Universal fluorescent tri-probe ligation equipped with capillary electrophoresis for targeting SMN1 and SMN2 genes in diagnosis of spinal muscular atrophy

Chun-Chi Wang a, Chi-Jen Shih b, Yuh-Jyh Jong c,d,e, Shou-Mei Wu a,f,*

a School of Pharmacy, College of Pharmacy, Kaohsiung Medical University, Kaohsiung, Taiwan
b Department of Fragrance and Cosmetic Science, College of Pharmacy, Kaohsiung Medical University, Kaohsiung, Taiwan
c Graduate Institute of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan
d Department of Pediatrics, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan
e College of Biological Science and Technology, National Chiao Tung University, Hsinchu, Taiwan
f Department of Chemistry, College of Sciences, National Sun Yat-Sen University, Kaohsiung, Taiwan

HIGHLIGHTS
- UFTPL-CE was used for detection of SMN1 and SMN2 genes in SMA.
- UFTPL was performed by adding three probes to differentiate SMN1 and SMN2.
- Of the 48 samples, the data of gene dosages were corresponding to DHPLC methods.

ARTICLE INFO
Article history:
Received 24 February 2014
Received in revised form 2 May 2014
Accepted 5 May 2014
Available online 13 May 2014

Keywords:
Universal fluorescent tri-probe ligation
Capillary electrophoresis
SMN1/SMN2
Spinal muscular atrophy

GRAPHICAL ABSTRACT

Abstract
This is the first ligase chain reaction used for diagnosis of spinal muscular atrophy (SMA). Universal fluorescent tri-probe ligation (UFTPL), a novel strategy used for distinguishing the multi-nucleotide alternations at single base, is developed to quantitatively analyze the SMN1/SMN2 genes in diagnosis of SMA. Ligase chain reaction was performed by adding three probes including universal fluorescent probe, connecting probe and recognizing probe to differentiate single nucleotide polymorphisms in UFTPL. Our approach was based on the two UFTPL products of survival motor neuron 1 (SMN1) and SMN2 genes (the difference of 9 mer) and analyzed by capillary electrophoresis (CE). We successfully determined various gene dosages of SMN1 and SMN2 genes in homologous or heterologous subjects. By using the UFTPL-CE method, the SMN1 and SMN2 genes were fully resolved with the resolution of 2.16 ± 0.37 (n = 3). The r values of SMN1 and SMN2 regression curves over a range of 1–4 copies were above 0.9944. Of the 48 DNA samples, the data of gene dosages were corresponding to that analyzed by conformation sensitive CE and denatured high-performance liquid chromatography (DHPLC). This technique was found to be a good methodology for quantification or determination of the relative genes having multi-nucleotide variants at single base.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Single-nucleotide polymorphisms (SNPs) are the most frequent genetic variant in humans, and often occur biallelically at a rate of approximately once every 100–300 bases [1,2]. A lot of individual SNPs located in coding regions have been already known to affect...
gene biological functions and cause several genetic diseases [3–6]. Therefore, highly selective and sensitive detection of SNPs is a requisite for heredity-related risk assessment, disease diagnostics and drug development. Several techniques have been developed for detecting SNPs, typically including polymerase chain reaction (PCR) [7,8], primer extension [9,10], strand displacement amplification (SDA) [11,12], rolling circle amplification (RCA) [13–15] and ligase chain reaction (LCR) [16–18]. LCR offers a sensitive and specific alternative platform for detection of SNPs [16–21]. In this study, a novel method of universal fluorescent tri-probe ligation (UFTPL) is developed for analyzing the multi-nucleotide difference at single base. The mechanism of this method is similar to LCR. LCR is a method of DNA amplification requiring the nucleic acid as the probes for each of the two DNA strands and then two partial probes are ligated to form the actual one. The LCR needs two enzymes of DNA polymerase and DNA ligase, and is more specific compared to PCR [22]. Our UFTPL method utilized three probes to provide higher specificity than LCR, and was carried out by only using the DNA ligase without the DNA polymerase. It is superior to LCR due to specific recognition of SNPs and more sensitive fluorescent labeling. This method was applied for quantitative determination of survival motor neuron genes (SMN) in diagnosis of spinal muscular atrophy (SMA).

SMA is an autosomal recessive disease characterized by degeneration of motor neurons which results in progressive muscular atrophy and weakness [23,24]. The severe disease has high incidence of 1 in 6000–10,000 live births and the carrier rate is 1:35 [25,26]. Two SMA-determining genes, a telomeric SMN1 (MIM# 600354) and a centromeric SMN2 gene (MIM# 601627), were located at chromosome 5q13. Most of SMA patients (95%) contract the disease by either deletion or conversion of at least exons 7 and 8 of the telomeric SMN1 gene [27,28]. The sequences of SMN1 and SMN2 genes are extremely similar with only five nucleotide differences, including cDNAs-substitution of single nucleotides in exons 7 (c.840 C>T) and 8 (G>A). The mutation in exon 7 renders the SMN2 gene incapable of compensating for the SMN1 gene, but allows the SMN1 gene to be distinguished from the SMN2 gene. Thus, detection of the absence of SMN1 exon 7 is a powerful tool for pre- and post-natal diagnosis of SMA.

Until now, many strategies have been established for determination of SMN1/SMN2 genes, such as quantitative real-time PCR [24,29], DHPLC [30], MALDI-TOF mass spectrometry [31] or capillary electrophoresis (CE) [32,33]. These methods are simple, but misdiagnosis may occur when an extra nucleotide variant happened on the genetic fragments. In addition to utilization of these strategies, some other helpful techniques have been employed for resolution of SMN1/SMN2 genes. Single-base extension (SBE) or primer extension, known as minisequencing, is widely used for SNP typing [34,35]. The SBE primer is only extended one base because the ddCTP or ddTTP acts as a terminator in the extension reaction [34]. Primer extension was performed in the presence of a mixture of dATP, dGTP, dTTP and ddCTP, resulting in 4 base pairs of different lengths, between SMN1 and SMN2 [35]. However, SBE or primer extension may cause the difficult recognition between SMN1 and SMN2. On the other hand, UFTPL could easily identify SMN1 and SMN2 by designing probes with different base numbers. In this study, a new methodology of UFTPL coupled with CE was developed to determine SMN1/SMN2 gene dosage.

2. Experimental

2.1. Materials

Hydroxypropyl cellulose (HEC, 1% in H2O ~145 mPa.s) was purchased from Fluka (Fluka Chemie GmbH, Switzerland). Hydroxyethyl cellulose (HEC, 1% in H2O ~145 mPa.s) was purchased from Fluka (Fluka Chemie GmbH, Switzerland). Sigma–Aldrich (Sigma, St. Louis, MO, USA). Methanol and urea were purchased from Merck (Merck, Darmstadt, Germany). 5× TBE buffer was purchased from Protech Technology Enterprise Co., Ltd. and diluted with dd-water before use. The dd-water was obtained from a Milli-Q water system (Millipore, Bedford, MA, USA).

2.2. Genomic DNA samples

The ethics approval for this study was obtained from the Institutional Review Board at Kaohsiung Medical University Hospital where participants were recruited and experiments on humans were conducted. Written informed consents were obtained from all participants. Total of 48 DNA samples were analyzed by this UFTPL method, which included 10 SMA patient samples, 12 parent samples of these 10 patients, and 26 normal controls. The phenotypes were differentiated by the doctors in the hospital. Genomic DNA was collected from peripheral whole blood through using GFX™ Genomic Blood DNA Purification Kit. The SMN1/SMN2 ratios of all DNA samples were quantitatively analyzed by the UFTPL method.

2.3. Amplification of SMN fragments

In this research, the SMN gene fragments including SMN1 and SMN2 and a KRIT1 gene fragment serving as internal standard (IS) for determining the relative gene dosage of SMN1/SMN2 were simultaneously amplified by PCR. The primers used in PCR were as shown in Table 1. A fluorescent dye was used to label the primer of KRIT1 gene in order to make it detectable by a fluorescent detector. The total volume of 50 μL for PCR contained 200 ng of genomic DNA, 0.16 μM of each primer for the KRIT1 gene and 0.2 μM of each primer for the SMN gene, 200 μM dNTPs, 0.2 μL of e2TAK DNA polymerase (TaKaRa Biotechnology Co., Ltd.), and 10 μL of 5× e2TAK PCR buffer as provided by the manufacturer. The PCR amplification was performed in a Px2 thermo cycler (Thermo Electron Co.) with an initial denatured step at 95 °C for 10 min, followed by 25 cycles consisting of denaturation at 95 °C for 45 s, annealing at 53 °C for 45 s, extension at 72 °C for 1 min and a final extension step at 72 °C for 10 min. After completion of PCR, the products were cleaned up by a DNA Clean/Extraction kit (GeneMark, Hopegen Biotechnology Development Enterprise). The purified DNA was stored at −20 °C prior to the UFTPL reaction.

2.4. Universal fluorescent tri-probe ligation (UFTPL)

In the UFTPL reaction, a total volume of 25 μL contained 5 μL of the purified DNA product, 2.5 μL of 10× Taq DNA ligase reaction buffer (200 mM Tris–HCl, 250 mM potassium acetate, 100 mM magnesium acetate, 10 mM NAD, 100 mM dithiothreitol and 1% Triton X-100 at pH 7.5), 20 nM universal fluorescent probe (UF

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Oligonucleotide sequence (5′#5–4#)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KRIT1-F</td>
<td>FAM-TTCCAGATGGTCTACTCCCTG</td>
</tr>
<tr>
<td>KRIT1-R</td>
<td>AAAAAATCTTTAATTTTACGAGCG</td>
</tr>
<tr>
<td>SMN-F</td>
<td>ATATAAGCTAGCTAGCAAGTGTTTCCTTTTATTTTCTACAGGGT</td>
</tr>
<tr>
<td>SMN-R</td>
<td>TTTTACACAAAATGATAGTTT</td>
</tr>
<tr>
<td>UF probe</td>
<td>FAM-TGACGCTACAGCAGG</td>
</tr>
<tr>
<td>Connecting</td>
<td>PO4-TCTTCTTTTATTTTTCACTAGGGT</td>
</tr>
<tr>
<td>SMN1 probe</td>
<td>PO4-CAGACACAAATCAAAGAAGAG</td>
</tr>
<tr>
<td>SMN2 probe</td>
<td>PO4-TAGACAAATGACACAGAGGCTTCT</td>
</tr>
</tbody>
</table>

FAM: 6-Carboxy-fluorescein was the fluorescent dye for LIF detection. PO4: Phosphorylation on the 5′-end.
20 nM connecting probe, 20 nM SMN1 probe, 2 nM SMN2 probe and 40 U of Taq DNA ligase (NEW ENGLAND Biolabs®, 240 County Road, Ipswich). The ligation was carried out at 95 °C for 3 min, 65 °C for 2 h and 95 °C for 10 min. The products were directly analyzed in CE system after 10-fold dilution with dd-water.

2.5. CE system

UFTPL products were analyzed in a coated DB-17 capillary (Agilent Technologies Co.) of 100 μm internal diameter and 30 cm effective length. Before analysis, capillary was pre-washed with MeOH for 10 min and dd-water for 20 min. Then the separation matrix consisting of a mixture of 1.5% HEC, 2.0% HPC in 2.0× TBE buffer was injected at 40 psi for 50 min. Between runs, the separation matrix was re-rinsed at 40 psi for 10 min. The CE instrument was the Beckman P/ACE MDQ system (Fullerton, CA, USA) equipped with a laser induced fluorescence detector which utilized a laser energy (488 nm) for exciting the electrons of materials to obtain the fluorescence (excitation wavelength: 488 nm; emission wavelength: 520 nm). Sample injection was carried out at −10 kV for 30 s and then the separation was accomplished at 8 kV in the reverse polarity mode. The temperature was normally set at 15 °C.

2.6. Determination of SMN1/SMN2 copy numbers

The KRIT1 gene fragment (IS, 327 bp) was used to obtain the number of copies of SMN1 and SMN2 genes in exons 7 through comparing the peak height of SMN1 and SMN2 with that of IS genetic fragments. Standards of known copies were applied to establish the calibration curve which was further used for calculation of SMN1/SMN2 ratios of the blind samples. Because the UFTPL of SMN gene is integral duplication, thus, the calibration curves were established by using of standards having one, two,
three and four copy of SMN1 and SMN2 genes, respectively. The detection limit of this method was one copy of SMN1 or SMN2 gene. Data were taken for all participants, including normal phenotype controls, carriers and SMA patients.

3. Results and discussion

3.1. UFTPL assay

This study used universal fluorescent probe (UF probe), connecting probe and recognizing probe for universally fluorescent labeling and rapidly determining SMN1/SMN2. The mechanism was as shown in Fig. 1. Initially, the SMN genes including SMN1 and SMN2 were amplified by a pair of primers where one contained a part of universal sequences. After PCR, all SMN1/SMN2 fragments possessed the universal sequence and the products were purified by a clean-up kit for UFTPL reaction. In UFTPL, three probes were used to establish the individual fluorescent fragments with different lengths. By designing unique sequences, the universal fluorescent probe and SMN1/SMN2 probes attached to the universal sequence and the specific sequence of SMN1/SMN2, respectively. The SMN1 and SMN2 probes having different lengths (difference of 9 mer) were used to individually recognize the SMN1 and SMN2 genes. Then, the three probes were linked by the ligation enzyme. The procedures of hybridization and ligation increase the specificity for detection of SMN1 and SMN2 and FAM labeling enhanced the sensitivity of DNA detection. Finally, fluorescently labeled single-strand SMN1 and SMN2 fragments with different lengths could be easily separated and detected by CE-LIF.

3.2. Optimization of UFTPL reaction

The ligation of oligonucleotides plays an important role in the UFTPL reaction. In this study, a case of SMN1/SMN2 equal to 2/2 was used for investigating the amount of ligase and time for ligation, respectively (Fig. 2). Fig. 2A shows the peak height ratios of SMN1 and SMN2 fragments compared to IS when adding 10, 20, 40, 60 and 80 units of ligase in 25 μL reaction volume. Considering the reaction efficiency, 40-unit of ligase has reached the maximum reaction. Therefore, 40-unit of ligase was chosen as the optimal condition. The data of reaction time for ligation was as shown in Fig. 2B. Finally, the ligation reaction was accomplished within 120 min. The peak of SMN2 was higher than SMN1, when the same amounts of SMN1 and SMN2 probes were added for ligation. In order to unify the peak height of SMN1 and SMN2 on the cases possessing equivalent gene ratio (1:1 or 2:2), the amount of SMN2 probe was reduced to obtain the same peak height of SMN1 and SMN2. At last, the amount of SMN2 was lower than SMN1 by 10-fold.

3.3. Evaluation of CE system

In order to obtain efficient resolution of SMN1 and SMN2 single-strand fragments, different compositions of the polymer solution and various capillary temperatures were investigated. Fig. 3A shows the effect of the polymers on the resolution of SMN1/SMN2 genes. Single polymer solution of 2% HPC could not be used for analysis of SMN1/SMN2, when compared to the mixture of 2% HPC and 1.5% HEC. The combination of polymers formed a new separation matrix taking advantage of the best qualities of each polymer and generated a quasi-interpenetrating network which could provide more efficacy than single polymer solutions [36]. Finally separation was performed by using the mixed polymer solution of 2% HPC and 1.5% HEC. As to the capillary temperature, low temperature provided a better resolution of SMN1/SMN2 than others (Fig. 3B), due to the viscosity and the rigid structure of the polymers at low temperature. Therefore, 15 °C was used.

3.4. UFTPL with CE analysis

The UFTPL is able to easily differentiate between SMN1 and SMN2 genes. Therefore, when compared to other methods used to produce the variance of SMN1 and SMN2 amplicons, such as SBE or primer extension [34,35], UFTPL is able to more easily recognize and resolve SMN1/SMN2 genes by CE analysis. Under the optimal CE conditions, the SMN1 and SMN2 peaks were completely baseline separated and resolution was 2.16 ± 0.37 (n = 3), calculated through Eq. (1).

$$Rs = \frac{2(t_b - t_A)}{W_A + W_b}$$  (1)

The migration time of SMN1 peak was about 17.23 ± 0.25 min and that of SMN2 was about 16.53 ± 0.17 min. Although the UFTPL product of SMN2 was longer than SMN1 by 9 mer, the SMN2 fragment migrated faster than SMN1. The single-strand structures of UFTPL products of SMN1 and SMN2 twist irregularly and thus the migration order could not be judged by the length. Fig. 4 displays the electropherograms and sequencing data of individuals
possessing only SMN1 gene, only SMN2 gene and both genes. The data indicated that the front peak was referred to SMN2 gene and the back peak was referred to SMN1 gene.

3.5. Various ratios of SMN1/SMN2

In this research, various copy numbers of SMN1 and SMN2 genes were analyzed by the UFTPL-CE method. Fig. 5A displays the electropherograms of subjects having different ratios of SMN1/SMN2. The UFTPL-CE technique was successfully applied to resolve and quantify the SMN1 and SMN2 genes. After calibrating the SMN1 and SMN2 copy numbers versus the peak height ratio by using the standard samples, the regression curves were established. Fig. 5B shows a linear relationship between the observed peak height ratios and the known SMN1 and SMN2 copy numbers. Regression equations of SMN1 and SMN2 were $y = 0.3119x - 0.0372$ ($r = 0.9944$) and $y = 0.3625x - 0.1422$ ($r = 0.9949$), respectively ($n = 3$). The formulas were further applied for calculating the SMN1/SMN2 gene ratios of the analyzed population.

3.6. Applications

In this research, a total of 48 DNA samples were analyzed by the UFTPL-CE method. The various SMN1/SMN2 ratios among the 48

---

Fig. 3. Electropherograms of (A) different polymers of (a) 2.0% HPC in 2× TBE; (b) 1.5% HEC in 2× TBE and (c) 1.5% HEC + 2.0% HPC in 2× TBE. (B) Various capillary temperatures of (a) 15°C; (b) 20°C and (c) 25°C. The analyte was the UFTPL product obtained from a subject possessing SMN1/SMN2 ratio of 2:2. Other CE conditions were as follows: sample injection, −10 kV for 30 s; separation voltage, −8 kV.
participants were as shown in Table 2. In the group of SMA patients, 2 had the \( \text{SMN1/SMN2} \) ratio of 0:4, 5 had 0:3 and 3 had 0:2. Among the carriers who were the parents of these SMA patients, the number of subjects possessing \( \text{SMN1/SMN2} \) ratio of 1:1 was 2, of 1:2 was 5 and of 1:3 was 5. Of the 26 normal controls, \( \text{SMN1/SMN2} \) ratios of 2:1, 2:2, 3:0 and 4:0 included 8, 16, 1 and 1 subjects, respectively. The other two methods of conformation sensitive capillary electrophoresis (CSCE) [32] and DHPLC [37] were also used for verifying, and the data were corresponding to those of this UFTPL-CE method. This method was demonstrated successfully to be a viable tool for quantification of \( \text{SMN1/SMN2} \) in diagnosis of SMA.

Fig. 4. CE electropherograms and sequencing analysis of UFTPL products of (A) an individual with only \textit{SMN1} gene; (B) an individual with only \textit{SMN2} gene and (C) an individual with a gene ratio of \( \text{SMN1/SMN2} \) equaled to 2:2.
(A) Various SMN1/SMN2 ratios

(a) SMN2 only (0:2)  (b) SMN1/SMN2 = 1:3  (c) SMN1/SMN2 = 1:2  (d) SMN1/SMN2 = 1:1

(e) SMN1/SMN2 = 2:2  (f) SMN1/SMN2 = 2:1  (g) SMN1/SMN2 = 3:1  (h) SMN1 only (2:0)

(B) Linear regression

(a) SMN1 regression curve

\[ y = 0.3199x - 0.0372 \]
\[ r = 0.9944 \]

(b) SMN2 regression curve

\[ y = 0.3625x - 0.1422 \]
\[ r = 0.9949 \]

Fig. 5. (A) CE electropherograms for analysis of UFTPL products obtained from different DNA samples with various SMN1/SMN2 gene ratios. (B) The linear relationship between (a) SMN1 and (b) SMN2 peak height ratios (compared with IS) versus the known copy number of (a) SMN1 and (b) SMN2.
Table 2
Data of different SMN1/SMN2 ratios analyzed by the UFTPL-CE method among the 48 subjects.

<table>
<thead>
<tr>
<th>Genotype (SMN1/SMN2)</th>
<th>Copy number detected by UFTPL-CE</th>
<th>Number of subjects</th>
<th>Status</th>
<th>CSCE method</th>
<th>DHPLC method</th>
</tr>
</thead>
<tbody>
<tr>
<td>0:4</td>
<td>0.00 ± 0.00:3.88 ± 0.10</td>
<td>2</td>
<td>SMA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0:3</td>
<td>0.00 ± 0.00:3.15 ± 0.16</td>
<td>5</td>
<td>SMA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0:2</td>
<td>0.00 ± 0.00:2.15 ± 0.10</td>
<td>3</td>
<td>SMA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:1</td>
<td>1.05 ± 0.04:1.11 ± 0.08</td>
<td>2</td>
<td>Carrier</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:2</td>
<td>1.04 ± 0.06:2.04 ± 0.07</td>
<td>5</td>
<td>Carrier</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1:3</td>
<td>1.02 ± 0.10:2.96 ± 0.08</td>
<td>5</td>
<td>Carrier</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2:1</td>
<td>1.98 ± 0.13:1.16 ± 0.09</td>
<td>8</td>
<td>Normal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2:2</td>
<td>2.02 ± 0.13:1.92 ± 0.08</td>
<td>16</td>
<td>Normal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3:0</td>
<td>2.93:0</td>
<td>1</td>
<td>Normal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2:0</td>
<td>2.07:0</td>
<td>1</td>
<td>Normal</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total 48

4. Conclusions
A novel UFTPL-CE method was successfully established for determination of SMN1 and SMN2 genes. The important advantages of this method are better resolution by adding an extra probe for differentiation, and using universal fluorescent primer. This method is expected to be extended for quantification or separation of the relative genes having multi-nucleotide variants at single base, such as K-ras, EGFR or others.

Acknowledgements
We deeply extend our sincere thanks to the volunteers who kindly contributed samples that were crucial to this study. We gratefully acknowledge the support of the Ministry of Science and Technology of Taiwan (MoST), Kaohsiung Medical University, and NSYSU-KMU 103-P023 JOIN RESEARCH PROJECT (#NSYSU-KMU103-P023) by way of funding of this work, and the help of Kaohsiung Medical University Chung-Ho Memorial Hospital, Kaohsiung, Taiwan.

References