BIOCELL
2018 42(3): 87-91

ISSN 1667-5746 (on-line)
ISSN 0327-9545 (printed)

The impact of paralog genes: detection of copy number variation in spinal muscle atrophy patients

Sergio LAURITO[1, 2]; Juan A. CUETO[1, 3]; Jimena PEREZ[1]; María ROQUÉ[1, 2]

[1] IHEM, Universidad Nacional de Cuyo, CONICET
[2] Departamento de Biología, Universidad Nacional de Cuyo, Facultad de Ciencias Exactas y Naturales
[3] Instituto de Fisiología, Universidad Nacional de Cuyo, Facultad de Ciencias Médicas

Key words: SMA, CNV, MLPA, Paralog gene

Abstract: Spinal muscular atrophy (SMA) is caused by dysfunction of the alpha motor neurons of the spinal cord. It is an autosomal recessive disease associated to the SMN1 gene, located in the subtelomeric region of 5q13. A paralog SMN2 gene is located at the centromeric region of the same chromosome, which apparently originated by an ancestral inverted duplication occurring only in humans. The exon sequence differs in two nucleotides in exon 7 and exon 8, which leads to an SMN2 transcript that lacks exon 7 and results in a truncated protein. Part (10%) of the SMN2 transcripts avoids the splicing of exon 7 but most of the copies are dysfunctional. In a disease scenario, the more SMN2 copies the higher possibility to restore at least partly the effects of SMN1 deficiency. Some therapeutic approaches are being developed to increase the expression of SMN2. To determine the number of SMN1 and SMN2 copies, the methodology must distinguish accurately between both genes. In this work, we present the results obtained using multiplex ligation-dependent probe amplification (MLPA) in 60 SMA suspected patients/carriers derived from different regions of Argentina. In 32 of these DNA samples we found alterations in SMN1. Among these, 16 presented a heterozygous deletion (carrier status) and 14 an homozygous deletion (patient status) in exon 7 and 8 of SMN1. In one case, exon 7 was found homozygously deleted but exon 8 presented a single copy, and in another case, exon 7 was found heterozygously deleted while exon 8 was normal. Almost half of the patients (7/15) presented a normal diploid number of SMN2 while the other half (8/15) presented an increased number. In this work we showed how a probe-based methodology such as MLPA was able to distinguish between the paralog genes and determine the amount of copies in DNA samples from suspected patients/carriers of SMA.

Introduction

Spinal muscular atrophy (SMA) is the most common debilitating neuromuscular disease in children, together with Duchenne muscular dystrophy. It is caused by dysfunction and death of the alpha motor neurons of the spinal cord. This provokes in the patients a generalized muscle atrophy and weakness. The clinical phenotype is very heterogeneous, and ranging from severe to mild, it is divided in SMA type I, type II, type III (Prior *et al.,* 2010) and for some authors even in type IV (Prior *et al.,* 2010; Arnold and Fischbeck 2018). In 1997, the first description of a genetic cause was published, reporting the disease as hereditable in an autosomal recessive way (Lefebvre *et al.,* 1997). The majority of the patients (92%) have homozygous deletions in the survival motor neuron 1 (SMN1) gene located on the telomeric region of chromosome 5q13.2 (Verhaart *et al.,* 2017).

* Address correspondence to: María Roqué,
mroque@mendoza-conicet.gob.ar

This paper belongs to the 60th Anniversary Collection of the Instituto de Histología y Embriología de Mendoza (IHEM)

Only few available studies mention incidence and prevalence of the disease. The incidence of all types of SMA is cited as 1 in 10-12.000 live births (Lunn and Wang 2008; Verhaart *et al.,* 2017). The prevalence of the disease, which is the number of living diseased individuals, is observed as 1-5 per 100.000 individuals, even though a broad variability is reported among different geographic regions and a possible sample bias has to be considered because a small error in the number of cases has a large impact on the estimated prevalence (Lunn and Wang 2008). Besides, another difficulty to assess the exact prevalence is the median life expectancy of the patients, which differs from around one year for type I patients to more than 20 years for type II patients, and figure close to the normal population life expectancy for type III patients (Zerres *et al.,* 1997; Chung *et al.,* 2004; Finkel *et al.,* 2014).

Since SMA is a genetic recessive disease, there are also heterozygous carriers who are clinically unaffected but can transmit the mutation to their descendants. The carrier frequency is estimated to be higher among Caucasian and Asian populations, as compared to the African and Hispanic populations (Verhaart *et al.,* 2017), however the exact determination has not been established, as very severe forms

(death in utero) and very mild forms (symptom-free in adults) interfere with an accurate estimation (Verhaart *et al.,* 2017). Carriers can be divided based on the location of normal and mutated copies of SMN1: the most frequent form is to have a functional SMN1 allele in one chromosome 5 and a deleted allele in the homolog chromosome. Less common is to present two functional genes on the same chromosome 5 and no corresponding alleles on the homolog chromosome. Some cases, very rarely present three, four or more copies of SMN1, unequally distributed among the two homolog chromosomes.

SMN1 is located in an unstable subtelomeric region of 5q13 and it has an almost identical copy gene called SMN2, at the centromeric region of the same chromosome. This paralog gene has been originated ancestrally by an inverted duplication, which apparently occurred in humans, but not in chimpanzees (Rochette *et al.,* 2001). Interestingly, chimpanzees have many copies of the SMN gene, though they are all identical in their sequence. Instead, the human SMN2 gene has a slightly different sequence and has been detected in populations from Europe, the Central African Republic, and the Congo. There is evidence that SMN gene duplication occurred more than 5 million years ago, before the separation of human and chimpanzee lineages, but that SMN2 appears for the first time in *Homo sapiens* (Rochette *et al.,* 2001). The exon sequence of both genes differs only in two nucleotides in the coding region (one in exon 7 and one in exon 8), which alters the splicing pattern, leading to an SMN2 transcript that lacks exon 7 and results in a truncated, nonfunctional protein. However, part of the SMN2 transcripts manages to avoid the splicing of exon 7, but since this occurs in a very low rate (10%), most copies are dysfunctional. Therefore, in a scenario of deleted SMN1 gene, the absence cannot be entirely compensated by SMN2 (Arnold and Fischbeck 2018). The more SMN2 copies, the higher the possibility to partly restore the effects of SMN1 deficiency. Therefore, in the last years, it has become crucial in clinic to determine the SMN2 copy number in SMA patients, because a correlation has been established between the severity of the disease and SMN2 copy number: type I patients have a normal or reduced number of SMN2 copies, but most of the type II and III patients show an increased number of SMN2 copies (3-4 copies; Calucho *et al.,* 2018).

In the last two decades, substantial advances have been made in the understanding of the genetic basis of SMA (Finkel *et al.,* 2018; Mercuri *et al.,* 2018), which has encouraged to propose new therapeutic approaches that could contribute to delay the progression of the disease. Most of them are focused on the existence of the paralog SMN2 gene. Some of these worthy to mention approaches are the efforts to increase the expression of SMN2 by HDAC inhibitors (such as valproic acid; Piepers *et al.,* 2011), the antisense-oligonucleotide that binds to the SMN2 pre-mRNA downstream of exon 7, promoting its incorporation into the mRNA and leading to the translation of a fully functional SMN protein (Hua *et al.,* 2010), and small molecules that modulate the splicing of SMN2 (Messina 2018). These findings are demonstrating that therapeutic tools that act on SMN2 RNA are able to ameliorate the SMA phenotype (Nizzardo *et al.,* 2015; Wang *et al.,* 2018), and are therefore suggesting the importance of studying the amount of copies of the paralog genes in patients.

To determine the copy number of SMN1 and SMN2, the applied methodology must be able to distinguish accurately between both genes. Multiplex Ligation-dependent Probe Amplification (MLPA) assay is based on different probes which hybridize specifically and simultaneously on distinct genomic regions, and after a fluorescence-based PCR amplification the number of targeted probes can be separately quantified. In this work, we present the MLPA results obtained in our laboratory on SMA patients and carriers derived from different regions of Argentina.

Materials and Methods

Patients and samples

Blood samples of 60 SMA suspected patients/carriers were derived to our service of genetic analyses. All samples were accompanied by a medical order for SMA testing by MLPA. Clinical criteria were not always informed in the medical order. DNA was extracted from whole blood as mentioned before (Laurito and Roqué 2018). In brief, 3 ml of each EDTA-anticoagulated blood sample was washed with $T_{10}E_{10}$ Buffer (Tris-HCl 10 mM and EDTA 10 mM). Pellets were resuspended in 3 ml of CTAB solution (2% CTAB; 700 mM NaCl; 50 mM Tris-HCl, pH 8; 10 mM EDTA; and 2% β-mercapto-ethanol) and incubated at 60°C. The solutions were further subjected to chloroform/isoamyl alcohol extraction. Finally, the aqueous phase was transferred to a clean tube and DNA was precipitated with chilled absolute ethanol.

MLPA assay

This assay for the analysis of copy number variation (CNV) was performed as communicated before for other diseases (Laurito and Roqué 2018; Mayorga *et al.,* 2016; Gomez *et al.,* 2009; Marzese *et al.,* 2008; Laurito *et al.,* 2013; 2015). In brief, 50-100 ng of DNA was denatured at 95°C for 5 min and then incubated for 16 hours at 60°C with the P021/P060 probe mix (P021 for the first 4 analyses, and P060 for the remaining 56 samples). The difference between P021/P060 resides in the probe targets and probe design for SMN2, since P021 has probes flanking SMN1 and SMN2. Following probe hybridization, DNA ligase was added and incubated for 15 min at 54°C. The ligation products were then amplified by the PCR according to the manufacturer's protocol using one primer labeled with 6-FAM. PCR products were subsequently resolved using a CEQ 8000 Capillary Sequencer (Beckman Coulter). The data were analyzed by the Gene Marker version 1.75 software.

Results

Mutation detection

Among the 60 derived samples, we found alterations in SMN1 by MLPA in 32 (54%). Among these, 16 (50%) presented a heterozygous deletion in exons 7 and 8 of SMN1 (carrier status) and 14 (43%) an homozygous deletion (patient status). In one case, exon 7 was found homozygously deleted, while exon 8 presented a single copy (3%), and in another case exon 7 was found heterozygously deleted while exon 8 was found normal (3%) (Tab. 1).

TABLE 1
Types of alterations detected by MLPA in the SMA derived samples

		N	%
CNV Detected		32	100
SMN1	0 copies 0 copies	14	43.75
SMN1	0 copies 1 copy	1*	3.125
SMN1	1 copy 2 copies	1*	3.125
SMN1	1 copy 1 copy	16	50

TABLE 2
Double entrance table showing the copy numbers of both paralogous SMN genes in SMA patients and carriers

		SMN2				
		0 copies	2 copies	>2 copies	ND	TOTAL
SMN1	0 copies		7	8		15
	1 copy	3	4	8	2	17

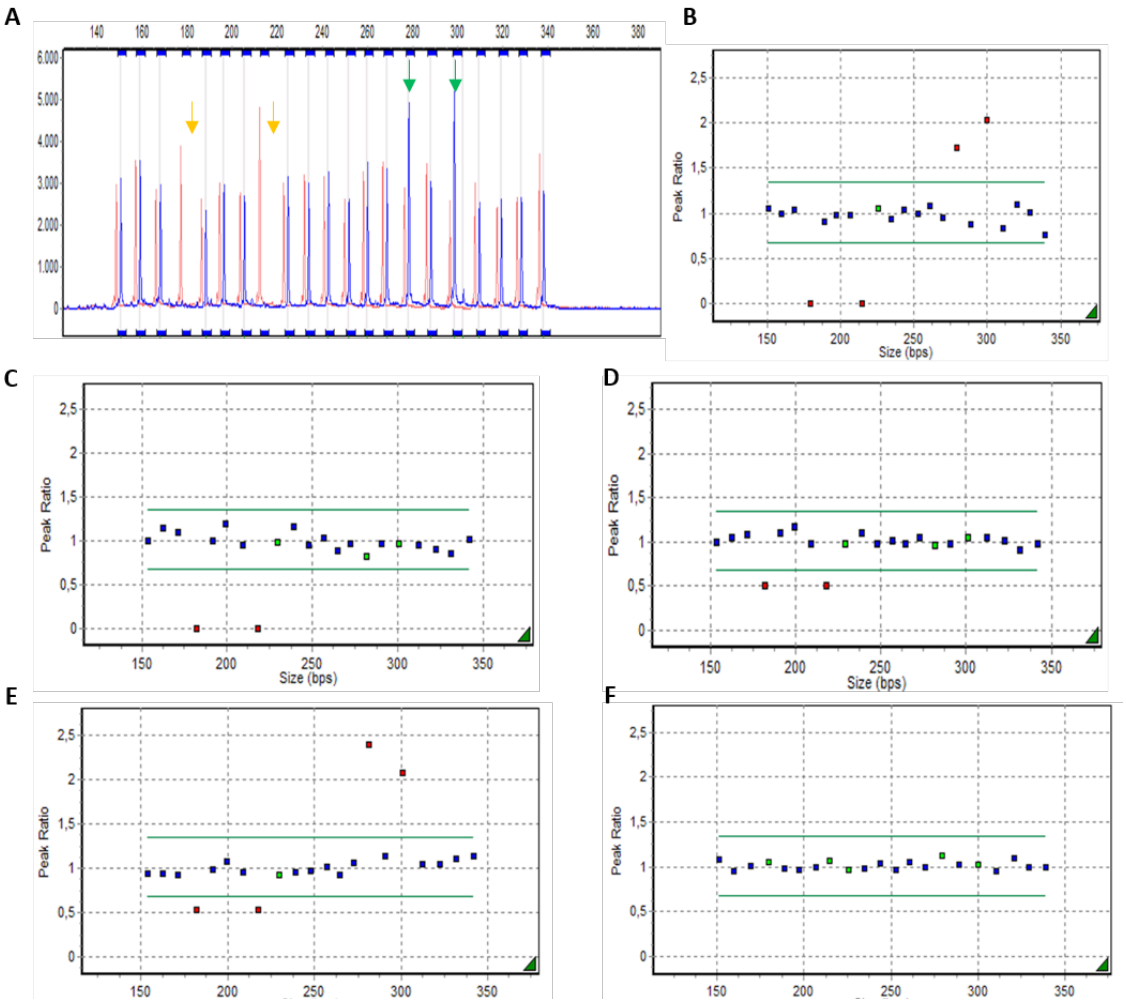

FIGURE 1. A. MLPA electropherogram comparing a DNA of a patient sample (blue peaks) and DNA control sample (red peaks). The orange arrows indicate a total absence of fluorescence for the probes corresponding to exons 7 and 8 in SMN1 gene in the patient sample, and an increase of fluorescence for the probes for exon 7 and 8 of SMN2 gene (green arrows). B-F. Ratio Analysis between the study/control sample, B shows a homozygous deletion for exon 7 and 8 of SMN1 gene (patient status), and an increment of copy numbers for exon 7 and 8 of SMN2 gene, C a homozygous deletion of SMN1 gene (patient status), with normal copy number of SMN2 gene. D a heterozygous deletion of SMN1 gene (carrier status) and normal copy number of SMN2 gene. E a heterozygous deletion of SMN1 gene (carrier status) and an increment of copy numbers of SMN2. F shows a normal copy number for SMN1 and SMN2.

Only exon 7 should be considered to determine the number of copies of SMN1 gene, because gene conversion between SMN1 and SMN2 exon 8 can occur in 5% of the samples, resulting in a chimeric gene containing the SMN1 exon 7 and the SMN2 exon 8 sequences. Such a hybrid gene results in a protein functionally identical to SMN1 protein (MLPA® MRC-Holland Product Description Probemix P060-B2 SMA). We suspect that both cases marked with an asterisk present an increased number of copies of exon 8, probably due to gene conversion.

Given the clinical importance of the amount of copies of SMN2, in Tab. 2 we crossed the SMN1 status in patients and carriers (0 copy and 1 copy) with the amount of copies of SMN2. Only exon 7 of both genes was considered, to avoid the possible interference of conversion between exons 8. As it can be observed, almost half of the patients (7, which have 0 copies of SMN1) have normal diploid number of SMN2, while the remaining 8 present an increased number. In these latter cases, a less severe phenotype and a possibly better response to treatments based on SMN2 RNA modulation is expected.

The electropherograms and ratio analyses obtained by the software GeneMarker v1.75 are shown in Fig. 1 to represent the different SMN1/2 statuses. Only 4 probes (2 for exon 7 and 8 of SMN1, and 2 for exon 7 and 8 of SMN2) were considered for the analysis, because the remaining ones are reference probes which hybridize in genomic regions not involved in the disease and serve as normal copy number controls.

Discussion

Gene duplication is an evolutionary source of new genes and variability, which has been maintained in nature over time (Zhang, 2013). Duplication of a genomic region can occur by error during DNA duplication, and once duplicated the copies can evolve differently, generating paralog genes and new proteins. Unlike ortholog genes, a paralog gene is a new gene that holds a new function. After duplication, two sub-functional proteins may be produced in some cases, which need from each other to cooperate (Baker *et al.,* 2013), while in other cases, a competitive interference may arise between the duplicates (Baker *et al.,* 2013).

In the case of SMN, the SMN2 gene produces 90% of a nonfunctional protein, but a small portion has a functionality similar to that of SMN1. Depending on the amount of SMN2 copies in the genome, the absence of SMN1 in SMA patients may be partially compensated. It has been reported that several variants in SMN2 exon 7 could act as a positive modifier in several patients who display milder clinical phenotypes than other patients with the same SMN2 copy number (Alías *et al.,* 2018). SMN protein is a 38-kDa protein that is ubiquitously expressed in the cytoplasm and nuclei of all tissues, involved in the metabolism and processing of RNA (Coovert *et al.,* 1997)

In this work we showed how a probe-based methodology such as MLPA was able to distinguish between the paralog genes and to determine the amount of copies in 60 studied DNA samples from suspected patients or carriers of SMA. The interpretation of the results is easy, and the costs of the study are reasonable for developing countries such as Argentina. The limit of the technique is the detection over more than 3-4 copies, since the PCR amplification there reaches a plateau. However, this seems not to be a limitation for its clinical uses, since more than 4 copies are rare events.

Therefore, MLPA can be considered a useful tool for diagnosis and copy number determination of the paralogous genes SMN1/2 in DNA obtained from circulating blood samples.

Authors Contributions

LAURITO (first author) and CUETO carried out the experiments. LAURITO and ROQUE carried out the data analysis. ROQUE wrote the manuscript with support of LAURITO, CUETO and PEREZ. All authors discussed the results and contributed to the final manuscript.

References

Alías L, Bernal S, Calucho M, Martínez E, March F, Gallano P, Fuentes-Prior P, Abuli A, Serra-Juhe C, Tizzano EF (2018). Utility of two *SMN1* variants to improve spinal muscular atrophy carrier diagnosis and genetic counselling. *European Journal of Human Genetics* **26**: 1554-1557**.**

Arnold ES, Fischbeck KH (2018). Spinal muscular atrophy. *Handbook of Clinical Neurology*, **148**: 591-601.

Baker CR, Hanson-Smith V, Johnson AD (2013). Following gene duplication, paralog interference constrains transcriptional circuit evolution. *Science* **342**: 104-108.

Calucho M, Bernal S, Alías L, March F, Venceslá A, Rodríguez-Álvarez FJ, Aller E, Fernández RM, Borrego S, Millán JM, Hernández-Chico C, Cuscó I, Fuentes-Prior P, Tizzano EF (2018). Correlation between SMA type and *SMN2* copy number revisited: An analysis of 625 unrelated Spanish patients and a compilation of 2834 reported cases. *Neuromuscular Disorders* **28**: 208-215.

Chung BHY, Wong VCN, Ip P (2004). Spinal muscular atrophy: survival pattern and functional status. *Pediatrics* **114**: e548-553.

Coovert DD, Le TT, McAndrew PE, Strasswimmer J, Crawford TO, Mendell JR, Coulson SE, Androphy EJ, Prior TW, Burghes AHM (1997). The survival motor neuron protein in spinal muscular atrophy. *Human Molecular Genetics* **6**: 1205-1214.

Finkel RS (2014). Observational study of spinal muscular atrophy type I and implications for clinical trials. *Neurology* **83**: 810-817.

Finkel RS (2018). Diagnosis and management of spinal muscular atrophy: Part 2: Pulmonary and acute care; medications, supplements and immunizations; other organ systems; and ethics. *Neuromuscular Disorders* **28**: 197-207.

Gomez LC, Marzese DM, Adi J, Bertani D, Ibarra J, Mol B, Vos IJ, De Marchi G, Roqué M (2009). MLPA mutation detection in Argentine HNPCC and FAP families. *Familial Cancer* **8**: 67-73.

Hua Y, Sahashi K, Hung G, Rigo F, Passini MA, Bennett CF, Krainer AR (2010). Antisense correction of SMN2 splicing in the CNS rescues necrosis in a type III SMA mouse model. *Genes & Development* **24**: 1634-1644.

Laurito S, Branham T, Herrero G, Marsa A, Garro F, Roqué M (2013). Detección de un caso de síndrome de Williams Beuren por MLPA [Detection of a Williams Beuren syndrome case by MLPA]. *Medicina (Buenos Aires)* **73**: 47-50.

Laurito S, Di Pierri J, Roqué M (2015). Neurofibromatosis tipo I. Mutación de *splicing* detectada por MLPA y secuenciación en la

Argentina [Neurofibromatosis type 1. splicing mutation detected by MLPA and DNA sequencing in Argentina] *Medicina (Buenos Aires)* **75**: 91-94.

Laurito S and Roqué M (2018). Análisis de variación del número de copias y de patrones de metilación en la región 15q11-q13 [Variation analysis of the number of copies and methylene patterns in region 15q11-q13]. *Medicina (Buenos Aires)* **78**: 1-5.

Lefebvre S, Burlet P, Liu Q, Bertrandy S, Clermont O, Munnich A, Dreyfuss G, Melki J (1997). Correlation between severity and SMN protein level in spinal muscular atrophy. *Nature Genetics* **16**: 265-269.

Lunn MR, Wang CH (2008). Spinal muscular atrophy. *The Lancet* **371**: 2120-2133.

Marzese DM, Mampel A, Gomez LC, Echeverría MI, Vargas Al, Ferreiro V, Giliberto F, Roqué M (2008). Detection of deletions and duplications in the Duchenne muscular dystrophy gene by the molecular method MLPA in the first Argentine affected families. *Genetics and Molecular Research* **7**: 223-233.

Mayorga L, Laurito S, Loos MA, Eiroa HD, Pinho S de, Lubieniecki F, Arroyo HA, Pereyra MF, Kauffman MA, Roqué M (2016). Mitochondrial DNA deletions detected by Multiplex Ligation-dependent Probe Amplification. *Mitochondrial DNA* **27**: 2864-2867.

Mercuri E, Finkel RS, Muntoni F, Wirth B, Montes J, Main M, Mazzone ES, Vitale M, Snyder B, Quijano-Roy S, Bertini E, Hurst Davis R, Meyer OH, Simonds AK, Schroth MK, Graham RJ, Kirschner J, Iannacone S, Crawford TO, Woods S, Qianf Y, Sejersen T (2018). Diagnosis and management of spinal muscular atrophy: Part 1: Recommendations for diagnosis, rehabilitation, orthopedic and nutritional care. *Neuromuscular Disorders* **28**: 103-115.

Messina S (2018). New directions for SMA therapy. *Journal of Clinical Medicine* **7**: 251 abstr.

MRC-Holland. Product Description SALSA®MLPA® Probemix P060-B2 SMA. *http://www.mrc-holland.com/webforms.*

Nizzardo M, Simone C, Dametti S, Salani S, Ulzi G, Rizzo F, Frattini E, Pagani F, Bresolin N, Comi G, Corti S (2015). Spinal muscular atrophy phenotype is ameliorated in human motor neurons by SMN increase via different novel RNA therapeutic approaches. *Scientific Reports* **5**: 11746.

Piepers S, Cobben JM, Sodaar P, Jansen MD, Wadman RI, Meester-Delver A, Pol-The BT, Lemmink HH, Wokke JHJ, van der Pol W-L, van den Berg LH (2011). Quantification of SMN protein in leucocytes from spinal muscular atrophy patients: effects of treatment with valproic acid. *Journal of Neurology, Neurosurgery and Psychiatry* **82**: 850-852.

Prior TW, Snyder P, Rink BD, Pearl DK, Pyatt RE, Mihal DC, Conlan T, Schmalz B,Montgomery L, Ziegler K, Noonan C, Hashimoto S, Garner S (2010). Newborn and carrier screening for spinal muscular atrophy. *American Journal of Medical Genetics Part A* **152A**: 1608-1616.

Rochette C, Gilbert N, Simard L (2001). *SMN* gene duplication and the emergence of the *SMN2* gene occurred in distinct hominids: *SMN2* is unique to *Homo sapiens*. *Human Genetics* **108**: 255-266.

Verhaart IEC, Robertson A, Wilson IJ, Aartsma-Rus A, Cameron S, Jones CC, Cook SF, Lochmüller H (2017). Prevalence, incidence and carrier frequency of 5q-linked spinal muscular atrophy-a literature review. *Orphanet Journal of Rare Diseases* **12**: 124.

Wang J, Schultz PG, Johnson KA (2018). Mechanistic studies of a small-molecule modulator of SMN2 splicing. *Proceedings of the National Academy of Sciences of the United States of America* **115**: E4604-4612.

Zerres K, Rudnik-Schöneborn S, Forrest E, Lusakowska A, Borkowska J, Hausmanowa-Petrusewicz I (1997). A collaborative study on the natural history of childhood and juvenile onset proximal spinal muscular atrophy (Type II and III SMA): 569 Patients. *Journal of the Neurological Sciences* **146**: 67-72.