Multicapillary Electrophoresis Analysis of Single-Nucleotide Sequence Variations in the Deoxycytidine Kinase Gene

**Eszter Szantai,**1 Zsolt Ronai,3 Maria Sasvari-Szekely,3 Günther Bonn,2 and András Guttman1*

**Background:** Investigation of the genetic background of complex traits is the focus of recent interest, as several common diseases or the individual response to treatments of various illnesses have not yet been explored. These studies require the development and implementation of reliable and large-scale genotyping methods. In this report, we introduce an efficient technique based on PCR–restriction fragment length sequence variation technique for the analysis of the −360CG and −201CT single-nucleotide sequence variations in the deoxycytidine kinase gene.

**Methods:** A multicapillary gel electrophoresis instrument was used for the size determination of the generated DNA fragments. A healthy Hungarian population of 100 individuals was investigated to determine allele and genotype frequencies for the 2 sequence variations of interest.

**Results:** We found that the occurrence of the minor allele is rather low, i.e., the frequency of both the −360G and −201T variants is 1%.

**Conclusions:** Our technique can readily facilitate the analysis of these important sequence variations in other ethnic groups to clarify the role of these sequence variations in conjunction with arabinosylcytosine treatment in acute myeloid leukemia.

© 2006 American Association for Clinical Chemistry

---

1 Horváth Laboratory of Bioseparation Sciences and 2 Institute of Analytical Chemistry and Radiochemistry, Institute of Analytical Chemistry and Radiochemistry, Leopold-Franzens University, Innsbruck, Austria.

3 Department of Medical Chemistry, Molecular Biology and Pathobiology, Semmelweis University, Budapest, Hungary.

*Address correspondence to this author at: Horváth Laboratory of Bioseparation Sciences, Institute of Analytical Chemistry and Radiochemistry, Leopold-Franzens University, A-6020 Innsbruck, Innrain 66, Austria. Fax 43-512-507-2857; e-mail Andras.Guttman@uibk.ac.at.

Received March 31, 2006; accepted June 7, 2006.

Previously published online at DOI: 10.1373/clinchem.2006.071159

---

1 Nonstandard abbreviations: SNV, single-nucleotide variation; Ara-C, arabinofuranosylcytosine; dCK, deoxycytidine kinase; RFLP, restriction fragment length sequence variation; AP2, activator protein 2.
in most instances (15), novel higher resolution and automated approaches have also been introduced recently. High-performance capillary gel electrophoresis is an instrumental approach with the capability of a single-base resolution (16, 17). Electrophoretic microdevices were also demonstrated to be useful in high-throughput genotyping studies (18–20). Ultrathin-layer gel electrophoresis combines the multilane format of horizontal slab gel electrophoresis with the high resolving power of capillary- and microchip-based approaches (21). This technique has been used by our group for the investigation of several genetic sequence variations of neurotransmitter systems in psychogenetic association studies (22, 23). Recently, multicapillary gel electrophoresis systems, such as the eGene HDA-GT12, were introduced to accommodate rapid DNA fragment analysis (24).

One of the most effective treatments of acute myeloid leukemia is 1-β-arabinofuranosylcytosine (Ara-C) therapy, which blocks the DNA synthesis in its activated triphosphate form (25, 26). Unfortunately, response failure or relapse occurs in almost one third of patients treated with Ara-C (27, 28). The most significant enzyme that takes part in Ara-C phosphorylation is deoxycytidine kinase (dCK)\(^5\) (29). Under physiological conditions, the role of dCK is the phosphorylation of deoxycytidine, deoxyadenine, and deoxyguanosine to their monophosphate forms. The gene that encodes this enzyme is located at the 4th chromosome (4q13.3) and contains 7 exons. It is mostly expressed in lymphoid cells, such as in thymus and solid tumors; however, it can apparently be detected in all tissues (30, 31). The presence of an inactive transcript, attributable to alternative splicing, was observed in acute myeloid leukemia patients who were resistant to chemotherapy (29, 31). However, this kind of nonfunctional form was not yet verified in therapy-resistant patients by other studies (32). In a Chinese investigation, 5 SNVs were described in the regulatory region of the dCK gene (−125GT, −201CT, −289TA, −360CG, and −740GC) and 1 in the 2nd exon (138AG) that does not cause any changes in the amino acid sequence (32). Two sequence variations, the −201CT and the −360CG, with relatively high allele frequencies were identified in acute myeloid leukemia patients. The occurrence of the minor allele (−201T and −360G) of both SNVs was measured as 15.6%, with a strong linkage disequilibrium between the 2 sequence variations. The −360CG allele was demonstrated to be associated with good clinical response to Ara-C treatment of acute myeloid leukemia. Moreover, this variant in a haplotype combination with the −201T form was shown to cause higher mRNA expression rates detected by a luciferase reporter system (32). On the other hand, the consequence of the −360C/−201C combination was poor response to Ara-C therapy and lower gene expression at the mRNA level. A possible explanation of the observation was that the −360C allele did not form a transcription factor binding site, whereas in the presence of the −360G form, an activator protein 2 (AP2) site was developed. Similarly, the −201C allele created an Sp1 and an AP2 binding site, both of which were diminished in case of the T variant.

Here, we report on the determination of allele and genotype frequencies of 2 SNVs (−360CG and −201CT) in the promoter region of the dck gene in a representative group of healthy Hungarian population. PCR–restriction fragment length polymorphism (PCR-RFLP)-based techniques were developed in conjunction with the use of a multicapillary gel electrophoresis system and optimized for efficient genotyping of these SNVs from clinical samples.

**Materials and Methods**

**CHEMICALS**

The DNA isolation reagent set was obtained from Gentra. The HotStarTaq DNA polymerase reagent set, including the Q-solution, was purchased from Qiagen, the primers were from MWG-Biotech AG, and the deoxyribonucleotide-triphosphates were from Promega GmbH. The BioRad MyCycler thermocycler instrument was used for the PCRs. Restriction endonucleases and the 10× NEBuffers and 100× bovine serum albumin solution for the RFLP reactions were bought from New England Biolabs.

**PARTICIPANTS AND NONINVASIVE DNA SAMPLING**

DNA was extracted from 100 healthy unrelated individuals in a Caucasian population of Hungarian origin. Signed informed consent was obtained from all the participants. The study was approved by the Hungarian Research Ethics Committee. Buccal cells were collected by cotton swabs from the inner surface of the mouth. DNA was isolated with the Gentra DNA isolation reagent set.

**GENOTYPING PROTOCOL**

The Qiagen\(^\textregistered\) HotStarTaq\(^\textsuperscript{TM}\) DNA polymerase reagent set was used for PCR. The reaction mixtures contained 200 μmol/L dATP, dCTP, dTTP, and dGTP; 1 μmol/L of forward (5’ GCC TTC TCC CCA GAT GAG TT 3’) and reverse (5’ GTG GCC ATT CCT TAG TCT G 3’) primers, ∼1.5 ng of DNA template, 0.4 units DNA polymerase (lacking 3’ exonuclease activity, considering the basic term of allele-specific amplification), 1× reaction buffer, and 1× Q solution in a total volume of 15 μL. The primers were designed by the Oligo 5.0 software, considering (when possible) the restriction endonuclease and the primer pair such that the PCR product contained a nonvariant recognition site (i.e., a control digestion site), thus improving the reliability of the technique.

Thermocycling was initiated at 95 °C for 15 min to activate the hot start polymerase and to denature genomic DNA, followed by 40 cycles of 1 min of denaturation at 94 °C, 30 s of annealing at 63 °C, and 1 min of extension at 72 °C. A final 10-min extension step at 72 °C was followed by cooling the samples to 8 °C. The same thermocycling

\(^5\) Human genes: DCK, deoxycytidine kinase.
conditions were applied for the allele-specific amplification reactions (carried out in 10-μL quantities), with the exception of the annealing temperature, which was 64 °C. Also, the same forward primer was applied in conjunction with either the −201C-specific (5′ CAC TGG CGG GCC TGC GGG 3′) or the −201T-specific (5′ CAC TGG CGG GCC TGC GGA 3′) reverse primers in 2 separate reactions. The PCR was followed by Bgl II or Kas I restriction endonuclease digestion to determine the genotype of the −201CT and −360CG SNVs, respectively. Restriction enzymes were chosen with the NEBcutter web-based software tool (http://tools.neb.com/NEBcutter2/index.php) for identifying endonucleases applicable for the genotyping of the 2 SNVs. In case of Bgl II digestion, 2 μL of reaction mixture was added to 10 μL of PCR product, containing 1 mmol/L dithio-dl-treithol, 10 mmol/L MgCl2 (final respective concentrations) and 0.2 units/μL restriction endonuclease enzyme. The Kas I reaction was carried out by adding 8 μL of reaction mixture to 2 μL of PCR product containing 1× NEBuffer2, 1× bovine serum albumin (provided with the enzyme), and 0.2 units/μL restriction endonuclease enzyme. Digestions were performed overnight at 37 °C.

MULTICAPILLARY ELECTROPHORESIS SYSTEM

The PCR-RFLP products were analyzed on an eGene HDA-GT12 system with a multicapillary (12) gel cartridge (GCK5000F). The eGene HDA-GT12 system is an automated DNA fragment analyzer offering high resolution and short analysis time. The integrated system includes chemistry and software that analyzes and interprets data, controls scheduling, and performs integration and data management of bioanalytical test results. The system is capable of rapid and simultaneous analysis of 12 DNA samples and offers hands-free sample analysis from a 96-well plate. Ethidium bromide served as intercalator dye that fluoresces intensely in the presence of dsDNA by means of the illumination of a green light-emitting diode (peak wavelength, 524 nm) as excitation source. Separations were performed in an array of 13-cm-long (effective) fused silica capillary columns by method AM270, with a separation voltage of 6000 V and an injection time of 60 s at ambient temperature.

RESULTS AND DISCUSSION

Genotyping

A PCR-RFLP-based method was developed for the investigation of the 2 SNVs of interest (−360CG and −201CT) in the dCK gene. Carefully designed flanking primers were used in the amplification reaction of the 5′ region of the dCK gene possessing both the −360CG and −201CT SNVs. The generated PCR products were divided into 2 aliquots for genotype determination of the 2 sequence variations. Bgl II restriction endonuclease was used for the analysis of the −201CT SNV (recognition site, GCCNNNNGGC; N = A, C, G, T) (Fig. 1A). The boldface, underlined letter indicates the SNV position. It can be

**Fig. 1.** PCR-RFLP analysis of the −201CT single-nucleotide sequence variation.

(A), schematic representation of genotype determination of the DCK −201CT SNV by Bgl II RFLP. GCCNNNNGGC, recognition site of the Bgl II restriction endonuclease ((N), A, C, G, T). (B), multicapillary gel electrophoresis traces of the generated DNA fragments. (Upper trace), homozygous CC; (lower trace), heterozygous CT sample. Separation was performed by method AM270 with a 60-s injection time. Capillary length, 13 cm (effective); applied voltage, 6000 V at ambient temperature.
seen that the cleavage site was present in the PCR product only if there was a cytosine at position −201 of the gene. Consequently, the 583-bp PCR amplicon was cleaved into 2 fragments (204 and 379 bp) in case of a −201C allele, whereas the undigested product (583 bp) demonstrated the −201T variant.

The same logic was used for the investigation of the −360CG SNV, as shown in Fig. 2. The KasI restriction endonuclease was readily applicable to the analysis of this SNV, because the GGCGCC recognition sequence was present at the variant site only in case of the −360C allele (Fig. 2A). Note that in this instance, the PCR amplicon also contained a nonvariant cleavage site (Fig. 2A, open star). Guanine at the SNV site (−360G) altered the recognition site of the enzyme; thus, in this instance, the PCR amplicon was digested only at the nonvariant site, generating a 93-bp product and a 490-bp product. On the other hand, in case of the −360C form, digestion occurred at the variant site as well; therefore, 3 fragments were generated, with sizes of 93, 214, and 276 bp.

MULTICAPILARY GEL ELECTROPHORESIS

The final step of the genotyping protocol was electrophoresis-based size determination of the generated PCR-RFLP products. Fig. 1B depicts the traces of multicapillary gel electrophoresis analysis of 2−201CT SNV samples. The internal calibration markers (lower and upper markers of 50 and 1000 bp) were co-injected with each sample to attain more accurate fragment sizing (Fig. 1B, C1 and C2). These internal calibration markers were also used to align the separation traces of all 12 capillaries. In the upper trace of Fig. 1B, fragments of 204 and 379 bp can be seen, suggesting that this sample was a −201CC homozygote. The undigested 583-bp fragment also shows up in the lower trace of Fig. 1B, indicating that this individual was heterozygous for the −201CT SNV. Note that in Figs. 2 and 3, we show the separations in gel view format, which for a larger number of samples was better in lane-to-lane comparisons. Fig. 2B demonstrates the genotype determination of the −360CG SNV for 12 samples. The benefit of the automatically co-injected sizing calibration markers is...
also shown in Fig. 2B. The separation of the 100-bp ladder was aligned with the sample lanes to ensure highly accurate size determination. The appearance of the 93-, 214-, 276-, and 490-bp products in lane 4 demonstrated that this sample was a \(-360\)CG heterozygote. The genotypes of the other individuals (Fig. 2B, lanes 1–3 and 5–12) were \(-360\)CC homozygous. Note that the electrophoresis analysis was accomplished in <5.5 min, corresponding to an analysis time of 27 s per sample; 12 samples were analyzed simultaneously by the 12-capillary cartridge. This throughput is comparable to the recently introduced microelectrophoresis devices.

### ALLELE FREQUENCY AND HAPLOTYPE DETERMINATION

The multicapillary gel electrophoresis technique described above was applied to allele frequency determination of the \(-201\)CT and the \(-360\)CG sequence variations in a healthy Hungarian population of 100 individuals. The allele frequency of the Hungarian group was as low as 1% for both the \(-201\)T and \(-360\)G alleles, with only 2 samples being heterozygous for each sequence variation. The presence of the minor allele of both SNVs in the same 2 samples appeared to confirm a strong linkage disequilibrium of the 2 sequence variations described earlier (32), although theoretically, in case of double heterozygotes, 2
haplotype combinations (i.e., relative chromosomal localization of the variant alleles) are possible.

The combination of allele-specific amplification and PCR-RFLP helped us to decide whether the −201T and −360G alleles are present on the same chromosome (haplotype determination) in the samples examined. In the 1st step, the 2 homologous chromosome pairs containing the −201C and −201T variants were subject to allele-specific amplification in 2 separate reaction mixtures by the use of the outer forward primer in combination with either a C- or a T-specific reverse primer (see sequences in Materials and Methods), respectively, as shown in Fig. 3A. The generated 392-bp product was then digested by the restriction endonuclease KsaI. Because the recognition site of this enzyme was present in the amplicon only in case of the −360C variant, formation of 2 digested fragments (178 and 214 bp) unambiguously defined the presence of cytosine at position −360. On the other hand, the lack of digestion products (intact 392-bp fragment) suggested the presence of the −360G allele. Fig. 3B clearly demonstrates the haplotype structure of the 2 double-heterozygous samples. It can be seen that no digestion occurred in case of the −201T-specific PCR amplicons, which suggested that the −360G variant was together with the −201T allele on the same chromosome in both samples, i.e., the haplotype was −360G − −201T (Fig. 3B, lanes 1 and 2). Confirming this finding, the KsaI enzyme digested the PCR product containing −201C, indicating that the haplotype of the other chromosome was −360C to − −201C (Fig. 3B, lanes 3 and 4). Although the significance of this linkage disequilibrium analysis was somewhat weak, owing to the low allele frequency values, we suggest that the −360C allele was colocalized with the −201C variant, whereas the −360G form was together in all cases with the −201T form.

CONCLUSION
Pharmacogenetic studies have recently focused on investigating the genetic sequence variations and haplotype structure of systems that play a key role in the efficacy of medical treatments. This approach aims at developing, on the basis of the individual genetic constellation of the patient, more personalized therapeutic protocols and consequently more efficient medications. It is important to note that these studies (as well as the investigation of any complex traits) suffer from a few shortcomings. Because the efficiency of a medication is influenced by several environmental and genetic variables, the effect of a single sequence variation is rather small and often hard to detect. On the other hand, in case–control studies, spurious associations can also appear, leading to false positives (33). Another major complication of complex trait genetics is that the allele frequencies of a given sequence variation can vary over an extremely wide range in different ethnic groups.

One of our goals was to develop an affordable and reliable method for the investigation of 2 SNVs of interest (−360CG and −201CT) in the 5′ region of the dCK gene. The technique that we introduce in this report is based on a PCR-RFLP technique designed and optimized specifically for the analysis of these variant sites. This technique is easy to use and relatively cheap to carry out in typical clinical laboratories. Although both restriction endonucleases and capillary electrophoresis have been used for a wide range of applications, their combination for genotyping is not very common. In fact, there are only a few studies reported in the literature describing methodologies with a similar approach (34, 35). Flexible capillary-illar instrument for electrophoresis analysis allowed the use of unlabeled PCR-primers, i.e., the DNA-fragments were labeled “in situ” by ethidium bromide during the separation process. The lower limit of detection was <0.1 ng/μL, which was more than sufficient to analyze digested PCR-amplified DNA fragments. Rapid genotyping of 100 samples clearly demonstrated that these SNVs are not widespread in the healthy Caucasian population of Hungary. The allele frequency of the −201T and −360G alleles was found to be as low as 1% among Hungarians. The method described here provides the opportunity to further investigate these genetic variations in the dCK gene to clarify the role of these SNVs.

In summary, we report here the allele frequency values of the −360CG and −201CT SNVs of the dCK gene in a Hungarian population. Our findings establish an essential basis for the investigation of the role of these genetic variations in other ethnic groups. In addition, our demonstrated method of PCR-RFLP combined with multiphase gel and haplotype analysis in clinical settings.

This work was supported by the European Community Marie Curie Chair 006733, the Hungarian Országos Tudományas Kutatási Alapprogramok T035203, and the Szent-Györgyi Albert Professorship of the Hungarian Ministry of Education (0018). E.S. gratefully acknowledges a grant from the GenAu program of the Austrian Ministry of Arts, Science and Education. The eGene HDA-GT12 system was a kind gift of eGene Inc.

References


