Amino acid motifs in both the major and minor capsid proteins of HPV51 impact antigenicity and infectivity

Anna Godi,1 Ilaria Epifano,1,2 Sara L. Bissett,1 Tiziana Dell’Anna,3 Andrea Piana,4 Clementina Cocuzza2 and Simon Beddows1

1Virus Reference Department, Public Health England, London, UK
2Dipartimento di Chirurgia e Medicina Traslazionale, Università di Milano-Bicocca, Milan, Italy
3Division of Gynaecology, San Gerardo Hospital, Monza, Italy
4Dipartimento di Scienze Biomediche, Università degli Studi di Sassari, Sassari, Italy

Persistent infection with oncogenic human papillomavirus (HPV) is a prerequisite for cervical disease development, yet data regarding the host immune response to infection at the genotype level are quite limited. We created pseudoviruses bearing the major (L1) and minor (L2) capsid proteins and L1 virus-like particles representing the reference sequence and a consensus of 34 European sequences of HPV51. Despite the formation of similarly sized particles, motifs in the reference L1 and L2 genes had a profound impact on the immunogenicity, antigenicity and infectivity of these antigens. The antibody status of women exhibiting low-grade disease was similar between HPV16 and the consensus HPV51, but both demonstrated discrepancies between binding and neutralizing antibody responses. These data support the use of pseudoviruses as the preferred target antigen in studies of natural HPV infection and the need to consider variation in both the L1 and L2 proteins for the appropriate presentation of antibody epitopes.

The non-enveloped capsid of HPV comprises the major (L1) and minor (L2) capsid proteins. The L1 protein facilitates virus attachment to host cells, whilst the L2 protein is essential for subsequent virus infectivity (Buck et al., 2013; Raff et al., 2013; Wang & Roden, 2013).

Current L1 virus-like particle (VLP)-based HPV vaccines target the most prevalent genotypes, HPV16 and HPV18, and elicit type-specific neutralizing antibodies thought to confer the high levels of efficacy observed in clinical trials (Lehtinen & Dillner, 2013; Schiller et al., 2014). A limited number of serological assays are available for measuring vaccine type-specific antibody responses, including an L1 VLP ELISA, a mAb competitive VLP assay and an L1L2 pseudovirus neutralization assay. Despite some discrepancies, overall inter-assay agreements are good (Dessy et al., 2008; Krajden et al., 2014).

Natural infection (NI) studies usually compare the HPV DNA status of individuals with their respective antibody status in order to better understand the pathogen–host relationship. Most of these studies have examined HPV16 and/or HPV18 (Castellsague et al., 2014; Lin et al., 2013; Pastrana et al., 2004; Safaeian et al., 2010; Xi et al., 2002), whilst some have expanded their investigations to evaluate other genotypes including HPV6, HPV11, HPV31, HPV33, HPV35, HPV45, HPV52 and HPV58 (Carter et al., 2000; Ochi et al., 2008; Syrjänen et al., 2009; Wilson et al., 2013). Most of these studies have made use of an immobilized L1-based target (Carter et al., 2000; Castellsague et al., 2014; Safaeian et al., 2010; Syrjänen et al., 2009; Wilson et al., 2013; Xi et al., 2002), whilst a few have used an L1L2 pseudovirus neutralization assay (Lin et al., 2013; Ochi et al., 2008; Pastrana et al., 2004). Inter-assay agreements for measuring NI antibody responses appear to be quite poor (Safaeian et al., 2012), in contrast to studies of vaccine antibody responses (Dessy et al., 2008; Krajden et al., 2014). There also appears to be a poor correlation between the presence of HPV DNA and type-specific
antibody in contemporary samples (Carter et al., 2000; Ochi et al., 2008; Pastrana et al., 2004; Syrjänen et al., 2009; Wilson et al., 2013). This is perhaps not surprising given the comparatively low rates of seroconversion following incident HPV infection (Carter et al., 2000; Syrjänen et al., 2009; Xi et al., 2002) and observations that NI antibodies are generally of low titre and may wane over time (Syrjänen et al., 2009; Wilson et al., 2013). Despite these potential shortcomings, individuals with high levels of type-specific NI antibodies may have a reduced risk of subsequent reinfection by that genotype (Castellsagué et al., 2014; Safaeian et al., 2010; Wilson et al., 2014).

HPV51 is an oncogenic genotype (Bernard et al., 2010; Bouvard et al., 2009) belonging to the alpha-5 species group. It is associated with 6–12 % of cervical disease cases (Guan et al., 2012) and approximately 1 % of invasive cervical cancer cases worldwide (de Sanjose et al., 2010), with a potential for geographical bias (Piana et al., 2013). Little is known about the biological properties of HPV51 except that it has transformation capability (Lungu et al., 1991), that motifs in the L2 protein can facilitate nuclear localization of an encapsidated reporter (Kondo et al., 2009) and that two broadly cross-reactive HPV16 L2 mAbs can recognize their cognate HPV51 L2 epitopes (Nakao et al., 2012).

Our objectives for this study were: (i) to generate HPV51 L1 and L2 sequences to inform the creation of representative L1 VLP and L1L2 pseudovirus antigens; (ii) to compare the antigenicity, immunogenicity and infectivity profiles with antigens based on the reference sequence; and (iii) to use appropriate antigens to compare the antibody responses generated during HPV16 and HPV51 infection. Such data should inform the structure–function relationship between the major and minor capsid proteins of an important oncogenic HPV genotype.

Twenty-five L1 and L2 sequences were generated from residual genital samples from women in England (Howell-Jones et al., 2012). Paired cytology and serum samples were available from women (n=151; Gynaecology Outpatients Clinic, San Gerardo Hospital, Monza, Italy) following a cytological diagnosis of atypical squamous cells of undetermined significance (ASCUS) or low-grade squamous intraepithelial lesion (LSIL) (Schiffman et al., 2007), generating another nine paired sequences. The study protocol (08/UNIMIB-HPA/HPV1) was approved by San Gerardo Hospital’s Ethical Committee and written informed consent was obtained from all participants. Amplicons were generated using Platinum Taq High Fidelity DNA polymerase (Life Technologies) (see Table S1, available in the online Supplementary Material, for primer sequences). The sequences determined were aligned with available sequences from the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/) database including the HPV51 reference (GenBank accession no. M62877) and an isolate from Japan (GQ487711 and GQ487712) (Kondo et al., 2009). Sequences representing HPV51 genome lineages A1 (reference M62877), A2 (KF436870), A3 (KF436873), A4 (KF436875), B1 (KF436883) and B2 (KF436886) were kindly provided by R. D. Burk, Albert Einstein College of Medicine, NY, USA (Burk et al., 2013). Alignments and neighbour-joining tree reconstruction (500 iterations) were carried out using MEGA6 (Tamura et al., 2013).

The majority of the sequences (29/34; 85 %) generated in this study clustered within the A1 sublineage, which contains the HPV51 reference (Fig. 1a). The remaining sequences (5/34; 15 %) clustered adjacent to, but in some cases distinct from, sequences representing sublineages A2, A3, A4, B1 and B2. It is unclear at this time whether sequences that do not segregate into these sublineages are due to a lack of sufficient diagnostic motifs in the L1 and L2 genes, compared with the full genome sequences from which they were derived (Burk et al., 2013), or whether additional sublineages exist.

Site-specific amino acid variation was evaluated using Shannon entropy (http://www.hiv.lanl.gov/content/sequence/ENTROPY/entropy_one.html) wherein a value of zero reflects site-specific conservation and higher values indicate increasing site-specific variation (Fig. 1b). Residue numbering was based on the reference sequence. Site-specific entropy was of low magnitude, with only seven sites (two in L1 and five in L2) exhibiting scores of >0.2. This is considerably lower than is found in RNA viruses such as human immunodeficiency virus type 1 (Rhee et al., 2008), reflecting the excellent proof-reading capability of host-cell polymerases (Duffy et al., 2008).

There were nine sites that displayed any degree of variation within the L1 protein (Fig. 1c). Two of these sites (adjacent FG loop residues, V264G and G265S) highlighted a motif (264VG265) present only in the reference sequence. Two sites [P181Q (EF loop) and D273N (FG loop)] were present in two sequences each (2/34; 6 %), whilst the remaining five variant sites were each present in one sequence only (1/34; 3 %). The L2 sequences were more variable, containing 24 variant residue sites. Two motifs (95DLWHH99 and 178DIYLLVHY186) were present only in the reference sequence, corroborating an assumption based on a single Japanese isolate (Kondo et al., 2009). Of the other variant residues, only K43T/N (2/34, 6 %), E109D (2/34, 6 %), D270N (4/34, 12 %), T283K (30/34, 88 %) and T400P (2/34, 6 %) were found in more than 5 % of sequences. These data corroborate (Kondo et al., 2009) and extend observations about how representative the reference sequence is within the diversity of HPV51 genomes and suggest that the reference may include sequence artefacts.

Consensus L1 and L2 sequences (Fig. 1c) were used to create L1 VLP (VLP_{CON}) and L1L2 pseudovirus (HPV51_{CON/CON}) constructs to compare with those based on the A1 reference sequence (VLP_{REF} and HPV51_{REF/REF}). L1 VLPs were expressed from Sf21 insect cells using the Bac-to-Bac Baculovirus System (Life Technologies), whilst L1L2
pseudoviruses were expressed from 293TT mammalian cells using the bicistronic pShELL vector (Buck & Thompson, 2007), as described previously (Draper et al., 2011, 2013; Huo et al., 2012). All HPV51 constructs made use of codon-optimized inserts (GeneArt; Life Technologies). Purified pseudovirus stocks were titrated on 293TT cells and the equivalent of the TCID_{50} was estimated using the Spearman–Karber equation (Draper et al., 2011). Particle: infectivity (PI) ratios were estimated using the L1 content determined by SDS-PAGE (Draper et al., 2013), assuming an estimated 3 \times 10^{10} particles (ng L1 protein)^{-1} (http://home.ccr.cancer.gov/lco/production.asp) and the ratio normalized for input volume and TCID_{50}. L1 VLP (VLP51_{CON} and VLP51_{REF}) and L1L2 pseudoviruses (HPV51_{CON/CON} and HPV51_{REF/REF}) displayed similarly sized (approx. 50 nm) particles to each other and those antigens based on HPV16 (Fig. 2). However, the HPV51_{REF/REF} pseudovirus exhibited a very low relative infectivity resulting in a substantially higher PI ratio (8.4 \times 10^5) compared with the ratios of approximately 10^2 observed for HPV16 and HPV51_{CON/CON} (Fig. 2).

Our simplistic interpretation for comparison purposes is that a PI ratio of 10^2 equates to approximately 1 % of particles able to transduce the target cell with the reporter plasmid. A chimeric HPV51 pseudovirus containing the reference L2 (295DLWHH109 and 178DIYLLVHY186) and an L1 sequence from a Japanese isolate [including the 264G{GS}{265} motif and two other substitutions from the reference: I52L (BC loop) and N272T (FG loop)] has been shown to be essentially non-infectious due to restricted nuclear localization of the encapsidated reporter bestowed by the reference L2 motifs (Kondo et al., 2009). We generated a chimaeric pseudovirus comprising the reference L1 but consensus L2 (HPV51_{REF/CON}), and, although it formed particles of a similar size to the other constructs (Fig. 2), it was also found to have a substantially higher PI ratio than the HPV51_{CON/CON} pseudovirus.

BALB/c mice were immunized with HPV51 reference (VLP_{REF}) and consensus (VLP_{CON}) L1 VLPs to determine whether the 264G{GS}{265} motif affected the immunogenicity and/or antigenicity of the L1 protein. Mice (n=5 per group) were immunized intramuscularly on days 0 and 14 with 2 μg VLPs [adsorbed onto aluminium hydroxide (Alhydrogel; Brenntag Biosector) then admixed with monophosphoryl lipid A (Sigma Adjuvant System)] before a terminal bleed was taken at day 21. Procedures were in accordance with UK Home Office guidelines and performed under licences PPL 70/7412 and 70/7414. ELISA and neutralization assays were performed as described elsewhere (Draper et al., 2011, 2013).

VLP_{REF} elicited a median fivefold [interquartile range (IQR) 4–5] higher antibody-binding titre against the VLP_{REF} antigen (median, 25 461; IQR, 23 050–55 337; n=5) than against the VLP_{CON} (5651; 4779–9413). Conversely, VLP_{CON} antisera bound the VLP_{REF} (42 327; 38 964–45 926) and VLP_{CON} (41 415; 41 012–51 100) with equivalent titres and higher than those elicited by VLP_{REF}. Mice immunized with denatured VLP_{CON} or adjuvant alone generated titres of <100. Although the HPV51_{REF/REF} pseudovirus was essentially non-infectious, the chimaeric HPV51_{REF/CON} construct retained sufficient infectivity to be used in neutralization assays. VLP_{REF} antisera neutralized the HPV51_{REF/CON} (4927; 3971–5541) pseudovirus to a threefold (IQR 2–4) higher median titre than the HPV51_{CON/CON} (872; 150–1143) pseudovirus. The VLP_{CON} antisera, however, neutralized the HPV51_{REF/CON} (14 164; 8507–16 271) and HPV51_{CON/ CON} (13 376; 9240–20 493) to a similar extent and to higher titres than antisera generated using VLP_{REF}. One possible explanation of these data is that the VLP_{REF} generates motif-specific (264G{GS}{265}) antibodies that are less well recognized by consensus-based (264G{GS}{265}) antigens. However, it is also possible that more distal immunogenic domains are differentially presented in, and recognized by, the reference-based reagents due to more global changes incurred by the 264G{GS}{265} motif. That the converse is not also true suggests that any potential HPV51 L1 VLP immunogen should be based on a consensus or similarly representative sequence.

Serum samples collected from women following a diagnosis of ASCUS or LSIL were tested against antigens based on the reference or consensus sequence to ascertain whether the choice of antigen would have an impact on the interpretation of serological data generated in NI studies. Of those NI sera (n=23) that were positive against either VLP, the antibody responses against VLP_{REF} (median titre 98; IQR 70–172) and VLP_{CON} (83; 25–171) were similar (P=0.484, Wilcoxon paired signed rank test; Stata 13.1). However, of 48 sera capable of neutralizing the HPV51_{CON/CON}-pseudovirus (80; 30–121), only 21 sera (44 %) were able to
neutralize the chimaeric HPV51REF/CON pseudovirus and to a much lower titre (10; 10–24).

Together, these data suggest that both the L1 and L2 sequences of the reference are not representative of circulating HPV51 sequences and that motifs in both the L1 (264VG265) and L2 (95DLWHH99 and 178a DIYLLVHY186) proteins profoundly impact the infectivity, immunogenicity and antigenicity of the resulting antigens. Two significant studies have examined the impact of capsid protein variation on L1 (Pastrana et al., 2001) and L2 (Seitz et al., 2012) antibody epitopes for a few genotypes (HPV16, HPV18 and HPV31), but these studies did not consider variation in the reciprocal capsid protein. The present data support the need to carefully consider both L1 and L2 sequences in context when generating HPV pseudoviruses.

For these reasons, we used the VLP\textsubscript{CON} and HPV51\textsubscript{CON/CON} antigens to represent the HPV51 genotype in a small comparison study of HPV16 and HPV51 NI (Table S2). Sixteen women were positive for HPV51 DNA [16/151; 10.6 %, 95 % confidence interval (CI) 6.2–16.6 %] which was similar to the number with currently detectable HPV16 DNA (19/151; 12.6 %, 7.7–19.0 %) (P=0.720; Fisher’s exact test). Only two women were positive for both HPV16 and HPV51 DNA (2/151; 1.3 %, 0.2–4.7 %). The median HPV51 viral load (1.5 copies per cell; IQR 0.1–13.6; n=16) was similar to that for HPV16 (4.9; 0.7–21.5; n=19) (P=0.529; Mann–Whitney U test). Sixteen women were positive for HPV51 VLP-binding antibodies (16/151; 11 %, 95 % CI 6–17 %), which was somewhat lower than the number positive for HPV16 VLP antibodies (26/151; 17 %, 12–24 %) (P=0.135). The number of women positive for antibodies against HPV51 (48/151; 32 %, 24–40 %) or HPV16 (43/151; 28 %, 21–36 %) pseudoviruses was similar (P=0.616) and higher than those positive for binding antibodies against their respective VLP for both HPV16 (P=0.028) and HPV51 (P<0.001). A high proportion (20/151; 13 %, 8–20 %) of individuals was positive for antibodies against both HPV16 and HPV51 pseudoviruses, probably due to historical independent infections given the low rate of concurrent HPV16 and HPV51 DNA detection.

Recent infection was defined as being DNA positive and antibody negative, and prior infection was defined as DNA negative and antibody positive, whilst concurrency was defined as being both DNA and antibody positive (Table 1). Such NI categories allowed the simple segregation of individuals based on their current host infection status. For example, the highest proportion of individuals was assigned to the prior infection category with 10 % of individuals displaying recent infection or concurrent detection (Syrjän¨ en et al., 2009). This was more marked when their status was assigned using L1L2 pseudovirus neutralization data, with 27 % of individuals classified as having prior HPV51 infection compared with 6 % (P<0.001) who exhibited a recent infection and 5 % (P<0.001) demonstrating concurrent detection of HPV51 DNA and antibody. Although the proportions of individuals were similar between the HPV16 and HPV51 NI status categories, neutralizing antibody titres against HPV16 tended be higher than those against HPV51. This may indicate true differences in the host response to

\begin{tabular}{|l|c|c|c|c|}
\hline
Pseudovirus & HPV16 & HPV51\textsubscript{CON/CON} & HPV51\textsubscript{REF/REF} & HPV51\textsubscript{REF/CON} \\
\hline
\textbf{Dimension (nm)} & 52 (50–55) & 52 (50–54) & 53 (51–56) & 52 (49–56) \\
\textbf{Infectivity (TCID\textsubscript{50})} & 3.8×10\textsuperscript{7} & 8.3×10\textsuperscript{6} & 4.1×10\textsuperscript{2} & 6.8×10\textsuperscript{3} \\
\textbf{L1 concentration (\textmu g ml\textsuperscript{-1})} & 3.34 & 1.37 & 0.24 & 0.15 \\
\textbf{PI ratio} & 1.6×10\textsuperscript{2} & 1.5×10\textsuperscript{2} & 8.4×10\textsuperscript{5} & 6.7×10\textsuperscript{4} \\
\hline
\end{tabular}

\begin{tabular}{|l|c|c|}
\hline
\textbf{VLP} & HPV16 & HPV51\textsubscript{CON} & HPV51\textsubscript{REF} \\
\hline
\textbf{Dimension (nm)} & 51 (46–53) & 49 (47–51) & 48 (45–54) \\
\textbf{L1 concentration (\textmu g ml\textsuperscript{-1})} & 30 & 29 & 96 \\
\hline
\end{tabular}

\textbf{Fig. 2.} HPV16 and HPV51 L1 VLP and L1L2 pseudovirus antigens. (a) Dimensions [median (interquartile range)], relative infectivity, L1 protein content and structural integrity of indicated 293TT mammalian cells expressing pseudovirus preparations. (b) Dimensions [median (interquartile range)] and L1 content of indicated baculovirus insect cell-derived VLP preparations. Data represent the means of two independent experiments.
infection between these genotypes and differences in persistence (Rositch et al., 2013) and/or highlight subtle differences in the ability of HPV pseudoviruses to mimic their archetypal strain.

These data improve our understanding of a little-studied oncogenic HPV genotype, HPV51, and demonstrate the importance of selecting representative HPV sequences, including consideration of variation in both the major and minor capsid proteins, for the construction of appropriate antigens. These data also corroborate recent observations of poor concordance between the VLP ELISA and the pseudovirus neutralization assay for measuring NI antibodies (Safaiean et al., 2012) in contrast to the generally good concordance for measuring vaccine-induced type-specific antibody (Dessy et al., 2008; Krajden et al., 2014). On a technical level, these data suggest that pseudoviruses should be considered the preferred antigen for such studies. Although the contribution of cell expression system-specific protein modifications or other technical confounding factors cannot be ruled out without more direct investigation, discrepancies between binding and neutralization specificities are perhaps not unexpected given the apparent structural alterations of the L1 protein when L2 is incorporated (Chen et al., 2011; Culp et al., 2007), and the possibility of antibodies elicited against L1 epitopes that are surface exposed only during the entry process and perhaps also including antibodies against L2 epitopes. These data appear to highlight differences in the host immune response to NI compared with vaccination by suggesting differential composition of the antibody repertoire. The specificity and importance of such antibodies would require further study.

Acknowledgements

We are indebted to Professor John T. Schiller and Dr Chris Buck (National Cancer Institute, Bethesda, MD, USA) for access to the pSheLL backbone used to make the HPV51 pseudovirus and for the HPV16 pseudovirus clone, p16shell. We thank Professor Robert D. Burk (Albert Einstein College of Medicine, NY, USA) for the lineage-specific HPV51 reference sequences. This work was funded in part by Fondazione Cariplo, Milan (Italy). I. E. was in receipt of an exchange student scholarship award from the Universita di Milano-Bicocca, Milan, Italy. The material produced by Public Health England is Crown copyright and is reproduced with the permission of PHE under delegated authority from the Controller of HMSO.

References


