TaqI is a single cut enzyme for HPV-18

John Meissner

Department of Microbiology, Duke University Medical Center, Durham, NC 27710, USA

Received December 18, 1992; Accepted January 5, 1993

EMBL accession no. L07917

The published nucleotide sequence of the human papillomavirus type 18 (HPV-18) genome (1) contains two *TaqI* restriction enzyme recognition sites (TCGA) at nucleotides 838 and 3083. However, a *PvuII*–*AvrII* subclone of HPV-18 extending from nucleotides 854–4599 (p18E1E2, which includes the E1, E2, E5, and a 5' portion of the L2 ORFs), was resistant to *TaqI* cleavage at the latter site (Figure 1). *TaqI* is inhibited by N6 methylation of the adenine in its recognition sequence (2, 3), but the sequence flanking the *TaqI* site at 3083 in HPV-18 is not a target for *E. coli* Dam methylation (See Figure 3).

Sequencing of the sense strand of the E2 ORF in p18E1E2 using a synthesized oligonucleotide primer revealed a compression artifact in the region of the predicted *TaqI* site. The substitution of dITP analogs resolved the artifact, and indicates that the sequence starting at nucleotide 3083 is TGCA, instead of the published TCGA (Figure 2). This was confirmed by opposite strand sequencing (data not shown). To eliminate the possibility that this difference was simply a spontaneous mutation in p18E1E2, both *TaqI* digestion and double stranded sequencing of the original recombinant plasmid used to determine the HPV-18 sequence, (pHPV18) was performed, with similar results (Figures 1 and 2).

The correction in the nucleotide sequence reduces *TaqI* to a single cut enzyme for HPV-18, with its sole recognition site at nucleotide 838. This is in agreement with the restriction map published for the original HPV-18 isolate (4). The correction also adds a third *DraIII* site (Figure 3), confirmed by pHPV18 *DraIII* digestion (data not shown). As well, the reversal of nucleotides 3084 and 3085 changes the identity of one residue in the deduced amino acid sequence of the putative HPV-18 E2 protein, from arginine to alanine (Figure 3). This residue lies within the so-called A2 α helix region of the E2 *trans*-activation domain (5).

ACKNOWLEDGEMENTS

The p18E1E2 plasmid was a gift of Dr Peter Howley, and Dr Olivier Danos generously provided pHPV18. I would also like to thank Dr Inder Patel for synthesizing the oligonucleotide primers.

REFERENCES

1. Cole, S.T. and Danos, O. (1987) *J. Mol. Biol.* **193**, 599–608.
2. Nelson, M. and McClelland, M. (1991) *Nucleic Acids Res.* **19** (Suppl.), 2045–2071.
3. Backman, K. (1980) *Gene* **11**, 169–171.
4. Boshart, M., Gissmann, L., Ikenberg, H., Kleinheinz, A., Scheurlen, W. and zur Hausen, H. (1984) *EMBO J.* **3**, 1151–1157.
5. Giri, I. and Yaniv, M. (1988) *EMBO J.* **7**, 2823–2829.

Figure 1. lane 1: p18E1E2 *TaqI* digest; lane 2: p18E1E2 *TaqI* – *Tth111I* double digest; lane 3: pHPV18 *EcoRI* fragment (linearized, full-length HPV-18 genome); lane 4: pHPV18 *EcoRI* fragment digested with *TaqI*; lane 5: Lambda DNA *HindIII* molecular size standard. *Tth111I* is a single cut enzyme for HPV-18, with a recognition site 300 nucleotides upstream of the predicted HPV-18 *TaqI* site at 3083. The 1400 bp band in lanes 2 and 3 is from the pUC vector, which has no *Tth111I* recognition sites.

Figure 2. Sequence of the sense strand of p18E1E2 (left) and pHPV18 (right) around the region of the predicted *TaqI* recognition site at nucleotide 3083, with lanes as indicated.

HPV-18 Published Sequence

Gln	Gly	Leu	Ala	Gln	Ser	Arg	Tyr	Lys	Thr
...CAA	GGC	CTT	GCA	CAA	AGT	CGA	TAC	AAA	ACC...
	3			3	<i>TaqI</i>		3		
	0			0	site		0		
	7			8			9		
	0			0			0		

HPV-18 Corrected Sequence

Gln	Gly	Leu	Ala	Gln	Ser	Ala	Tyr	Lys	Thr
...CAA	GGC	CTT	GCA	CAA	AGT	GCA	TAC	AAA	ACC...
	3				<i>DraIII</i>		3		
	0				site		0		
	7						9		
	0						0		

Figure 3. Nucleotide and deduced amino acid sequence within the A2 region of the HPV-18 E2 protein as published (top) and from corrected nucleotide sequence data (bottom).