Molecular and Cellular Probes 25 (2011) 260–265

Contents lists available at ScienceDirect

Molecular and Cellular Probes

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

The HPV16 transcriptome in cervical lesions of different grades

Markus Schmitt [a,*], Véronique Dalstein [b], Tim Waterboer [a], Christine Clavel [b], Lutz Gissmann [a,c], Michael Pawlita [a]

[a] Department of Genome Modifications and Carcinogenesis (F020), Research Program Infection and Cancer, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 242, 69120 Heidelberg, Germany
[b] INSERM UMR-S 903, CHU Reims (University Hospital of Reims), Laboratory of Histology, Cytology and Cellular Biology, 51100 Reims, France
[c] Department of Botany and Microbiology, King Saud University, Riyadh, Saudi Arabia

ARTICLE INFO

Article history:
Received 23 February 2011
Received in revised form
25 May 2011
Accepted 25 May 2011
Available online 2 June 2011

Keywords:
HPV
HPV16
Human papillomavirus
Cervical cancer (CxCa)
Transcriptome
NASBA
BSGP5+/6+-PCR/MPG
Transcription
mRNA

ABSTRACT

Infections with high-risk human papillomaviruses (HPV), mainly HPV type 16, can cause malignant transformation of the human cervical epithelium and cervical cancer (CxCa). Very little is known about the quantitative expression of HPV16 transcripts in cervical lesions of different grades.

We have analysed the viral transcriptome in 80 HPV16 DNA positive cervical smears including lesions of different cytological grades, using nucleic acid sequence-based amplification (NASBA)–Luminex hybridisation assays quantifying spliced and unspliced HPV16 transcripts.

Based on the quantitative analysis of single transcripts, highly significant changes in transcript levels were observed between different grades of cervical lesions.

In conclusion, quantitative expression changes of HPV16 transcript markers may be involved in tumour progression. This study provides a basis for selection of candidate RNA markers for diagnostics of HPV16-related disease.

1. Introduction

Human papillomavirus type 16 (HPV16) is present in about 55% of all cases of cancer of the uterine cervix (cervical cancer, CxCa) [1] and plays a central role in the development of this disease. During carcinogenesis, integration of viral DNA with disruption of the E2 *open reading frame* (ORF) occurs in about 62% of cases, leading to an upregulation of the viral oncogenes E6 and E7 [2–4]. CxCa develops from persistent high-risk (hr) HPV infections through stages histologically classified as cervical intraepithelial neoplasia 1 to 3 (CIN1 to CIN3), or cytologically defined as low- or high-grade squamous intraepithelial lesion (LSIL or HSIL).

HPV16 gene expression is complex. Transcription is regulated by at least two promoters, multiple splice donor and acceptor sites, at least two polyadenylation signals and further post-transcriptional mechanisms. All transcripts are synthesised from only one DNA strand and are polycistronic. Depending on splicing events, up to three different reading frames can be employed for translation. Most prominent promoters are the early promoter (p97) located in the *long control region* (LCR) and the differentiation-inducible promoter (p670) important for transcription of late genes encoding L1, L2 and E1^E4 [5]. Additional promoter sites may exist in the LCR, but also in the E4 and E5 region for generation of late transcripts [6]. The early and late polyadenylation signals are located at nucleotides (nt) 4215 and 7321, respectively. Viral RNA maturation involves complex splicing. The HPV16 transcriptome exhibits several splice donor (nt 226, 880, 1302 and 3632) and acceptor sites (nt 409, 526, 742, 2582, 2709, 3358 and 5639) giving rise to at least 11 different splice junctions [7–11].

Spliced transcripts may lead to differential use of upstream codons ("leaky scanning" [12,13]) and allow more efficient translation of downstream ORFs [14–16]. For example, the E2 protein is mainly expressed from spliced 226^2709 mRNA (E6*IV) but also from the 880^2709 and 880^2582 mRNAs [13]. In contrast, the E5 protein has been reported to be expressed from an unspliced E2/E5 transcript, but not from the E1^E4/E5 transcript [17].

Moreover, spliced transcripts may be translated into truncated proteins exhibiting functions important for regulation of the viral

* Corresponding author. Tel.: +49 6221 424935.
E-mail address: markus.schmitt@dkfz.de (M. Schmitt).

0890-8508/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.mcp.2011.05.003

life cycle [18]. For instance, the N-terminally truncated E1 (E1C) protein encoded by the 880^2582 spliced mRNA has been suggested to act as a transactivator of the LCR [18]. A shorter form of E2 (E2C) translated from the 1302^3358 mRNA, is a putative replication and transcription repressor that inhibits the function of the full-length E2 protein [6]. Splicing in E6 can generate four either truncated (E6*I and *II) or fused forms (E6*III and *IV). The mode of action of putative proteins generated from these transcripts is still in question. While Alloul and Sherman reported that E6*I may transactivate and E6*IV may transrepress the viral LCR [18], others argued that these protein fragments are not themselves important, but that splicing shifts the translation balance in favour of E7 or other viral proteins [14].

HPV transcript patterns and their splice junctions have been studied in HPV16 positive cervical carcinomas and high-grade lesions [19], keratinocytes immortalised by HPV16 [10,19], and cervical cancer-derived cell lines [2,11,20–22]. First attempts to quantify expression profiles from E6*I, E6*II, and E6/E7 full-length (fl) mRNA, have been described for pre-malignant lesions [21–23] and tumour-derived cell lines [21,24]. Increased levels of E6/E7 mRNA were found to correlate with lesion progression [25].

However, very little is known about the quantitative expression changes of HPV16 transcripts in smears of cervical lesions of different grades. To this end, 80 RNAs originating from HPV16 DNA positive smears, 25 with no intraepithelial lesion or malignancy (NIL/M), 24 LSIL, 24 HSIL, and 7 CxCa cytological samples, were quantitatively analysed for 15 spliced and full-length HPV16 transcripts, and cellular $p16^{INK4A}$ as well as two housekeeping transcripts. We recently described the key features of the technology, showed the prevalence and expression data of four viral transcripts (E6*II, E1C, E1^E4, L1 fl) and identified a potential diagnostic HPV16 RNA pattern [26]. Here, we focus on describing the quantitative changes of all 15 HPV16 mRNAs in different stages of cervical lesions including NIL/M, LSIL, HSIL and CxCa using the same set of clinical samples.

2. Material and methods

2.1. Ethics statement

The study has been approved by the "Comité de Protection des Personnes" (CPP-Est III, Nancy, France), statement no DC-2008-374. All women were informed and gave their written consent to participate in the study.

2.2. Clinical samples

We selected 80 HPV16 DNA positive cervical scrapes stored in PreservCyt™ medium (Cytyc Corp., Marlborough, MA, USA) obtained from French patients, harbouring (i) no cytological abnormality (25 NIL/M) and (ii) lesions ranging from LSIL (24) over HSIL (24) to CxCa (7) [26]. Cytological classification was done according to the Bethesda system. All CxCa cases were histologically confirmed. As histological data were available for most HSIL, for only few LSIL and for none of NIL/M cases (which is in accordance with French guidelines for the management of cytological abnormalities), cytological classification alone was retained for all of these samples.

2.3. Nucleic acid extraction

Exfoliated cervical cells were stored in PreservCyt™ medium at room temperature for at most 3 weeks before freezing at −80 °C. After vigorous homogenisation, 3–12 ml of cell suspension was pelleted for 10 min at 10 °C and 4000 rpm (300 × *g*). Absolute ethanol (2.5 ml) was added and the mixture was homogenised by pipetting. The cell suspension was divided in two tubes. DNA isolation (1.0 ml) was performed using the EZ1 DNA kit (Qiagen, Hilden, Germany) and RNA isolation (1.5 ml) with the EZ1 RNA kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. DNase treatment was omitted during RNA extraction due to RNA specificity of NASBA. RNA concentration was measured and purified RNA was stored at −80 °C for up to two years. The median yield of total RNA was 41 ng/μL (range 3.4–341 ng/μL). For NASBA amplification, RNA was four-fold diluted in H_2O to dilute residual ethanol.

2.4. Quantitative BSGP5+/6+-PCR/MPG assay

HPV genotyping and β-globin assessment by the BSGP5+/6+-PCR/Multiplex HPV Genotyping (MPG) assay was performed as described [27–29] with some modifications. HPV amplification was carried out using the Multiplex PCR Kit (Qiagen, Hilden, Germany). 0.2–0.5 μM each of the BSGP5+ and 5′-biotinylated BSGP6+ primers, and 0.15 μM each of the β-globin primers MS3 and 5′-biotinylated MS10 were added to the PCR mixture. The cutoff value (5 net median fluorescence intensity (MFI)) to define HPV DNA positivity was applied as described previously [27]. Quantification of HPV signals was accomplished by computing for each positive reaction the relative HPV MFI signal (%) by dividing the measured HPV MFI value with the maximum value detected of this HPV type using colony PCR products. Finally, the relative MFI (%) was divided by the measured β-globin MFI value to form a non-descriptive viral load value (BS viral load) (% HPV MFI/β-globin MFI) (Schmitt et al., in preparation).

2.5. HPV16 RNA transcriptome analysis

Recently, we have described novel singleplex assays [26] that quantify 10 spliced HPV16 RNA sequences (226^409 (coding potential: E6*I), 226^526 (E6*II), 226^3358 (E6*III), 226^2709 (E6*IV), 880^2582 (E1C), 880^2709 (E2*), 880^3358 (E1^E4), 1302^3358 (E2C), 1302^5639, 3632^5639 (L1*)), 5 unspliced HPV16 RNA sequences (E6 fl, E7 fl, E1 fl, E5 fl, L1 fl) (Fig. 1), cellular $p16^{INK4A}$ mRNA and RNA of two housekeeping genes (Ubiquitin C (UbC) and U1A). 226^742 spliced mRNA was insufficiently amplified and thus excluded. The detection limits (DL) ranged from 25 to 2500 copies for all transcripts with the exception of $p16^{INK4A}$ reaching a DL of 25,000 copies per reaction [26]. Briefly, specific sequences were amplified from purified mRNA using the principle of competitive nucleic acid sequence-based amplification (NASBA) [30–32], in combination with simultaneous co-amplification of calibrator (Q) RNA (primer binding sites and amplimer size identical to wild-type (wt) RNA, but unique probe binding site), followed by hybridisation to oligonucleotide probes coupled to Luminex beads [28]. For each probe, MFI values in reactions with no amplimer added to the hybridisation mixture were considered background values. Net MFI values of hybridised amplimers were computed by subtracting 1.2 times the median background value from the raw MFI value. Net MFI values above 3 MFI were defined as positive reactions. For each probe, this cutoff definition was above the mean background plus three times the standard deviation.

For quantification of RNA amplimers, net MFI values above cutoff were normalised by the MFI signal obtained with the Q-RNA. Due to the non-linear relationship of the dose–response curve for some transcripts, quantitative categorisation of wt- versus Q-RNA ratios was achieved using standard curves generated from 25 to 2.5×10^6 copies of *in vitro*-transcribed wt-RNA per NASBA reaction allowing the detection of 10-fold expression differences.

2.6. Statistics

The respective wt- to Q-RNA ratios were correlated with cytological diagnosis and the significance of differences evaluated using