Methodology Report

Simple Detection of Large InDeLS by DHPLC: The ACE Gene as a Model

Renata Guedes Koyama, Rosa M. R. P. S. Castro, Marco Túlio De Mello, Sergio Tufik, and Mario Pedrazzoli

1 Departamento de Psicobiologia, Universidade Federal de São Paulo (UNIFESP), Rua Napoleão de Barros, 925/1 andar Vila Clementino, São Paulo 04024-002, SP, Brazil
2 Centro de Estudos de Psicobiologia do Exercício (CEPE), Universidade Federal de São Paulo (UNIFESP), Rua Marselhesa, 500/9 andar Vila Clementino, São Paulo 04020-060, SP, Brazil

Correspondence should be addressed to Rosa M. R. P. S. Castro, castro.rosamaria@psicobio.epm.br

Received 4 August 2007; Revised 23 December 2007; Accepted 2 February 2008

Recommended by Pierre Lehn

Insertion-deletion polymorphism (InDel) is the second most frequent type of genetic variation in the human genome. For the detection of large InDeLs, researchers usually resort to either PCR gel analysis or RFLP, but these are time consuming and dependent on human interpretation. Therefore, a more efficient method for genotyping this kind of genetic variation is needed. In this report, we describe a method that can detect large InDeLs by DHPLC (denaturating high-performance liquid chromatography) using the angiotensin-converting enzyme (ACE) gene I/D polymorphism as a model. The InDel targeted in this study is characterized by a 288 bp Alu element insertion (I). We used DHPLC at non-denaturing conditions to analyze the PCR product with a flow through the chromatographic column under two different gradients based on the differences between D and I sequences. The analysis described is quick and easy, making this technique a suitable and efficient means for DHPLC users to screen InDeLs in genetic epidemiological studies.

Copyright © 2008 Renata Guedes Koyama et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. INTRODUCTION

Insertion-deletion (InDel) polymorphisms are an important and abundant form of human genome variation; they are the second most frequent type of polymorphisms in the human genome and may be observed at the level of single-base pairs, multibase pair expansions of repeat units, random DNA sequence insertions and deletions, and as transposon insertions. This kind of DNA polymorphism can be used for the purpose of genetic mapping and diagnostics [1].

It is well known that translocation insertions and gross deletions (>100 pb) are important causes of both cancer and inherited diseases. Microdeletions and microinsertions (≤20 bp) account for 17% of all inherited diseases, as reported in the May 2007 release of the human gene mutation database (www.hgmd.org/).

An example of such polymorphism is the insertion/deletion (I/D) in the angiotensin-converting enzyme (ACE) gene that has been associated with many diseases, particularly of the cardiovascular system. Reports of such associations include risk of myocardial infarction and cardiovascular disease, ischemic stroke, effects on response of human muscle to strength training [2–6], and hypertension in subjects with mild to moderate degrees of sleep apnea (5 ≤ apnea-hypopnea index ≤ 30) [7].

In the renin-angiotensin system, ACE is a well-known zinc metallopeptidase that is widely distributed on the surface of endothelial and epithelial cells. ACE plays a role in the conversion of the inactive decapeptide, angiotensin I (Ang I or Ang 1–10), into the active octapeptide and potent vasoconstrictor angiotensin II (Ang II or Ang 1–8) as well as in the inactivation of the vasodilator bradykinin [8, 9].

The ACE InDel polymorphism is characterized by the presence (insertion, I) or absence (deletion, D) of a 288 bp DNA sequence in intron 16 of the gene. This insertion sequence is an Alu element (three Alu-repeat).

Although several genetic studies of the ACE D allele have associated polymorphism with cardiovascular disease, these data are questionable due to probable mistyping of the ACE I allele [10]. This problem occurs because of the different sizes
of I and D allele PCR products, which are 490 bp and 190 bp, respectively. But it happens that the D (shorter) allele possess the property of being preferentially amplified compared to the I allele, and this differential response to amplification may lead to mistyping of the ID sample.

Conflicting reports of ACE I/D frequencies have also been described, given the difficult nature of this genotyping. The diversity of findings has been attributed to methodological and technical variations in detection of the polymorphisms [11].

The first description of ACE I/D polymorphism genotyping proposed by Rigat et al. [12] was a PCR method using a set of primers flanking the insertion. However, as mentioned earlier, the D allele is preferentially amplified, so the ID heterozygote can be mistyped as DD, thus causing a higher frequency for this genotype. The probability of this mistyping has been estimated to be 5–10% [13, 14]. Therefore, careful control is required and repeated testing is also often necessary, especially when verifying the ID heterozygote.

This method has been modified several times, and a confirmatory PCR has been proposed by Shanmugan et al. [15] to minimize mistyping of the I allele as a D allele. In this method, a new sense primer is used inside the Alu sequence, resulting in amplification of a 408 bp fragment of allele I. The D allele shows no amplification under this condition due to lack of an annealing site for the new sense primer, so it works as an additional amplification after the conventional method. The authors reported 100% accuracy with this methodology. However, use of a primer annealing inside an Alu sequence may cause nonspecific annealing due to structural differences in the I and D allele in which there are three Alu sequences.

Therefore, problems with preferential amplification by PCR are not entirely excluded with this method [16]. Eight human genes specific to Alu insertion polymorphisms have been described (ACE, TPA25, PV92, APO, FXIIIB, D1, A25, and B65) [17], and theoretically, any of them can be amplified by PCR using a primer that anneals in the Alu sequence. Moreover, this methodology generally involves various steps and is time consuming.

Other modifications have included a step-down PCR, as described by Chiang et al. [16], which aims to increase the product and the detection rate of the I allele. This method involves initial PCR annealing temperatures higher than the melting point of the primers, followed by annealing temperatures reduced stepwise to the melting point. In this technique with high amplification failure, the results should be interpreted by two different observers, and consensus opinions have to be obtained from a third observer blinded to the results from the other two.

The low resolution of these methods can lead to difficulties in data interpretation, and also requires various time-consuming steps. Therefore, a more efficient means of ACE genotyping is particularly desirable for clinical and epidemiologic investigations. Other techniques have been proposed such as real-time PCR [18], to detect ACE ID polymorphism, but this has not been used in genetic association studies, likely because of the need for specific and more expensive reagents. The multiplex approach, proposed by Evans et al. [19], improves the accuracy of ACE genotyping, but presents difficulties in the post-PCR handling, such as agarose diagonal gel electrophoresis. Moreover, to our knowledge, the multiplex approach is specific to ACE InDel genotyping and cannot be extended to other systems.

Here, we report a protocol using DHPLC (denaturating high-performance liquid chromatography) to genotype large InDels employing the ACE gene I/D polymorphism as a model. The major advantage of this method is the ease it provides in screening a large number of samples while avoiding electrophoresis and gel analysis.

The technique is based on automated detection of DNA segments by ion-pair reverse-phase high performance liquid chromatography [20]. This approach normally compares two alleles after denaturing and reannealing PCR amplicons. In this methodology, a preliminary quality analysis of the amplicon is always done in non-denaturing conditions at 50°C in order to quantify and verify the quality of the PCR product. Under such non-denaturing conditions, the presence of InDels in single-base pairs or small repeat units can be detected [21–24]. This method does not work well, however, in detecting large InDels (>50 pb) because the gradient flow is designed for small (deletion) sequences [25].

In the present study, we chose to analyze the amplicon using the DHPLC instrument at a non-denaturing temperature, using the same PCR product to flow through the chromatography column under two different gradients, which were based on the predicted D and I allele sequences. We compared this DHPLC method with the conventional and confirmatory methods for ACE I/D polymorphism genotyping, and we propose this methodology for genotyping InDels throughout the genome.

2. MATERIALS AND METHODS

2.1. Samples

Genomic DNA was directly extracted from 3 mL of whole blood [26] from 335 volunteers with mixed ethnic backgrounds, after obtaining their written informed consent.

2.2. Conventional PCR

Genotyping of the ACE gene was performed in all 335 samples, as described by Rigat et al. [12]. The primers anneal outside the insertion/deletion region in intron 16 of the ACE gene and yield a PCR product of 490 base pairs (bp) in the case of the insertion allele or 190 bp in the case of deletion allele. Depending on the presence or absence of the insertion allele, the genotype of the subjects was classified as II (homozygote for the insertion allele), DD (homozygote for the deletion allele), or ID (heterozygote).

2.3. Confirmatory PCR

Each sample found to be of the DD genotype using the conventional method was subjected to a second independent PCR amplification with a set of primers that recognize the Alu insertion-specific sequence, as described by Shanmugan et al. [15]. A PCR product of 408 bp indicated the insertion allele.
DHPLC analysis was performed in each sample found to have the DD genotype using the conventional method [12].

Genomic DNA samples were subjected to PCR using 10 mM of forward primer and 10 mM of reverse primer, as described by Rigat et al. [12], in a solution containing 1.0 mM MgCl2, 2.5 mM Tris-HCl (pH 9.0), 1.0 mM each of dNTP, and 1 U of Platinum Taq DNA polymerase (Invitrogen, SP, Brazil).

We entered the two predicted sequences, representing the ACE gene I and D alleles (NCBI ref X62855, SNP: rs13447447-Figure 1), into the system to create two different gradients of buffers and acetonitrile. The system control software (Transgenomic Navigator Software version 1.5.4, Transgenomic Inc., USA) gave the flow rate of the reagents based on the predicted sequences. The reverse phase gradient was performed under two different conditions, one for each allele. In condition 1, for the D sequence, the gradient was initially set to 47.2% 0.1 M TEA, pH 7.0 (buffer A) mixed with 52.8% 0.1 M TEA, pH 7.0, v/v 25% acetonitrile (buffer B). At the end point (5 minutes), the gradient was 38.2% buffer A and 61.8% buffer B. The gradient used for condition 2, for the large sequence (I), started with 39.2% of buffer A mixed with 60.8% of buffer B, and at the end point (5 minutes) the gradient was composed of 30.2% buffer A and 69.8% buffer B.

Eight µL of PCR product was applied twice to a DNASep column (Transgenomic-Wave 3500A DHPLC system, ref DNA 99 3510, 4.6 mm × 50 mm, Transgenomic Inc., USA). Elution of DNA was detected by 260 nm UV absorbance, and the chromatograms were analyzed by the presence or absence of ampiclon at 50°C with the two gradients and flows (Figure 1).

3. RESULTS

From the initial 335 samples, 95 were found to be of the DD genotype. These were reevaluated using the conventional genotyping method, as well as the confirmatory and DHPLC methods for comparison. We confirmed that 81.05% of the 95 samples were the DD genotype. However, in using the DHPLC and confirmatory methods, 18.95% of the samples proved to be the ID genotype. In two cases (2.1%) only the DHPLC detected the ID genotype and in another two cases (2.1%) only confirmatory PCR detected the ID genotype (Table 1).

An example of the pattern found using the DHPLC method is shown in Figure 1. The DD genotype was only detected by the specific DHPLC gradient and flow designed for the shorter PCR product (see Mat&Meth DHPLC condition 1), and the ID genotype was detected only by the specific DHPLC gradient and flow designed for the larger PCR product (see Mat&Meth DHPLC condition 2). When we had heterozygosity (I/D genotype sample), we observed peaks in the two different gradient and flow conditions.

4. DISCUSSION

In this report, we established the use of DHPLC for rapid screening of large InDels such as ACE I/D polymorphisms. Compared to other genotyping techniques, DHPLC offers several technical advantages. In a large-scale genotyping setting such as population screening, the adaptability of DHPLC, along with its high throughput, should significantly reduce overall processing time. Also, the autorun mode of DHPLC significantly decreases handling time without the loss of assay specificity [27–29], and screening is relatively quick and easy (8 minutes for total run per sample after PCR, including sample injection, column equilibration, and cleaning).

The method presented here offers several specific advantages over the other methods used to genotype ACE I/D polymorphisms, in which the false-positive mistyping of the DD genotype is a concern. The probability of mistyping
has been cited by some authors to range from 5–10% [13, 14], but in our samples we detected a discrepancy in 19% of cases. Certainly, this accuracy depends on the observer’s expertise in performing PCR gel electrophoresis analyses. The confirmatory PCR method, used to overcome this problem, is time consuming since two PCRs must be performed, and some believe that the results (the PCR bands in the gel) should be assessed by more than one observer to avoid misinterpretation [15]. This method has been reported to have 100% accuracy in genotyping the ID heterozygote. However, in our study, the ID heterozygote was not detected in 2% of all samples. Chiang et al. [16] also showed decreased detection of the I/D heterozygote using this method relative to the step-down approach. DHPLC has the advantage of genotyping many samples without supervision and the DHPLC can run overnight with automatically generated results. If the PCR conditions is carefully optimized and the injected amplicon is unique and pure, the exit pattern shows only one clear and recognizable peak, thus avoiding the necessity of analysis by two independent observers.

DHPLC methodology makes genotyping of slightly modified DNA, like SNPs and small InDels (1 to 50 bp) feasible for numerous samples and has been used in many studies [30–32]. Traditionally, DHPLC has not been recommended for large InDels, because the gradient flow is designed for one sequence (deletion, e.g.), thus rendering the other (insertion) unrecognizable [24, 25]. We observed this phenomenon when we analyzed the ACE amplicon in a single DHPLC run, but in this study we demonstrated that it is possible to overcome this problem with two separate runs using different gradient flows for each allele. Therefore, DHPLC may be used to detect large InDels (up to ∼200 bp) if one analyzes the amplicon using two nondenaturating runs with two buffer gradients and two flow adjustments.

To the best of our knowledge, this is the first report of a DHPLC selective approach to genotyping large InDels using two sequences input for two gradients of buffers and two flow adjustments to analyze both sequences separately. A multiplex-PCR coupled to HPLC analysis under non-denaturating conditions for detection of large InDels [33] has already been proposed, but small nonspecific peaks might hamper the analysis.

We have shown in this study that this DHPLC method is highly efficient and reproducible in the detection and genotyping of the ACE I/D polymorphism, which has been a challenging goal to achieve. In our study, two cases with the ID genotype were not detected by DHPLC but were detected by confirmatory PCR. In another two cases, only DHPLC, but not confirmatory PCR, detected the ID genotype. There is no clear explanation for this, but we conclude that identification of large InDels, like the ACE ID, are very difficult to genotype using only one methodology. Moreover, in confirmatory PCR a primer inside the Alu sequence is used, and the two samples not detected by DHPLC could have amplified other genes to Alu insertion polymorphisms.

Comparing DHPLC and the confirmatory PCR methodology used for the same purpose, DHPLC has the advantages of ease of handling and significantly fewer problems with data interpretation. Besides its application in ACE I/D polymorphism, DHPLC is a powerful tool for clinical and population analysis of InDels in general, which may greatly increase laboratory throughput.

## 5. CONCLUSIONS

InDels may be observed as one- or multibase pair expansions of repeat units, besides random DNA sequences as transposon insertions and deletions and their detection is used for the purpose of genetic mapping and diagnostics. The DHPLC method usually can detect small repeat units under nondenaturing conditions. With the approach described in the present study, DHPLC can be used to genotype large InDels, as demonstrated in the detection of ACE gene I/D polymorphisms. This method could be useful in clinical applications, and carries the advantage of avoiding possible gel misinterpretation and, thus, genotype misinterpretations. In conclusion, the DHPLC design described in this study is an alternative approach for large InDel detection and could be used in studies that evaluate this kind of polymorphism.

## ACKNOWLEDGMENTS

We are grateful to CNPQ, FAPESP (Fellowships no. 05/57504-4 (R. G. Koyama) and no. 06/58 104-2 (R. M. R. P. S. Castro), CEPID Grant no. 98/143003-3, and AFIP for financial support of this project.

## REFERENCES


### Table 1: Frequencies of DD only genotype using DHPLC and confirmatory methods.

<table>
<thead>
<tr>
<th>Genotype</th>
<th>% (n)</th>
<th>Confirmatory % (n)</th>
<th>DHPLC % (n)</th>
<th>DHPLC + Confirmatory % (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Concordances</td>
</tr>
<tr>
<td>DD</td>
<td>100 (95)</td>
<td>81.05 (77)</td>
<td>81.05 (77)</td>
<td>78.95 (75)</td>
</tr>
<tr>
<td>ID</td>
<td>0</td>
<td>18.95 (18)</td>
<td>18.95 (18)</td>
<td>16.84 (16)</td>
</tr>
</tbody>
</table>


