

## Rapid microarray-based typing of forensic SNPs

M. Bogus<sup>a,\*</sup>, B. Sobrino<sup>b</sup>, K. Bender<sup>a</sup>,  
A. Carracedo<sup>b</sup>, P.M. Schneider<sup>a,c</sup>,  
SNPforID Consortium

<sup>a</sup> *Institute of Legal Medicine, Johannes Gutenberg University of Mainz, Germany*

<sup>b</sup> *Institute of Legal Medicine, University of Santiago de Compostela, Spain*

<sup>c</sup> *Institute of Legal Medicine, University of Cologne, Germany*

---

**Abstract.** The single base extension-tag array (SBE-Tag Array) method is carried out on glass slides and combines the specificity of minisequencing for SNP typing with the high throughput capacity of microarrays. Following multiplex PCR, a single tube SBE reaction is carried out, and the fluorescently labelled extension products are hybridized to the complementary DNA sequence tag (cTag) immobilized on a glass slide for locus-specific laser scan analysis. The aim is to prove and optimise the conventional microarray reaction on accuracy and efficiency for forensic applications. © 2005 Elsevier B.V. All rights reserved.

*Keywords:* Single nucleotide polymorphism (SNP); Microarray; Single base extension (SBE)

---

### 1. Introduction

We are developing a rapid multiplex genotyping SNP-assay suitable for forensic identification of individuals, based on the SBE method [2,3]. The multiplex PCR reaction with amplicon sizes ranging from 59 to 115 bp for 23 SNPs selected by the SNPforID consortium is followed by an allele-specific primer extension reaction with fluorescently labelled dideoxynucleotides (Fig. 1). The elongated fragments, carrying specific tags at their 5' ends, are hybridized to their complementary 20 mer tag sequences (cTags) immobilized on a glass slide. Detection is carried out at four or two wavelengths.

### 2. Materials and methods

PCR primers and conditions were comparable to the adapted conditions from the validated multiplex for SNaPshot analysis. 29/23 non-crossreacting tag sequences were chosen [1], and the complementary sequences were spotted as capture probes in duplicates on glass slides (MWG, CodeLink). To increase the typing throughput, multiple arrays were spotted with 4 to

---

\* Corresponding author. Tel.: +49 61313932688; fax: +49 61313936676.

E-mail address: bogus@uni-mainz.de (M. Bogus).

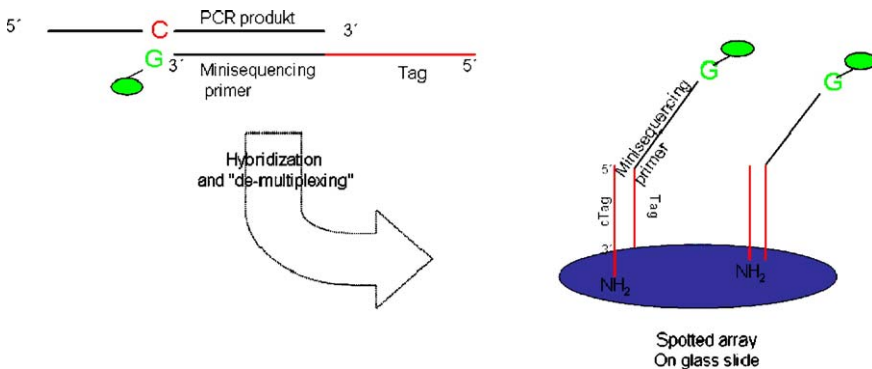


Fig. 1. Principle of the SBE reaction in solution and following hybridization based demultiplexing on glass slide.

56 arrays on the same slide and hybridized with individual samples, using a design called “array of arrays” [3].

SBE reaction was performed in 15  $\mu\text{l}$  total volume, containing  $1\times$  buffer, 1–4 U of sequencing enzyme (*Thermosequenase*, Amersham or *Theminator* DNA polymerase, New England Biolabs), 1  $\mu\text{M}$  reaction control (RC) tag, 0.041  $\mu\text{M}$  RC template, 1% Triton-X, 0.2  $\mu\text{M}$  tag-SBE primer mix, 0.2–3  $\mu\text{M}$  fluorochrome-labelled ddNTPs (Tamra-ddGTP, Cy5-ddUTP, Fluorescein-12-ddCTP, Texas Red-ddATP, Perkin Elmer) and 7  $\mu\text{l}$  purified PCR product. Labelling was performed for 25 cycles (96  $^{\circ}\text{C}$  for 30 s, 60  $^{\circ}\text{C}$  for 5 s and 60  $^{\circ}\text{C}$  for 10 s). In Santiago, each of four 10  $\mu\text{l}$  SBE reactions contained 2.5  $\mu\text{l}$  of concentrated template,  $1\times$  buffer, 0.25  $\mu\text{mol}$  of each tag-SBE primer, 2U of *Thermosequenase* and 0.5  $\mu\text{M}$  ddNTPs (one Tamra-labelled, three unlabelled). Labelling was performed for 30 cycles (96  $^{\circ}\text{C}$  for 30 s, 50  $^{\circ}\text{C}$  for 30 s, and 60  $^{\circ}\text{C}$  for 30 s). The hybridization was performed in 22  $\mu\text{l}$  total volume.

The SBE primers hybridize by their Tag-tails with spotted cTags and the reaction demultiplexes on the array revealing the genotypes (Fig. 1). Fluorescence detection was carried out at four (Mainz, GeneTAC<sup>TM</sup> LMIV, Genomic Solutions, access to instrument generously provided by D. Tautz, Institute of Genetics, University of Cologne) or a single wavelength (Santiago, Affymetrix 418 microarray scanner). Signal quantification was carried out using the ImaGene<sup>TM</sup> 5.5/4.1 software (BioDiscovery). Final allele calling was performed by the SNPSnapper software (Juha Saharinen, National Public Health Institute, Helsinki; website [www.Bioinfo.helsinki.fi/snpsnapper/start.htm](http://www.Bioinfo.helsinki.fi/snpsnapper/start.htm)) [4].

### 3. Results and discussion

Present experiments (Mainz/Santiago) show promising results. In total 21/22 out of 22/23 SNPs in reaction were detected and 16/21 were typed successfully. Optimization was conducted in terms of nucleotide and polymerase concentrations. To increase reaction stringency different temperature ranges were tested in minisequencing reactions. Comparisons were drawn with two sequencing enzymes, referring to imbalanced incorporation of dideoxynucleotides in previous experiments.

Annealing temperature, ddNTP and enzyme concentrations are critical factors to obtain correct genotypes. In Mainz, increasing the temperature to 60  $^{\circ}\text{C}$  for both enzymes resulted in signal improvement. In Santiago good results were obtained at 50  $^{\circ}\text{C}$ . One

constant dropout occurred (SNP14), and non-specific signals for ddGTP appeared at two loci (SNP3/SNP10a). In Mainz dropouts were depending on experimental conditions.

A high rate of misincorporation was observed for ddGTP. This led to a high background and false positive results for some of the allele calls. In contrast, the signals for A and C were generally weak and led to imbalanced intensities. Better distribution of dye intensities resulted after adjusting the dye concentrations. For further optimization two sequencing enzymes were compared. Results obtained with *Thermosequenase* showed to be the best at a concentration of 3 U and the addition of 4  $\mu\text{M}$   $\text{MgCl}_2$ . First tests with *Therminator* DNA Polymerase gave best allele calls at a concentration of 2 U per reaction without additives. In some assays allelic dropout occurred. Signal strengths were higher with *Thermosequenase* but the signal proportions were comparable for most alleles. Significant differences were observed within results for GA and AT polymorphisms. *Therminator* appeared to incorporate ddATP more efficiently, and ddTTP at a remarkably weak range. *Thermosequenase* showed the opposite characteristics. Both resulted in higher homozygote rates for the affected loci. No background signals occurred in described experiments.

The minisequencing reaction is the most critical factor to get balanced signal calls. A significantly higher incorporation rate for ddGTP has been observed. This is congruent with other studies [5]. Critical are also the preferences in incorporation of dyes. A possible explanation could be different steric dye properties. The varying preferences of polymerases are yet not fully understood. Data from Santiago using separate labelling reactions indicate that the unsatisfyingly high rate of wrong allele calls using all four ddNTPs in a single reaction is most likely due to the preferential activity of the polymerase regarding particular ddNTP/fluorochrome combinations. Reactions with separately amplified ddNTPs showed similar results for the described loci. For single dye reactions, it is possible to use less polymerase to obtain good results.

Some alleles showed weaker intensities than others. Compared to electropherograms from SNaPshot analysis, these loci exhibited similar proportions and probably correlate to the amount of PCR product. Further progress in the method described above is the ongoing optimization of SBE reaction conditions with the prospect to increase the markers in the multiplex and in parallel to increase the number of spotted tags on slide. The SBE-Tag array on glass slides is a promising and cost-efficient genotyping technology, which can be further extended with respect to the number of simultaneously analysed individuals and the size of the multiplex PCR reaction.

### Acknowledgment

Supported by a grant from the European Community, GROWTH programme, contract G6RD-CT-2002-00844.

### References

- [1] J.N. Hirschorn, et al., SBE-TAGS: an array-based method for efficient single nucleotide polymorphism genotyping, *Proc. Natl. Acad. Sci.* 97 (22) (2000) 12164–12169.
- [2] T. Pastinen, et al., Array-based multiplex analysis of candidate genes reveals two independent and additive genetic risk factors for myocardial infarction in the Finnish population, *Hum. Mol. Genet.* 7 (9) (1998) 1453–1462.
- [3] T. Pastinen, et al., A system for specific, high-throughput genotyping by allele-specific primer extension on microarrays, *Genome Res.* 10 (2000) 1031–1042.
- [4] J. Saharinen, et al., SNPSnapper application for allele calling, genotype assignment and database storing of SNP genotypes produced by microarray technology, Poster at ISMB, 2001.
- [5] M. Jain, et al., Genotyping African Haplotypes in ATM using a co-spotted single-base extension assay, *Human Mutat.* 22 (2003) 214–221.