Type-Specific Multiple Sequencing Primers

A Novel Strategy for Reliable and Rapid Genotyping of Human Papillomaviruses by Pyrosequencing Technology

Baback Gharizadeh,* Maria Oggionni,[†]
Biyang Zheng,[‡] Edit Akom,[§] Nader Pourmand,*
Afshin Ahmadian,[¶] Keng-Ling Wallin,^{‡||} and
Pål Nyrén[¶]

From the Stanford Genome Technology Center, Stanford University, Palo Alto, California; the Department of Pathology,[†] Unit of Experimental Molecular Pathology, Istituto Nazionale per lo Studio e la Cura dei Tumori, Milan, Italy; the Université du Québec à Montréal,[§] Montréal, Canada; the Department of Molecular Medicine,[‡] Center for Molecular Medicine, Karolinska Hospital, Stockholm, Sweden; the Department of Biotechnology,[¶] Royal Institute of Technology, Stockholm, Sweden; and the Department of Public Health and Clinical Medicine,^{||} Nutritional Research, Umeå University Hospital, Umeå, Sweden*

DNA sequencing is the gold standard method for accurate microbial and viral typing. However, DNA sequencing techniques have been facing limitations in typing of human papillomaviruses when the specimen harbors multiple genotypes and yields nonspecific amplification products, resulting in nonspecific and noninterpretable sequence data. To address these limitations we have developed a type-specific multiple sequencing primer DNA-sequencing method. This new strategy is suitable for sequencing and typing of samples harboring different genotypes (co-infections with multiple genotypes) and yielding nonspecific amplifications, thus eliminating the need for nested polymerase chain reaction (PCR), stringent PCR conditions, and cloning. The new approach has also proved useful for amplicons containing low PCR yield or subdominant types, avoiding reperforming of amplifications. We have applied the multiple sequencing primer method for genotyping of clinically relevant human papillomaviruses in a clinical test panel by using a combined pool of seven type-specific sequencing primers for HPV-6, -11, -16, -18, -31, -33, and -45. Furthermore, we introduced a sequence pattern recognition approach when there was a plurality of genotypes in the sample to facilitate typing of more than one target DNA in the sample. The multiple sequencing primer method has proved to be a multifaceted approach for typing of human papillomaviruses by DNA sequencing technologies. (*J Mol Diagn* 2005, 7:198–205)

DNA sequencing has revolutionized many scientific fields since its advent and the DNA sequence information has provided resourceful genetic information, which is beneficial for a large number of different sciences. DNA sequencing is the gold standard method for accurate typing of microorganisms and viruses, and is more reliable than the hybridization and serological-based techniques, which might be subjected to false hybridization and non-specific binding during discrimination of closely related species and/or types. In addition, DNA sequencing provides nucleic acid sequence information, which is versatilely informative and is the core of every organism. However, DNA sequencing faces limitations when the microbial amplification products contain more than one type and/or species or nonspecific amplification products (or both). Because sequence signals from all available types/species/nonspecific products are detected, this phenomenon makes microbial and viral typing a challenging task. Moreover, sequence signals representing a minority of the total amplification product may remain unnoticed and only the dominant type will be detected.

In a previous work, we have used pyrosequencing technology for detection of human papillomaviruses (HPVs) in a clinical study showing that genotyping could be achieved by sequencing 14 to 21 bases¹ but sequencing of co-infections and nonspecific amplification were a challenge. In a following study, we introduced the basic principle of the multiple sequencing primer method in DNA-sequencing technologies on HPV whole DNA plasmids.² The technique is based on using a number of type- or species-specific sequencing oligonucleotide primers in a combined pool for detection of microorganisms and viruses. The work presented here is an extensive study of this technique on a clinical test panel for detection of HPV. Nevertheless, HPV is representing a model system here, while the method is in fact a general

Supported by grants from the Cancerfonden (Swedish Cancer Society), the Tore Nilsons Stiftelse för Medicinsk Forskning, and the Swedish Research Council.

M.O. and B.Z. contributed equally to this study.

Accepted for publication October 22, 2004.

Address reprint requests to Professor Pål Nyrén, Department of Biotechnology, Royal Institute of Technology (KTH), Roslagstullbacken 21, SE-106 91 Stockholm, Sweden. E-mail: paaln@biotech.kth.se.

approach and is applicable to many other microorganisms, be it bacteria, viruses, or fungi.

HPVs represent a large family of more than 100 identified and sequenced genotypes, classified based on their DNA sequence homology, including some 30 types that usually infect the genital region.³ These HPVs can be categorized as low-risk, intermediate-risk, and high-risk types. Although the first category causes productive infection, leading to development of genital condylomata (warts), the others cause transformative infection that, in certain conditions, can lead to carcinogenesis. The main cancer associated with HPVs is cervical cancer,^{4,5} the second most frequent cause of cancer fatality in women worldwide,⁶ other HPV-related malignancies are genital tract cancers (vulvar, vaginal, and penile cancer), anal cancer, and a subset of head and neck cancer, especially oropharyngeal carcinoma of the tonsillar region.⁷ In fact high-risk HPVs have been shown to be present in up to 99.7% of cervical cancer specimens worldwide.^{8,9} More specifically, HPV-16 and HPV-18, and to lesser extent, HPV-31, -33, -35, -39, -45, -51, -52, -56, -58, -66, and -68 play a central role in the development of cervical intraepithelial neoplasia and, subsequently, cervical carcinoma.^{8,10,11} The HPV types most frequently associated with genital warts are HPV-6 and HPV-11.^{8,12} Thus, specific HPV detection and typing is important for epidemiological studies, assessment of the clinical behavior of particular genotypes, clinical management of women, clinical follow-up studies, evaluating prevention strategies, and development of new therapies and of prophylactic and/or therapeutic vaccines. HPV co-infections with multiple genotypes/variants are a common phenomenon and they range to from a few percent up to 50% depending on the group studied,⁵ higher prevalence being found mostly in pathological specimens and less in healthy carriers. Consensus and degenerative primer sets, such as GP5+/6+ and MY09/11, are used for HPV amplification and may result in amplification of non-specific products from genomic DNA for different reasons.^{1,13}

Here we report in detail the use of a multiple sequencing primer method, which can be of great benefit for DNA-sequencing technologies for microbial and viral typing. We have used a seven-primer pool for genotyping of the most prevalent HPV types (HPV-6, -11, -16, -18, -31, -33, and -45). The new method is especially suited for detecting and typing of different DNA targets harbored in the same sample, such as HPV co-infections with multiple genotypes/variants or nonspecific amplifications as well as specific typing of the relevant types regardless of presence of other types/species. The method is also suitable for specific genotyping of microbial and viral DNA in samples with low polymerase chain reaction (PCR) yield and in samples containing subdominant types/species. This new strategy is flexible, eliminates the need for nested PCR, stringent PCR conditions, and cloning, and, in addition, increases the sequence quality. We also introduce efficient sequence pattern recognition in pyrosequencing technology, when there is a multitude of types/species in the sample, which facilitates typing of more than one target DNA in the sample.

Materials and Methods

HPV DNA Plasmids, Cell Lines, and Clinical Samples

HPV whole genomic plasmids for types 6, 11, 16, 18, 33, 40, 45, 72, and 73 (kindly donated by Dr. E.-M. de Villiers of the Deutsches Krebsforschungszentrum, Heidelberg, Germany) were used for evaluation of the multiple sequencing primer method. HeLa, Caski, and SiHa cell lines were tested in this study as positive HPV controls. Eighty-two cervical and oropharyngeal samples were selected for this study. Sixty-five cervical samples were from Sweden and 17 oropharyngeal samples were from Italy. Forty of these cervical samples were biopsies (collected between 1969 to 1995) from Umeå University Hospital, 5 were biopsies (collected between 1997 to 2000) from Huddinge University Hospital, and 20 samples were cervical scrapes (collected between 1999 to 2000) from Karolinska Hospital in Sweden. The 17 Italian samples were surgical specimens (collected between 1990 to 1997) from the National Cancer Institute of Milan in Italy. All biopsy samples were paraffin-embedded material except for the cervical scrapes.

DNA Preparation

The cervical scrape samples were collected using Cyto-brush (Medscan Medical AB, Malmö, Sweden) and the cells were collected in 1 ml of sterile phosphate-buffered saline (pH 8.0). DNA was extracted from the entire cervical sample and the extraction was performed using the QIAmp system (Qiagen Inc., Valencia, CA) according to the instructions described in the kit. The final volume of DNA obtained was 200 μ l and the concentration was ~50 to 270 μ g/ml. The buffered formalin-fixed, paraffin-embedded samples (40 cervical and 17 oropharyngeal) were prepared according to the following procedure. Methylene blue-stained sections from buffered formalin-fixed, paraffin-embedded tissues of oropharyngeal samples were subjected to a careful microdissection under the microscope to obtain malignant tissues. Genomic DNA extraction was performed as previously described.¹⁴ For cervical cancer biopsies, DNA was extracted from paraffin sections taken from each cervical biopsy. The paraffin sections were treated as previously described.¹⁵ The DNA was dissolved in 200 μ l of Tris-ethylenediamine tetraacetic acid buffer, pH 7.4.

Consensus PCR Conditions

HPV one-step PCR using GP5+/GP6+ and MY09/11 consensus primer sets were performed for clinical samples, cell line controls, and HPV whole genomic plasmid DNA as previously described.^{1,16} Nested PCR, using both sets of consensus HPV primers MY09/MY11 (outer primers) and GP5+/GP6+ (inner primers), was performed in two tubes on four samples that were not amplified by one-step PCR reactions.¹

Design of Multiple Sequencing Primers

For specific oligonucleotide design and with consideration for cross/false-hybridizations, the DNA sequence of the L1 region of HPV-6, -11, -16, -18, -26, -31, -32, -33, -34, -35, -39, -40, -42, -43, -44, -45, -51, -52, -53, -54, -56, -58, -59, -61, -66, -67, -68, -69, -70, -71 (CP8061), -72 (LVX100), -73 (MM9), -81 (CP8304), -82 (MM4), -83 (MM7), -84 (MM8), CP6108, and IS324 genotypes were aligned. These genotypes are all amplified by HPV consensus/general primer sets. Seven specific oligonucleotide primers, showing complete matches for clinically high-risk HPV-16, -18, -31, -33, and -45 and low-risk HPV-6 and -11, respectively, were designed. The designed sequencing primers had no sequence similarity with other HPV genotypes and a database search (BLAST from National Center for Biotechnology Information) showed complete matches with corresponding HPV types. The sequencing oligonucleotide primers HPV-16 (5'-GCTGCCATATCTACTTCAGA), HPV-18 (5'-GCTTCTACACAGTCTCCTGT), HPV-31 (5'-GTGCTGCAATTGCAACAGT), HPV-33 (5'-ACACAAGTAACTAGTGACAG), HPV-45 (5'-TATGTGCCTCTACACAAAAT), HPV-6 (5'-GTGCATCCGTAACATCATCTT), and HPV-11 (5'-GTGCACTCTGTGTCTAAATCTG) were synthesized and HPLC-purified by Thermo Hybaid (Germany).

Robotic Single-Strand Separation

Single-stranded DNA sample preparation was performed automatically on a Magnatrix 1200 robot (Magnetic Biosolutions, Stockholm, Sweden; www.magbio.com). In brief, 45 μ l of biotinylated PCR product was immobilized onto 200 μ 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 μ l of 0.1 mol/L NaOH for 5 minutes. The immobilized strand was suspended in 40 μ l of annealing buffer (10 mmol/L Tris-acetate, pH 7.75, 5 mmol/L Mg-acetate). Single-stranded DNA, corresponding to 45 μ l of PCR product, was hybridized to 10 pmol of general sequencing primer (GP5+) or specific seven-sequencing primer pool (10 pmol of each primer) at 80°C for 4 minutes followed by incubation at room temperature for 5 minutes. The excess unbound sequencing primers were removed by an extra wash.

Pyrosequencing Technology and Dideoxy DNA Sequencing

Sequencing by pyrosequencing technology was performed on an automated plate-based bench-top PSQ96 system (Pyrosequencing AB, Uppsala, Sweden), at a dispensing pressure of 650 mbar with 8 msec open time and 65 seconds cycle time. The sequence results were obtained in pyrogram formats. The seven-sequencing primer pool was used as sequencing primers in MY09/11 HPV plasmid amplicons for DNA sequencing on an ABI 3700 DNA analyzer (Applied Biosystems, Foster City,

CA) using the BigDye terminator chemistry according to the manufacturer's instructions.

Principle of the Multiple Sequencing Primer Method

The principle of the method is illustrated in Figure 1. The target-specific multiple sequencing primer strategy is a versatile molecular tool for DNA sequencing. It can be used for detection of specific types/species regardless of presence of co-infections or nonspecific amplification products, eliminating the need for nested PCR, cloning, or reperforming unsuccessful PCRs. When the PCR product contains HPV co-infections with multiple genotypes or/and nonspecific amplification products (Figure 1, a and c), the general primer will give rise to sequence signals from all of the present HPV genotypes and nonspecific amplification products. By using a set of HPV-specific multiple sequencing primers, we can avoid these problems and important relevant genotypes can be easily sequenced and identified (Figure 1, b and d). For co-infection of relevant HPV genotypes, pattern recognition is used for genotyping.

Primer-Dimer Evaluation and Use of Single-Stranded DNA-Binding Protein (SSB)

To test if the designed primers form primer dimers, GP5+ and the seven-primer pool were each added to pyrosequencing sequencing mixture in separate reactions in the absence of amplified DNA. Both GP5+ and the seven-primer pool resulted in sequence signals in the absence of DNA template (data not shown). The same experiments were performed in the presence of 2.5 μ g of SSB¹⁷ (Amersham Biosciences, Uppsala, Sweden) and no sequence signals were observed for GP5+ and the seven-primer pool (data not shown). To avoid primer-dimer (cross-hybridization of primers to one another, which results in extension by DNA polymerase in presence of nucleotides) during the course of DNA sequencing, an extra wash was performed in the template preparation process after the primer-annealing step to remove unbound and excess primers.

Results

DNA Sequencing by Specific Multiple Sequencing Primers on PCR Products from HPV Plasmids

All of the sequencing experiments on GP5+/6+ and MY09/11 amplification products from HPV DNA plasmids as well as the simulated co-infections were sequenced both in presence and absence of SSB (with the extra wash performed before the pyrosequencing reaction). The sequence signals were strongly improved by applying SSB. Figure 2 shows two HPV amplicons sequenced both in the presence and in the absence of SSB; the

Figure 1. Illustrative demonstration of sequencing of amplicons containing HPV co-infections with multiple genotypes and nonspecific amplification products. **a:** HPV sample containing nonspecific amplification products and sequenced by the GP5+ sequencing (general) primer, resulting in sequence signals from the HPV and nonspecific products. **b:** Sequencing of the amplified DNA by multiple sequencing primers. Only HPV-16 will result in specific sequence signals. **c:** HPV sample containing co-infections of HPV-16 and HPV-18, and nonspecific amplification products, resulting in sequence signals from all present genotypes and the nonspecific amplification product by using the general primer. **d:** Sequencing of the same co-infected sample as described above by multiple sequencing primers. HPV-16 and HPV-18 both result in sequence signals and are genotyped easily by pattern recognition.

Figure 2. Effects of SSB on improvement of sequence signal quality. **a:** Pyrograms from sequencing of HPV-72 by GP5+ both in the absence and in the presence of SSB. **b:** Sequencing of HPV-31 in a simulated co-infection of HPV-31/40/73 sequenced by the seven-primer pool both in the absence and in the presence of SSB.

sequence results for HPV-72 sequenced by the GP5+ primer and HPV-31 (simulated amplicon mixture of HPV-31/40/73) sequenced by the seven-primer pool were remarkably improved by the presence of SSB as shown. Consequently, SSB was used in all further pyrosequencing reactions.

The HPV primer pool for HPV-6, -11, -16, -18, -31, -33, and -45 was evaluated for sequencing of amplicons derived from HPV plasmids by the pyrosequencing method (a known clinical sample was used for HPV-31). The PCR products were amplified by MY09/11 and GP5+/6+ primer sets separately and sequenced by the general primer GP5+ and the HPV seven-primer pool in separate sequence reactions. The sequence data obtained were in concordance with expected sequence results (data not shown). Table 1 shows the HPV sequence information for specific genotyping (based on GenBank search) of HPV-6, -11, -16, -18, -31, -33, and -45 by the seven-primer pool. The HPV seven-primer pool was also used for sequencing of simulated mixed triple infections (mixtures) by mixing one relevant (HPV-6, -11, -16, -18, -31, -33, or -45) type and two low-risk genotypes (HPV-40,

Table 1. Type-Specific HPV Sequence Information for Each Sequencing Primer

HPV types	Number of bases needed for genotyping on GenBank	HPV sequence downstream of sequencing primer pool
HPV-6	22	CCACATACACCAATTCTGATTA
HPV-11	22	CTACATACACTAATTCAGATTA
HPV-16	18	AACTACATATAAAAAATAC
HPV-18	18	ACCTGGGCAATATGATGC
HPV-31	23	GATACTACATTTAAAAGTAGTAA
HPV-33	20	TACATATAAAAAATGAAAATT
HPV-45	17	CCTGTGCCAAGTACATA

-72, or -73) for GP5+/6+- and MY09/11-derived fragments. High-quality results were obtained for each mixture containing one of the seven relevant HPV types and two low-risk types (data not shown). Mixed MY09/11-derived amplicons, presenting simulated triple infections (each relevant type mixed by two low-risk types), were also sequenced by Sanger dideoxy sequencing, applying the seven-sequencing primer pool as sequencing primer. Accurate results were obtained (data not shown).

Clinical Specimens Sequenced by General Sequencing Primer and Seven-Primer Pool

A total number of 82 samples were sequenced by pyrosequencing technology. The clinical specimens were amplified separately by MY09/11 and GP5+/6+ primer sets in one-step PCR. Four cervical samples were not amplified in the one-step PCR but were amplified by nested PCR. Amplification products from clinical specimens were sequenced in the same sequence run while using different sequencing primers. The MY09/11-derived fragments (450 bp) were sequenced by MY11 and the seven-primer pool in two separate reactions. The GP5+/6+-derived fragments (150 bp) were sequenced by GP5+ and the seven-primer pool separately as well.

The sequencing results for MY09/11-amplified fragments that were sequenced by MY11 primer (general primer) and the seven-primer pool showed different sequencing quality. The MY11 primer sequence results showed that 72 of 78 samples (four were not amplified by the MY09/11 primer set) contained nonspecific amplification products, giving rise to nonspecific sequence signals. These nonspecific amplification products in clinical specimens amplified by MY09/11 were observed in ethidium bromide-stained agarose gel (Figure 3). The same amplicons, containing nonspecific amplification products or co-infections, were sequenced and genotyped correctly by the seven-primer pool (data not shown).

The sequence results for GP5+/6+-derived fragments sequenced by the GP5+ primer (general primer) and the seven-primer pool also showed different sequence quality. The GP5+ primer sequence results indicated that 22 of the 82 amplicons could not be typed properly due to either nonspecific amplification products or HPV co-infections or both (data not shown). On the other hand, the amplicons were sequenced and genotyped correctly by the multiple sequencing primer pool. A total number of 10

double infections were found, all in the cervical samples. All of the double-infected samples contained HPV-16 and HPV-18, as previously confirmed by type-specific PCR in an earlier study.¹⁵ The double infections were amplified by both the GP+/GP6+ and MY09/11 protocols and genotyped by the multiple sequencing primer method.

Sequence Pattern Recognition Approach for Genotyping of Co-Infections

The multiple sequencing primer approach could specifically genotype all of the samples tested. The overall sequence quality and sequence signal performance was enhanced compared to using the general/consensus primer as sequencing primer. Samples with co-infections and nonspecific amplification products were easily genotyped as well as samples containing relevant subdominant types or samples with low specific DNA yield.

Figure 4 (also intended as a comparison tool with Figure 5) demonstrates the pyrograms of HPV-16 and HPV-18, sequenced by the seven-primer pool. Figure 5 shows three cervical cancer samples of our test panel harboring co-infections with HPV-16 and HPV-18. The HPV co-infections with multiple genotypes were genotyped by pattern recognition. By pattern recognition is meant comparison-by-alignment of at least two sequence pattern results and determining the characteristic sequence for each type present in the sample in the se-

Figure 3. Ethidium bromide agarose-stained gel of HPV amplicons amplified by MY09/11 primer set resulting in nonspecific amplification products. The specific HPV band is at 450 bp. All these samples were easily genotyped by the multiple sequencing primer strategy.

Figure 4. Pyrograms of HPV-16 and HPV-18 amplicons sequenced by the seven-primer pool as reference for co-infection comparisons in Figure 5.

quence pattern combination. Thus, seven sequence patterns are expected because there are seven sequencing primers and one characteristic pattern is reserved for each genotype. For example, HPV-18 is characterized by GGG after the seventh nucleotide addition and HPV-16 is characterized by the sequence peaks for A, C, and A, after the fifth, sixth, and ninth nucleotide addition order, respectively. Figure 5 demonstrates clearly these characteristics in HPV-16 and HPV-18 and at the same time the dominance of each genotype could easily be noticed in the PCR products. The common/shared and specific bases for each type are noted on top of each peak, characterizing each type, which facilitates genotyping of each DNA target. The dominant type could be easily

Multiple infection of HPV-16 and HPV-18 in 3 unrelated clinical samples sequenced by the multiple sequencing primer method and genotyped by sequence-pattern-recognition

Figure 5. Pyrograms of co-infections of HPV-16 and HPV-18 in three different clinical samples sequenced by the seven multiple sequencing primer pool and genotyped by pattern recognition. The samples contained double infections of HPV-16 and HPV-18 and were genotyped by comparing characteristic sequence signals specific for each type. The common and specific bases for each type are noted on **top** of each peak, characterizing each type, facilitating genotyping. Sequence peaks on positions 1, 2, 5, 6, and 9 represent HPV-16 and sequence peaks on positions 1, 2, 7, and 10 represent HPV-18. The **solid** and **dotted arrows** are for demonstration of the dominance of HPV-16 and HPV-18. The dominant type could be easily observed by comparison of single bases shown by **arrows**. **a:** HPV-16 and HPV-18 almost equal in dominance; **b:** HPV-18-dominant; and **c:** HPV-16-dominant.

A Challenging template sequenced by multiple sequencing primers

Figure 6. Genotyping of a challenging clinical sample containing dominant co-infections or nonspecific amplification products by using general primer resulting in nonspecific sequence signals (**a**) and resulting in specific sequencing signals (genotyping accurately HPV-33) (**b**). **c:** Pyrogram of HPV-33 from plasmid as a reference for comparison with **b**.

observed by comparison of single bases shown by arrows. Figure 5a shows HPV-16 and HPV-18 almost in equal concentration while HPV-18 is dominant in Figure 5b and HPV-16 is dominant in Figure 5c.

Figure 6a shows a challenging clinical sample containing the clinically important HPV-33, which is in minority (low DNA concentration) and not detectable by the general primer, due to high sequence signals from either nonspecific amplification products or co-infections, but is genotyped by multiple sequencing primers (Figure 6b). Figure 6c shows the pyrogram of an HPV-33 amplicon as a reference for comparison with Figure 6b. The challenging template is genotyped correctly despite its extremely low concentration and presence of HPV co-infection/nonspecific amplification products.

Discussion

DNA sequencing is an accurate method for microbial detection because it reveals the nucleotide sequence of the species/types. However, DNA sequencing has held some drawbacks in microbial and viral typing when the sample has harbored nonspecific amplification products or co-infections. In the cases of co-infections, cloning has been performed before DNA sequencing, which is a laborious task. In the cases of nonspecific amplification products from genomic DNA, nested PCR has been performed, which requires an additional step with no promise for high-quality results. We have addressed these

issues by using a multiple sequencing primer strategy (Figure 1).

Seven genotypes were chosen for design of multiple sequencing primers according to their clinical relevance and prevalence. High-risk HPV-16, -18, -31, -33, and -45 were selected based on their oncogenic potential. These types together are found in more than 80% of the cervical cancers worldwide.^{8,11} In addition, low-risk HPV-6 and HPV-11 were chosen for their causative role in the pathogenesis of genital warts. The multiple sequencing primers designed for this method are applicable to both MY09/11- and GP5+/6+-derived amplicons.

Our results demonstrate that the target-specific multiple sequencing primer method was able to specifically sequence and identify HPV amplicons containing co-infections and nonspecific amplification products, which were amplified with consensus GP5+/6+ or degenerate primers MY09/11. Using consensus/degenerate primers may amplify other regions of the genome especially in one-step PCR. The multiple sequencing primer pool, however, eliminates the need for nested PCR, stringent PCR conditions, or cloning of PCR products before DNA sequencing.

One of the foremost advantages of this method is that primers can be designed for specific regions, with the shortest possible sequence-read. This is particularly advantageous for pyrosequencing technology because the order of nucleotide dispensation can be easily programmed, and more importantly, sequencing is performed directly on the first base on the DNA template downstream of the sequencing primer, making primer design more flexible. This also eliminates long reads and increases accuracy. In addition, specific and shorter reads will enhance the sequence quality and detection ability, and it will be time- and cost-effective. As for our assay, HPV-16 and HPV-18 could be easily detected with the first two nucleotide additions (A and C). If the sequence starts with AAC, it distinguishingly characterizes HPV-16, and sequence ACC characterizes HPV-18 (Figure 4), which are sufficient for genotyping and the sequencing time is minimized to a few minutes. For all of the seven HPV genotypes, three bases is sufficient to characterize each type (Table 1). Above all, these results are obtained by DNA sequencing, which is by far more reliable than the hybridization-based results.

Furthermore, the target-specific sequencing primer method could be easily tailored and adapted according to the desired applications or clinical settings based on the regional prevalence of microorganisms and viruses and other associated factors. As the cost for DNA sequencing is dropping, the same sample could be sequenced in parallel with two or three different target-specific primer pools covering a broader spectrum of species/types, or the primer pool of choice could be used together with the general primer. In the case of HPV, the general primer set GP5+/6+ amplifies up to 35 genotypes, and the GP5+ primer can sequence these different genotypes. By parallel sequencing reactions of the same clinical sample, using the general primer and the multiple sequencing primers, a clinical laboratory can easily deal with samples containing HPV co-infections

with multiple genotypes/variants, nonspecific amplification products, low-yield amplicons, and subdominant genotypes.

In cases with more than one type/species in a sample that naturally give rise to sequence signals, sequence pattern recognition could be used for differentiation of the genotypes. Pattern recognition is comparison-by-alignment of at least two sequence pattern signals from the same specimen in a pyrogram. By identifying the characteristic sequence for each type/species present, it is possible to differentiate between different species or genotypes. Pattern recognition also reveals the dominance or subdominance of species/types in multiple infections in the amplified product. Figure 5 demonstrates these characteristics in HPV-16 and HPV-18 and the dominance of each HPV type in the amplification products. In our test panel, we had only co-infections of HPV-16 and HPV-18. Pattern recognition could be simplified by a clever choice of primers and selection of proper downstream upcoming sequence for each primer. Pattern differentiation software could be designed to differentiate type sequence patterns for relevant HPVs. In addition, this can also serve as a general principle for microbial and viral typing.

The multiple sequencing primer approach could be used for the Sanger dideoxy sequencing method, although when there is a plurality of relevant types/species in the specimen, pattern recognition is not applicable to electropherograms. However, in such cases, the sequence pattern of the electropherogram reveals co-infections, which is clinically important. Separate sequencing reactions by each primer are recommended for accurate genotyping. Pattern recognition is suitable for pyrosequencing technology because it is a nonelectrophoretic, bioluminometric method in which the sequence signals are detected in real time and there is the possibility of alignment of different sequence signals.

Interestingly, the multiple sequencing primer method has shown remarkable results for samples with low PCR yield or when the genotype is in minority in presence of sequence signals from nonspecific amplification products. Sequence pyrogram of each type will reveal easily the genotype as shown in Figure 6b. For pyrosequencing technology, the DNA concentration limit is 0.2 to 0.4 pmol of DNA for DNA sequencing with the Pyrosequencer PSQ96 system in the absence of background.

In brief, the genotyping results obtained by the new technique have been more than satisfactory. The multiple sequencing primer approach evidently enhances the discriminatory ability and accuracy of DNA sequencing technologies. Accurate results were achieved for amplicons derived from the widely used GP5+/6+ and MY09/11. Moreover, HPV co-infections with multiple genotypes were detected easily, which is of significant clinical importance, because a significant number of HPV carriers are co-infected with different HPV types. As mentioned earlier, HPV infection is the main cause of cervical cancer, but the risk associated with the various types has not been adequately assessed.¹¹ When we started working on this project, HPV-72 and HPV-73 were not being referred as clinically significant until Munoz and col-

leagues¹¹ in a recent broad epidemiological study classified HPV-73 as high-risk type and HPV-72 as significant low-risk type. This suggests strongly that accurate and specific genotyping is of major importance for epidemiological studies, for understanding the natural history and the clinical characteristics linked to each genotype, for prophylaxis measurements, follow-up patient studies, and monitoring of treatments.

This study demonstrates that the target-specific multiple sequencing primer approach could be used for different purposes such as: 1) specific detection and identification of amplicons containing co-infections; 2) target-specific detection of genotypes of interest in amplicons containing nonspecific amplification products; 3) specific detection of species/types/target DNA of interest with designed specific sequencing primers in amplicons amplified with consensus/degenerate primers or multiplex PCR; 4) specific detection of amplicon(s) containing a subdominant type, which could not be detected due to low sequence signals in presence of dominant sequence signals from nonspecific amplification products or multiple types/species; and 5) specific detection of amplicons with low PCR yield, which are considered as failed PCR and are reperfomed.

In conclusion, we have developed and evaluated an accurate and practical target-specific multiple sequencing primer approach improving the performance of DNA sequencing technologies in HPV genotyping; although HPV is a model system and the approach is applicable to other microorganisms and viruses and is suitable for large-scale clinical settings for microbial detection.

Acknowledgment

We thank Dr. Silvana Pilotti for her valuable comments on this work.

References

- Gharizadeh B, Kalantari M, Garcia CA, Johansson B, Nyren P: Typing of human papillomavirus by pyrosequencing. *Lab Invest* 2001, 81:673–679
- Gharizadeh B, Ghaderi M, Donnelly D, Amini B, Wallin KL, Nyren P: Multiple-primer DNA sequencing method. *Electrophoresis* 2003, 24:1145–1151
- zur Hausen H: Papillomaviruses causing cancer: evasion from host-cell control in early events in carcinogenesis. *J Natl Cancer Inst* 2000, 92:690–698
- Milde-Langosch K, Riethdorf S, Loning T: Association of human papillomavirus infection with carcinoma of the cervix uteri and its precursor lesions: theoretical and practical implications. *Virchows Arch* 2000, 437:227–233
- van Doorn LJ, Kleter B, Quint WG: Molecular detection and genotyping of human papillomavirus. *Expert Rev Mol Diagn* 2001, 1:394–402
- Parkin DM, Laara E, Muir CS: Estimates of the worldwide frequency of sixteen major cancers in 1980. *Int J Cancer* 1988, 41:184–197
- Jenkins D: Diagnosing human papillomaviruses: recent advances. *Curr Opin Infect Dis* 2001, 14:53–62
- Bosch FX, Manos MM, Munoz N, Sherman M, Jansen AM, Peto J, Schiffman MH, Moreno V, Kurman R, Shah KV: Prevalence of human papillomavirus in cervical cancer: a worldwide perspective. International Biological Study on Cervical Cancer (IBSCC) Study Group. *J Natl Cancer Inst* 1995, 87:796–802
- Walboomers JM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah KV, Snijders PJ, Peto J, Meijer CJ, Munoz N: Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol* 1999, 189:12–19
- Koutsky LA, Holmes KK, Critchlow CW, Stevens CE, Paavonen J, Beckmann AM, DeRouen TA, Galloway DA, Vernon D, Kiviat NB: A cohort study of the risk of cervical intraepithelial neoplasia grade 2 or 3 in relation to papillomavirus infection. *N Engl J Med* 1992, 327:1272–1278
- Munoz N, Bosch FX, de Sanjose S, Herrero R, Castellsague X, Shah KV, Snijders PJ, Meijer CJ: Epidemiologic classification of human papillomavirus types associated with cervical cancer. *N Engl J Med* 2003, 348:518–527
- Nobbenhuis MA, Walboomers JM, Helmerhorst TJ, Rozendaal L, Remmink AJ, Risse EK, van der Linden HC, Voorhorst FJ, Kenemans P, Meijer CJ: Relation of human papillomavirus status to cervical lesions and consequences for cervical-cancer screening: a prospective study. *Lancet* 1999, 354:20–25
- Fernandez-Contreras ME, Sarria C, Nieto S, Lazo PA: Amplification of human genomic sequences by human papillomaviruses universal consensus primers. *J Virol Methods* 2000, 87:171–175
- Hagmar B, Kalantari M, Skyldberg B, Moberger B, Johansson B, Walaas L, Warleby B: Human papillomavirus in cell samples from Stockholm gynecologic health screening. *Acta Cytol* 1995, 39:741–745
- Wallin KL, Wiklund F, Angstrom T, Bergman F, Stendahl U, Wadell G, Hallmans G, Dillner J: Type-specific persistence of human papillomavirus DNA before the development of invasive cervical cancer. *N Engl J Med* 1999, 341:1633–1638
- Gravitt PE, Peyton CL, Alessi TQ, Wheeler CM, Coutlee F, Hildesheim A, Schiffman MH, Scott DR, Apple RJ: Improved amplification of genital human papillomaviruses. *J Clin Microbiol* 2000, 38:357–361
- Ronaghi M: Improved performance of pyrosequencing using single-stranded DNA-binding protein. *Anal Biochem* 2000, 286:282–288