

Hypothetical genes are frequent targets of HPV DNA integration in cervical carcinomas



Irene Kraus (1,2)*, Corina Ziegert (3), Tatiana Koch (3), Achim Schneider (4), Matthias Dürst (3)

(1) Institute of Pathology, The Norwegian National Hospital, Oslo, Norway. (2) NorChip AS, Klokkestrua, Norway. (3) Gynäkologische Molekularbiologie, Frauenklinik der Friedrich-Schiller-Universität Jena, Germany. (4) Frauenklinik mit Hochschulambulanz Campus Benjamin Franklin, Charité Universitätsmedizin, Berlin, Germany.

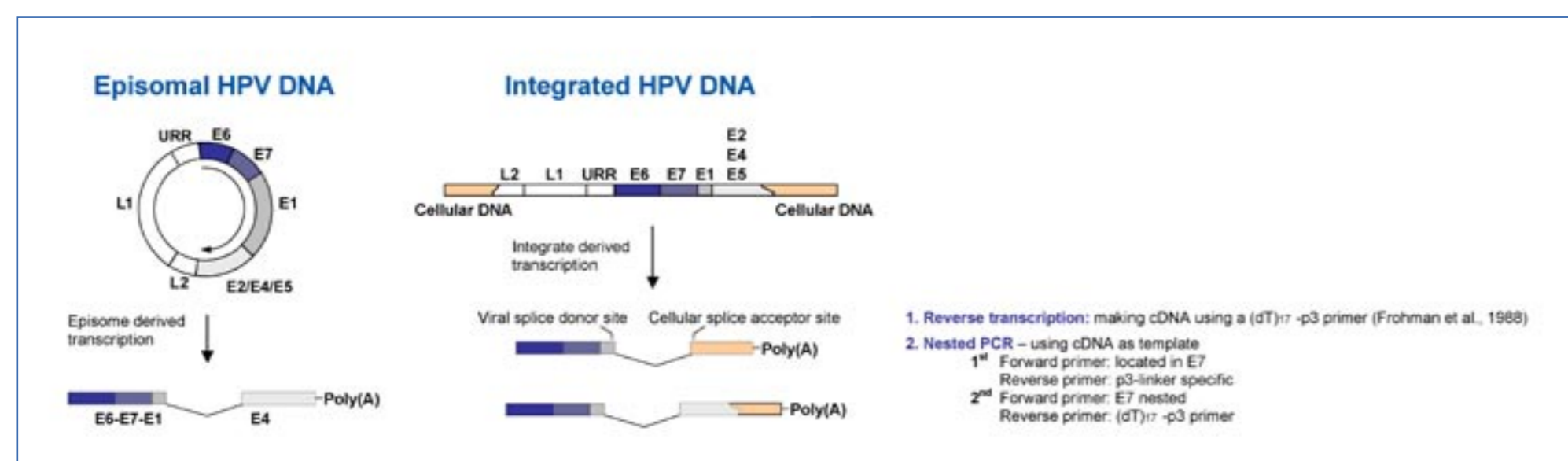
Objective

Integration of HPV DNA into the host genome is a frequent event in cervical carcinogenesis. Viral integration sites seem to be randomly distributed among all chromosomes with a clear preference for genomic fragile sites. To date about 200 HPV DNA integration sites have been mapped (reviewed by Wentzensen et al., 2004). However, this number is still too small to fully comprehend the different mechanisms by which HPV DNA integration contributes to carcinogenesis.

Methods

A total of 56 cervical carcinoma samples of FIGO stage IA to IIIB positive for either HPV16 (n=33) or HPV18 (n=23) were analysed. Episomal and viral-cellular fusion transcripts were detected using the APOT assay (Figure 1), a specific 3'-RACE PCR (Klaes et al., 1999). All putative viral-cellular fusion transcripts were cloned into the pCR4-TOPO™ vector and analysed by sequencing. The existence of the fusion transcripts was confirmed by RT-PCR using integration site specific viral and cellular primers. The genomic localisation of the integration sites were determined by database query using NCBI Blast and UCSC Blat. For cases with integration within a hypothetical gene, specific primers were designed and an extended expression analysis was performed to investigate the expression pattern of the respective gene in tumours and CIN-lesions. To assess a possible up or down-regulation of the genes, 25 tumour samples, 10 CINIII samples, and 10 normal tissue samples were included in a quantitative expression analysis using SYBR-green real-time PCR. GAPDH was used as housekeeping gene.

Figure 1. APOT - Amplification of Papillomavirus Oncogene Transcripts



Results

The detection rate of viral-cellular fusion transcripts from HPV16 positive tumours was significantly lower ($p=0.012$, two-sided t-test) than that of HPV18 positive tumours (Table 1). A BLASTN comparison of the fusion transcripts with the NCBI Database found nine fusion transcripts containing known genes; 22 integration sites were mapped within hypothetical genes listed in the UCSC Genome Browser Database (Table 2). For the cases of integration into known genes, seven out of nine samples showed integration within an intron. Chromosomal integration loci were distributed over the whole genome, with a preferential site in the chromosomal bandings 8q24 and 13q22.1. Nine hypothetical genes showed expression also in normal tissue or CIN-lesions (Figure 2A), for which a comprehensive expression profile was performed using real-time quantitative PCR (Figure 2B).

Table 1. Frequency of viral-cellular fusion transcripts

	HPV 16 (Frequency %)	HPV 18 (Frequency %)
Episomal	16 (48,5%)	5 (21,7%)
Episomal+ Integrated	4 (12,1%)	1 (4,3%)
Integrated	13 (39,4%)	17 (74%)
Total	33 (100%)	23 (100%)

Discussion

- a high proportion of viral integration sites within in silico predicted genes (71%).
- a relatively low degree of integration sites located within known genes (29%), being not in line with previous reports showing a higher frequency of integration into genes with known gene function (Ziegert et al., 2003; Wentzensen et al., 2004).
- a high prevalence of integration into intronic regions, being in accordance with previous reports.
- three cases of integration into the chromosome banding 8q24, harboring i.a. the MYC gene, and 13q22.1, verifying the finding of these loci being common targets for HPV-integration (Wentzensen et al., 2004).
- upregulation or downregulation of a hypothetical gene in CxCa compared to normal tissue or CIN-lesions, suggests the respective gene as a candidate gene involved in the process of carcinogenesis.
- HPV integration analysis by the APOT assay may be a useful tool for the detection of new oncogenes and tumour-suppressor genes.

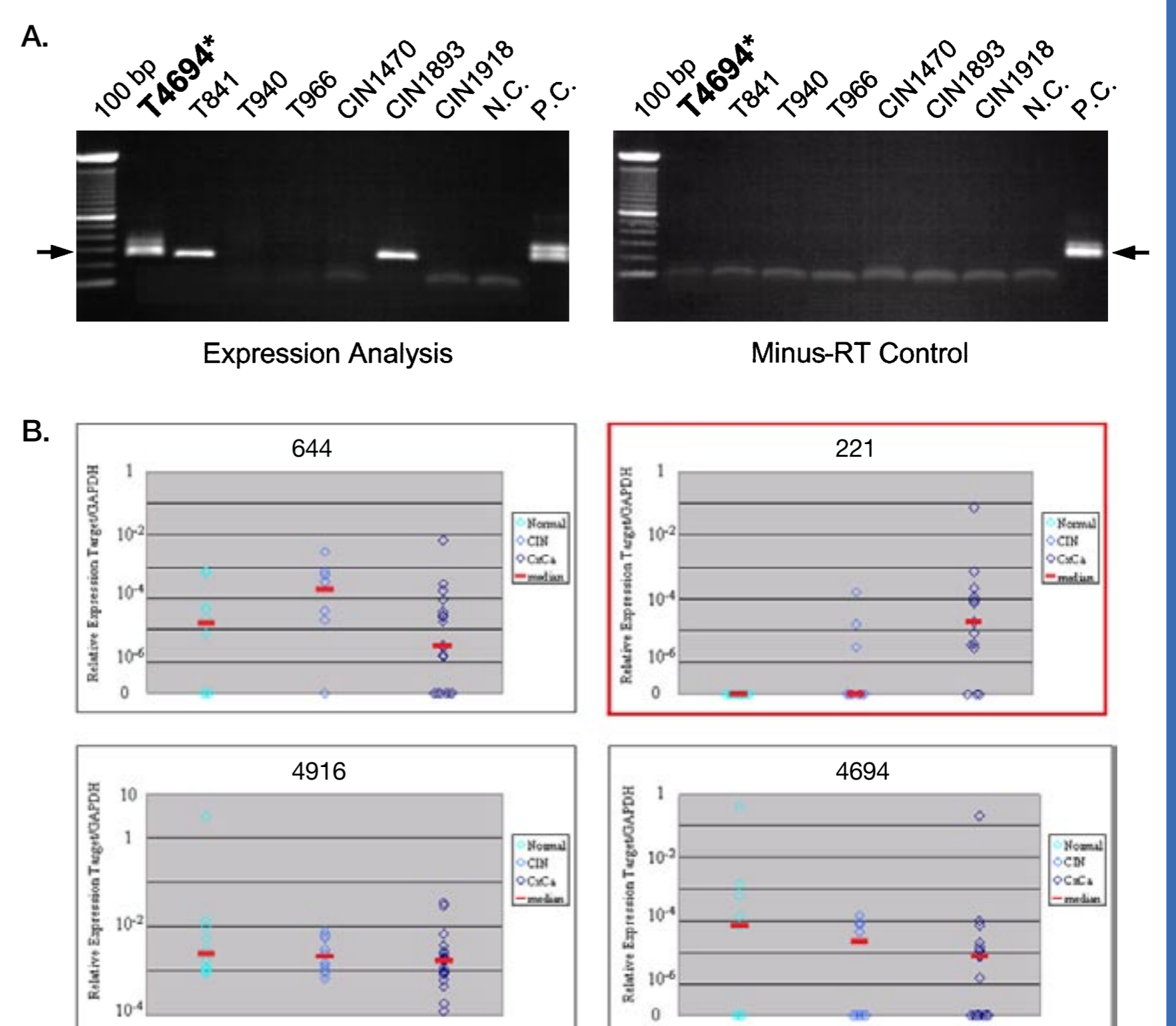
Table 2. HPV 16 and 18 integration sites – summary of the transcripts analysed

KNOWN GENES							
Sample ID	HPV type	HPV donor site (E1)	Algorithm used for prediction	Map	HUGO gene name	Location of cellular acceptor site	Orientation of HPV relative to cellular gene
T2147	16	880	NCBI Blast	17q21	SUJ1	Intron	Sense
T4743	16	880	NCBI Blast	19q13	CEACAM6	Intron	Antisense
T4177	18	929	NCBI Blast	2q24.3	ITGB6	Intron 4/7	Antisense
T2341	18	929	NCBI Blast	4p15.32	KCNIP4	Intron 1	Antisense
T4251	18	929	NCBI Blast	20p11	FLRT3	Exon 2	Sense
T4290	18	929	NCBI Blast	14q11.2	C14orf94	Intron 2	Antisense
DNR-3835	18	929	NCB Blast	7q31	IMMP2L	Exon 5	Sense
DNR-3918	18	929	NCB Blast	2q21.2	LRP1B	Intron 7	Antisense
DNR-4026	18	929	NCB Blast	5q31	ARHGAP	Intron 20	Antisense

HYPOTHETICAL GENES							
Sample ID	HPV type	HPV donor site (E1)	Algorithm used for prediction	Map	Predicted full-length mRNA (bp)	Location of cellular acceptor site	Orientation of HPV relative to cellular gene
T174	16	880	AceView	2p24.3	1045	Intron 2	Antisense
T183	16	880	Swiss-prot	1p36.11	790	Intron 1	Antisense
T315	16	880	ECgene	8q24.1	581	Exon	Sense
T644	16	880	ECgene	2p13.1	644	5'UTR+Exon1	Antisense
T1013	16	880	AceView	20p11.21	679	5'UTR	Sense
T1051	16	880	Gene id	4q23	2247	Intron 2	Sense
T2594	16	880	Swiss-prot	8q12.1	2971	Intron 1	Sense
T3283	16	880	NCBI-Blast	13q22.1	588	Intron 2	Antisense
T3706	16	880	Genescan	13q22.1	915	Intron 5	Sense
T4653	16	880	AceView	19p13.11	878	Intron 3	Sense
T4694	16	880	Gene id	8q24.21	457	Intron 1	Sense
T3817	16	880	Swiss-prot	9p24	3623	Intron 3	Antisense
T1981	18	929	Swiss-prot	2p25.1	779	Intron 2	Sense
T221	18	929	NCBI-Blast	15q21.2	498	Intron 1	Antisense
T1729	18	929	SGP	15q25.3	279	Intron 1	Sense
T4916	18	929	Genescan	19q13.11	1239	Exon 1	Sense
DNR-3464	18	929	AceView	8q24.21	660	Exon 3	Sense
DNR-3772	18	929	SGP	13q22.1	654	Intron 1 + Exon 2	Antisense
DNR-3899	18	929	Genescan	4q12	3903	Intron 4	Sense
DNR-3923	18	929	AceView	Xq25	3656	Intron 18	Antisense
DNR-4034	18	929	Genescan	2q34	273	Intron 1	Sense
DNR-4045	18	929	Genescan	8q21.3	246	Intron 2	Sense

Definition of hypothetical genes: Expressed sequence tags (ESTs) and sequences found containing an open reading frame, for some also with a predicted function based on homology to known proteins.

Figure 2. Expression analysis of hypothetical genes in different tissue material



A. Expression of one candidate hypothetical gene (sample T4694) in tumours (T) and CIN-lesions (picture left), including a minus RT-control to exclude the possibility of being a product of DNA-amplification (picture right); expected size is shown by arrows. The results clearly indicate that the predicted gene is expressed also in lesions in which this locus is not affected by virus DNA integration. Notably, in several lesions this gene appears to be silent. NC: negative control, PC: positive control. B. Quantitative expression of four candidate genes. The hypothetical gene found targeted in sample T221, shows an upregulation in CxCa compared to normal tissue and CIN-lesions (highlighted).



References

- Klaes, R., Woerner, S.M., Ridder, R., Wentzensen, N., Dürst, M., Schneider, A., Lotz, B., Melsheimer, P., and von Knebel Doeberitz, M. (1999). Detection of high-risk cervical intraepithelial neoplasia and cervical cancer by amplification of transcripts derived from integrated papillomavirus oncogenes. *Cancer Res* 59, 6132-6.
- Ziegert, C., Wentzensen, N., Vinokurova, S., Kisselev, F., Einankel, J., Hoeckel, M., and von Knebel, D.M. (2003). A comprehensive analysis of HPV integration loci in anogenital lesions combining transcript and genome-based amplification techniques. *Oncogene* 22, 3977-3984.
- Wentzensen, N., Vinokurova, S., and von Knebel, D.M. (2004). Systematic review of genomic integration sites of human papillomavirus genomes in epithelial dysplasia and invasive cancer of the female lower genital tract. *Cancer Res* 64, 3878-3884.