1.91	0.534
Reference	8	24	11–42	15	19	11–30	0.76	0.26–2.34	0.613
G7520A alone	1	3	0–16	4	5	1–13	1.75	0.16–89.00	1.000
A7316C alone	4	12	3–28	7	9	4–18	0.73	0.17–3.65	0.731
G7193T/G7520A	3	9	2–24	12	16	8–26	1.85	0.45–10.88	0.546
G7193T/G7520A/T350G	7	21	9–39	18	23	14–34	1.13	0.39–3.61	1.000
G7193T/G7520A/T109C/T350G	2	6	1–20	2	3	0–9	0.41	0.03–5.99	0.582
G7193T/T7495C/G7520A	2	6	1–20	9	12	5–21	2.05	0.39–20.50	0.500
G7193T/T7449C/G7520A/T350G	6	18	7–35	8	10	5–19	0.52	0.14–2.02	0.349
G7193T/T7449C/G7520A/T109C/T350G	0	0	0–11	2	3	0–9	N/A	N/A	1.000

[a] Distribution of HPV LCR–E6 reference sequence, individual site variants regardless of other site-specific variation (G7193T, A7316C, T7449C, T7495C, G7520A, T109C, T350G) and distribution of combinatorial variants for these seven major sites in cases and controls. OR, crude Odds Ratio; 95%CI, 95% confidence intervals; N/A, not applicable.

ent single HPV16 infection is a significant improvement on the use of unselected samples, wherein the potential impact of such masking is likely to be far higher. Nevertheless, there exists the possibility that a minority of samples in this study contained low levels of other, non-HPV16, genotypes that could be the genotype(s) causally associated with the reported cervical disease phenotype. Finally, although cytology samples are the primary sample type collected during cervical screening and are readily amenable to such an evaluation, the use of cytology samples to evaluate bias in HPV variant distribution is potentially problematic given sampling heterogeneity and the possibility of reporting inconsistency between the overlapping cytology and histology stage designations (Sherman, 2003).

4. Conclusions

The purpose of this study was to evaluate any potential relationship between HPV16 site-specific variants and cervical disease stage in the context of singly infected samples, wherein the association with disease can be assumed with some confidence. Overall, there appeared to be no individual LCR or E6 sites, or combinations thereof, including the T350G (L83V) E6 substitution, that demonstrated any significant association with cervical disease stage. Although these data suggest that LCR–E6 sequence variation *per se* may not be a significant driving force in the development of cervical disease, improving the resolution of sequence variation worldwide will nevertheless lead to a better understanding of the diversity of HPV and may further illuminate important aspects of HPV biology.

Funding

This work was funded by a PhD student project award from Public Health England.

Author contributions

L.M. conceived and designed the experiments, performed the experiments, analyzed the data and contributed to drafting the manuscript. A.G. conceived and designed the experiments and contributed to drafting the manuscript. S.B. conceived and designed the experiments, analyzed the data and contributed to drafting the manuscript. J.V.P. contributed to drafting the manuscript. All authors read and approved the final manuscript.

Acknowledgment

We thank E.M. de Villiers from the German Cancer Research Centre, Heidelberg, Germany for the use of the full genome plasmid for HPV16.

References

Andersson, S., Alemi, M., Rylander, E., Strand, A., Larsson, B., Sallstrom, J., Wilander, E., 2000. Uneven distribution of HPV 16 E6 prototype and variant (L83V) oncoprotein in cervical neoplastic lesions. Br. J. Cancer 83, 307–310.

Bernard, H.U., 2013. Regulatory elements in the viral genome. Virology 445, 197–204.

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.

Burk, R.D., Harari, A., Chen, Z., 2013. Human papillomavirus genome variants. Virology 445, 232–243.

Chan, S.Y., Ho, L., Ong, C.K., Chow, V., Drescher, B., Durst, M., ter Meulen, J., Villa, L., Luande, J., Mgaya, H.N., et al., 1992. Molecular variants of human papillomavirus type 16 from four continents suggest ancient pandemic spread of the virus and its coevolution with humankind. J. Virol. 66, 2057–2066.

Chan, P.K., Lam, C.W., Cheung, T.H., Li, W.W., Lo, K.W., Chan, M.Y., Cheung, J.L., Xu, L.Y., Cheng, A.F., 2002. Human papillomavirus type 16 intratypic variant infection and risk for cervical neoplasia in southern China. J. Infect. Dis. 186, 696–700.

Chang, Y.J., Chen, H.C., Pan, M.H., Lee, B.H., You, S.L., Lin, C.Y., Chou, Y.C., Hsieh, C.Y., Cheng, Y.J., Liaw, K.L., Hsing, A.W., Schiffman, M., Chen, C.J., 2013. Intratypic variants of human papillomavirus type 16 and risk of cervical neoplasia in Taiwan. J. Med. Virol. 85, 1567–1576.

Chen, Z., Terai, M., Fu, L., Herrero, R., DeSalle, R., Burk, R.D., 2005. Diversifying selection in human papillomavirus type 16 lineages based on complete genome analyses. J. Virol. 79, 7014–7023.

Chen, Z., DeSalle, R., Schiffman, M., Herrero, R., Burk, R.D., 2009. Evolutionary dynamics of variant genomes of human papillomavirus types 18, 45, and 97. J. Virol. 83, 1443–1455.

Cornet, I., Gheit, T., Franceschi, S., Vignat, J., Burk, R.D., Sylla, B.S., Tommasino, M., Clifford, G.M., 2012. Human papillomavirus type 16 genetic variants: phylogeny and classification based on E6 and LCR. J. Virol. 86, 6855–6861.

Cornet, I., Gheit, T., Clifford, G.M., Combes, J.D., Dalstein, V., Franceschi, S., Tommasino, M., Clavel, C., 2013a. Human papillomavirus type 16 E6 variants in France and risk of viral persistence. Infect. Agent Cancer 8, 4.

Cornet, I., Gheit, T., Iannacone, M.R., Vignat, J., Sylla, B.S., Del Mistro, A., Franceschi, S., Tommasino, M., Clifford, G.M., 2013b. HPV16 genetic variation and the development of cervical cancer worldwide. Br. J. Cancer 108, 240–244.

Eklund, C., Forslund, O., Wallin, K.L., Dillner, J., 2013. Global improvement in genotyping of human papillomavirus DNA: the 2011 HPV LabNet International Proficiency Study. J. Clin. Microbiol. (Epub ahead of print).

Gheit, T., Cornet, I., Clifford, G.M., Iftner, T., Munk, C., Tommasino, M., Kjaer, S.K., 2011. Risks for persistence and progression by human papillomavirus type 16 variant lineages among a population-based sample of Danish women. Cancer Epidemiol. Biomarkers Prev. 20, 1315–1321.

Guan, P., Howell-Jones, R., Li, N., Bruni, L., de Sanjose, S., Franceschi, S., Clifford, G.M., 2012. Human papillomavirus types in 115,789 HPV-positive women: a meta-analysis from cervical infection to cancer. Int. J. Cancer 131, 2349–2359.

Howell-Jones, R., Bailey, A., Beddows, S., Sargent, A., de Silva, N., Wilson, G., Anton, J., Nichols, T., Soldan, K., Kitchener, H.C., 2010. Multi-site study of HPV type-specific prevalence in women with cervical cancer, intraepithelial neoplasia and normal cytology, in England. Br. J. Cancer 103, 209–216.

Kammer, C., Warthorst, U., Torrez-Martinez, N., Wheeler, C.M., Pfister, H., 2000. Sequence analysis of the long control region of human papillomavirus type 16 variants and functional consequences for P97 promoter activity. J. Gen. Virol. 81, 1975–1981.

Kammer, C., Tommasino, M., Syrjanen, S., Delius, H., Hebling, U., Warthorst, U., Pfister, H., Zehbe, I., 2002. Variants of the long control region and the E6 oncogene in European human papillomavirus type 16 isolates: implications for cervical disease. Br. J. Cancer 86, 269–273.

Korber, B.T., Kunstman, K.J., Patterson, B.K., Furtado, M., McEvilly, M.M., Levy, R., Wolinsky, S.M., 1994. Genetic differences between blood- and brain-derived viral sequences from human immunodeficiency virus type 1-infected patients: evidence of conserved elements in the V3 region of the envelope protein of brain-derived sequences. J. Virol. 68, 7467–7481.

Korber, B., Muldoon, M., Theiler, J., Gao, F., Gupta, R., Lapedes, A., Hahn, B.H., Wolinsky, S., Bhattacharya, T., 2000. Timing the ancestor of the HIV-1 pandemic strains. Science 288, 1789–1796.

Lace, M.J., Isacson, C., Anson, J.R., Lorincz, A.T., Wilczynski, S.P., Haugen, T.H., Turek, L.P., 2009. Upstream regulatory region alterations found in human papillomavirus type 16 (HPV-16) isolates from cervical carcinomas increase transcription, ori function, and HPV immortalization capacity in culture. J. Virol. 83, 7457–7466.

Moody, C.A., Laimins, L.A., 2010. Human papillomavirus oncoproteins: pathways to transformation. Nat. Rev. Cancer 10, 550–560.

Niccoli, S., Abraham, S., Richard, C., Zehbe, I., 2012. The Asian-American E6 variant protein of human papillomavirus 16 alone is sufficient to promote immortalization, transformation, and migration of primary human foreskin keratinocytes. J. Virol. 86, 12384–12396.

Nindl, I., Rindfleisch, K., Lotz, B., Schneider, A., Durst, M., 1999. Uniform distribution of HPV 16 E6 and E7 variants in patients with normal histology, cervical intraepithelial neoplasia and cervical cancer. Int. J. Cancer 82, 203–207.

Richard, C., Lanner, C., Naryzhny, S.N., Sherman, L., Lee, H., Lambert, P.F., Zehbe, I., 2010. The immortalizing and transforming ability of two common human papillomavirus 16 E6 variants with different prevalences in cervical cancer. Oncogene 29, 3435–3445.

Schiffman, M., Castle, P.E., Jeronimo, J., Rodriguez, A.C., Wacholder, S., 2007. Human papillomavirus and cervical cancer. Lancet 370, 890–907.

Schiffman, M., Rodriguez, A.C., Chen, Z., Wacholder, S., Herrero, R., Hildesheim, A., Desalle, R., Befano, B., Yu, K., Safaeian, M., Sherman, M.E., Morales, J., Guillen, D., Alfaro, M., Hutchinson, M., Solomon, D., Castle, P.E., Burk, R.D., 2010. A population-based prospective study of carcinogenic human papillomavirus variant lineages, viral persistence, and cervical neoplasia. Cancer Res. 70, 3159–3169.

Schmidt, M., Kedzia, W., Gozdzicka-Jozefiak, A., 2001. Intratype HPV16 sequence variation within LCR of isolates from asymptomatic carriers and cervical cancers. J. Clin. Virol. 23, 65–77.

Sherman, M.E., 2003. Chapter 11: future directions in cervical pathology. J. Natl. Cancer Inst. Monogr., 72–79.

Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24, 1596–1599.

Tornesello, M.L., Losito, S., Benincasa, G., Fulciniti, F., Botti, G., Greggi, S., Buonaguro, L., Buonaguro, F.M., 2011. Human papillomavirus (HPV) genotypes and HPV16 variants and risk of adenocarcinoma and squamous cell carcinoma of the cervix. Gynecol. Oncol. 121, 32–42.

van Alewijk, D., Kleter, B., Vent, M., Delroisse, J.M., de Koning, M., van Doorn, L.J., Quint, W., Colau, B., 2013. A human papilloma virus testing algorithm comprising a combination of the L1 broad-spectrum SPF10 PCR assay and a novel E6 high-risk multiplex type-specific genotyping PCR assay. J. Clin. Microbiol. 51, 1171–1178.

Veress, G., Szarka, K., Dong, X.P., Gergely, L., Pfister, H., 1999. Functional significance of sequence variation in the E2 gene and the long control region of human papillomavirus type 16. J. Gen. Virol. 80 (Pt 4), 1035–1043.

Villa, L.L., Sichero, L., Rahal, P., Caballero, O., Ferenczy, A., Rohan, T., Franco, E.L., 2000. Molecular variants of human papillomavirus types 16 and 18 preferentially associated with cervical neoplasia. J. Gen. Virol. 81, 2959–2968.

Xi, L.F., Koutsky, L.A., Galloway, D.A., Kuypers, J., Hughes, J.P., Wheeler, C.M., Holmes, K.K., Kiviat, N.B., 1997. Genomic variation of human papillomavirus type 16 and risk for high grade cervical intraepithelial neoplasia. J. Natl. Cancer Inst. 89, 796–802.

Xi, L.F., Koutsky, L.A., Hildesheim, A., Galloway, D.A., Wheeler, C.M., Winer, R.L., Ho, J., Kiviat, N.B., 2007. Risk for high-grade cervical intraepithelial neoplasia associated with variants of human papillomavirus types 16 and 18. Cancer Epidemiol. Biomarkers Prev. 16, 4–10.

Yamada, T., Manos, M.M., Peto, J., Greer, C.E., Munoz, N., Bosch, F.X., Wheeler, C.M., 1997. Human papillomavirus type 16 sequence variation in cervical cancers: a worldwide perspective. J. Virol. 71, 2463–2472.

Zehbe, I., Lichtig, H., Westerback, A., Lambert, P.F., Tommasino, M., Sherman, L., 2011. Rare human papillomavirus 16 E6 variants reveal significant oncogenic potential. Mol. Cancer 10, 77.

Zuna, R.E., Moore, W.E., Shanesmith, R.P., Dunn, S.T., Wang, S.S., Schiffman, M., Blakey, G.L., Teel, T., 2009. Association of HPV16 E6 variants with diagnostic severity in cervical cytology samples of 354 women in a US population. Int. J. Cancer 125, 2609–2613.