

## ARTICLES

# Typing of Human Papillomavirus by Pyrosequencing

Baback Gharizadeh, Mina Kalantari, Carlos A. Garcia, Bo Johansson, and Pål Nyren

*Department of Biotechnology (BG, CAG, PN), Royal Institute of Technology, Stockholm; and Department of Immunology, Microbiology, Pathology and Infectious Diseases (MK, BJ), Division of Clinical Virology, Karolinska Institutet, Huddinge University Hospital, Huddinge, Sweden*

**SUMMARY:** The possibility of using a new bioluminometric DNA sequencing technique, called pyrosequencing, for typing of human papillomaviruses (HPV) was investigated. A blinded pyrosequencing test was performed on an HPV test panel of 67 GP5+/GP6+ PCR-derived amplification products. The 67 clinical DNA samples were sequenced up to 25 bases and sequences were searched using BLAST. All of the samples were correctly genotyped by pyrosequencing and the results were unequivocally in accordance with the results obtained from conventional DNA sequencing. Pyrosequencing was found to be a fast and efficient tool for identifying individual HPV types. Furthermore, pyrosequencing has the capability of determining novel HPV types as well as HPV sequence variants harboring mutation(s). The method is robust and well suited for large-scale programs. (*Lab Invest* 2001, 81:673-679).

Human papillomaviruses (HPV) belong to the *Papovaviridae* family. HPV carry circular double-stranded DNA, approximately 8 kb in length (Godfroid et al, 1998), which encodes for several regulatory and structural proteins known as early (E), E1, E2, and E4 to E7, and late (L), L1 and L2, proteins. These proteins are involved in viral replication and have transforming (oncogenic) properties (Poljak et al, 1998; Schneider, 1993).

More than 100 different HPV types, more than 30 of which infect the cervical mucosa, have been identified on the basis of DNA sequence homology (Chan et al, 1995; Vernon et al, 2000). Because all HPV types are closely related, assays can be designed to target conserved regions of the genome or to target regions whose sequences can best be used to discriminate between different HPV types. Consensus assays use primers directed at relatively conserved regions of the HPV genome. The consensus amplicon is then typed by methods such as dot blot hybridization with type-specific probes, restriction fragment length polymorphism analysis, or gel-electrophoretic DNA sequencing (Vernon et al, 2000).

An alternative approach to specifically typing HPV is sequencing by a new bioluminometric, nonelectro-

phoretic DNA-sequencing method, called pyrosequencing (Ronaghi et al, 1998). This technique employs a cascade of coupled enzymatic reactions, using DNA polymerase, ATP sulfurylase, and luciferase to monitor DNA synthesis, along with a nucleotide-degrading enzyme in the system enabling iterative nucleotide dispensation to the reaction mixture (Fig. 1). The technique has the advantages of accuracy, flexibility, parallel processing, and simple automation. Additionally, it avoids the use of labeled primers (except in template preparation), labeled nucleotides, and gel electrophoresis.

Because of the high stability and conservation of HPV genomes over evolutionary times (Chan et al, 1995), even short segments of the genome can be used for reliable typing. The nucleotide sequences of the first 50 nucleotide bases of the L1 "consensus region" immediately downstream of the GP5+ primer site are type-specific enough for genotyping of most common genital HPV types. In this report, we discuss the application of pyrosequencing for sequencing 20 to 40 bases of amplicons for HPV genotyping. In this blinded clinical test, we used pyrosequencing to genotype 67 clinical specimens from different individuals. The results were compared with conventional DNA-sequencing data for confirmation. The samples were previously genotyped by type-specific PCR.

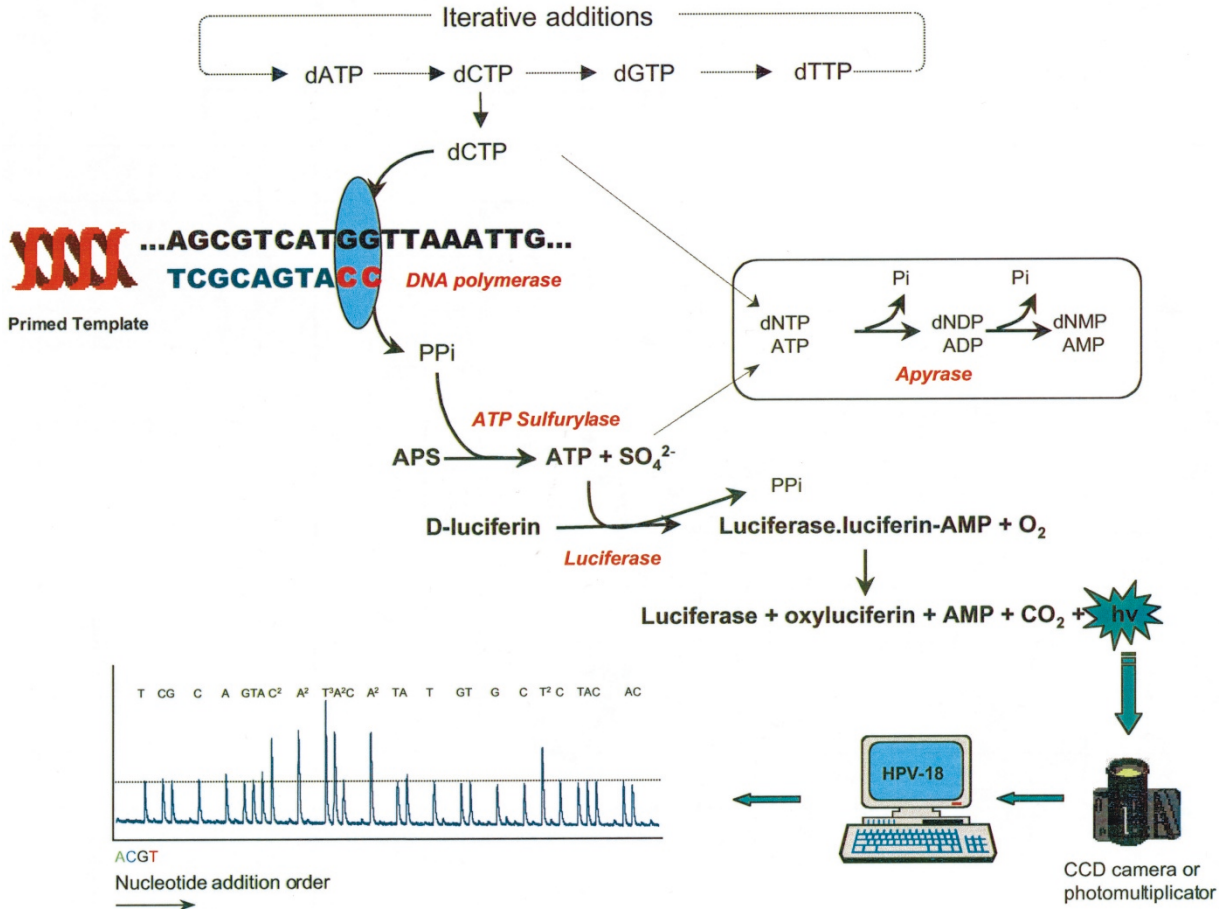
## Results

An HPV test panel, consisting of 67 samples from different individuals, was amplified in 150 bp fragments from the L1-conserved region in a blinded test. The general primers, GP5+ and biotinylated GP6+,

Received November 13, 2000.

*This work was supported by grants from the Swedish Research Council for Engineering Sciences (TFR), and the Swedish National Board for Industrial and Technical Development (NUTEK).*

*Address reprint requests to: Dr. Pål Nyren, Department of Biotechnology, The Royal Institute of Technology, Teknikringen 34, SE-100 44 Stockholm, Sweden. E-mail: paaln@biochem.kth.se*



**Figure 1.** Schematic representation of the automated pyrosequencing system.

were used. Twelve of the 67 samples were directly amplified from cell lysates, dealt with separately in this report. The other 55 samples were amplified from extracted DNA. The efficiency of the PCR amplification was evaluated by gel electrophoresis and ethidium bromide staining.

The immobilized single-stranded templates hybridized to the GP5+ sequencing primer were sequenced by pyrosequencing to determine the HPV type of each sample. The objective was to obtain sequence data from 20 to 25 bases. As shown in Figure 2, the nucleotide sequences of the first 14 to 40 nucleotides



**Figure 2.** Color-coded sequence alignment of the first fifty nucleotides downstream of the human papillomavirus (HPV) L1 GP5+ consensus primer site. The alignment was created using the Omega program, version 1.1.3 (Oxford Molecular Group, The Medawar Centre, Oxford, United Kingdom).

of the L1 consensus region immediately downstream of the GP5+ primer site are sufficiently type-specific among the most common HPV types to allow for typing.

The sequence data obtained by pyrosequencing were analyzed by BLAST search (<http://www.ncbi.nlm.nih.gov/BLAST/>) for HPV genotyping. To confirm the accuracy of the results, the PCR products were also sequenced by conventional DNA sequencing. The samples had also been previously typed by type-specific PCR.

As shown in Table 1, there was 100% concordance between pyrosequencing and conventional DNA sequencing. Figure 3 shows sequence results from seven different HPV types; the height of each peak is proportional to the number of nucleotides incorporated.

All of the 55 DNA extracts amplified with GP5+/6+ were sequenced in 20 to 25 bases by pyrosequencing and typed correctly. Only one sample (IS324, a cellular sample from a healthy individual) required further sequencing. This sample had six HPV type/isolate hits with 18 bases in the BLAST search. The differentiation for IS324 was made with data from 38 sequence bases.

In one case, an amplicon that had been previously (mis)typed as HPV-31 by type-specific PCR, was correctly genotyped as HPV-52 by both pyrosequencing and conventional DNA sequencing Table 1. We also detected a single T to C substitution in HPV-type 31 (Fig. 4).

In the cell lysate samples, only 4 of the 12 amplicons had a single band after gel staining and could be sequenced by pyrosequencing. Mixed pyrosequence signals were observed in the other eight samples. This could be due to unspecific amplifications because multiple bands were observed in the gel staining from those amplicons. Nested PCR was performed on the cell lysate samples using MY09/11 and GP5+/6+

primer sets. All of the nested amplicons were sequenced by pyrosequencing, demonstrating high signal intensity, which indicates large quantities and specific amplification products. The acquired sequences from the nested PCR agreed with conventional DNA sequencing.

In the samples from cervical cancer patients, HPV-16 was found in 20 of 35 cases (57%), HPV-18 in 6 of 35 (17%), and HPV-31 in 3 of 35 (9%). The remaining six samples contained the following HPV types: 6, 33, 35, 45, 52, and 59. Of the samples from dysplastic individuals, HPV-16 was found in 3 of 11 cases (27%), HPV-31 in 3 of 11 (27%), HPV-18 in 2 of 11 (18%), and HPV-66 in 1 of 11 (9%). The following HPV types were found in the samples from screened healthy individuals: HPV-31 in 6 of 21 cases (29%), HPV-18 in 4 of 21 (19%), HPV-6 in 3 of 21 (14%), HPV-16 in 3 of 21 (14%), HPV-33 in 2 of 21 (10%), HPV-30 in 1 of 21 (5%), plus two isolates (CP8304 and IS324).

## Discussion

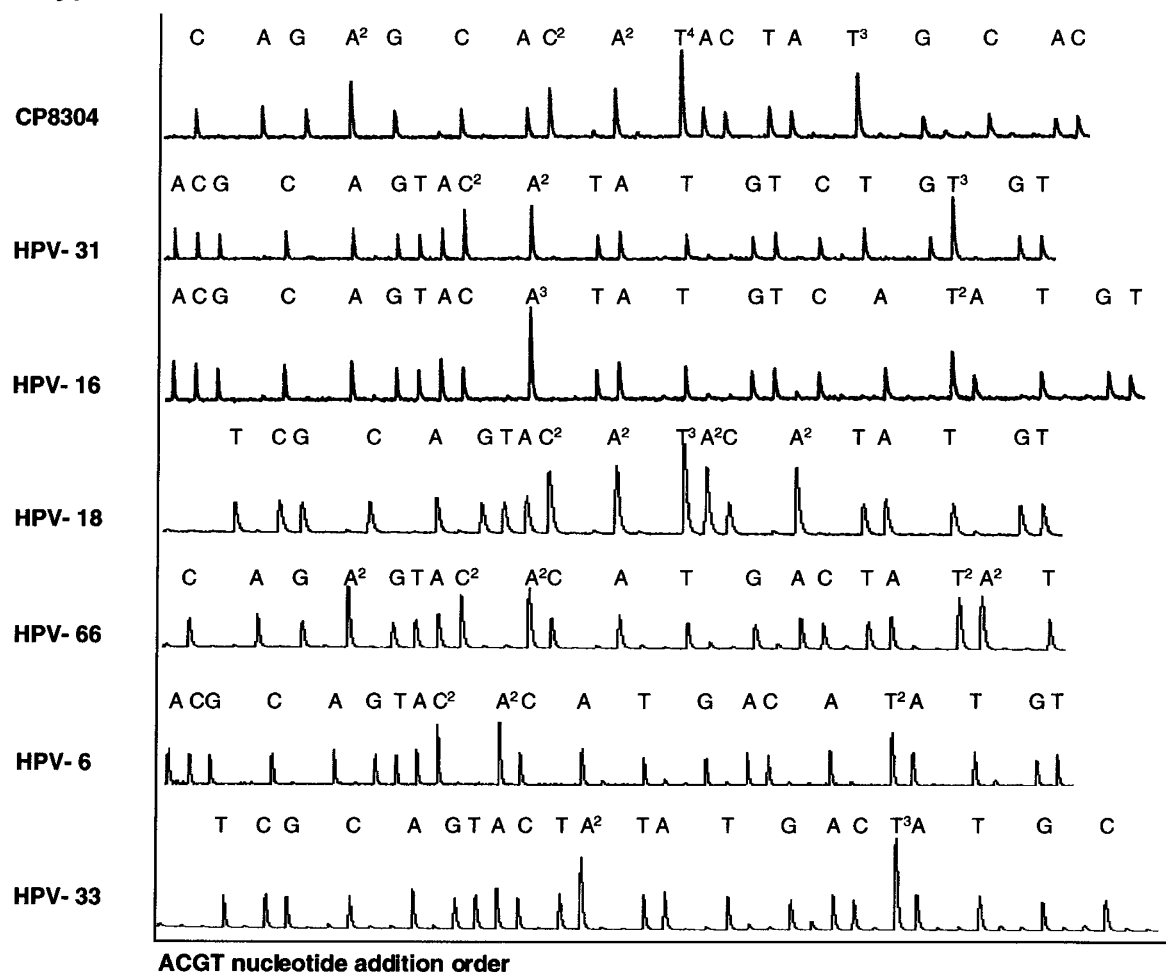
Genital HPV are commonly detected from clinical samples by consensus PCR methods (Gravitt et al, 2000). Two commonly used primer systems, the MY09/11 primers and the GP5+/6+ primers, amplify a broad spectrum of HPV genotypes (de Roda Husman et al, 1995; Resnick et al, 1990). A number of other consensus primers specific for the L1, E6, E6/E7, E7/E1, and E1 HPV regions have been described, each of which allows for the detection of a wide spectrum of HPV genotypes (Poljak et al, 1998). In our assay, the GP5+/6+ and MY09/11 primer sets were used for amplification of clinical specimens by PCR. However, depending on the purpose of the task, other consensus primer sets or type-specific primers could be applied for pyrosequencing.

**Table 1. Pyrosequencing HPV Genotyping Results Compared with Results from Conventional DNA Sequencing and Type-Specific PCR**

Number of samples	Number of bases needed for HPV genotyping by pyrosequencing	HPV typing by pyrosequencing	HPV typing by conventional DNA sequencing (ABI 310)	HPV typing by type-specific PCR
4	14	HPV-6	HPV-6	HPV-6
26	17	HPV-16	HPV-16	HPV-16
12	17	HPV-18	HPV-18	HPV-18
1	18	HPV-30	HPV-30	HPV-30
12	21	HPV-31	HPV-31	HPV-31
5	15	HPV-33	HPV-33	HPV-33
1	18	HPV-35	HPV-35	HPV-35
1	19	HPV-45	HPV-45	HPV-45
1	15	HPV-52	HPV-52	HPV-31 <sup>a</sup>
1	18	HPV-59	HPV-59	HPV-59
1	18	HPV-66	HPV-66	HPV-66
1	18	CP8304	CP8304	CP8304
1	38	IS324	IS324	IS324

HPV, human papillomavirus.

<sup>a</sup> The sample that was mistyped by type-specific PCR as HPV-31 was typed correctly as HPV-52 by pyrosequencing and conventional DNA sequencing.

**HPV type****Figure 3.**

Raw pyrogram sequence data from seven HPV amplicons amplified with GP5+/6+ primers.

Here we report, for the first time, the use of pyrosequencing, a novel DNA sequencing method, for detection and genotyping of HPV. The objective was to sequence and genotype HPV samples with pyrosequencing, and compare the results with type-specific PCR method and conventional gel-electrophoretic DNA sequencing.

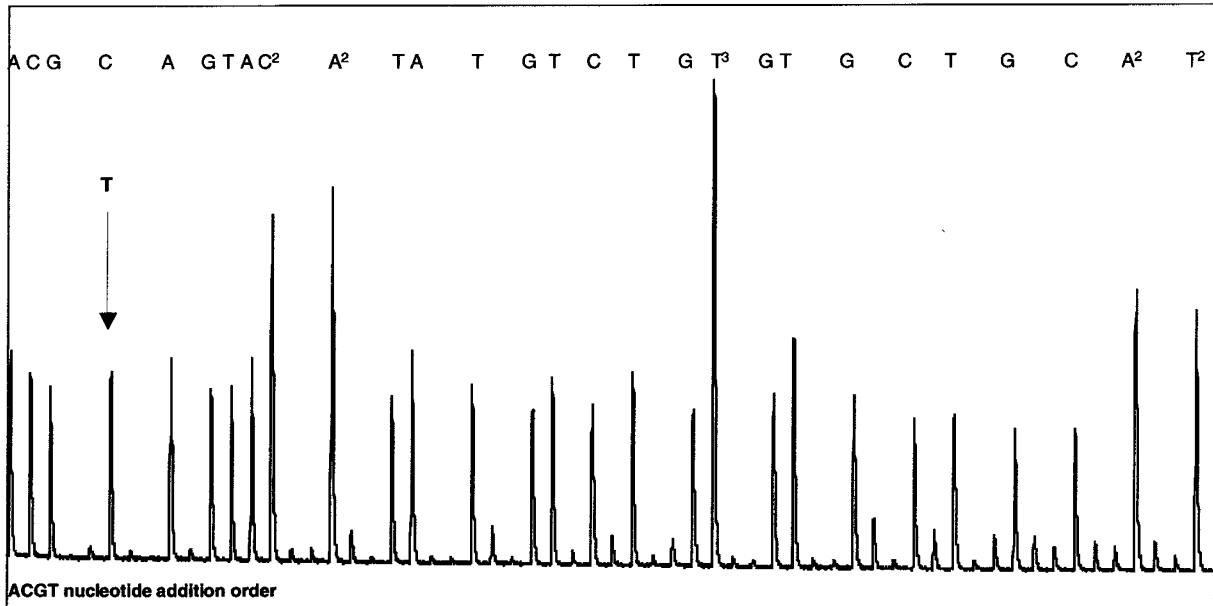
The sequences obtained by pyrosequencing and conventional DNA sequencing were in complete agreement, indicating the ability of the system to type HPV samples with high accuracy and throughput. Between 14 and 21 bases generally were required to genotype our samples, with the exception of one case, where 38 sequence bases were required (the sample was revealed at 18 bases as HPV). Longer sequencing (up to 50 bases) could be performed by pyrosequencing to cover all of the genotypes. Table 1 demonstrates the number of sequence bases needed for genotyping of each type that was sequenced by pyrosequencing.

The sample that had been earlier (mis)typed as HPV-31 by primer specific PCR was correctly genotyped as HPV-52 by both pyrosequencing and con-

ventional DNA sequencing. This indicates that pyrosequencing is a highly precise tool for HPV genotyping.

Pyrosequencing is suitable also for mutation detection in HPV sequence variants harboring mutation(s); eg, a single T to C substitution was detected in HPV-31 (Fig. 4).

Furthermore, we performed PCR on cell lysates to investigate the possibility of one-step amplification directly from cytobrush cell specimens. Multiple bands were visible after gel staining from most of the amplicons. Only 4 of 12 samples had a single band after gel staining, giving a low signal intensity in pyrosequencing. The sequence data obtained from the remaining amplicons suggests the existence of unspecific amplification along with a low amount of amplified DNA. Unspecific amplifications with the consensus primer set My09/11 has previously been described (Fernandez-Contreras et al, 2000). The cell lysate samples were amplified using nested PCR, which yielded remarkably high signal intensities, characteristic of high amounts and specific PCR products. Thus, nested PCR is recommended for direct amplifi-



**Figure 4.**

Pyrogram for HPV-31. A single base T to C substitution was detected by pyrosequencing.

cation from cell lysates or in cases of unspecific amplification from DNA extracts. To maintain the performance characteristics (the sensitivity and specificity of the system) in pyrosequencing, samples should be sequenced when a clear and specific product band is seen on ethidium bromide-stained agarose gels after PCR amplification.

Use of pyrosequencing as a tool for HPV genotyping offers a relatively simple technology that eliminates the misclassification in broad-spectrum HPV genotypes that can occur with nonsequencing methods (Poljak et al, 1998). Additionally, pyrosequencing was found to be a reliable technique to identify novel HPV types. In hybridization methods, these novel types either are not detected or are misclassified because of cross-hybridization with a closely related type (Vernon et al, 2000).

As with other available methods, multiple infections present in one specimen might be problematic to detect, depending on the proportional dominance and number of genotypes present in the amplicon. At present, pyrosequencing might not be particularly useful for identifying infection with more than one HPV genotype, because multiple infections give sequence signals from all of the available types in the specimen. Typing may be possible, provided one type is solidly dominant, with low background signal(s) from other existing genotype(s). However, this information alone may be insufficient. For example, it would be problematic to identify the presence of a low-risk type of HPV, but fail to identify the presence of a (subdominant) high-risk HPV in a mixed infection. A possible solution to this issue is to use high-risk HPV-specific sequencing primers for sequencing.

In conclusion, pyrosequencing is a rapid, reliable, and robust system for detection of HPV. It is appropriate for routine clinical screening with large numbers

of samples, and is easily adapted to laboratory automation. Pyrosequencing also enables detection of a broader spectrum of HPV, including putative novel types and mutations.

## Materials and Methods

### HPV Samples

Sixty-seven cervical samples were used in our test. Thirty-five samples were from cervical cancer patients, 21 from screened healthy individuals, and 11 from dysplastic individuals. Fifty-five of the samples were amplified from DNA extract (Hagmar et al, 1995) and the remaining 12 samples were amplified directly from cell lysates. All dysplasia and cancer samples were fresh-frozen biopsies, whereas the normal (screening) samples were cellular samples (cytobrush). Extraction was performed with phenol/chloroform followed by ethanol precipitation.

### HPV PCR

The DNA amplifications were performed in 50  $\mu$ l mixtures consisting of 5  $\mu$ l of DNA sample, 5  $\mu$ l of PCR buffer (Perkin-Elmer, Norwalk, Connecticut), 3.5 mM MgCl<sub>2</sub>, 0.2 mM dNTP, 25 pmol of GP5+/6+ (de Roda Husman et al, 1995; Snijders et al, 1990) primer set, and 1 U of AmpliTaq (Perkin Elmer). The GP6+ primer was biotinylated. The thermocycler temperature program consisted of denaturation at 94° C for 1 minute, annealing at 38° C for 1 minute, and extension at 71° C for 2 minutes during 40 cycles. Each PCR was initiated with a 4-minute denaturation step at 94° C and finished by a 4-minute extension step at 71° C. A Perkin-Elmer 9700 thermocycler was used for all amplifications.

### HPV Nested PCR

Cervical cells from 12 cytobrush samples were suspended in 1 ml of 10 mM Tris-HCl (pH 7.4). After one-time freezing and thawing, the samples were boiled for 10 minutes at 100° C. DNA from cervical cell lysates was amplified by two sets of general primers within the L1 open-reading frame. The MY09/11 consensus primer set was used in combination with the GP5+/6+ general primer set in a nested, two-step amplification. The DNA amplifications were performed in 50  $\mu$ l mixtures containing 5  $\mu$ l of prepared DNA sample from the cell lysate, 5  $\mu$ l of PCR buffer, 2 mM MgCl<sub>2</sub>, 0.2 mM dNTP, 2 pmol of primer MY09/11, and 1 U of AmpliTaq. The thermocycler temperature program consisted of denaturation at 95° C for 30 seconds, annealing at 45° C for 30 seconds, and extension at 72° C for 1 minute during 35 cycles. Each PCR was initiated by a 5-minute denaturation step at 95° C and finished by a 10-minute extension step at 72° C. Five microliters of amplified DNA was used as the template for the second PCR with the GP5+/6+ primer pair. The procedure for the second PCR was as described above for HPV PCR.

### Single-Strand Template Preparation for Pyrosequencing

Fifty microliters of biotinylated PCR product was immobilized onto 200  $\mu$ g of streptavidin-coated super paramagnetic beads (Dynabeads M-280-streptavidin, Dynal AS, Oslo, Norway) by incubation at 43° C for 30 minutes. Single-stranded DNA was obtained by incubating the immobilized PCR product in 5  $\mu$ l of 0.1 M NaOH for 4 minutes. The immobilized strand was suspended in 8  $\mu$ l of H<sub>2</sub>O plus 1  $\mu$ l of annealing buffer (100 mM Tris-acetate pH 7.75, 20 mM Mg-acetate). Single-stranded DNA corresponding to 50  $\mu$ l of PCR product was hybridized to 10 pmol of GP5+ sequencing primer at 70° C for 3 minutes, and incubated at room temperature for 5 minutes.

### Pyrosequencing

The primed PCR product was added to the pyrosequencing reaction mixture containing 0.1 M Tris-acetate pH 7.75, 0.05% Tween 20, 10 U of exonuclease deficient (exo<sup>-</sup>) Klenow DNA polymerase, 50 mU of apyrase (Sigma, St. Louis, Missouri), 0.8  $\mu$ g of purified luciferase (BioThema, Dalarö, Sweden), 15 mU of recombinant ATP sulfurylase (Karamohamed et al, 1999), 0.5  $\mu$ g of single-stranded DNA-binding protein (Amersham Pharmacia Biotech, Uppsala, Sweden), 0.5 mM EDTA, 5 mM Mg-acetate, 0.1% bovine serum albumin (BioThema), 1 mM dithiothreitol, 5  $\mu$ M adenosine 5'-phosphosulfate (Sigma), 0.4 mg/ml of polyvinylpyrrolidone (360 000), and 100  $\mu$ g/ml of D-luciferin (BioThema) in a total volume of 50  $\mu$ l. Pyrosequencing was performed at 28° C on an automated pyrosequencer PSQ 96 (Pyrosequencing AB, Uppsala, Sweden; www.pyrosequencing.com) at a dispensing pressure of 600 mbar with 8-msec open times and 65-second cycle times. The sequencing procedure was carried out by stepwise elongation of the primer strand

upon cyclic dispensation of the different deoxynucleoside triphosphates (Amersham Pharmacia Biotech). A CCD camera detected the light output resulting from nucleotide incorporation. The data were obtained in Microsoft Excel and graphic format.

### Conventional DNA Sequencing

The sequencing data obtained from pyrosequencing were confirmed by DNA sequencing on an ABI 310 (Perkin Elmer), using BigDye terminator chemistry, as described previously (Elfgren et al, 2000).

### Type-Specific PCR

The samples were genotyped by the type-specific PCR method. The consensus PCR procedures with My09/11 general primers and the type-specific PCR detection were previously described in detail (Hagmar et al, 1995; Skyldberg et al, 1991).

### Acknowledgement

The authors thank Afshin Ahmadian and Keng-Ling Wallin for critical reading of the paper.

### References

- Chan SY, Delius H, Halpern AL, and Bernard HU (1995). Analysis of genomic sequences of 95 papillomavirus types: Uniting typing, phylogeny, and taxonomy. *J Virol* 69:3074–3083.
- de Roda Husman AM, Walboomers JM, van den Brule AJ, Meijer CJ, and Snijders PJ (1995). The use of general primers GP5 and GP6 elongated at their 3' ends with adjacent highly conserved sequences improves human papillomavirus detection by PCR. *J Gen Virol* 76:1057–1062.
- Elfgren K, Kalantari M, Moberger B, Hagmar B, and Dillner J (2000). A population-based five-year follow-up study of cervical human papillomavirus infection. *Am J Obstet Gynecol* 183:561–567.
- Fernandez-Contreras ME, Sarria C, Nieto S, and Lazo PA (2000). Amplification of human genomic sequences by human papillomaviruses universal consensus primers. *J Virol Methods* 87:171–175.
- Godfroid E, Heinderyckx M, Mansy F, Fayt I, Noel JC, Thiry L, and Bollen A (1998). Detection and identification of human papilloma viral DNA, types 16, 18, and 33, by a combination of polymerase chain reaction and a colorimetric solid phase capture hybridisation assay. *J Virol Methods* 75:69–81.
- Gravitt PE, Peyton CL, Alessi TQ, Wheeler CM, Coutlée F, Hildesheim A, Schiffman MH, Scott DR, and Apple RJ (2000). Improved amplification of genital human papillomaviruses. *J Clin Microbiol* 38:357–361.
- Hagmar B, Kalantari M, Skyldberg B, Moberger B, Johansson B, Walaas L, and Warleby B (1995). Human papillomavirus in cell samples from Stockholm Gynecologic Health Screening. *Acta Cytol* 39:741–745.
- Karamohamed S, Nilsson J, Nourizad K, Ronaghi M, Pettersson B, and Nyrén P (1999). Production, purification, and luminometric analysis of recombinant *Saccharomyces cerevisiae* MET3 adenosine triphosphate sulfurylase expressed in *Escherichia coli*. *Protein Expr Purif* 15:381–388.

Poljak M, Seme K, and Gale N (1998). Detection of human papillomaviruses in tissue specimens. *Adv Anat Pathol* 5:216–234.

Resnick RM, Cornelissen MT, Wright DK, Eichinger GH, Fox HS, ter Schegget J, and Manos MM (1990). Detection and typing of human papillomavirus in archival cervical cancer specimens by DNA amplification with consensus primers. *J Natl Cancer Inst* 82:1477–1484.

Ronaghi M, Uhlén M, and Nyrén P (1998). A sequencing method based on real-time pyrophosphate. *Science* 281:363–365.

Schneider A (1993). Pathogenesis of genital HPV infection. *Genitourin Med* 69:165–173.

Skyldberg B, Kalantari M, Karki M, Johansson B, Hagmar B, and Walaas L (1991). Detection of human papillomavirus infection in tissue blocks by in situ hybridization as compared with a polymerase chain reaction procedure. *Hum Pathol* 22:578–582.

Snijders PJ, van den Brule AJ, Schrijnemakers HF, Snow G, Meijer CJ, and Walboomers JM (1990). The use of general primers in the polymerase chain reaction permits the detection of a broad spectrum of human papillomavirus genotypes. *J Gen Virol* 71:173–181.

Vernon SD, Unger ER, and Williams D (2000). Comparison of human papillomavirus detection and typing by cycle sequencing, line blotting, and hybrid capture. *J Clin Microbiol* 38:651–655.