Available online at www.sciencedirect.com
ScienceDirect
Diagnostic Microbiology and Infectious Disease 59 (2007) 101 – 104
DIAGNOSTIC MICROBIOLOGY AND INFECTIOUS DISEASE
www.elsevier.com/locate/diagmicrobio

Synthetic genes: a tool for identifying human papillomavirus genotypes by hybridization and polymerase chain reaction-based assays

Karen Orjuela[a,1], Catalina Amador[a,1], Manuel A. Patarroyo[a,b,*]

[a]*Molecular Biology Department, Fundacion Instituto de Inmunologia de Colombia, Bogota, Colombia*
[b]*Universidad Nacional de Colombia, Bogota, Colombia*

Received 19 December 2006; accepted 11 April 2007

Abstract

Obtaining positive polymerase chain reaction (PCR) controls for human papillomavirus (HPV) diagnostic tests has been difficult because of prevalence variation in different geographic regions of each high-risk viral type. Overlapping oligonucleotides were designed for HPV-18, HPV-31, HPV-45, and HPV-58 type-specific (TS) sequences. Synthetic HPV viral genes were constructed by 2-step assembly PCR for accurately diagnosing TS HPV infection.

Keywords: Synthetic genes; Human papillomavirus; Cervical cancer; Genotyping

Human papillomavirus (HPV) has been linked to cancer development at mucosal sites (Bernard, 2005) and identified as a causal agent for cervical squamous neoplasia and adenocarcinoma (Walboomers et al., 1999; Bosch et al., 2002; Tjalma et al., 2005). Only high-risk HPV (HR-HPV) types play a leading role in pathogenesis of cervical cancer, which is the second most common cancer in women. New cases (471 000) are reported annually, 80% of which occur in developing countries, where 83% of total deaths due to cervical cancer are also occurring (Parkin, 2001).

Nearly 118 different HPV genotypes have been described (de Villiers et al., 2004). According to epidemiologic and phylogenetic classification, 15 HR-HPV types have been associated with the development of cervical cancer (16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68, 73, and 82) (Munoz et al., 2003). Each viral type's characteristic variation pattern has led to the selection and discrimination of their genomic regions (Molijn et al., 2005).

HR-HPV has variable worldwide prevalence, making it difficult to find some specific types in certain geographic regions. For example, HPV-16 is found twice as frequently in Asian, South American, and European cytologically normal women than other HR-HPV; however, HPV-42 is more common in sub-Saharan African women. Although HPV-18 worldwide prevalence is consistently high, it varies in some European and South American regions (Clifford et al., 2005). HPV-33 and 56 show high prevalence in Asia, likewise for HPV-58 in South America and HPV-31 in Europe. Globally, HPV-45 has 5.9% overall prevalence, with only 0.8% and 0.6% prevalence in Spain and Colombia, respectively (Munoz et al., 2003).

Polymerase chain reaction (PCR)-based methods and Hybrid Capture 2 (Digene, Gaithersburg, MD) have been validated in large-scale trials (Munoz et al., 2003) and epidemiologic studies for identifying HR-HPV DNA. PCR uses the general PCR primers GP5+/GP6+ (Jacobs et al., 1997), MY09/11 (Hildesheim et al., 1994), and SPF_{10} (Kleter et al., 1998), which contain a conserved L1 gene region in each HR-HPV genotype. HPV DNA identification tests do not provide specific information about viral type (Denny and Wright, 2005); type-specific (TS) PCR primers and oligohybridization methods (Munoz et al., 2003) are required for successful genotyping (Molijn et al., 2005).

Our group is currently studying HR-HPV prevalence in a female Colombian population. Cytology samples of these women underwent PCR HR-HPV genotyping using TS

Abbreviations: HPV, Human papillomavirus; HR-HPV, High-risk human papillomavirus; PCR, Polymerase chain reaction; TS, Type specific.

* Corresponding author. Tel.: +57-13244672x141; fax: +57-14815269.
E-mail address: mapatarr@fidic.org.co (M.A. Patarroyo).

[1] These authors contributed equally to this work.

0732-8893/$ – see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.diagmicrobio.2007.04.003

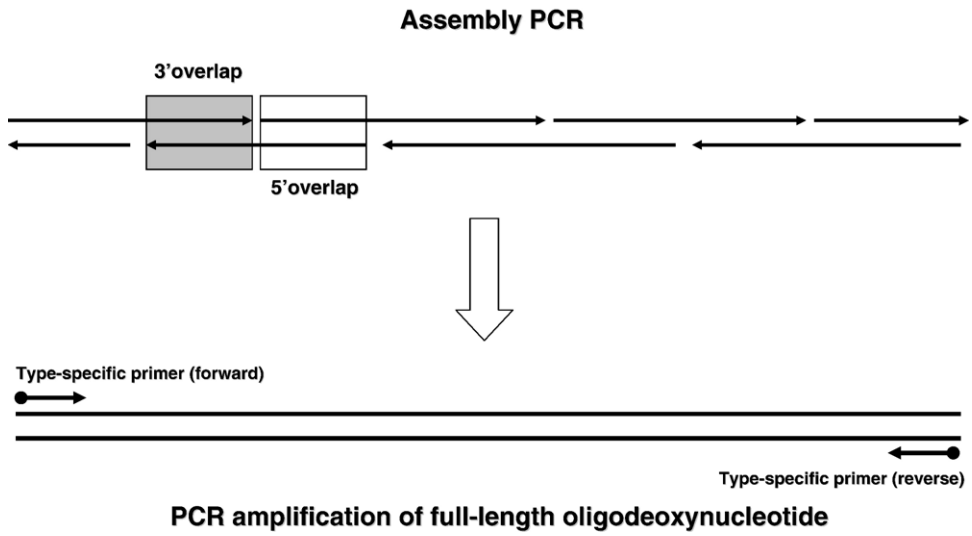

Fig. 1. Schematic representation of the synthetic gene assembly process. Between 8 and 12 oligonucleotides (arrows) were synthesized for the assembly PCR depending on the viral type (see Supplementary Table for sequences). Boxes represent hybridization units that can occur either at the 5′ or at the 3′ region of the interacting oligonucleotides in the assembly PCR step. TS primers are then used to amplify the entire assembled region.

primers. Because of low prevalence in Colombia, HPV-18, HPV-31, HPV-45, and HPV-58–positive controls needed for the assays were developed by gene synthesis.

Oligonucleotide ligation (Heyneker et al., 1976; Itakura et al., 1977; Goeddel et al., 1979), the "FokI method" (Mandecki and Bolling, 1988), DNA shuffling (Stemmer, 1994; Stemmer et al., 1995), and PCR (Dillon and Rosen, 1990; Prodromou and Pearl, 1992; Chen et al., 1994; Hayashi et al., 1994) have been proposed for in vitro gene synthesis. The recently described 2-step assembly PCR has also been very successful (Smith et al., 2003). The first step requires assembly PCR, where multiple oligodeoxynucleotides containing overlapping regions anneal and DNA polymerase elongates the primers to fill in discontinuous regions (Fig. 1). Variable length products result from annealing combinations involving less than the total number of oligodeoxynucleotides. The second step involves a pair of full-length oligodeoxynucleotide-specific primers; the full-length product is selectively amplified from the mixture (Young and Dong, 2004). This 2-step method produced the positive controls needed for the PCR HPV genotyping we performed.

TS primer-amplified sequences were synthesized to obtain positive typing controls. Table 1 describes their size and location in the viral genome. Gene2Oligo software was used for designing the oligonucleotides used in gene synthesis (http://berry.engin.umich.edu/gene2oligo/) (Rouillard et al., 2004), taking the following GenBank viral sequences as templates: HPV-18 (NC_001357), HPV-31 (J04353), HPV-45 (NC_001590), and HPV-58 (D90400). Oligonucleotides were synthesized by IDT DNA, Coralville, IA. Sequences of all oligonucleotides used are available in the Supplementary Table.

The first assembly PCR reaction was carried out in 50-μL final volume containing 200 μmol/L of each deoxyribonucleotide triphosphate (dNTP), 0.2 μmol/L of each oligonucleotide, 1 U Taq DNA polymerase (Promega, Madison, WI), and buffer containing 1.5 mmol/L $MgCl_2$. The following cycling temperatures were used: denaturing at 95 °C for 5 min, 30 cycles at 95 °C for 1 min, variable

Table 1
TS primers and amplified sequences

TS primer	Sequence	Amplified product (bp) and synthetic gene GenBank accessions	Gene	Reference
18 F	CCGAGCACGACAGGAACGACT	173 DQ863706	E6-E7	Karslen et al. (1996)
18 R	TCGTTTTCTTCCTCTGAGTCGCTT			
31 F	CTACAGTAAGCATTGTGCTAT	155 DQ863707	E5	Karslen et al. (1996)
31 R	ACGTAATGGAGAGGTTGCAATAACCC			
45 D	ACGGCAAGAAAGACTTCGCA	134 DQ863709	E6-E7	[a]
45 R	CACAACAGGTCAACAGGATC			
58 D	CGAGGATGAAATAGGCTTGG	109 DQ863705	E7	Walboomers et al. (1999)
58 R	ACACAAACGAACCGTGGTGC			

[a] This primer set was designed by us based on the HPV type 45 sequence reported in GenBank under accession number NC_001590.

annealing temperature for 1 min, 72 °C for 1 min, and final extension at 72 °C for 5 min. Each HPV sequence PCR annealing temperature was calculated using Gene2Oligo software for averaging temperatures (68 °C for HPV-18, 61 °C for HPV-31, 64 °C for HPV-45, and 66 °C for HPV-58).

The second reaction was carried out using HPV TS primers (Table 1) in a 25-μL reaction containing 2 μL of assembly PCR product, 100 μmol/L of each dNTP, 40 pmol of each primer, 1 U Taq DNA polymerase (Promega), and buffer containing 3 mmol/L $MgCl_2$. PCR amplification conditions for HPV-18, HPV-33, and HPV-58 have been previously described (Karlsen et al., 1996; Walboomers et al., 1999), whereas TS primers for HPV-45 were designed by us and amplified using the following conditions: denaturing at 95 °C for 5 min followed by 35 cycles at 95 °C for 1 min, 52 °C for 1 min, 72 °C for 1 min, and final extension at 72 °C for 5 min.

PCR mixtures were analyzed by agarose gel electrophoresis and their products excised from the gel by scalpel. Ligation of the products into the pGEM-T Easy vector was conducted according to the manufacturer's instructions (Promega). Amplification products were not purified from the solution because of their small size but rather ligated directly from the gel into the vector (Struhl, 1985) and transformed to *Escherichia coli* competent cells (dh5-α) where, if successfully transferred, they would grow as white colonies to later be analyzed.

Recombinant colonies were screened by α-complementation (Ausubel, 1987), and resulting white colonies were further screened by PCR using TS primers. Two positive clones from each gene were sequenced by MACROGEN, Seoul, Korea, and aligned with the originals to confirm correct gene assembly; 100% identity with template sequences was obtained for all synthetic constructs. Fig. 2 shows the amplification products for each viral type constructed using TS primers. Each TS primer pair was tested for specificity using the remaining viral synthetic genes as targets. No amplification bands were observed (data not shown).

Specific molecular tools will be required given substantial HPV genetic heterogeneity and specific subtypes' possible clinical relevance. Choosing a genomic region for typing viral isolates is important; it must show sufficient intertype variation discrimination for distinguishing a worldwide range of genotypes, whereas intratype variation must also be determined for reliable clinical follow-up and in vaccination studies (Molijn et al., 2005). Two-step genetic synthesis is a simple, reliable, low-cost technique for obtaining long DNA sequences. Creating synthetic DNA sequences can lead to the development of molecules that have diagnostic value when used as positive controls, specifically when determining HPV viral types. The gene synthesis technique applied here will lead to the creation of a repository of different HR-HPV viral controls available to other research groups. Currently, the controls described here can be obtained by contacting the authors directly and in the near future through ATCC.

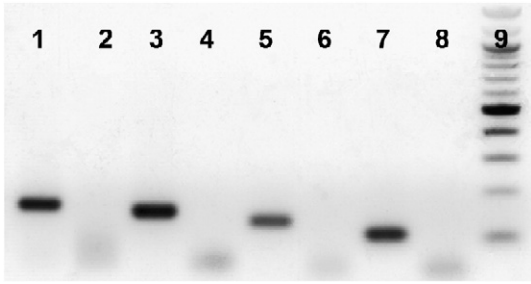

Fig. 2. PCR amplification of newly synthesized HPV DNA sequences. Lanes 1 and 2, HPV-18 sequence and its negative control. Lanes 3 and 4, HPV-31 and its negative control. Lanes 5 and 6, HPV-45 and its negative control. Lanes 7 and 8, HPV-58. Lane 9, 100-bp ladder.

Acknowledgments

We would like to thank the President of Colombia's office, Tricia Hadley, for patiently reviewing the manuscript, Sandra Milena Camargo for technical support, and Professor Manuel Elkin Patarroyo for invaluable comments. This work was supported by Colciencias (contract #RC-0060-2006).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.diagmicrobio.2007.04.003.

References

Ausubel FM (1987) Current protocols in molecular biology. Greene Publishing Associates; J. Wiley, order fulfillment, Brooklyn, NY, Media, PA.

Bernard HU (2005) The clinical importance of the nomenclature, evolution and taxonomy of human papillomaviruses. *J Clin Virol* 32(Suppl 1): S1–S6.

Bosch FX, Lorincz A, Munoz N, Meijer CJ, Shah KV (2002) The causal relation between human papillomavirus and cervical cancer. *J Clin Pathol* 55:244–265.

Clifford GM, Gallus S, Herrero R, Munoz N, Snijders PJ, Vaccarella S, Anh PT, Ferreccio C, Hieu NT, Matos E, Molano M, Rajkumar R, Ronco G, de Sanjose S, Shin HR, Sukvirach S, Thomas JO, Tunsakul S, Meijer CJ, Franceschi S (2005) Worldwide distribution of human papillomavirus types in cytologically normal women in the International Agency for Research on Cancer HPV prevalence surveys: a pooled analysis. *Lancet* 366:991–998.

Chen GQ, Choi I, Ramachandran B, Gouaux JE (1994) Total gene synthesis: novel single step and convergent strategies applied to the construction of a 779 base pair bacteriorhodopsin gene. *J Am Chem Soc* 11:8799–8800.

de Villiers EM, Fauquet C, Broker TR, Bernard HU, zur Hausen H (2004) Classification of papillomaviruses. *Virology* 324:17–27.

Denny LA, Wright Jr TC (2005) Human papillomavirus testing and screening. *Best Pract Res Clin Obstet Gynaecol* 19:501–515.

Dillon PJ, Rosen CA (1990) A rapid method for the construction of synthetic genes using the polymerase chain reaction. *Biotechniques* 9:298, 300.

Goeddel DV, Kleid DG, Bolivar F, Heyneker HL, Yansura DG, Crea R, Hirose T, Kraszewski A, Itakura K, Riggs AD (1979) Expression in *Escherichia coli* of chemically synthesized genes for human insulin. *Proc Natl Acad Sci U S A* 76:106–110.

Hayashi N, Welschof M, Zewe M, Braunagel M, Dubel S, Breitling F, Little M (1994) Simultaneous mutagenesis of antibody CDR regions by overlap extension and PCR. Biotechniques 17:310, 312, 314–315.

Heyneker HL, Shine J, Goodman HM, Boyer HW, Rosenberg J, Dickerson RE, Narang SA, Itakura K, Lin S, Riggs AD (1976) Synthetic lac operator DNA is functional in vivo. *Nature* 263:748–752.

Hildesheim A, Schiffman MH, Gravitt PE, Glass AG, Greer CE, Zhang T, Scott DR, Rush BB, Lawler P, Sherman ME, et al (1994) Persistence of type-specific human papillomavirus infection among cytologically normal women. *J Infect Dis* 169:235–240.

Itakura K, Hirose T, Crea R, Riggs AD, Heyneker HL, Bolivar F, Boyer HW (1977) Expression in *Escherichia coli* of a chemically synthesized gene for the hormone somatostatin. *Science* 198: 1056–1063.

Jacobs MV, Snijders PJ, van den Brule AJ, Helmerhorst TJ, Meijer CJ, Walboomers JM (1997) A general primer GP5+/GP6(+)-mediated PCR-enzyme immunoassay method for rapid detection of 14 high-risk and 6 low-risk human papillomavirus genotypes in cervical scrapings. *J Clin Microbiol* 35:791–795.

Karlsen F, Kalantari M, Jenkins A, Pettersen E, Kristensen G, Holm R, Johansson B, Hagmar B (1996) Use of multiple PCR primer sets for optimal detection of human papillomavirus. *J Clin Microbiol* 34:2095–2100.

Kleter B, van Doorn LJ, ter Schegget J, Schrauwen L, van Krimpen K, Burger M, ter Harmsel B, Quint W (1998) Novel short-fragment PCR assay for highly sensitive broad-spectrum detection of anogenital human papillomaviruses. *Am J Pathol* 153:1731–1739.

Mandecki W, Bolling TJ (1988) FokI method of gene synthesis. *Gene* 68:101–107.

Molijn A, Kleter B, Quint W, van Doorn LJ (2005) Molecular diagnosis of human papillomavirus (HPV) infections. *J Clin Virol* 32(Suppl 1): S43–S51.

Munoz N, Bosch FX, de Sanjose S, Herrero R, Castellsague X, Shah KV, Snijders PJ, Meijer CJ (2003) Epidemiologic classification of human papillomavirus types associated with cervical cancer. *N Engl J Med* 348:518–527.

Parkin DM (2001) Global cancer statistics in the year 2000. *Lancet Oncol* 2:533–543.

Prodromou C, Pearl LH (1992) Recursive PCR: a novel technique for total gene synthesis. *Protein Eng* 5:827–829.

Rouillard JM, Lee W, Truan G, Gao X, Zhou X, Gulari E (2004) Gene2Oligo: oligonucleotide design for in vitro gene synthesis. *Nucleic Acids Res* 32:W176–W180.

Smith HO, Hutchison III CA, Pfannkoch C, Venter JC (2003) Generating a synthetic genome by whole genome assembly: phiX174 bacteriophage from synthetic oligonucleotides. *Proc Natl Acad Sci U S A* 100: 15440–15445.

Stemmer WP (1994) DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. *Proc Natl Acad Sci U S A* 91:10747–10751.

Stemmer WP, Crameri A, Ha KD, Brennan TM, Heyneker HL (1995) Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides. *Gene* 164:49–53.

Struhl K (1985) A rapid method for creating recombinant DNA molecules. *Biotechniques* 3:452–453.

Tjalma WA, Van Waes TR, Van den Eeden LE, Bogers JJ (2005) Role of human papillomavirus in the carcinogenesis of squamous cell carcinoma and adenocarcinoma of the cervix. *Best Pract Res Clin Obstet Gynaecol* 19:469–483.

Walboomers JM, Jacobs MV, Manos MM, Bosch FX, Kummer JA, Shah KV, Snijders PJ, Peto J, Meijer CJ, Munoz N (1999) Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. *J Pathol* 189:12–19.

Young L, Dong Q (2004) Two-step total gene synthesis method. *Nucleic Acids Res* 32:e59.