Acta Biochimica Polonica

Vol. 55 No. 4/2008, 815–819

on-line at: www.actabp.pl

Communication

SNP-minisequencing as an excellent tool for analysing degraded DNA recovered from archival tissues★

Katarzyna Bąbol-Pokora✉ and Jarosław Berent

Department of Forensic Medicine, Medical University of Lodz, Łódź, Poland

Received: 27 April, 2008; revised: 08 July, 2008; accepted: 08 October, 2008
available on-line: 4 November, 2008

SNP-minisequencing has become common in forensic genetics, especially for analysing degraded or low copy number DNA (LCN DNA). The aim of this study was to examine the usefulness of five SNP (single nucleotide polymorphism) markers for analyzing degraded and LCN DNA recovered from archival samples. DNA extractions of eight formalin-fixed paraffin-embedded (FFPE) tissues were performed and DNA fragments were amplified in one multiplex PCR (polymerase chain reaction). SNPs were identified in a minisequencing reaction and a gel electrophoresis in ABI Prism® 377 Sequencer. The research confirmed the usefulness of SNP-minisequencing for analysing FFPE tissues.

Keywords: SNP, low copy number DNA, minisequencing, formalin-fixed paraffin-embedded tissues

INTRODUCTION

Archival biological materials, such as formalin-fixed paraffin-embedded tissues (FFPE), represent an abundant source of DNA for forensic applications as well as for disease investigations, especially for genomic studies on human cancer (Gilbert *et al.*, 2007). A large number of FFPE tissue archives have been established during many years of studies (Li *et al.*, 2008). They are common research objects, because developing high-throughput screening methods for exploiting archival tissues is an important area of investigation (van Beers *et al.*, 2006). However, chemical modifications during the fixation process degrade DNA, which makes the genetic analyses difficult (Li *et al.*, 2008). Formaldehyde interacts with nucleotides causing the formation of labile methylol derivatives. Moreover, DNA often undergoes further damage as a result of secondary effects and pH of reagents used (Feldman *et al.*, 1973). This results in poor yields and low quality of extracted DNA, being degraded to fewer than 300 bp (Cronin *et al.*, 2004). Hence, standard methodologies are not sufficient for investigating DNA obtained from such samples. Forensic laboratories use short tandem repeats (STR) as the standard DNA identification method (Collins *et al.*, 2004). All markers available in commercial multiplexes are ideal tools for paternity testing and personal identification of high quality DNA. However, since DNA extracted from FFPE tissues is heavily degraded, archival sample analyses using STR are impossible to carry through. It is because of the fact that STR markers have a relatively large amplicon size (Gill, 2002). This generates the necessity of applying new markers and methods for analysing FFPE tissues. Single nucleotide polymorphisms (SNPs) seem to be ideal tools for such purposes (Gill, 2001). They serve as excellent biological markers for a variety of applications. SNP is a small genetic variation that occurs within a human DNA every 100 to 300 bases along the 3-billion base human genome. For a variation to be considered a SNP, it must occur in at least 1% of the population. SNPs appear in coding and noncoding regions of the genome, and, in both nuclear and mitochondrial DNA. SNP analyses have become common in personal identification

✉Corresponding author: K. Babol-Pokora, Departament of Forensic Medicine, Medical University of Lodz, Sedziowska 18a, 91-304 Łódź, Poland; phone (+48 42) 654 45 36; e-mail: pokorak@gmail.com

★Presented at the XXXV Winter School "The Structure and Function of Protein and Nucleic Acids" organized by Faculty of Biochemistry, Biophysics and Biotechnology, Jagiellonian University, 23–27 February, 2008, Zakopane, Poland.

Abbreviations: FFPE, formalin-fixed paraffin-embedded; LCN, low copy number; PCR, polymerase chain reaction; SNP, single nucleotide polymorphism; STR, short tandem repeats.

or genotyping markers of disease. SNPs amplicons are shorter in length than other markers, thus enable efficient amplification of fragmented DNA. It concerns both evidence from crime scenes and FFPE tissues that represent often degraded DNA or trace amounts, described as Low Copy Number DNA and refers to less than 100 pg DNA per sample (Gill *et al.*, 2000). SNPs are therefore of great importance, especially in forensic genetics, where DNA is often degraded to short fragments and commonly used STR markers range in length between 150 and 450 bp, which makes them to large to be successfully amplified in degraded DNA samples. In contradistinction to that, SNPs allow designing PCR target sequences to be shorter than 150 bp (Alonso *et al.*, 2003). One of the methods ideal to perform SNP analyses is a minisequencing reaction (Sobrino *et al.*, 2005). It is a high-throughput and sensitive reaction which enables amplification and detection of very minute amounts of DNA and thus makes genotyping of FFPE tissues possible.

The aim of this study was to examine if the minisequencing reaction in a range of five SNP loci (rs2294067, rs2282160, rs2070764, rs2277216, rs1063739) is proper for analysing highly degraded DNA obtained from formalin-fixed paraffin-embedded tissues and to compare the results with those obtained from standard analyses using STR markers.

MATERIALS AND METHODS

DNA extraction. The study material comprised eight formalin-fixed paraffin-embedded tissues of six patients and their reference blood samples. DNA was extracted from small shavings of FFPE tissues by means of the phenol/chloroform extraction and using the Sherlock AX kit (A&A Biotechnology). All FFPE samples were dewaxed in xylene and rinsed in ethanol in order to remove the paraffin. The reference materials were isolated using the Sherlock AX kit and quantified by means of Quantiblot™ (Applied Biosystems). Since FFPE tissues were heavily degraded it was impossible to quantify them using the same method, because Quantiblot™ is proper only for high quality DNA and the sensitivity of this method, according to producer, is 150 pg DNA per sample. Quantitative assay using PicoGreen dye also failed, due to the fact that this method requires high quality dsDNA (Koba *et al.*, 2007). Hence, the quantity of DNA extracted from FFPE tissues was estimated on the basis of their minisequencing results compared to control DNA dilution series (not presented).

STR analyses. Samples were subjected to standard forensic analyses in a range of 15 simultaneously amplified STR loci (D3S1358, TH01, D21S11, D18S51, D5S818, D2S1338 D13S317, D7S820, D16S539, CSF1PO, D19S433, vWA, D8S1179, TPOX and FGA), as well as the sex typing marker Amelogenin, contained in the AmpFlSTR Identifiler™ Kit (Applied Biosystems). All STR assays were run in accordance with the manufacturer's recommendation. The PCR contained 10 μl AmpFlSTR PCR Reaction Mix, 5 μl AmpFlSTR Identifiler™ Primer Set, 2.5 U of AmpliTaq® Gold DNA Polymerase and 5 μl of DNA template (0.1–1 ng/μl). Thermal cycling was performed in a Biometra T1 Thermocycler (Biomedizinische Analytik), with the following cycling parameters: an 11-min incubation at 95°C followed by 30 cycles of 45 s at 94°C, 1 min at 59°C, and 1 min at 72°C and concluded with a 60-min incubation at 60°C. A slab-gel electrophoresis was performed in the ABI Prism® 377 Sequencer using GeneScan™ LIZ® 500 as an internal lane standard. The raw data were compiled and analysed using the accessory software – ABI Data Collection Software and GeneScan™ Programme (Applied Biosystems).

SNP analyses. DNA samples were amplified in one multiplex PCR in a range of five SNP biallelic loci (rs2294067, rs2282160, rs2070764, rs2277216, rs2101039), selected from a large number of SNPs, representing blood groups antigens, which in the past were applied to paternity testing (Babol-Pokora *et al.*, 2006). The size of the SNP amplicons was less than 150 bp, i.e. 123 bp, 99 bp, 93 bp, 85 bp and 71 bp. PCR reaction mix (25 μl) contained: 0.1–10 ng of DNA, 15 mM Tris/HCl buffer, (pH=8.0), 50 mM KCl, 0.2 mM of each dNTP, 2.5 mM $MgCl_2$, 4–8 pmol of each primer, and 1.25 U AmpliTaq® Gold DNA polymerase (Applied Biosystems). DNA was first denatured at 95°C for 11 min followed by 37 cycles of 94°C for 60 s, 61°C for 60 s, and at 70°C for 120 s. The final extension was carried out at 70°C for 30 min. The amplified SNP loci, after purification using the MiniElute® kit (Qiagen), were identified by a minisequencing reaction that is a single-base extension method using SNaPshot™ Kit (Applied Biosystems) according to the manufacturer's instructions. The reaction mix contained 2 μl of purified PCR product, 5 pmol of each primer (Babol-Pokora *et al.*, 2006), 5 μl of ABI Prism® SNaPshot™ Multiplex Kit (ddATP-dR6G™, ddCTPdTAMRA™, ddUTP-dROX™, ddGTP-dR110™, AmpliTaq® FS TACS/Core and Multiplex Reaction Buffer). The reaction was performed in a total volume of 10 μl and subjected to 25 cycles of 96°C for 10 s, 50°C for 5 s and 60°C for 30 s. Minisequencing products were next purified with the use of the DyeEx® kit (Qiagen) to remove unincorporated fluorescent ddNTPs, and resolved through a slab-gel electrophoresis in the ABI Prism®™ 377 Sequencer using

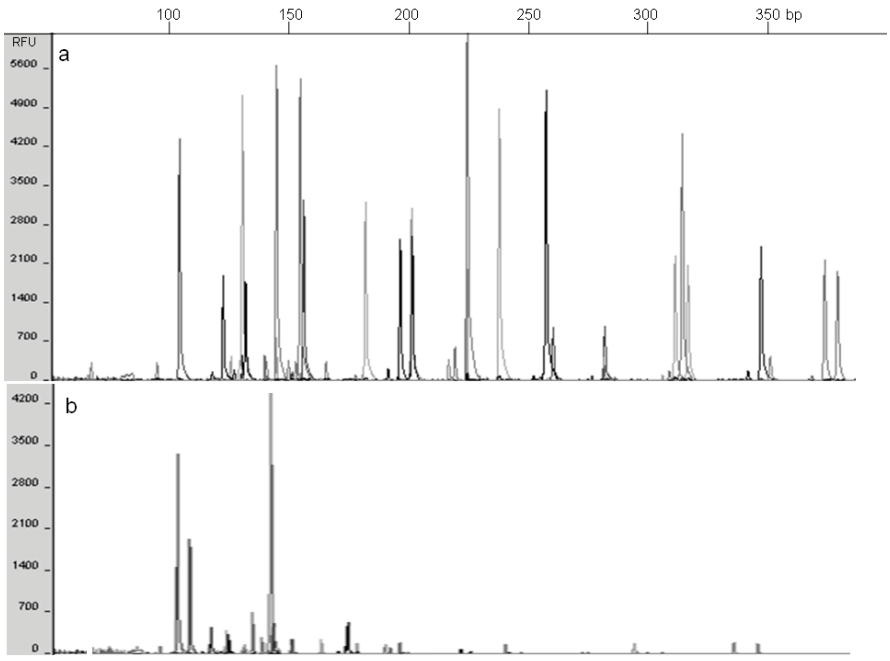

Figure 1. An example of multiplex analysis in a range of 15 STR loci and amelogenin locus: a) full genetic profile of a reference sample; b) partial genetic profile of FFPE sample; lack of PCR products with amplicons larger than 200 bp. RFU, relative fluorescence unit.

Gene Scan™ 120 LIZ® as an internal lane standard. The results were analysed by means of Gene Scan™ programme. Statistical analyses were performed using GDA (Lewis, 2001) and DLP (Berent & Szram, 2003) programmes.

RESULTS

Eight FFPE tissues were subjected to standard analyses with the use of 15 STR markers and amelogenin locus with simultaneous analyses of reference samples in the same range. FFPE tissues analyses were, however, impossible to carry through due to highly degraded DNA (Fig. 1). Only partial genetic profiles were obtained for all samples. There was no amplification product for markers with the largest amplicon sizes, i.e. CSF1PO (301–337 bp) and D2S1338 (305–358 bp) in all cases. Moreover, the amplification failed for D18S818 marker (264–346 bp) in four cases, for D16S539 (250–290 bp) and D7S820 (256–292 bp) in three cases and for D13S317 (204–244 bp) in one case. For other markers the yield of amplification products was poor. It was impossible to match the investigated FFPE tissues with reference samples on the basis of the obtained results. Archival samples were therefore subjected to analyses in a range of five SNP markers. After analysis by means of minisequencing reaction, adequate results were obtained for all FFPE tissues. The fluorescent signals for particular loci, after comparison with those obtained from serial control dilutions, suggest that concentrations of DNA extracted from FFPE tissues averaged 20 pg/µl. Despite employing highly degraded DNA, full genetic profiles were obtained in every case. Reference blood samples were typed using the same method. After analysing FFPE tissues, four different profiles were obtained (Table 1) and three of them matched the reference samples. The matching reference profiles were identical with those obtained from archival samples. Statistical calculations were made on the basis of a population database of 500 alleles, which was published earlier (Babol-Pokora *et al.*, 2006). The average profile frequency for the five SNP was f = 0.0089 (f_{max} = 0.033; f_{min} = 0.00012). On the basis of the results the following conclusions were drawn: four out of eight FFPE tissues originated from the same patient, three other samples came from the other patients and the origin of the last sample was unknown. Figure 2 shows the minisequencing results of two randomly chosen

Table 1. Genetic profiles and their frequencies obtained from analysis of eight FFPE tissues for five SNP markers

Locus	Profile I	Profile II	Profile III	Profile IV
rs2294067	G/G	G/G	G/G	G/C
rs2070764	A/T	A/A	A/T	A/A
rs1063739	A/C	A/C	C/C	A/C
rs2282160	A/G	A/G	G/G	G/G
rs2277216	C/T	C/T	T/T	C/T
f_{total}	**0.0146**	**0.005286**	**0.0005659**	**0.003654**
frequency	1/68	1/189	1/1767	1/273

f_{total} – profile frequency

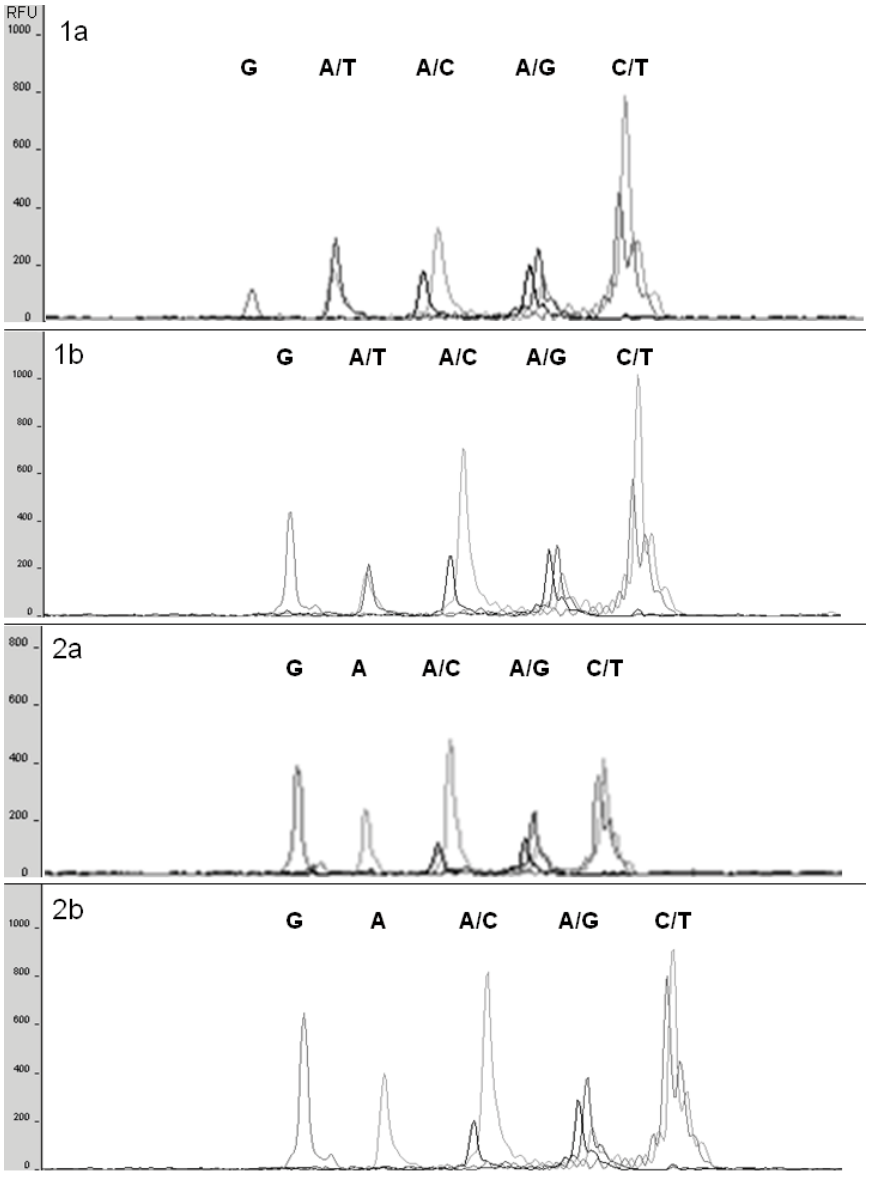

Figure 2. Results of minisequencing reactions presented for two samples (1a, 2a) which were randomly chosen out of eight investigated FFPE tissues and their references (1b, 2b). In both examples FFPE samples share the same genotypes with their references.

FFPE samples under investigation with their references.

DISCUSSION

During our study we compared the usefulness of standard STR markers with a recently elaborated SNP-pentaplex by applying them to FFPE tissue analyses. The results revealed that the SNP-minisequencing reaction is much more effective for LCN-DNA analyses as compared to commercial STR multiplexes, which are frequently used for such purpose. Since archival samples represented by FFPE tissues became a quite common source of DNA for a variety of applications, SNP-minisequencing seems to be the method of choice in future medical and forensic genetic studies of degraded samples. This is the first report, to our knowledge, concerning the application of SNP minisequencing to forensic identification of FFPE tissues in Poland. A similar study was conducted for a greater range of SNP loci by Gilbert *et al.* (2007). They, however, focused on SNP analyses only, without any comparative studies. Several genetic studies exploiting DNA extracted from FFPE samples have also been made with the use of different markers (van Beers *et al.*, 2006). Cawkwell and Quirke (2000) proposed the use of two microsatellite loci for direct analysis, without the DNA extraction step. In a different study three BCR-ABL fusion transcripts were applied to multiplex RT-PCR as a time and cost-sparing diagnostic tool (Bock *et al.*, 2003). Population studies, made on the basis of a central Poland population database composed of 500 alleles, revealed slight differences in allele frequencies between the target population and two Japanese populations composed of 204 alleles (Doi *et al.*, 2004) and 1276 alleles (http://snp.ims.u-tokyo.ac.jp/), ex-

cept in the locus rs2277216, where the allele frequencies were considerably different, i.e. the frequencies of the T allele in the rs2277216 locus among the Japanese populations were 0.520 and 0.483, whilst in the Polish population it was 0.256, and the frequencies of the C allele among the Japanese populations were 0.480 and 0.517, whilst in the Polish population it was 0.744.

The number of simultaneously amplified markers is significant for the power of discrimination of the multiplex. However, the general tendency in implementing new markers is to increase the chance of amplifying highly degraded DNA, using even less polymorphic markers, like SNPs, rather than to increase the discriminating power of the current techniques. There are several low-discriminating SNP genotyping sets composed of six SNP loci (Doi *et al.*, 2004) or eight SNP loci (Turchi *et al.*, 2004). There are also some sets with an increased number of markers, i.e. 21 (Dixon *et al.*, 2005) or even 52 SNP loci (Sanchez *et al.*, 2006). Those, however, during analysis are divided into several groups and run separately, because large multiplexes cannot be amplified simultaneously during one minisequencing reaction. There are, of course, lots of different methods for analysing SNP markers, based on modern high-throughput technologies, i.e. FRET (Lareu *et al.*, 2001), Single Base Extension-Tag microarrays (Balogh *et al.*, 2006) or pyrosequencing (Harrison *et al.*, 2006). They, however, call for special, very expensive equipment and some of them also require large amounts of DNA. In contrast to the above, minisequencing is a high-throughput and sensitive method, which enables amplification and detection of very minute amounts of DNA. It also does not require additional analyzers besides standard genetic laboratory equipment. The multiplex capability is, however, the most important feature of SNP-minisequencing, which explains its frequent application for SNP genotyping in forensic laboratories all over the world (Doi *et al.*, 2004; Quintáns *et al.*, 2004). Therefore, creating multiplexes with a capability of analysing degraded DNA seems more important for a variety of applications. SNPs have turned out to be excellent markers for analysing heavily degraded DNA recovered from archival samples. The discrimination power of SNP-pentaplex is not enough to obtain sufficient evidential value, however, the set has been successfully applied as an ideal screening method for personal identification. Although the SNP-pentaplex used in this study was designed predominantly for forensic applications, it undoubtedly demonstrates an important step forward in analysing heavily degraded FFPE specimens.

REFERENCES

Alonso A, Martin P, Albarrán C, García P, Primorac D, García O, Fernández de Simón L, García-Hirschfeld J, Sancho M, Fernández-Piqueras J (2003) *Croat Med J* **44**: 273–280.

Babol-Pokora K, Prosniak A, Jacewicz R, Berent J (2006) *Arch Med Sad Kryminol* **56**: 228–231.

Balogh MK, Bender K, Schneider PM (2006) *Prog Foren Genet* **11**: 43–45.

Berent J, Szram S (2003) *Forensic Sci* **1**: 46–47.

Bock O, Reising D, Keripe H (2003) *Diagn Mol Pathol* **12**: 119–123.

Cawkwell L, Quirke P (2000) *J Clin Pathol* **53**: 51–52.

Collins PJ, Hennessy LK, Leibelt CS, Roby RK, Reeder DJ, Foxall PA (2004) *J Forensic Sci* **49**: 1265–1277.

Cronin M, Pho M, Dutta D, Stephans JC, Shak S, Kiefer MC, Esteban JM, Baker JB (2004) *Am J Pathol* **164**: 35–42.

Dixon LA, Murray CM, Archer EJ, Dobbins AE, Koumi P, Gill P (2005) *Forensic Sci Int* **154**: 62–77.

Doi Y, Yamamoto Y, Inagaki S, Shigeta Y, Miyaishi S, Ishizu H (2004) *Legal Med* **6**: 213–223.

Feldman MY (1973) *Prog Nucleic Acid Res Mol Biol* **13**: 1–49.

Gilbert MT, Sanchez JJ, Haselkorn T, Jewell LD, Lucas SB, Van Marck E, Børsting C, Morling N, Worobey M (2007) *Electrophoresis* **28**: 2361–2367.

Gill P (2001) *Int J Legal Med* **114**: 204–210.

Gill P (2002) *Biotechniques* **32**: 366–368.

Gill P, Whitaker J, Flaxman C, Brown N, Buckleton J (2000) *Forensic Sci Int* **112**: 17–40.

Harrison C, Musgrave-Brown E, Bender K (2006) *Prog Forensic Genet* **11**: 52–54.

Koba M, Szostek A, Konopa J (2007) *Acta Biochim Polon* **54**: 883–886.

Lareu M, Puente J, Sobrino B, Quintáns B, Brión M, Carracedo A (2001) *Forensic Sci Int* **118**: 163–168.

Lewis PO, Zaykin D (2001) Genetic Data Analysis,Version 1.0 (d16c), http://lewis.eeb.uconn.edu/lewishome.

Li J, Smyth P, Cahill S, Denning K, Flavin R, Aherne S, Pirotta M, Guenther SM, O'Leary JJ, Sheils O (2008) *BMC Biotechnology* **8**: 10.

Quintáns B, Alvarez-Iglesias V, Salas A, Phillips C, Lareu MV, Carracedo A (2004) *Forensic Sci Int* **140**: 251–257.

Sanchez J, Phillips C, Børsting C, Balogh K, Bogus M, Fondevila M, Harrison CD, Musgrave-Brown E, Salas A, Syndercombe-Court D, Schneider PM, Carracedo A, Morling N (2006) *Electrophoresis* **27**: 1713–1724.

Sobrino B, Brion M, Carracedo A (2005) *Forensic Sci Int* **154**: 181–194.

Turchi C, Pesaresi M, Presciuttini S, Alessandrini F, Sassaroli C, Tagliabracci A (2004) *Prog Forensic Genet* **10**: 213–215.

van Beers EH, Joosse SA, Ligtenberg MJ, Fles R, Hogervorst FB, Verhoeff S, Nederlof PM (2006) *Brit J Cancer* **94**: 333–337.