Molecular analysis of SMN1, SMN2, NAIP, GTF2H2, and H4F5 genes in 157 Chinese patients with spinal muscular atrophy

Jin He, Qi-Jie Zhang, Qi-Fang Lin, Ya-Fang Chen, Xiao-Zhen Lin, Min-Ting Lin, Shen-Xing Murong, Ning Wang, and Wan-Jin Chen

1 Department of Neurology and Institute of Neurology, First Affiliated Hospital, Fujian Medical University, China
2 Center of Neuroscience, Fujian Medical University, China

Abstract

Spinal muscular atrophy (SMA) is a common and lethal autosomal recessive neurodegenerative disorder, which is caused by mutations of the survival motor neuron 1 (SMN1) gene. Additionally, the phenotype is modified by several genes nearby SMN1 in the 5q13 region. In this study, we analyzed mutations in SMN1 and quantified the modifying genes, including SMN2, NAIP, GTF2H2, and H4F5 by polymerase chain reaction–restriction fragment length polymorphism (PCR-RFLP), multiplex ligation-dependent probe amplification (MLPA), TA cloning, allele-specific long-range PCR, and Sanger sequencing in 157 SMA patients. Most SMA patients (94.90%) possessed a homozygous SMN1 deletion, while 10 patients demonstrated only the absence of exon 7, but the presence of exon 8. Two missense mutations (c.689 C>T and c.844 C>T) were identified in 2 patients who both carried a single copy of SMN1. We found inverse correlations between SMN2, the NAIP copy number, and the clinical severity of the disease. Furthermore, 7 severe type I patients possessed large-scale deletions, including SMN1, NAIP, and GTF2H2. We conclude that SMN1 gene conversion, SMN1 subtle mutations, SMN2 copy number, and the extent of deletion in the 5q13 region should all be considered in the genotype–phenotype analysis of SMA.

1. Introduction

Spinal muscular atrophy (SMA), first described in 2 infant brothers by Guido Werdnig in 1891, is a common autosomal recessive neurodegenerative disorder characterized by symmetric and progressive myasthenia and amyotrophy owing to the dysfunction and loss of motor neurons in the anterior horn of the spinal cord. The prevalence of SMA is 1 in 6000 to 1 in 10,000 live births, with a carrier frequency of 1 in 60 to 1 in 1000 live births, with a carrier frequency of 1 in 40 in the general population (Iannaccone, 1998; Kolb and Kissel, 2011) these figures are 13.44 in 100,000 and 1 in 42, respectively, in mainland China (Sheng-Yuan et al., 2010). Childhood onset SMA can be classified into 3 types on the basis of age at onset and severity of clinical course (Munsat and Davies, 1992). In the most severe form, patients (type I) suffer from muscular weakness before 6 months and cannot sit without aid, and approximately 60–70% of affected individuals die from respiratory failure within the first 2 years (Meldrum et al., 2007; Munsat and Davies, 1992; Prior et al., 2011). Type II SMA is an intermediate form with onset before the age of 18 months, and patients of this type are able to sit, but unable to stand or walk without aid. Type III SMA (Kugelberg-Welander disease) is the mildest form, with onset during the juvenile period; these patients can walk and have a normal life span (Kolb and Kissel, 2011; Munsat and Davies, 1992).

The SMA-causative gene, the survival of motor neuron gene (SMN), was first identified by Lefebvre in 1995 (Lefebvre et al., 1995). In human beings, SMN is present in 2 highly homologous copies, SMN1 and SMN2, both of which are located in an inverted repeat area on chromosome 5q13. The functional difference between these 2 genes is a C-T variation in exon 7 (Burghes and Beattie, 2009; Lefebvre et al., 1995). About 95% of SMA cases are attributed to the homozygous loss of SMN1, while the remaining 5% have intragenic subtle mutations, such as nonsense mutations, small deletions, and splice site mutations (Alías et al., 2009; Lefebvre et al., 1995). Compared with the full-length SMN protein encoded by SMN1 gene, 85% of protein from SMN2 transcripts is truncated and unstable, which fails to compensate for the loss of SMN1 in SMA patients (Kolb and Kissel, 2011). However, the severity and duration
of SMA largely depends on the copy number variations of SMN2 (Feldkötter et al., 2002). In a previous study, we analyzed the SMN2 copy number of 51 SMA patients by real-time PCR and found that there was a negative correlation between SMN2 copy number and severity of the disease (Chen et al., 2005). Moreover, within the 5q13 region, NAIP, GTF2H2, and H4F5 are regarded as disease-modifying genes, correlating with disease severity and duration of survival. (Bürglen et al., 1997; Ma et al., 1999; Roy et al., 1995; Scharf et al., 1998).

Because of the lack of specific manifestations, a definite diagnosis of SMA largely relies on genetic testing. In standard genetic testing, polymerase chain reaction–restriction fragment length polymorphism (PCR-RFLP) is only capable of detecting homozygous deletion of SMN1; hence, a series of quantitative techniques, including denaturing high-performance liquid chromatography (DHPLC), real-time PCR, and multiplex ligation-dependent probe amplification (MLPA), have been developed to detect copy number variants related to SMA phenotype (Arkblad et al., 2006; Chen et al., 2007; Feldkötter et al., 2002). In particular, MLPA, which possesses the ability to analyze up to 50 DNA sequences in a single reaction, has become a widely used technique for copy number variation analysis (Schouten et al., 2002).

Here, we adopted PCR-RFLP, sequencing, and MLPA to analyze the correlation between SMN1, SMN2, NAIP, GTF2H2, and H4F5 genes variants and the SMA phenotype in 157 Chinese patients.

2. Materials and methods

2.1. Subjects and DNA isolation

Our study was approved by the Ethics Committee of the First Affiliated Hospital of Fujian Medical University and all subjects offered written consent.

From September 10, 1998, to April 30, 2012, 157 SMA patients from 143 unrelated families of Chinese origin were recruited, including the 87 patients who had been analyzed by PCR-RFLP and DHPLC in our previous study (Chen et al., 2007). These included 80 men and 77 women. The average age at onset is 5.795 years (SD = 7.833, range from 3 days to 30 years). All of the SMA cases fulfilled the diagnostic criteria defined by the International SMA Consortium (Munsat and Davies, 1992). Genomic DNA was extracted from peripheral vein blood by using the QIAamp DNA Kit (German QIAGEN Company).

2.2. SMN1 homozygous deletion analysis

PCR-RFLP was applied for detection of homozygous deletion of SMN1. The details of this method have been described by van der Steege et al. in 1995 (van der Steege et al., 1995). The products were visualized on 3% agarose gels under UV light.

2.3. MLPA analysis of 5q13 region genes

MLPA analysis for genetic testing of SMA was based on the commercially available SALSA MLPA kit P021 (MRC-Hollyland, Amsterdam, the Netherlands, www.mlpa.com). The kit contains several probes for genes in the 5q13 region, including SMN1, SMN2, NAIP, GTF2H2, and H4F5, and 21 reference probes (Stuppia et al., 2012). After 4 reaction steps (denaturation, hybridization, ligation, and amplification), the products were separated on an ABI-3130 genetic analyzer (Applied Biosystems, California, USA), with 500 LIZ® (Applied Biosystems, California, USA) as the internal size standard. Then, data were collected by Genemapper 3.0 (Applied Biosystems, California, USA) and analyzed using Excel software (Arkblad et al., 2006).

2.4. SMN1 subtle mutation analysis

For those patients without homozygous deletion of SMN1, the following 3 steps were employed to screen for subtle mutations. First, the copy number of SMN1 was quantified by MLPA. Second, the patients with 1 copy of SMN1 were screened for subtle mutations by Sanger sequencing, according to our previously described report (Chen et al., 2007). Third, TA cloning and long-range PCR were employed to identify the mutations in SMN1 or SMN2. For mutation in exon 7, TA cloning was employed to discriminate SMN1 from SMN2 on the basis of the

Fig. 1. MLPA analysis and Sanger sequencing of patients with c.689 C>T. A. MLPA results in individuals with a single copy of SMN1 (a, c) and a single copy of SMN2 (b, d). The c.689 C>T mutation in exon 5 was identified in SMN by direct genomic DNA sequencing; however, this method does not allow distinction of whether the mutation was present in SMN1 or SMN2. C. The mutation was identified as being present in SMN1 by allele-specific long-range PCR and sequencing.
nucleotide differences between SMN1 and SMN2 in exon 7 and intron 6. For exon 2a-6, allele-specific long-range PCR was performed on the basis of the nucleotide difference in exon 7; the procedure was in accordance with that described in the previous study (Clermont et al., 2004). For exon 1, considering that no mutations have been detected by direct sequencing, and that the intron region is more than 13 kb and does not differ between SMN1 and SMN2, further investigation was suspended.

3. Results

3.1. SMN1 gene mutation analysis

Using PCR-RFLP and MLPA, we found that 94.9% (149/157) of patients harbored homozygous deletions of SMN1; among these, 10 patients demonstrated absence of exon 7, but presence of exon 8. The remaining 5.1% (8/157) patients possessed 1 copy (5 patients) or 2 copies (3 patients) of SMN1. The 5 patients with a single copy of SMN1 were likely compound heterozygotes, with an SMN1 deletion on 1 allele and a subtle mutation on the other allele; 2 missense mutations were identified: c.689 C>T (p.S230L) and c.844 C>T (p.Q282X) (Figs. 1 and 2).

The patient with the c.844 C>T mutation had been diagnosed with SMA type I and died at 2 years of age, while the patient with the c.689 C>T mutation had been diagnosed with SMA type II, and was 6-years-old at the time of analysis. We then investigated the origin of the c.844 C>T mutation in the parents, and found that the mutation originated from the patient’s mother. MLPA results showed that the father possessed 2 copies of SMN1, whereas the father had only 1 copy (Fig. 2). The parents of the patient with the c.689 C>T mutation refused further genetic testing. In 300 healthy control individuals, neither c.689 C>T nor c.844 C>T was observed.

3.2. Copy number variants analysis of SMN2 gene

Among these 149 patients with SMN1-deletions, the distributions of SMN2 copy number were as follows: 1 (8.72%), 2 (73.83%), 3 (15.43%), and 4 (2.01%), while no case with homozygous SMN2 deletions was found (Table 1). An inverse correlation was observed between SMN2 copy number and SMA phenotype by Spearman correlation test \( r = 0.496, p = 0.000 \). Furthermore, based on the copy number of SMN2, we calculated the posterior probability of developing a SMA type I for those patients with SMN1 homozygous deletion, according to Bayes’s formula: the risk of developing type I for a patient with 1, 2, or 3 copies of SMN2 was 92.31%, 31.82%, and 4.35% respectively.

3.3. Genetic analysis of NAIP, GTF2H2, and H4F5

The distribution of NAIP copy number in SMN1-deleted patients were as follows (Table 2): homozygous deletion (10.07%), 1 copy present (84.56%), and 2 copies present (5.37%). A negative correlation was seen between NAIP copy number and SMA phenotype (Spearman

![Fig. 2. Pedigree chart, MLPA analysis, and Sanger sequencing of patients with c.844 C>T. A. MLPA results, “a” and “c” indicate SMN1, “b” and “d” indicate SMN2; the father possessed 1 copy of SMN1 and 2 copies of SMN2 (I1); the mother had 2 copies of SMN1 and SMN2 (I2); the patient had one copy of SMN1 and 2 copies of SMN2 (II1). B. Sanger sequence of SMN indicated a C/T difference between SMN1 and SMN2 (arrows in black), and a mutation, c.844 C>T, in exon 7 (arrows in red) (I1, I2, II1-1), which was shown to arise from SMN1 by TA cloning and Sanger sequencing (B, II1-2). C. The patient’s pedigree chart.](https://example.com/fig2.png)
Table 1
SMN2 copy number of three SMA clinical types.

<table>
<thead>
<tr>
<th>Clinical types</th>
<th>SMN2 copy number</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Type I</td>
<td>12 (25.00%)</td>
<td>35 (72.92%)</td>
</tr>
<tr>
<td>Type II</td>
<td>1 (1.00%)</td>
<td>64 (82.51%)</td>
</tr>
<tr>
<td>Type III</td>
<td>0 (0.00%)</td>
<td>11 (47.83%)</td>
</tr>
<tr>
<td>Total</td>
<td>13 (8.72%)</td>
<td>110 (73.83%)</td>
</tr>
</tbody>
</table>

Table 2
NAIP copy number of three SMA clinical types.

<table>
<thead>
<tr>
<th>Clinical types</th>
<th>NAIP copy number</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Type I</td>
<td>15 (31.25%)</td>
<td>32 (66.67%)</td>
</tr>
<tr>
<td>Type II</td>
<td>0 (0.00%)</td>
<td>74 (94.87%)</td>
</tr>
<tr>
<td>Type III</td>
<td>0 (0.00%)</td>
<td>20 (86.96%)</td>
</tr>
<tr>
<td>Total</td>
<td>15 (10.07%)</td>
<td>126 (84.56%)</td>
</tr>
</tbody>
</table>

correlation, \( r = 0.422, p = 0.000 \), and all the 15 affected individuals with no copies of NAIP had been diagnosed with SMA type I.

Additionally, 7 patients had a complete absence of GTF2H2. Interestingly, these 7 patients who demonstrated homozygous loss of NAIP, simultaneously possessed 1–2 copies of SMN2 (Table 3 and Fig. 3), and all of them had been diagnosed with type I SMA. Unexpectedly, we did not detect any deletion of H4F5 in any of 149 SMN1-deleted patients, as determined by MLPA.

4. Discussion

At present, no specific clinical features or examinations enable the diagnosis of SMA; hence, genetic testing plays an important role in identifying and classifying this condition. SMA is attributed to SMN1 mutations, and the absence of SMN1 includes homozygous deletion and conversion to SMN2 (Burghes, 1997). Here, we show that of 157 Chinese patients, 94.90% demonstrated homozygous loss of SMN1, which is consistent with that reported previously (Leffebvre et al., 1995). We also detected 10 patients who demonstrated an absence of exon 7 but the presence of exon 8 of SMN1, indicating the occurrence of a conversion event. According to the literature (Burghes, 1997), conversion events are related to mild SMA types (type II and type III); among these 10 patients in our study, only 1 had been diagnosed with type I.

Moreover, about 5% of SMA patients harbor compound mutations of SMN1, including heterozygous deletion of the gene, and in our study, among 5 patients with a single copy of SMN1, we identified 2 novel subtle mutations: c.689 C>T (p.S230L) and c.844 C>T (p.Q282X). Our study revealed that subtle mutations located in different domains may influence the phenotype of SMA. The p.S230L missense mutation, located in exon 5, which encodes the proline-rich domain, had no effect on cell viability and proliferation in an SMA animal model (Wang and Dreyfuss, 2001), and the patient with the p.S230L mutation in this study showed a relatively mild phenotype. The p.Q282X mutation, located in exon 7, affects a highly conserved tyrosine/glycine-rich sequence (Y/G box), which is important for oligomerization and stability of the SMN protein (Talbot et al., 1997). Several missense mutations other than p.Q282X that occur in and around the Y/G box, including p.Y272C, p.H273R, p.T274I, p.G275S, p.G279C, and p.G279V, have been associated with a severe type of SMA (Leffebvre et al., 1995; Wirth, 2000); in this study, the patient with p.Q282X also demonstrated a severe phenotype. In our study, the remaining three patients with one copy of SMN1 and three patients with two copies of SMN1 were also screened for subtle mutations of SMN1, and no mutation was found in coding region and intron-exon boundaries. We speculate that they may belong to other types of lower motor neuron disease, mutations in other genes or intron region of SMN1 need investigation in future.

In addition to SMN1, increasing emphasis is being placed on modifying factors—SMN2, NAIP, GTF2H2, and H4F5. In this population-based study of SMA, we tentatively analyzed the relationship between modifying genes and the severity of disease. Based on the distribution of SMN2 copies in 149 SMN1-deleted patients, an inverse relationship between SMN2 copy number and severity of disease was identified, which was consistent with previous reports in Caucasians and Tunisians (Amara et al., 2012; Arkblad et al., 2009; Feldkötter et al., 2002). The SMN2 copy number is also a credible prognostic factor; patients with a single copy of SMN2 are more likely to develop the severe type I SMA, with a high risk of 92.31%. Apparently, the SMN2 copy number plays an essential role in the variation of the SMA phenotype; however, there must be other disease-modifying factors in those SMA patients who possess the same number of SMN2 copies. In previous studies, the absence of NAIP, GTF2H2, or H4F5 was highly associated with the severe form of SMA (type I) (Bürglen et al., 1997; Ma et al., 1999; Scharf et al., 1998). In our study, all 7 patients who carried a large-scale deletion, including SMN1, NAIP, and GTF2H2, developed severe SMA at approximate 1–2-months of age and died from respiratory infection within 2 years, on the other hand, another 8 patients with deletion of SMN1 and NAIP also have developed a severe type of SMA. Unexpectedly, no patients with homozygous deletion of H4F5 were found. We compared the probe sequence of H4F5 (www.mlpa.com) with the genomic DNA sequence (Ensembl gene identification number: ENSG00000172058) and found that the probe was located upstream of H4F5, indicating that this MLPA kit needs to be improved.

In conclusion, after performing sequential genetic testing for 157 Chinese SMA-affected individuals, we propose the following 4 strategies as a reasonable approach for SMA screening. First, SMN1 homozygous deletion should be evaluated by PCR-RFLP. Second, the SMN1 copy number should be quantified by MLPA for those without complete loss of SMN1. Third, if necessary, a screen should be performed to identify subtle mutations in SMN1. Lastly, if possible, the adjacent modifying genes in the SMA-related region should be analyzed to assess the severity of the disease. Thus, SMN1 gene conversion, SMN1 subtle mutations, SMN2 copy number, and the extent of deletion in 5q13 region should all be considered during genotype-phenotype analysis of SMA.

Funding

This work was supported by grant 30900481 from the National Natural Science Foundation of China, grant 2012J06016 from the National Natural Science Foundation of Fujian Province of China, grant JA12129
from the Program for New Century Excellent Talents in Fujian Province University, grant 2009-CXZ-25 from the Fujian Medical Innovating Program, Program for Innovative Research Team in Science and Technology in Fujian Province University (Min Jiao Ke 2012 3), a program for clinical medical key discipline of Fujian Medical University (XK201108), and a key program of scientific research of Fujian Medical University (2009D064).

Acknowledgments

The authors sincerely thank the SMA families for their help and willingness to participate in this study.

References


Fig. 3. MLPA results of a normal control (A) and a patient (B) with a large-scale deletion in the 5q13 region. “a” indicates exon 10 of GTF2H2, “b” indicates NAIP, “c” and “e” indicate SMN1, and “d” and “f” indicate SMN2.