ORIGINAL ARTICLE / ОРИГИНАЛНИ РАД

SMN1 copy number as a modifying factor of survival in Serbian patients with sporadic amyotrophic lateral sclerosis

Miloš Brkušanin1, Irena Jeftović-Velkova2, Vladimir M. Jovanović3, Stojan Perić2, Jovan Pešović1, Goran Brajušković1, Zorica Stević1, Dušanka Savić-Pavičević1

1University of Belgrade, Faculty of Biology, Center for Human Molecular Genetics, Belgrade, Serbia; 2University of Belgrade, Faculty of Medicine, Clinical Center of Serbia, Neurology Clinic, Belgrade, Serbia; 3University of Belgrade, Siša Stanković Institute for Biological Research, Department of Genetic Research, Belgrade, Serbia

SUMMARY

Introduction/Objective Amyotrophic lateral sclerosis (ALS) is a devastating motor neuron disease. The majority of cases are apparently sporadic ALS (SALS) with variants in susceptibility genes or sometimes in high-risk ALS genes. Two ALS susceptibility genes are SMN1, whose functional loss causes spinal muscular atrophy (SMA), and a nearly identical SMN2 gene, which modulates SMA severity. In this study we examined the association of copy number variations (CNVs) of SMN1 and SMN2 genes and two additional genes, SERF1 and NAIP, residing in the same genomic region (i.e. 5q13.2 segmental duplication), with SALS in patients from Serbia.

Methods Multiplex ligation-dependent probe amplification was used to determine CNVs of each gene in a clinically well-characterised group of 153 Serbian SALS patients and 153 controls.

Results Individual association between SMN1, SMN2, SERF1 or NAIP CNVs and SALS susceptibility or survival was not found. Survival curves based on the multivariable Cox regression analysis showed that three SMN1 copies, lower ALS Functional Rating Scale Revised (ALSFRS-R) score at the time of diagnosis, faster decline of the ALSFRS-R score over time, and shorter diagnostic delay result in shorter survival of Serbian SALS patients.

Conclusion Clinical variables might be complemented with the SMN1 copy number to improve prediction of survival in Serbian SALS patients.

Keywords: survival motor neuron; amyotrophic lateral sclerosis; H4F5; NAIP; SMN1

INTRODUCTION

Amyotrophic lateral sclerosis (ALS) is an incurable and the most common late-onset motor neuron disease (MND) with an inexorably progressive course and mean survival around three to five years [1]. The incidence of ALS in Belgrade, Serbia, is 1.11/100,000 and worldwide crude incidence is 1.68/100,000 [2, 3]. Up to 10% of ALS cases are familial, caused by mutations in a growing number of genes [4]. The remaining ALS cases are considered sporadic (SALS), with no apparent family history of the disease. SALS is a multifactorial disease with multiple genetic factors of small effect size underlying its pathogenesis. In addition, environmental and stochastic factors and their interactions act as precipitating factors in genetically susceptible individuals [4]. Furthermore, up to 10% of SALS patients carry variants in ALS genes associated with the familial form (high-risk ALS genes) [5].

Candidate-gene and genome-wide association studies identified numerous single nucleotide variants, but with an unclear contribution to disease pathogenesis [4, 6]. Additionally, copy number variations (CNVs) represent important risk factors for a number of complex human disorders, including ALS [7]. Among them, 5q13.2 segmental duplication, harboring the survival of motor neuron (SMN) genes, was repeatedly studied due to its association with spinal muscular atrophy (SMA), the most common childhood-onset MND, and the involvement of SMN and some ALS-associated proteins in common molecular pathways [8].

Functional loss of the SMN1 gene in the telomeric part of the 5q13.2 duplication causes SMA, while a nearly identical gene, called SMN2, located in its centromeric part, represents the major genetic modifier of the SMA phenotype. Case-control studies of SMN1 CNVs as risk factors for SALS susceptibility and survival have delivered varying results [9–12]. Meta-analyses provided firm evidence that the SMN1 duplication (three copies) can be considered a risk factor for SALS susceptibility [13, 14]. Nevertheless, the mechanism behind the effect of SMN1 duplications on disease susceptibility remains elusive. One of the assumptions is that SMN1 duplication might be accompanied by duplications of surrounding genes, whose structural rearrangements and/or impaired functions of their protein products could confer
the risk for SALS independent of the SMN protein [13]. However, an association of CNVs of SMN-flanking genes in the 5q13.2 duplication with SALS has not been investigated. The closest gene to SMN is the SERF1 (small EDRK-rich factor 1, previous gene symbol H4F5) gene, which exists in two identical copies – SERF1A (telomeric) and SERF1B (centromeric). The SERF1 protein is a general regulator of protein aggregation and proteotoxicity. The NAIP (NLR family apoptosis inhibitory protein) gene exists as a full-length telomeric copy and additional pseudogene copies. The NAIP protein acts as a neuroprotector by the inhibition of motor neuron apoptosis. Deletions of telomeric copies of NAIP and particularly SERF1 mostly accompany deletions of the neighboring SMN1 gene in SMA patients [15]. However, duplications of NAIP were detected in ~7.5% of chromosomes carrying one SMN1 copy [15].

The aim of this study was to investigate whether the SMN1, SMN2, SERF1, and NAIP CNVs confer the risk for SALS. We also examined their modulatory effect on survival in a clinically well-characterized group of SALS patients from Serbia, either as independent variables, or as variables that could complement clinical variables in disease prognosis.

METHODS

Subjects

A total of 154 SALS patients (91 males and 63 females), who met the revised El Escorial criteria for probable or definite ALS and had no apparent family history of the disease [16], were recruited from January 2015 until June 2017 at the Neurology Clinic, Clinical Center of Serbia, Belgrade, Serbia. One patient was excluded from the study due to neoplasm of the spinal cord since this affected his functionality and course of the disease. All the remaining 153 patients (90 males and 63 females) were screened for two major causative mutations in Serbian ALS patients [17] in two high-risk ALS genes – p.Leu145Phe (c.435G>C, Ensembl transcript SOD1-201 ENST00000270142.10) mutation in the SOD1 gene and the expansion of the GGGGCC repeat in the C9orf72 gene (mutation = [GGGGCC] > 30). Demographic and clinical characteristics of patients are shown in Table 1. A total of 153 (90 males and 63 females) unrelated age-matched individuals from the general Serbian population were used as controls. The study was approved by the Ethics Committee of the Clinical Center of Serbia (number 339/4) in concordance with principles of the Declaration of Helsinki and informed consent was obtained from all participants.

Genetic analyses

Genomic DNA was extracted from blood samples using the QIAaamp DNA Blood Mini Kit (QIAGEN, Hilden, Germany) and quantified with the Qubit dsDNA BR Assay Kit (Thermo Fisher Scientific, Grand Island, NY, USA). Screening of the SOD1 p.Leu145Phe mutation was performed by PCR and CviKI-1 digestion (New England Biolabs, Ipswich, MA, USA), while screening of the C9orf72 expansion was done by conventional and repeat primed PCR [18]. Multiplex ligation-dependent probe amplification (MLPA), using the P021 A2 probe mix (MRC-Holland, Amsterdam, the Netherlands), was performed according to the manufacturer’s guidelines to determine CNVs of the SMN1, SMN2, SERF1, and NAIP genes. Approximately 200 ng of DNA was used for each reaction. Since MLPA is a relative quantification technique, all samples were compared with 15 reference samples using the Coffalyser.Net software (MRC-Holland, Amsterdam, the Netherlands). Ratios of MLPA probes specific for 5q13.2 genes and the assigned copy numbers were as previously described [15]. Reproducibility of the MLPA method was determined by running ~30% of randomly selected samples in independent duplicates and by obtaining identical results between the replicates. All samples were blindly analyzed regardless of clinical data.

Statistical analyses

All statistical analyses were carried out in the R ver. 3.3.3 with the significance level set at 0.05 [19]. In case of multiple comparisons, the Bonferroni correction was used to...
correct the significance level. The association of the SMN1, SMN2, SERF1, and NAIP gene copy numbers with SALS was tested by Fisher’s exact test.

Survival of patients was analyzed and visualized using the survival and survminer R packages. Cumulative survival probability was estimated using the Kaplan–Meier method. Survival curves of subgroups of patients (strata) within each genetic, demographic, and clinical variable were compared by the log-rank test and continuous variables were dichotomized with a cut-off to form strata. Age at onset, ALS Functional Rating Scale Revised (ALSFRS-R) at the time of diagnosis and diagnostic delay were dichotomized using the median split, whereas the cut-off value for Mini Mental State Examination (MMSE) and forced vital capacity (FVC), both at the time of diagnosis, was the one below which findings are indicative of mild cognitive impairment (< 24) or respiratory compromise (< 90%), respectively [20].

Assuming that effect of a single contributing factor may be of small size and obscured or confounded by other factors, the effect of multiple factors on survival was tested by the multivariable Cox regression analysis. The univariable Cox regression analysis was performed as a screening tool for the selection of clinical variables (covariates) for the multivariable Cox regression analysis. The overall significance of the Cox model was described by the likelihood ratio test (LRT) for the global null hypothesis. Percentage of the variance in survival in both Cox regression analyses was estimated by the R2 value. To verify the proportional hazards assumption, for each covariate in the obtained Cox model and for the model as a whole, the corresponding set of scaled Schoenfeld residuals was correlated with time to test for their independence. The proportional hazards assumption is violated if a statistically significant relationship between residuals and time is found.

RESULTS

In a total of 153 Serbian SALS patients, we identified 10 patients (6.54%) with the SOD1 p.Leu145Phe mutation and nine patients (5.88%) with the expansion of the GGGGCC repeat in the C9orf72 gene. No statistically significant association was found between CNVs of SMN1 (p = 0.650, Fisher’s exact test), SMN2 (p = 0.369), SERF1 (p = 0.257) or NAIP (p = 0.908) genes and SALS susceptibility in the examined group of Serbian patients (Table 2).

Median survival time was 22 months (Figure 1). The univariable Cox regression analysis revealed that the overall survival was statistically significantly influenced by the following protective factors: ALSFRS-R score (hazard ratio [HR] = 0.939, 95% CI 0.910–0.968, p = 1.415e-04) and FVC (HR = 0.975, 95% CI 0.964–0.987, p = 8.799e-05), both at the time of diagnosis, and diagnostic delay (HR = 0.955, 95% CI 0.921–0.990, p = 3.025e-03); and by the following risk factor: change of the ALSFRS-R score over time (ΔALSFRS: HR = 2.253, 95% CI 1.713–2.964, p = 4.886e-07) (Table 3). For the SMN1 copy number we observed a statistical trend (HR = 1.811, 95% CI 0.910–3.607, p = 0.091). However, ΔALSFRS showed the strongest effect explaining 24.3% of the survival variance.

In order to estimate the direction of each variable’s effect, the univariable Cox regression analysis was repeated by converting continuous data into categorical using the median split. The following protective strata were identified: ALSFRS-R score > 39 (HR = 0.299, 95% CI 0.161–0.555), FVC > 90% (HR = 0.249, 95% CI 0.138–0.450), diagnostic delay > 11 months (HR = 0.541, 95% CI 0.311–0.942), as well as the following risk stratum: ΔALSFRS > 0.9 (HR = 8.249, 95% CI 3.08–22.1).

Log-rank test showed that survival was significantly shorter in subgroups of older age (> 60 years, p = 1.7e-02), lower ALSFRS-R (≤ 39, p = 5e-05) and lower FVC.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Copy number</th>
<th>Patients</th>
<th>Controls</th>
<th>p-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMN1</td>
<td>1</td>
<td>6</td>
<td>9</td>
<td>0.650</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>134</td>
<td>134</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>13</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>6</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>SMN2</td>
<td>0</td>
<td>62</td>
<td>54</td>
<td>0.369</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>82</td>
<td>82</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>3</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>18</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>127</td>
<td>114</td>
<td></td>
</tr>
<tr>
<td>SERF1</td>
<td>0</td>
<td>5</td>
<td>6</td>
<td>0.257</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>11</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>119</td>
<td>115</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>20</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>3</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>NAIP</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0.908</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>119</td>
<td>115</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>20</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>3</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

*Fisher’s exact test
(≤ 90%, p = 7.1e-07) at the time of diagnosis, lower MMSE score (< 24, p = 4.3e-04), shorter diagnostic delay (< 11 months, p = 2.7e-02) and in a group with both upper and lower motor neurons affected (p = 2.6e-02). Even though differences in the survival of patients with different SMN1 copy number were not statistically significant (p = 0.242), their median survival times were different: those with one SMN1 copy did not reach the estimated median survival time of 50%, those with two SMN1 copies had median survival time of 22 months, while patients with three SMN1 copies had median survival time of 19 months.

The multivariable Cox regression analysis included only statistically significant variables from the univariable analysis (ALSFRS-R, FVC, ΔALSFRS, and diagnostic delay), with the addition of examined genetic factors (SMN1, SMN2, SERF1, and NAIP copy number). For a total of 90 SALS patients, out of whom 28 died, the values of all variables included in the model were available. The multivariable Cox regression analysis explained 41.2% of the survival variance in SALS patients and identified ALSFRS-R score at the time of diagnosis (HR = 0.885, 95% CI 0.834–0.938), change of the ALSFRS-R score over time (HR = 2.193, 95% CI 1.618–2.972), diagnostic delay (HR = 0.939, 95% CI 0.888–0.994) and SMN1 copy number (HR = 10.276, 95% CI 3.106–33.998) as significant independent predictors of survival (Table 4). Lower ALSFRS-R score at the time of diagnosis, faster decline of the overall functionality (ΔALSFRS), shorter diagnostic delay and higher number of SMN1 copies resulted in shorter survival of Serbian SALS patients. Patients with one SMN1 copy had survival probability of 100%, patients with two SMN1 copies had survival probability of 39.2%, while those with three SMN1 copies had the lowest survival probability of only 0.03% (Figure 2). The assumption of proportional hazards was not violated (for ALSFRS p = 0.325; for ΔALSFRS p = 0.088; for diagnostic delay p = 0.548; for SMN1 copy number p = 0.265; for the whole model p = 0.329), confirming the validity of the model.

**DISCUSSION**

About 12% of SALS patients had either SOD1 p.Leu145Phe or C9orf72 repeat expansion. High frequency of p.Leu145Phe mutation in our study and other cohorts from Southern Europe suggests a probable founder effect [17, 21]. On the other hand, the frequency of the C9orf72 repeat expansions in the Serbian SALS patients was in agreement with literature data [22]. Among 19 SALS patients with the SOD1 p.Leu145Phe mutation or the C9orf72 repeat expansion, three SMN1 copies were present in two SOD1-positive patients. Co-occurrence of a high-risk variant
with susceptibility variants has been reported in a small percent of patients from a large ALS cohort [22].

To our knowledge, this is the first study examining CNVs of genes in 5q13.2 duplication other than SMN as risk factors for SALS susceptibility and survival. Former studies took the NAIP gene into consideration; nonetheless, due to methodological limitations, they could only detect its homozygous deletion [23]. Despite the absence of association in our study, a recent study by Kano et al. [24] reported that the NAIP protein level in ALS patients was lower by nearly half compared to healthy controls. These latest findings and the implication of SERF1 CNVs in the SMA clinical outcome justify research efforts focused on the role of NAIP and SERF1 CNVs in SALS [15].

Lack of association between SMN1 and SMN2 copy numbers and SALS in Serbian patients compared to previous results might be, at least partially, explained by a different genetic architecture of diverse populations [12]. Meta-analyses have supported SMN1 duplication as a major risk factor for SALS susceptibility, but the analyzed samples mostly included patients from Western Europe [13, 14]. Additionally, in a recent paper, the SMN1 gene has been investigated as an ALS susceptibility gene with intermediate effect and its duplication was found in a larger number of sporadic cases (6.6%) compared to controls (3.4%) [22]. It is of note that the frequency of heterozygous deletion of SMN1 in our control group is twice as high as what is expected according to the literature (9/153 in our study, compared to 1/35) [25]. To our knowledge, this is the first study showing a higher frequency of one SMN1 copy in the controls than in SALS patients, which might explain the absence of association between one SMN1 copy and SALS susceptibility.

Clinical factors with the highest predictive strength on ALS survival are respiratory status and the ALSFRS-R score, particularly their change over time [26, 27]. However, prognosis based on these two clinical variables alone may be uncertain for clinical use given that motor neuron death may occur prior to the appearance of symptoms. Complementing clinical variables with objectively measurable time-invariant genetic factors may help develop better clinically useful algorithms.

Our results show that SMN1 copy number, together with the ALSFRS-R score at the time of diagnosis, its change over time, and diagnostic delay might be useful to predict survival in Serbian SALS patients. Survival curves demonstrate that three SMN1 copies result in a higher mortality rate compared to one copy. Previous studies also showed that SMN CNVs might influence the survival of SALS patients. Lower SMN1 copy number has been associated with an increased mortality rate in Dutch and French patients [10, 11]. The absence or decrease of SMN2 copy number has been identified as a negative prognostic factor for survival in Dutch patients [10, 28], while homozygous absence of SMN2 has been observed as a protective factor in Swedish patients [12]. These data additionally underline the significance of performing gene association studies in diverse populations. Although the mechanism by which the increased SMN1 copy number could contribute to shorter survival in ALS patients remains unknown, the study on zebrafish showed that the highest temporal requirement for SMN is in the earliest stages of development, while the study on mice showed that only moderate SMN levels are required in the adult central nervous system compared to neonatal mice [29, 30].

Our analysis revealed that the lower ALSFRS-R score at the time of diagnosis, faster change in the ALSFRS-R score over time, and shorter diagnostic delay result in higher mortality percentage, reflecting a greater overall functional impairment during the natural course of the disease. Additionally, our results support a better prognostic value of longitudinal than cross-sectional data for ALSFRS-R score regarding survival of SALS patients, given that the R² value of ΔALSFRS-R was twice as high as it was for ALSFRS-R at the time of diagnosis [26].

The limitation of our study is its relatively small sample size, which is a common occurrence in rare disease population studies, and the recruitment of a significantly larger SALS sample is hampered by the small size of the population of Serbia. Therefore, the solidity of our findings should be verified on a larger study sample from different ethnic origins, and meta-analyses. However, the present study puts an emphasis on a clinically well-characterized sample and on collecting follow-up data. It also offers insight into genetic architecture of the 5q13.2 duplication in SALS patients from a Southeast European population. Compared to qPCR mostly used in previous studies, MLPA utilizes one primer pair to amplify multiple genes of interest eliminating the error that might arise due to differences in PCR efficiency of target and normalization primers.

CONCLUSION

Our finding regarding the modulating role of SMN1 CNVs on the survival of Serbian SALS patients might lead to a modeling of a stronger, more reliable and practical prognostic algorithm in SALS, in which clinical predictive factors are complemented with more objective, time-invariant genetic factors. This may, in turn, enable more precise prognosis and more targeted clinical trials.

ACKNOWLEDGMENT

This study was supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia, under grant No. 173016.
REFERENCES


Број копија гена SMN1 као модификатор преживљавања код болесника из Србије са спорадичном формом амиотрофичне латералне склерозе

Милош Бркушанин1, Ирена Јефтовић-Велкова2, Владимир М. Јовановић1, Стојан Перић2, Јован Пешовић1, Горан Брајушковић1, Зорица Стевић2, Душанка Савић-Павићевић1
1Универзitet у Београду, Биолошки факултет, Центар за хуману молекуларну генетику, Београд, Србија;
2Универзitet у Београду, Медицински факултет, Клинички центар Србије, Клиника за неурологију, Београд, Србија;
Универзитет у Београду, Институт за биолошка истраживања „Синиша Станковић“, Одељење за генетичка истраживања, Београд, Србија

САЖЕТАК
Увод/Циљ Амиотрофична латерална склероза (АЛС) представља тешко обољење моторног неурона. Већина случајева је наизглед спорадична, са варијантама у високоризичним АЛС генима или генима асоциираним са АЛС. Два гена асоциирана са АЛС су SMN1, чије функционално одсуство узрокује спиналну мишићну атрофију (СМА), и њему готово идентичан ген SMN2, који је модifikатор фенотипа СМА. Оба гена смештена су у сегменталној дупликацији 5q13.2. Циљ студије био је испитивање асоцијације броја копија гена SMN1, SMN2, SERF1 или NAIP, смештенih у истом геномском региону, са ризиком за развој болести и преживљавањем код болесника из Србије са спорадичном формом АЛС (САЛС).

Методе Мултипла лигационо-зависна амплификација је коришћена за одређивање броја копија гена SMN1, SMN2, SERF1 и NAIP у клинички детаљно окарактерисаној групи од 153 болесника са САЛС из Србије и 153 контролне особе.

Резултати Асоцијација броја копија гена SMN1, SMN2, SERF1 или NAIP са ризиком за развој САЛС или преживљавањем није показана. Криве преживљавања заснована на мултиваријабилној Коксовој регресионој анализи показале су да три копије SMN1, нижи збир скале ALSFRS-R у тренутку дијагнозе, бржи пад збира ALSFRS-R током времена и краће дијагностичко кашњење резултују краћим преживљавањем болесника са САЛС из Србије.

Закључак Прогноза преживљавања болесника са САЛС заснована на клиничким параметрима може бити унапређена коришћењем генетичког параметра – броја копија SMN1. Кључне речи: преживљавање моторног неурона; амиотрофична латерална склерозе; H4F5; NAIP; SMN1

DOI: https://doi.org/10.2298/SARH180801069B