© 1998 Oxford University Press Nucleic Acids Research, 1998, Vol. 26, No. 21 **5013–5015**

# Primer specific and mispair extension analysis (PSMEA) as a simple approach to fast genotyping

**Yu-Wen Hu*, Evan Balaskas, Garry Kessler, Claudia Issid, Linda J. Scully, Donald G. Murphy[1], Aline Rinfret[2], Antonio Giulivi, Vito Scalia and Peter Gill**

The Canadian Red Cross Society, Ottawa, Ontario K1G 4J5, Canada, [1]Laboratoire de Santé Publique du Québec, Ste-Anne-de-Bellevue, Quebéc H9X 3R5, Canada and [2]The Canadian Red Cross Society and CHUM Research Centre, Saint-Luc Campus, 264 east René-Lévesque, Montréal, Québec H2X 1P1, Canada

Received June 29, 1998; Revised and Accepted September 9, 1998

## ABSTRACT

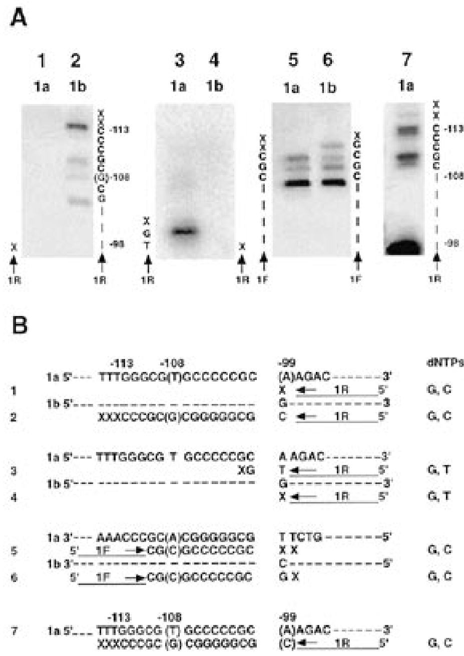
**A simple method, primer specific and mispair extension analysis (PSMEA) with *pfu* DNA polymerase was developed for genotyping. PSMEA is based on the unique properties of 3′→5′ exonuclease proofreading activity. In the presence of an incomplete set of dNTPs, *pfu* was found to be extremely discriminative in nucleotide incorporation and proofreading at the initiation step of DNA synthesis, completely preventing primer extension when mispair(s) are found adjacent to the 3′-end of the primer. This has allowed us to accurately detect nucleotide variations, deletions and insertions for fast genotyping.**

Mispair formations and extensions occur during DNA synthesis in a primer extension reaction with an incomplete set of dNTPs when using DNA polymerases (1,2). However, mispair extension frequencies are much higher with enzymes lacking 3′→5′ proofreading activity than with enzymes harboring this activity during the process of DNA synthesis. *pfu* (*Pyrococcus furiosus*) DNA polymerase exhibits 3′→5′ exonuclease activity, which peaks sharply at its standard polymerization temperature (3). We found that *pfu* was extremely discriminative in nucleotide incorporation and proofreading at the initiation step of DNA synthesis. The highly efficient 3′→5′ exonuclease proofreading of *pfu* could completely prevent primer extension when a mispair (or mispairs) is found at the initiation site of DNA synthesis. Taking advantage of this unique property of *pfu* 3′→5′ proofreading activity, nucleotide variations could be accurately identified by single primer specific and mispair extension analysis (PSMEA).

We used hepatitis C virus (HCV) genotyping as a test system for this new method. A single nucleotide variation site at position –99 in the 5′ untranslated region (5′ UR) of HCV is routinely used for differentiation between genotypes 1a and 1b with several current genotyping methods (4,5). Figure 1, lane 1 shows that, in the presence of dCTP and dGTP, primer 1R could not be extended on a genotype 1a template due to a mismatched pair involving an A nucleotide at the variation site immediately adjacent to the 3′-end of the primer on the template, leading to failure of primer extension. However, this primer extended on a genotype 1b template because a G instead of an A is found at the variation site (Fig. 1, lane 2). In contrast, primer 1R extended on genotype 1a (Fig. 1, lane 3), but not on 1b (Fig. 1, lane 4) when using dTTP and dGTP instead of dCTP and dGTP in the reaction. Thus, genotypes 1a and 1b were clearly differentiated by either one of the incomplete dNTP sets. However, *pfu* did not completely proofread as it did allow some single mispair formation and extensions at certain nucleotide positions that were several nucleotides away from the primer extension initiation site, for example, the mispair at position –108 on template 1b in the reactions with primer 1R or 1F (Fig. 1, lanes 2, 5 and 6). As with other DNA polymerases (1,2), any two or more consecutive mispairs located downstream of the primer could completely terminate primer extension by *pfu* (Fig. 1, lane 5). In addition, we found that two or more mispairs separated by one or two correct pairs also could terminate primer extension by *pfu* (Fig. 1, lane 6). Different lengths of extended primers represent the various termination points caused by the mispair(s), yielding genotype-specific band patterns in the DNA sequencing gel (Fig. 1, lanes 1–6). In contrast, a single mismatched pair at any nucleotide position, including the DNA synthesis initiation site, could be formed and extended in the presence of *Taq* (*Thermus aquaticus)* DNA polymerase, which lacks 3′→5′ exonuclease proofreading activity (6) (Fig. 1, lane 7). Thus, this enzyme was not suitable for PSMEA.

We found that at least two consecutive correct nucleotide pairings adjacent to the 3′-end of the primer were required to initiate primer extension with *pfu*. If only a single correct nucleotide pairing, followed by two or more mismatched pairs, existed at this position, the primer could not be extended with this single base by the enzyme. This provides a means to identify nucleotide deletions and insertions as well as multiple nucleotide variations using PSMEA. A unique CA insertion in the 5′ UR of HCV genotype 6a is used for discriminating it from the other genotypes (5). Figure 2, lanes 1 and 2 show that in the presence of either dCTP and dGTP, or dGTP and dTTP, primer 6R-1 could not be extended on genotype 1a because there was only a single nucleotide matched with one of the dNTPs used adjacent to the 3′-end of the primer. Furthermore, primer 6R1 was not extended

*To whom correspondence should be addressed at: The Canadian Blood Services, 1800 Alta Vista Drive, Ottawa, Ontario K1G 4JS, Canada. Tel: +1 613 739 2439; Fax: +1 613 739 2426; Email: yhn@redcross.ca



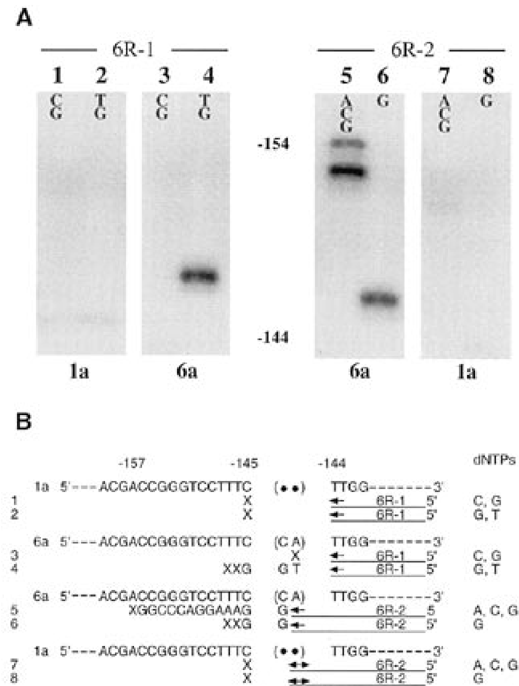




**Figure 1.** Characteristics of primer specific and mispair extension by *pfu* and *Taq* DNA polymerases. Primer extension reactions contained 20 ng of primer, 20–30 ng of PCR product purified with the QIAquick PCR Purification Kit (Qiagen), 20 µm of each of dNTPs, 10 µCi of each $^{32}$P-labelled dNTP used, 1.25 U of *pfu* DNA polymerase and 10 µl of 10× *pfu* reaction buffer (Stratagene). When 5′-end $^{32}$P-labelled primers were used, the $^{32}$P-labelled dNTPs were omitted, and 20 µm of each of non-labelled dNTPs were used in the reactions. Primer extensions were performed in a 100 µl reaction volume in a thermocycler (Perkin Elmer, GeneAmp 9600). Twenty cycles of 94°C for 20 s, 64°C for 20 s and 72°C for 35 s were performed. The primer extension products (1 µl) were mixed with 1 µl of the sequencing stop solution (Pharmacia Biotech) and electrophoresed on 8% polyacrylamide 8 M urea TBE gels for 40 min. (**A**) Primer extension products visualized by autoradiography. Primer 1R extension by *pfu* on template 1a and 1b in the presence of dCTP and dGTP (lanes 1 and 2) or dGTP and dTTP (lanes 3 and 4) with $^{32}$P-labelled dNTPs corresponding to the dNTPs used. Primer 1F extension by *pfu* on templates 1a and 1b (antisense strand) using dCTP and dGTP with corresponding $^{32}$P-labelled dNTPs (lanes 5 and 6). $^{32}$P-labelled primer 1R extension by *Taq* on template 1a showing that extension occurred in the presence of a mispair at the DNA initiation site (lane 7). (**B**) Sequences of templates, primers and extension products in reactions illustrated in (A). X and XX represent the sites of nucleotide mismatches that terminated primer extension. → represents a nucleotide at the 3′-end of the primer that is complementary to the opposite nucleotide in the template. A, C, G or T denote the position of the nucleotide when a mispair is produced. –113, –108 and –99 are the nucleotide positions in the 5′ UR of HCV. Underlined sequences indicate the primer binding site.

**Figure 2.** (**A**) Detection of nucleotide deletions and insertions by PSMEA illustrated with primer 6R-1 and 6R-2 extensions by *pfu* on templates 1a and 6a. Primer 6R-1 could not be extended with template 1a when using either dCTP plus dTTP, or dCTP plus dGTP as substrates due to a CA deletion (••), resulting in only a single matched pairing at position -145 adjacent to the 3′-end of the primer (lanes 1 and 2). Primer 6AR-1 could not be extended on template 6a when using dCTP and dGTP due to the mismatched pair at the position adjacent to the 3′-end of the primer (lane 3). However, the primer was extended by three bases with template 6a using dGTP and dTTP which matched the nucleotides in the CA insertion in template 6a (lane 4). Primer 6R-2 was extended on template 6a (lanes 5 and 6) but not on template 1a (lanes 7 and 8) using either dATP, dCTP and dGTP, or only dGTP. ↔ denotes the removal of the first nucleotide mismatched at the 3′-end of the primer. See Figure 1 for other symbols. (**B**) Nucleotide sequences of templates, primers and extension products in reactions illustrated in (A).

on template 6a in the presence of dCTP and dGTP due to a mispair existing adjacent to the 3′-end of the primer (Fig. 2, lane 3). However, this primer extended on template 6a when using dGTP and dTTP in the reaction (Fig. 2, lane 4) because three consecutive nucleotide pairings (i.e. A-T, C-G and C-G) were found adjacent to the 3′-end of the primer. Another primer, 6R-2, was designed to have its 3′-terminal nucleotide matched with the variation site (i.e. the A in the CA insertion of genotype 6a template). Thus, the primer extended 13 or 2 bases on template 6a, depending on the dNTPs used (Fig. 2, lanes 5 and 6), but not on template 1a due to the absence of CA (i.e. CA deletion) that resulted in a mismatched residue at the 3′-end of the primer being removed by the 3′→5′ exonuclease activity of *pfu* after the onset of DNA synthesis (Fig. 2, lanes 7 and 8). These data suggest that any small region with multiple point nucleotide variations, including insertions and deletions, can be identified by PSMEA with either manipulation of dNTP pools or the use of different genotype-specific primers.

Using PSMEA with five type-specific and subtype-specific primers, >200 HCV isolates have been genotyped, including 1a, 1b, 2a, 2b, 3a, 3b, 4, 5a and 6a. The genotyping results showed 100% concordance with those obtained with restriction fragment length polymorphism (RFLP) analysis or direct DNA sequencing.

PSMEA not only provides an accurate tool for genotyping, but also offers extraordinary sensitivity for the detection of mixed viral genotype infections. To compare the sensitivity of PSMEA to direct DNA sequencing, PCR products from HCV genotypes 1b and 2a isolates were mixed in different proportions to mimic mixed genotype infections. Figure 3A' shows that genotype 2a could be clearly identified by direct DNA sequencing only when it reached a proportion of 50% in the mix. Only some of the nucleotide variations could be recognized when the proportion of 2a molecules was ≤25%, but correct genotype identification was

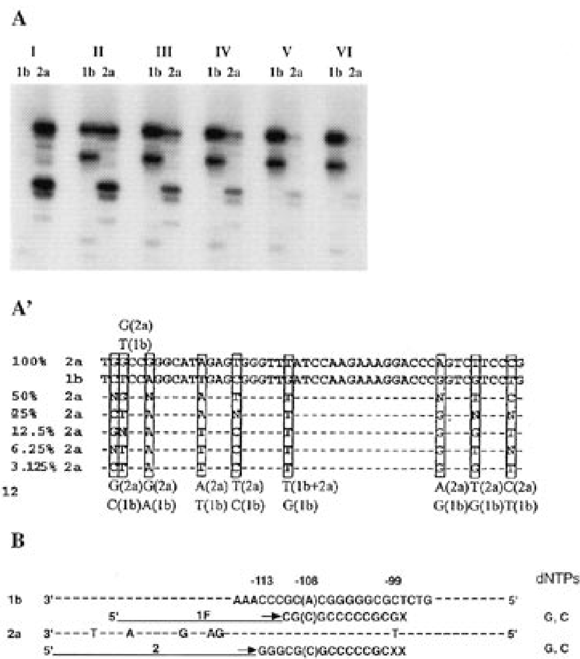



**Figure 3.** Comparison of the sensitivity between PSMEA and direct DNA sequencing for detection of mixed genotypes. Autoradiography results are presented in (**A**) and the data from a computer analysis of automated sequencing are presented in (**A'**). Different proportions of genotypes 1b and 2a (0–100%, I; 50–50%, II; 75–25%, III; 87.5–12.5%, IV; 93.75–6.25%, V; and 96.875–3.125%, IV) in the mix were analyzed with PSMEA (A) and direct DNA sequencing (A'). (**B**) Sequences of templates, primers and extension products in reactions illustrated in (A).

not possible. However, a proportion of genotype 2a as low as 3% in the mix was clearly detected by PSMEA (Fig. 3A), demonstrating an ~10-fold sensitivity improvement over direct DNA sequencing. Figure 4 shows a typical pattern of mixed infections with genotypes 1a and 2a identified in a sample by PSMEA. The result was confirmed by direct DNA sequencing.

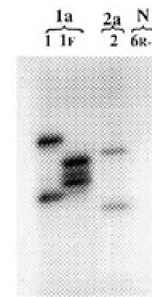



**Figure 4.** Two genotypes 1a and 2a were identified in a thalassaemia patient sample by PSMEA with: type-specific primer, 1; subtype-specific primer, 1F; and subtype-specific primer, 2. N represents the negative control with primer, 6R-1.

In conclusion, PSMEA is a simple and fast method for genotyping and for detecting low levels of mixed genotype viral infections.

## REFERENCES

1 Perinno,F., Bradley,W., Preston,D., Sandell,L.L. and Loeb,L.A. (1989) *Proc. Natl Acad. Sci. USA*, **86**, 8343–8347.
2 Reha-Krantz,L.J., Stocki,S., Nonay,R.L., Eimayuga,E., Goodrich,L.D., Konigsberg,W.H. and Spicer,E.K. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 2417- 2421.
3 Lundberg,K.S., Shoemaker,D.D., Adams,M.W.W., Short,J.M., Sorge,J.A. and Mathur,E.J. (1991) *Gene*, **108**, 1–6.
4 Simmonds,P. (1995) *Hepatology*, **21**, 570–582.
5 Stuyver,L., Rossau,R., Wyseur,A., Duhamel,M., Vanderborght,B., Heuverswyn,V.H. and Maertens,G. (1996) *J. Clin. Microbiol.*, **34**, 2259–2266.
6 Tindall,K.R. and Kunkel,T.A. (1988) *Biochemistry*, **27**, 6008–6013.