Identification of point mutations in mixtures by capillary electrophoresis hybridization

Konstantin Khrapko1,2,*, Hilary A. Coller3, John S. Hanekamp1 and William G. Thilly1

1Center for Environmental Health Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA, 2Division on Aging, Harvard Medical School, Boston, MA 02115, USA and 3Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA

Received August 22, 1998; Revised and Accepted October 27, 1998

ABSTRACT

We have developed a rapid method for unambiguous identification and mutant fraction determination of individual mutants in mixtures of DNA sequence variants each differing by one or a few nucleotides. This method has applications to such diverse areas as interpretation of mutational spectra, screening of populations for polymorphisms and identification of species in environmental mixtures. In our approach, a mixture of unknown sequences labeled with a fluorescent dye is combined with a set of predetermined sequences (standards) representing the variants to be assayed. Labeling the standards with another dye allows the two sets of variants to be measured independently. Using constant denaturing capillary electrophoresis, the sequence variants are separated as individual peaks on the basis of differential melting equilibria. The unknown sequence variants are initially identified based on co-migration with particular standards. This preliminary identification is verified by hybridization of the unknown variants with the co-migrating standards within the capillary. We demonstrate the use of capillary electrophoresis hybridization to dissect complex mutational spectra of human cells in culture.

The method described in this communication offers the capacity to identify and measure specific point mutants within a mixture of variants of the same prototype DNA sequence. Such measurements would be instrumental in several important applications. Mutational spectra, for example, are powerful tools in studying causes and mechanisms leading to mutation, cancer and aging in humans (1), and are represented by complex mixtures of somatic mutations (2). Another potential application is the analysis of spectra of mechanisms leading to mutation, cancer and aging in humans (1), and are represented by complex mixtures of somatic mutations (2). Another potential application is the analysis of spectra of somatic mutations (3). In this case, many DNA samples from a large number of individuals that represent a population of interest may be pooled together and processed as one. The distribution of polymorphisms within the population will be represented by the distribution of sequence variants in the mixture that results from such pooling and can be measured by our method. In a similar way, it may be possible to analyze the frequency distribution of particular species in environmental mixtures (such as the set of viral or bacterial strains infecting a patient). In this case one would measure the frequency distribution of the sequence variants corresponding to individual strains.

Herein we describe a method in which a DNA sample containing a mixture of nucleotide variants (‘sample’) can be rapidly probed for the presence of any of a predetermined set of variants (‘standards’). In our approach, we first separate sequence variants (both the sample and the standards) by constant denaturant capillary electrophoresis (CDCE). The sequence variants of the sample are initially identified based on co-migration with particular standards. This preliminary identification is verified by hybridization of the sample variants with the co-migrating standards within the capillary.

The principle of separation of sequence variants by CDCE and CDCE equipment are described elsewhere (4,5). Briefly, the principle of separation is as follows. DNA fragments to be separated contain two isomelting domains, one with a high and the other with a low melting temperature. At a certain temperature, the fragments are in a rapid equilibrium between the completely double-stranded and the ‘partially melted’ conformations. The DNA molecule in a partially melted conformation has greatly reduced electrophoretic mobility (6). When base pair changes in the sequence variants to be separated are located in the low melting domain, they affect the melting equilibrium and, hence, the mobility of the fragment. In CDCE, these differences in electrophoretic mobility are used to separate mutants by means of capillary electrophoresis.

The CDCE instrument (Fig. 1, scheme at the bottom) consists of a 40 cm-long coated fused silica capillary of 75 micron inner diameter, filled with a viscous replaceable matrix (high molecular weight polyacrylamide in electrophoresis buffer: 89 mM Tris–HCl, 89 mM borate, 1 mM EDTA, 30 mM sodium borate, pH 8.4). The ends of the capillary are bathed in two chambers filled with electrophoresis buffer, and the chambers are connected to a high voltage power supply via platinum electrodes. Fluorescently-labeled DNA is electrokinetically injected into the capillary and electrophoresed at 100 V/cm. The process is monitored at a single point by a two-wavelength laser-induced fluorescence detector capable of independently measuring DNA fragments labeled with fluorescein or TMR. The ‘fluorescence versus time’ output of the
detector is represented as 'fluorescence versus position along the capillary' plots. Prior to the detector, the DNA passes through a 10 cm-long separation jacket and a 15 cm-long hybridization jacket. The jackets are maintained at precise temperatures (±0.01°C) by water baths (e.g., Neslab, Portsmouth, NH).

In a typical experiment, sample DNA labeled with fluorescein and the standards labeled with tetramethylrhodamine (TMR) are mixed together and injected onto a capillary. These two dyes induce very similar mobility shifts, which allows a mutant labeled with fluorescein to co-migrate essentially exactly with the same mutant labeled with TMR. While in the cold portion of the capillary, the DNA species move as a single peak. When the combined DNA peak enters the separation jacket, the DNA is subjected to partially melting equilibrium. Each mutant attains a unique mobility, and the mixture is gradually resolved into separate peaks (Fig. 1B). The optimal temperature for the separation of mutants can be easily determined for each particular target sequence (4).

In Figure 1, the challenges of identification of sequence variants are illustrated by a simplified imaginary experiment, in which the sample and the standard set consist of two variants each. Even though the two peaks of the sample co-migrate with the two standards A and B, respectively, they may or may not be identical to them. The sample peaks are labeled A? and B? respectively, with the question marks emphasizing the uncertainty. To determine whether sample peaks A? and B? are indeed identical to the standards A and B or merely co-migrate with them, one needs to perform a hybridization.

For hybridization, the different sequence variants, now represented by individual peaks, are stopped in the hybridization jacket (Fig. 1B). The capillary is heated to 90°C to denature DNA, then cooled to a temperature that permits reannealing of the single strands (55°C). It is important to note that injection of DNA onto the capillary was adjusted so that the concentration of the TMR-labeled standards in the capillary was about two orders of magnitude higher than that of fluorescein-labeled sample DNA (the relative sizes of the peaks in the two channels shown in Figure 1 do not reflect this difference). As a result, by mass action, almost all (∼99%) of the fluorescently-labeled sample DNA strands are forced to form duplexes with complementary strands of TMR-labeled standards.

After reannealing, the temperature is raised to that used for the initial separation of sequence variants and separation is resumed. If a sample DNA variant co-migrating with a particular standard were not identical to that standard (variant B? in Fig. 1), then it would reanneal with the excess of the standard to form two heteroduplexes with mismatches at the sites of non-identity. Mismatches destabilize the low melting domain of the DNA fragment, so the melting equilibrium is shifted toward the partially melted form, and the electrophoretic mobility of the fragment is significantly decreased. As DNA migrates through the rest of the hybridization jacket, the two heteroduplexes are retarded and move as two peaks far behind peak B (B:B? heteroduplexes in Fig. 1C). Such ‘disappearance’ of peak B? from its pre-hybridization position in the fluorescein channel separation profile indicates that sequence variant B? is not identical to the standard variant B, even though the two species co-migrate during separation. The absence of any fluorescent peak co-migrating with standard B after hybridization indicates that standard B is not present in the test population of molecules.

If a sample peak and a standard were indeed identical (see peaks A and A? in Fig. 1B and C), then heteroduplexes containing one strand originating from the sample and one from the standard DNA will contain no mismatches. The mobility of the fragments will be the same as those before hybridization and fluorescently-labeled species will continue to co-migrate with the TMR-labeled standard after hybridization, unambiguously confirming their identity.

To illustrate the power of capillary electrophoresis hybridization, we show the identification of mutations in the mitochondrial mutational spectra of the human lymphoblastoid cell line TK6 grown in culture. This spectrum is the most complex of the mutational spectra we have measured and is therefore an appropriate target for testing the technique. The spectrum in question (Fig. 2A) consists of multiple peaks, many of which are not completely separated from each other, and are therefore difficult to identify or isolate.
Figure 2. Dissection of complex mutational spectra into sub-spectra by capillary electrophoresis hybridization. DNA from cultured human lymphoblastoid cells (TK-6) was isolated and subjected to mutational spectrometry procedure as described elsewhere (5). A mixture of enriched somatic mutants was subjected to CDCE separation with (B and C) or without (A) hybridization. Each peak represents a different mutant species and the area under the peak is proportional to the mutant fraction for that mutant. Mutant fractions are calculated based on comparison to internal standards (5). (A) The mutational profile of the sample as separated without using hybridization; (B and C) mutational profiles of the same sample (upper curves) when separation was followed by hybridization with each of the two sets of standards (lower curves). Standard mutants are identified by numbers; the corresponding sequences are listed in (2). X denotes a peak that does not have a counterpart among the standards. When comparing Figures 1 and 2, note that each panel of Figure 2 represents only the homoduplex portion of the separation profiles shown in Figure 1C (delimited by the box).

In order to identify and measure specific mutations comprising the spectrum shown in Figure 2A, the sample was hybridized in the capillary with several subsets of the standard set of mutations. Subsets of the standard set were used rather than the full set in order to achieve complete (to baseline) separation, which is not possible if all standard mutations are separated simultaneously. Separation to baseline is essential to prevent cross-hybridization of the standard mutations. Figure 2B and C represents the results of capillary hybridization of the sample with two different subsets of the standard variants.

As shown in Figure 2, capillary hybridization results in a significant simplification of the separation profile because all the peaks representing variants not present in the subset used for hybridization ‘disappear’, as explained above. Hybridization thus facilitates ‘dissection’ of the complex spectra into simpler and more manageable ‘sub-spectra’. The individual mutations comprising the mutational spectra can then be easily identified and measured. In the case where a mutation in a spectrum does not belong to the standard set of mutations, the corresponding peak will disappear after hybridization with each of the subsets used (e.g. peak X in Fig. 2A). This will indicate that the mutation must be isolated, sequenced, and added to the standard set of mutations.

In conclusion, capillary electrophoresis hybridization is an efficient method for rapid unambiguous identification of sequence variants in complex mixtures. Importantly, this approach enables determination of the sequences directly in a mixture, without time-consuming isolation of individual species as required for other sequencing methods. Given recent progress in microfabricated electrophoresis devices, which use the same replaceable matrix system as does CDCE, e.g. (7), the approach described herein has the potential to evolve to a fast automated method for analysis of mixtures of sequence variants.

ACKNOWLEDGEMENTS

This work was supported in part by NIH grants AG13314 and CA77044.

REFERENCES