

# Determination of the complete genome and functional analysis of HPV6 isolate VBD19/10 from a patient with aggressive recurrent respiratory papillomatosis

R. Y. SEEDAT<sup>1,2</sup>, C. E. COMBRINCK<sup>3</sup>, P. A. BESTER<sup>3</sup>, J.-Y. LEE<sup>3</sup> AND F. J. BURT<sup>3,4\*</sup>

<sup>1</sup> *Department of Otorhinolaryngology, School of Medicine, Faculty of Health Sciences, University of the Free State, Bloemfontein, South Africa*

<sup>2</sup> *Department of Otorhinolaryngology, Universitas Academic Hospital, Bloemfontein, South Africa*

<sup>3</sup> *Department of Medical Microbiology and Virology, School of Medicine, Faculty of Health Sciences, University of the Free State, Bloemfontein, South Africa*

<sup>4</sup> *Department of Medical Microbiology and Virology, National Health Laboratory Service, Bloemfontein, South Africa*

*Received 20 November 2015; Final revision 19 January 2016; Accepted 9 February 2016; first published online 2 March 2016*

## SUMMARY

Human papillomavirus (HPV) types 6 and 11 are the aetiological agent of recurrent respiratory papillomatosis (RRP). The complete genome of an HPV6 isolate with a 170 base pair (bp) duplication identified within the long control region (LCR) from a patient with aggressive recurrent respiratory papillomatosis was determined. The promoter sequence from the HPV LCR including the 170 bp duplication was placed upstream of a heterologous reporter gene and the activity of the reporter gene product determined using transfected cells. In total, mutations were observed at 157 nucleotide positions of the complete genome and included nucleotide substitutions, deletions and insertions, resulting in amino acid changes at 43 residue positions. Reporter gene activity using an HPV-derived LCR region with a 170 bp duplication was significantly higher than that using an HPV-derived LCR region with no duplication within this region. The results suggest that novel HPV variants warrant further investigation for potential biomarkers of aggressive disease.

**Key words:** Human papillomavirus, recurrent respiratory papillomatosis (RRP).

## INTRODUCTION

Human papillomaviruses (HPVs) are members of the Papillomaviridae family, with over 150 HPV types having been identified [1, 2]. They are small, non-enveloped DNA viruses with a double-stranded circular DNA molecule of ~8000 base pairs (bp) [1, 2]. HPVs, usually HPV types 6 or 11, are the aetiological agent of recurrent respiratory papillomatosis

(RRP) [3]. HPV6 is classified into two variant lineages (A and B), and five B sublineages (B1, B2, B3, B4, B5) [4]. HPV variant lineages and sublineages are defined based on the nucleotide differences across the complete HPV genome sequences of 1.0–10.0% and 0.5–1.0%, respectively [5].

The HPV genome is organized into early (E) and late (L) open reading frames (ORFs) and includes a non-coding region [6]. The early region ORFs code for proteins that play a role in viral regulation as well as for proteins that are essential in initiating viral DNA replication. The late ORFs code for the two structural capsid proteins [6, 7]. The non-coding region, also known as the long control region (LCR)

\* Author for correspondence: Professor F. J. Burt, Department of Medical Microbiology and Virology, Faculty of Health Sciences, University of the Free State, Bloemfontein, 9301, South Africa. (Email: burtfj@ufs.ac.za)

or the upstream regulatory region (URR), contains the crucial elements necessary for control of viral replication and transcription as well as the origin of DNA replication [6, 8]. The LCR extends from the termination of the L1 gene to the first methionine of the E6 gene. It is less conserved than one would expect these regions to be and contains a variety of transcriptional regulatory motifs, the early promoter, viral coded E2 regulator binding sites and the viral origin of DNA replication [8].

Nucleotide alterations in the HPV6 LCR have previously been described [4, 9–11]. The clinical significance of these variations is still not clear. Conflicting results have been described in studies investigating the effect of alteration on viral enhancer activity-based assays using reporter gene constructs [11–15]. While almost equal amounts of luciferase activities were demonstrated with HPV6 variants where point mutations were observed in the LCR, reporter gene activity using HPV-derived LCR regions from two variants that included duplications of 459 bp and 236 bp within the 3' region of the LCR was significantly raised [14]. Similarly, functional analysis of nucleotide substitutions in seven different HPV16 LCR isolates revealed enhanced transcriptional activity of the P97 promoter for two variants and no change in transcriptional activity for five of the seven variants [12]. It is unclear whether mutations within the LCR influence the transcriptional activity, and it is possible that the site where the mutation occurs, or the type of mutation, might play a role in the alteration of replication or transcription.

In a previous study, we identified a 170 bp duplication within the LCR of an HPV6 isolate obtained from a patient with aggressive RRP [10]. Based on uncertainty regarding the influence of alterations, further investigation of the duplication on function was warranted. Hence the complete genome of the isolate was determined and the promoter sequence from the HPV LCR including the 170 bp duplication was placed upstream of a heterologous reporter gene and the activity of the reporter gene product determined using transfected cells.

## MATERIALS AND METHODS

The complete nucleotide sequence data was determined for the HPV6 isolate laboratory number VBD19/10 with a 170 bp duplication within the LCR (Genbank accession no. JN573172.1). Sequence data were determined for HPV6 isolate

VBD02/10 (Genbank accession no. JN573167.1) that did not have alterations within the LCR relative to prototypic strain and was selected as a control for reporter gene studies.

The biopsies were collected from patients treated for RRP at the Universitas Academic hospital in Bloemfontein, South Africa in 2010. The clinical features of the patients are summarized in Table 1. Neither of the patients had spread distal to the subglottis or had required a tracheostomy. Written informed consent for inclusion in the study was obtained from the parents of these patients. The study was approved by the Ethics Committee of the Faculty of Health Sciences, University of the Free State (ECUFS 6/2011).

Two pairs of primers were designed that amplified two overlapping fragments of the full-length HPV6 genome. The complete genome sequence data of one HPV6a (L41216), two HPV6b (NC\_001355 and X00203) and one HPV6vc (AF092932) isolates were retrieved from GenBank. The sequence data was aligned in Clustal X version 2.1 [16]. Using the alignment file, a primer pair designated HPV6\_F1 (5'-CAG TTA TAG GGG AAG CAC CAG-3') at position 1826–1846 relative to HPV6b, and HPV6\_R1 (5'-CTG GTA ATA AGT TCT AAG GGC GG-3') at position 6354–6332 relative to HPV6b were designed to amplify a region of ~4529 bp. A second primer pair designated HPV6\_F2 (5'-GTA TCC AAA GTT GTT GCC ACG G-3') at position 5837–5858 relative to HPV6b, and HPV6\_R2 (5'-CTC ATA CAA AAG TAC GAT TTC CCA G-3') at position 2300–2276 relative to HPV6b were designed to amplify a region that overlapped of ~4460–4631 bp.

The primer pairs amplified two overlapping regions that targeted the complete HPV6 genome of each isolate. The amplification of each region was performed using the Phusion HotStart DNA polymerase-mediated PCR amplification kit (Finnzymes, Finland) according to the manufacturer's instructions using the following cycling conditions, denaturation at 98 °C for 30 s and 30 cycles of 98 °C for 10 s, 47 °C for 30 s, 72 °C for 135 s and a final extension at 72 °C for 10 min. PCR amplicons were purified for nucleotide sequence determination.

Purified amplicons were submitted to the National Institute for Communicable Diseases (NICD), Sandringham, Johannesburg for determination of nucleotide sequences of amplicons using the Roche GS Junior System (Roche 454 Life Sciences, USA). The sequence data obtained was mapped against the reference HPV6b using GS Reference Mapper Roche

Table 1. *Clinical features of patients with recurrent respiratory papillomatosis*

Laboratory no.	Age at diagnosis (years)	Average Coltera–Derkay score	Total procedures	Average procedures per year while having active disease	Procedures in first year after diagnosis	Average Coltera–Derkay score in first year after diagnosis	Status at last follow-up
2/10	6.3	19.4	15	2.3	4	12.5	Remission
19/10	6.9	26.5	20	5.9	7	26.1	Remission

software and assembled using the GS *De Novo* Assembler Roche software. The contiguous file obtained from the GS Reference Mapper and GS *De Novo* Assembler Roche software were assembled and aligned against the sequence of the corrected HPV6b reference genome (PAVE ID: HPV6REF), HPV6a (GenBank accession no. L41216) and HPV6c (GenBank accession no. AF092932) using Geneious Pro v. 4.8.5 (www.geneious.com) [17]. The aligned sequence data was edited in Geneious Pro v. 4.8.5 according to the phred confidence values observed in Gap4 v. 4.11 [18].

The complete genome sequence data of the isolate was aligned with complete sequence data of 193 HPV6 isolates from GenBank using MEGA v. 6.06 [19] using the ClustalW algorithm and a maximum likelihood phylogenetic tree was constructed.

### Functional analysis of LCR region

The purified DNA amplicons were cloned into pBlue-TOPO reporter vector (Invitrogen, USA) according to the manufacturer's instructions. Positive transformants were identified and confirmed by sequence analysis. Purified recombinant plasmid was prepared from overnight cultures and transfection grade DNA plasmid was purified using the Qiagen Plasmid Mini kit (Qiagen, USA) according to the manufacturer's instructions.

Baby hamster kidney (BHK) 21 (ATCC CCL-10) cells were grown to 90–95% confluency in growth media, Dulbecco's Modified Eagle's Medium (DMEM; Lonza, Belgium) with 5% fetal calf serum (FCS; Delta Bioproducts, South Africa), 1% L-glutamine (L-Glu; Sigma Aldrich, UK), 1% non-essential amino acids (NEAA; Lonza), and 1% penicillin/streptomycin (Pen/Strep) antibiotics (Sigma Aldrich). A ratio of 1.6–1.7  $\mu\text{g}$  DNA:5  $\mu\text{l}$  lipofectamine was determined as optimal based on transfection efficiencies and plasmid. Untransfected BHK cells were used as negative controls. Levels of active  $\beta$ -galactosidase expressed from BHK cells transfected with plasmids expressing

the *lacZ* gene were determined using the  $\beta$ -galactosidase assay kit (Invitrogen) according to the manufacturer's instructions. Absorbance values were measured at 420 nm using a Spectronic Genesys 5 spectrophotometer (Thermo Electron Corp., USA). The protein concentration of the cell lysate was determined using the Qubit protein assay (Invitrogen) according to the manufacturer's instructions. The specific activity of the cell lysate, determined in a total volume of  $8 \times 10^5$  nl, was calculated as follows:

specific activity

$$= \text{nmol ONPG hydrolysed}/t/\text{mg protein},$$

where nmol ONPG hydrolysed

$$= \frac{(\text{OD}_{420}) (8 \times 10^5 \text{ nl})}{(4500 \text{ nl/nmol} - \text{cm}) (1 \text{ cm})},$$

where ONPG = ortho-nitrophenyl- $\beta$ -galactoside; 4500 = the extinction coefficient,  $t$  = time of incubation (min) at 37 °C, and mg protein = the amount of protein assayed. Two independent transfection experiments were performed with duplicates within each experiment with three readings within each experiment. The ratio of pBlueTopo\_19/10 to pBlueTopo\_02/10 was determined for each experiment in order to normalize data from different experiments.

Data analysis was performed using IBM SPSS Statistics v. 22 (IBM Corp., USA). The specific activities of the cell lysate measured were analysed using the independent samples  $t$  test and 95% confidence intervals (CIs) were calculated for the ratio of pBlueTopo\_19/10 to pBlueTopo\_02/10.

### RESULTS

The complete genome sequence data for isolate VBD19/10 was aligned for identification of variants. Nucleotide positions were numbered relative to the corrected HPV6b reference genome (PAVE ID: HPV6REF). Genome data for HPV6b from lineage A, HPV6c (AF092932) from sublineage B1, and

Table 2. Schematic representation of the mutations observed within the complete genome of isolates VBD19/10

	E6										E7					E1																							
	221	251	323	365	368	392	473	479	594	684	791	823	1011	1012	1426	1494	1534	1535	1554	1669	1740	1743	1866	1878	1926	2091	2169	2193	2199	2207	2314	2362	2395	2529	2547	2556	2653	2709	
HPV6b	A	C	A	A	T	C	G	C	G	T	A	C	A	C	T	A	T	G	T	A	C	A	C	G	G	C	C	T	C	C	T	T	G	T	T	A	T		
HPV6a	T	G	C	T																																			
VBD19/10	T	G	C	T																																			
HPV6vc	T	G	T	T																																			

	E2																				E4																	
	2801	2813	2830	2973	3147	3152	3271	3387	3404	3433	3460	3518	3551	3604	3642	3692	3696	3734	3764	3781	3787	3793	3805	3815	3271	3387	3404	3433	3460	3518	3551							
HPV6b	C	T	G	G	A	A	T	T	A	G	C	C	C	C	G	T	A	G	G	C	C	C	A	C	T	T	A	G	C	C	C							
HPV6a	A																																					
VBD19/10	A																																					
HPV6vc	A	C	A	A	C	T	A	C	C	A	A	G	A	T	A	C																						

	E5a										E5b										NNCR3*	L2																											
	3898	3934	4003	4048	4123	4126	4136	4141	4147	4149	4194	4234	4251	4275	4296	4313	4322	4342	4345	4346	4352	4384	4386	4412	4455	4491	4530	4566	4578	4605	4645	4646	4731	4755	4863	4869	4975	5004	5020	5025	5077	5280	5406	5451	5489	5596	5619		
HPV6b	G	C	G	T	A	T	G	T	T	T	C	A	G	G	G	A	A	T	A	C	A	C	G	C	T	C	T	G	G	A	T	G	T	C	G	T	G	A	C	A	C	T	A	G	C	A	G		
HPV6a	A	A	C	C																																													
VBD19/10	A	A	C	C																																													
HPV6vc	A	A	C	C																																													

	L1										LCR																																			
	5923	5992	6052	6073	6104	6217	6598	6625	6661	7219	7282	7349	7356	7400	7452	7467	7513	7514	7515	7516	7517	7518	7519	7520	7616	7617	7618	7623	7628	7661	7696	7700	7747	7748	7812	7840	7884	7929	7954	20						
HPV6b	A	G	C	A	A	A	A	C	G	A	T	G	C	C	C	G	C	C	T	C	T	T	G	T	T	A	A	C	C	A	T	C	G	C	C	A	C	A	C	G	G	T				
HPV6a	T	A	T	G	C																																									
VBD19/10	T	A	T	G	C																																									
HPV6vc	T		T	G		T	T	A	A	C	C	T		A	A	T	A	II																												

Base positions where mutations were observed are indicated at the top vertically. The shaded blocks indicate no variation from the reference genome. Insertions are indicated by I and deletions are indicated by d.

I1 = TACATTATTGTATA.

I2 = ATATGTTTATTGCCACTGCA.

I3 = TCACCTGGCGCCAGGGTGC GG TATTGCCTTACTCATATGTTTATTGCCACTGCAATAAACCTGTCTTTGTGTTATACTTTTCTGCAC TGTAGCCA ACTCTTAAAAGCATT TTTGGCTTGTAGCAGAACATTTTTTGTCTTACTGTTTGGTATAACAATAACATAAAAATG.

I4 = T.

\* NNCR3 is a non-classic non-coding region between the stop codon of genomic region E5b and start codon of genomic region L2.

HPV6a (L41216) from sublineage B3, were included in the alignment for comparing genomic variation. Mutations observed in the full-length genome of VBD19/10 with base positions relative to the HPV6b reference genome are summarized in Table 2 with the amino acid substitutions and corresponding residue positions in Table 3.

The phylogenetic tree showed the isolate to form part of the B3 sublineage, with the closest isolates being ZA54/10 and ZA65/11 (Fig. 1). Pairwise distances relative to these isolates and to the corrected

reference HPV6b isolate from lineage A, HPV6vc from sublineage B1, and HPV6a from sublineage B3 are shown in Table 4.

### Functional analysis

The level of  $\beta$ -galactosidase activity expressed from the lacZ gene of cells transfected with the plasmid containing the duplication within the LCR (pBlueTopo\_19/10) was significantly higher than the level expressed from cells transfected with the plasmid

Table 3. Summary of the amino acid substitutions observed in the full-length genome of VBD19/10

Genomic region	Reference amino acid	Mutated amino acid	Residue position*	Nucleotide position†
E6	H	Q	50	251
E7	G	E	22	594
	N	D	88	791
E1	L	A	235	1534/1535
	L	V	280	1669
	E	D	303	1740
	T	S	459	2207
	D	E	575	2556
	T	A	608	2653
E2	H	N	27	2801
	M	I	36	2830
	R	K	84	2973
	N	T	142	3147
	T	S	144	3152
	H	Q	183	3271
	L	P	222	3387
	K	Q	228	3404
	R	K	307	3642
	S	P	324	3692
	D	H	338	3734
	D	N	348	3764
	S	R	357	3793
E4	I	N	6	3271
	G	E	60	3433
	P	H	69	3460
	D	E	99	3551
E5a	F	L	16	3934
	E	D	39	4003
	L	C	80	4125/4126
	Y	H	84	4136
	T	N	88	4149
E5b	M	V	26	4234
	K	N	46	4296
	N	T	52	4313
	D	A	55	4322
	Y	H	62	4342
	T	D	63	4345/4346
	D	A	65	4352
L2	Q	G	75	4645/4646
	V	I	185	4975
	V	I	200	5020
	G	D	356	5489
	L	I	392	5596

None of the mutations observed in the L1 gene modified the amino acid.

\* Residue positions are relative to the HPV6b prototype.

† Nucleotide positions are relative to the HPV6b prototype.

without the duplication (pBlueTopo\_02/10) in both experiments. In experiment 1, the range of pBlueTopo\_02/10 was 9.4–12.7 (mean 10.89) and that of pBlueTopo\_19/10 was 35.3–57.2 (mean 44.6)

( $P=0.000$ ) while in experiment 2, pBlueTopo\_02/10 ranged between 39.1 and 92.1 (mean 70.0) and pBlueTopo\_19/10 ranged between 113.5 and 287.8 (mean 193.6) ( $P=0.002$ ) (Fig. 2). The ratio of specific activity detected in cells transfected with plasmid pBlueTopo\_19/10 to that detected in cells transfected with plasmid pBlueTopo\_02/10 was 3.4 (95% CI 2.9–3.9). The difference in absolute values between the two experiments is probably because they were performed at different times using different passages of cells.

## DISCUSSION

The complete genome sequence was determined and *in vitro* functional analysis was performed on HPV6 isolate VBD19/10. In total, mutations were observed at 157 nucleotide positions and included nucleotide substitutions, deletions and insertions, resulting in amino acid changes at 43 residue positions.

Four amino acid mutations that have been described previously: at residue 222 in the E2, residue 6 in the E4, residue 65 in the E5b, and residue 55 in E5b regions resulting in changes in the hydrophobic property of the protein that are probably more significant than the two novel mutations that did not alter hydrophobicity index, as hydrophobic amino acids are likely to be located within the protein folds. The E2 gene regulates the transcription of the HPV genome and a difference in transcriptional activation can be a result of a difference in the specific binding of E2 to DNA. Thus, a modification in the amino acid properties could alter the folding of the protein and may influence the specific binding of E2 to DNA and ultimately influence transcription of the viral genome. Moreover, the E4 and E5 proteins modify the cellular environment to assist indirectly in amplification of the viral genome, with reports of transforming activities for the E5 gene. Therefore, mutations in the E4 and E5 proteins, especially non-synonymous mutations could have extensive effects on the structure of the proteins. The implications and significance of each amino acid change would require further investigation, possibly using mutagenesis studies.

The biological functioning of promoter sequences can be investigated using reporter vectors in which PCR-amplified regions are inserted into promoterless vectors upstream of a reporter gene. A 170 bp duplication was previously identified in the central segment of the LCR of an HPV6 isolate from a patient with

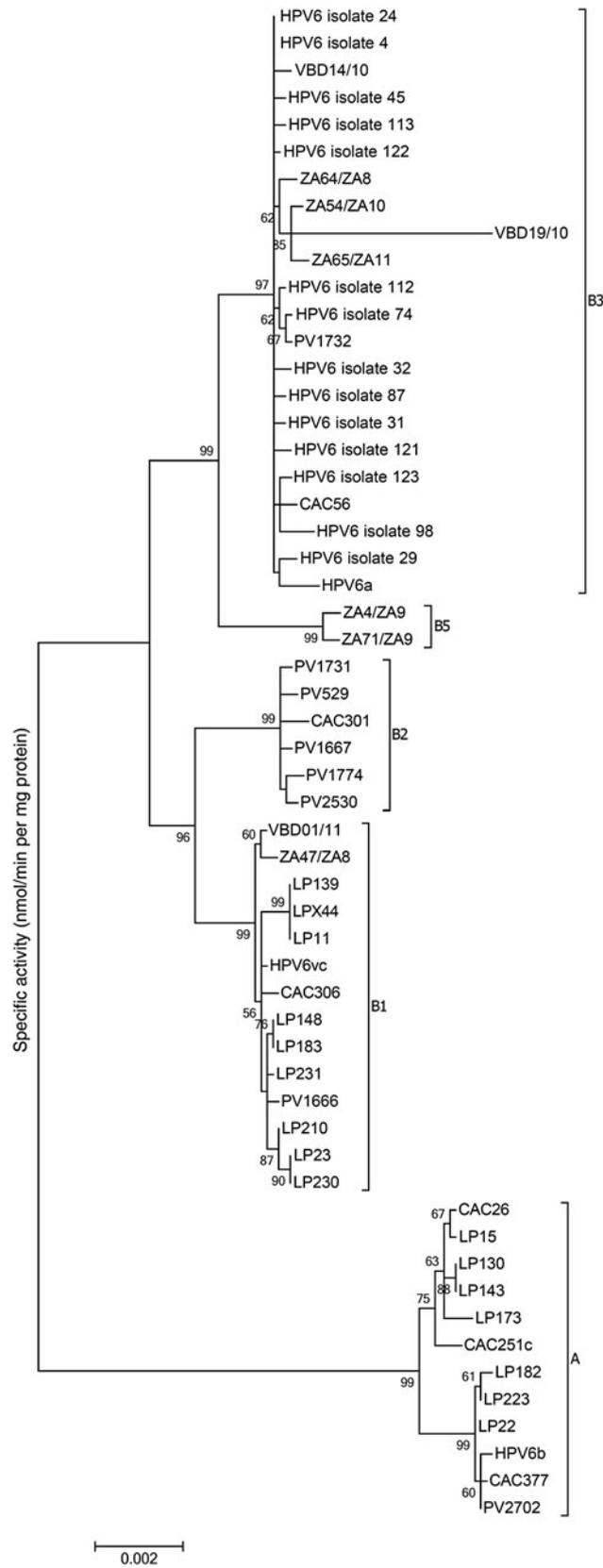
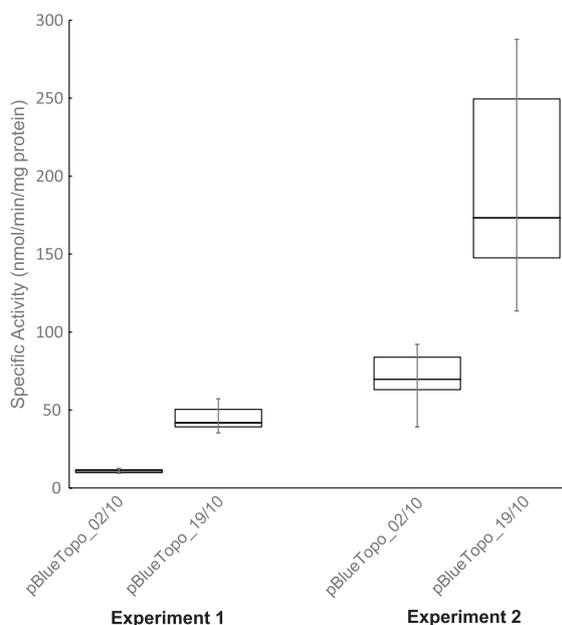


Fig. 1. Phylogenetic tree of representative HPV6 isolates.

Table 4. *Pairwise distances*

Isolate	Pairwise distance (%)
ZA54/10	0.4
ZA65/11	0.4
HPV6b	1.8
HPV6vc	0.9
HPV6a	0.5



**Fig. 2.** Level of  $\beta$ -galactosidase activity expressed from the *lacZ* gene of cells transfected with the plasmid containing the duplication within the long control region (pBlueTopo\_19/10) and without the duplication (pBlueTopo\_02/10).

RRP [10]. The duplication was located where the majority of transcriptional factor-binding sites are and between two E2-binding sites [8]. In HPV11, this central segment of the LCR has been shown to serve as an enhancer specific for epithelial transcription [20]. Quantitation of the enzyme  $\beta$ -galactosidase expressed by the reporter gene, *lacZ*, was used as an indicator of the functioning of the promoter sequences. Although the specific activity level of  $\beta$ -galactosidase expressed from cells transfected with the recombinant plasmid pBlueTopo\_19/10 was 3.4 times higher than the activity determined from cells transfected with recombinant plasmid pBlueTopo\_2/10, the conclusions drawn need to take into consideration the limitations of the assay. Estimations of transfection efficiencies based on staining cells for protein  $\beta$ -galactosidase showed higher transfection efficiencies from the recombinant plasmid pBlueTopo\_19/10. Higher transfection

efficiency could be responsible for detection of higher levels of specific activity. However, similarly, higher transfection efficiencies may have been detected by staining due to higher levels of protein expression. Hence the outcome suggests that the influence of the 170 bp duplication on the levels of transcription of downstream genes warrants further investigation.

Mutations that occur in the LCR could have an enhancing or inhibiting effect on the replication and transcription of the virus. Reporter gene technology has been used successfully in the analysis of transcriptional regulation by HPV6 E2 proteins and in the evaluation of the effect of sequence modifications in the LCR of HPV6 isolates on gene expression [14, 21]. Sequence rearrangements in the LCR did not seem to affect the promoter activity, but where large duplications in the LCR were identified, it resulted in enhanced promoter activity and thus it was stated that the duplication may have caused an increase in the oncogenic potential of HPV6 variants ascribed to overexpression of E6 and E7 [14]. Further investigations and long-term studies of patients are warranted to determine if it is possible to use duplications as prognostic indicators or biomarkers of disease. It may also depend on the site at which the mutation occurs, for instance at the origin of replication, at E2-binding sites, or whether additional transcription binding sites were produced. This, however, needs further investigation, namely introducing mutagenesis in the form of the same duplication in the LCR of the control in order to determine whether an increase in expression levels is due to an increase in promoter activity or due to a higher transfection efficiency. If the duplication enhanced the activity and consequentially the expression of the protein, it could be used as a molecular determinant in the prediction of aggressiveness of RRP disease. Isolate VBD19/10, which was obtained from a patient with aggressive RRP disease, showed additional mutations in other regions of the genome other than the duplication observed in the LCR. The role of these mutations also warrants investigation. The combination of the mutations could also contribute to the aggressive behaviour of the disease, but requires further investigation.

In conclusion, this study provides complete genome sequence data on a HPV6 isolate from a patient with aggressive RRP. We have possibly identified a molecular determinant that could influence promoter activity. However, further mutagenesis investigations will be required to substantiate the exact role of

duplications on promoter activity and possibly disease severity.

## ACKNOWLEDGEMENTS

We acknowledge Mr Stephanus Riekert for providing resources and assistance on the High Performance Computing Unit's infrastructure at the University of the Free State.

This project was funded by grants from the South African Society of Otorhinolaryngology-Head and Neck Surgery Research Fund, National Health Laboratory Service Research Trust and University of the Free State Cluster funding.

## DECLARATION OF INTEREST

None.

## REFERENCES

1. **de Villiers E-M, et al.** Classification of papillomaviruses. *Virology* 2004; **324**: 17–27.
2. **Bernard H-U, et al.** Classification of papillomaviruses (PVs) based on 189 PV types and proposal of taxonomic amendments. *Virology* 2010; **401**: 70–79.
3. **Seedat RY, Combrinck CE, Burt FJ.** HPV associated with recurrent respiratory papillomatosis. *Future Virology* 2013; **8**: 477–492.
4. **Jelen MM, et al.** Global genomic diversity of human papillomavirus 6 based on 724 isolates and 190 complete genome sequences. *Journal of Virology* 2014; **88**: 7307–7316.
5. **Burk RD, Harari A, Chen Z.** Human papillomavirus genome variants. *Virology* 2013; **445**: 232–243.
6. **Lazarczyk M, et al.** The EVER proteins as a natural barrier against papillomaviruses: a new insight into the pathogenesis of human papillomavirus infections. *Microbiology and Molecular Biology Reviews* 2009; **73**: 348–370.
7. **Zheng Z-M, Baker CC.** Papillomavirus genome structure, expression, and post-transcriptional regulation. *Frontiers in Bioscience* 2006; **11**: 2286–2302.
8. **O'Connor M, Chan S-Y, Bernard H-U.** Transcription factor binding sites in the long control region of genital HPV. In: Myers G, Bernard H-U, Delius H, et al., eds. *Human Papillomaviruses, 1995 Compendium, Part III*. Los Alamos, New Mexico: Los Alamos National Laboratory, 1995, pp. 21–40.
9. **Kocjan BJ, et al.** Pre-vaccination genomic diversity of human papillomavirus genotype 6 (HPV 6): a comparative analysis of 21 full-length genome sequences. *Infection, Genetics and Evolution* 2011; **11**: 1805–1810.
10. **Combrinck CE, et al.** Novel HPV-6 variants of human papillomavirus causing recurrent respiratory papillomatosis in southern Africa. *Epidemiology and Infection* 2012; **140**: 1095–1101.
11. **Heinzel PA, et al.** Variation of human papillomavirus type 6 (HPV-6) and HPV-11 genomes sampled throughout the world. *Journal of Clinical Microbiology* 1995; **33**: 1746–1754.
12. **Kämmer C, et al.** Sequence analysis of the long control region of human papillomavirus type 16 variants and functional consequences for P97 promoter activity. *Journal of General Virology* 2000; **81**: 1975–1981.
13. **Farr A, et al.** Relative enhancer activity and transforming potential of authentic human papillomavirus type 6 genomes from benign and malignant lesions. *Journal of General Virology* 1991; **72**: 519–526.
14. **Grassmann K, et al.** HPV6 variants from malignant tumors with sequence alterations in the regulatory region do not reveal differences in the activities of the oncogene promoters but do contain amino acid exchanges in the E6 and E7 proteins. *Virology* 1996; **223**: 185–197.
15. **Lace MJ, et al.** 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. *Journal of Virology* 2009; **83**: 7457–7466.
16. **Larkin MA, et al.** Clustal W and Clustal X version 2.0. *Bioinformatics* 2007; **23**: 2947–1948.
17. **Kearse M, et al.** Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 2012; **28**: 1647–1649.
18. **Bonfield JK, Smith KF, Staden R.** A new DNA sequence assembly program. *Nucleic Acids Research* 1995; **23**: 4992–4999.
19. **Tamura K, et al.** MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution* 2013; **30**: 2725–2729.
20. **Chin MT, Broker TR, Chow LT.** Identification of a novel constitutive enhancer element and an associated binding protein: implications for human papillomavirus type 11 enhancer regulation. *Journal of Virology* 1989; **63**: 2967–2976.
21. **Kovelman R, et al.** Human papillomavirus type 6: classification of clinical isolates and functional analysis of E2 proteins. *Journal of General Virology* 1999; **80**: 2445–2451.