UAB
Universitat Autònoma de Barcelona

# Adapted application of NGS to improve genetic characterization of tuberous sclerosis complex and spinal muscular atrophy patients

**Laura Blasco Pérez**

**ADVERTIMENT.** L'accés als continguts d'aquesta tesi queda condicionat a l'acceptació de les condicions d'ús establertes per la següent llicència Creative Commons: https://creativecommons.org/licenses/?lang=ca

**ADVERTENCIA.** El acceso a los contenidos de esta tesis queda condicionado a la aceptación de las condiciones de uso establecidas por la siguiente licencia Creative Commons: https://creativecommons.org/licenses/?lang=es

**WARNING.** The access to the contents of this doctoral thesis it is limited to the acceptance of the use conditions set by the following Creative Commons license: https://creativecommons.org/licenses/?lang=en

# Adapted application of NGS to improve genetic characterization of tuberous sclerosis complex and spinal muscular atrophy patients

Laura Blasco Pérez

DOCTORAL THESIS - 2023

THESIS DIRECTORS

**Dr. Eduardo Tizzano Ferrari**

**Dra. Elena García Arumí**

THESIS TUTOR

**Dr. Francisco Rodríguez Frías**

Vall d'Hebron Institut de Recerca

Medicine Genetics Group,
Vall d'Hebron Research Institute (VHIR)
Department of Clinical and Molecular Genetics,
Hospital Vall d'Hebron

UAB Universitat Autònoma de Barcelona

Programa de Doctorat en Bioquímica, Biologia Molecular i Biomedicina
Departament de Bioquímica i de Biologia Molecular, Universitat Autònoma de Barcelona

"Girls should never be afraid to be smart"

Emma Watson

A mis padres,

a mi hermano,

a David.

# ABSTRACT

Next-generation sequencing (NGS) has supposed a radical change in diagnosis of genetic diseases, as it allows a massively parallel sequencing of large DNA regions in multiple patients in the same experiment. In the field of tuberous sclerosis complex (TSC), the use of NGS achieves the genetic diagnosis in 85-90% of patients. To improve the diagnostic yield, we have developed a workflow focused on the identification of low-frequency and deep intronic variants. Its application has allowed us to diagnose 29 patients, including 5 patients with low frequency variants (allele frequency <20%) and 2 with deep intronic variants. In the field of spinal muscular atrophy (SMA), about 95% of patients present homozygous deletion of *SMN1* gene as the cause of the disease. However, there is a large phenotypic variability between patients. Copy number of *SMN2*, an almost identical paralog gene of *SMN1*, is considered the main phenotypic modifier of the disease as *SMN2* is partially functional and compensates to some extent the production of SMN protein. To explore the genetic variability of *SMN2* in SMA patients, we have developed a NGS-based method that enables the deep study of the SMN region, beyond *SMN1* and *SMN2* copy number determination. Its application in 368 SMA patients allow us to detect modifiers variants, *SMN2-SMN1* hybrid structures and analysis the consensus regions of *SMN2* modifier therapies. The adapted application of the NGS technique employed in this work, improved the genetic diagnostic yield in TSC disease and the genetic characterization of the SMN region in SMA patients.

# INDEX

# INTRODUCTION

# 1. Tuberous sclerosis complex (TSC)

Tuberous sclerosis complex (TSC, OMIM #191100 and #613254) is a multisystem neurocutaneous genetic disorder characterized by the presence of hamartomas in multiple organs[1].

## 1.1 Epidemiology

TSC disease was first described by Friedrich Daniel von Recklinghausen in 1862. He reported a baby with cardiac tumors and sclerotic areas in the brain. Almost 20 years later, in 1880, Désiré-Magloire Bourneville provided a more detailed neurological characterization of the disease and described it as "tuberous sclerosis of the cerebral convolutions", where the current name comes from. It took decades to describe all the multisystemic symptoms associated with the disease and, in 1988, the first TSC diagnostic criteria were developed. They were particularly relevant as molecular diagnosis was not yet possible[2,3].

As early as 1885 the disease was considered to have a genetic base, but it was not until 1987 that Fryer established the locus of the disease in chromosome 9. It was shortly demonstrated that many TSC patients did not show linkage to this locus and, in 1993, *TSC2* (OMIM #605284) gene was discovered on chromosome 16p13.3 and also considered responsible for the disease. Finally, in 1997, *TSC1* (OMIM #191092) gene was identified at chromosome 9q34. A few years later it became possible to study the *TSC1* and *TSC2* genes for routine molecular diagnosis[2,4,5]. Tuberous sclerosis complex incidence was recently estimated to range from 1:5800 to 1:13,520 live births[6,7].

## 1.2 Clinical manifestations

Tuberous sclerosis complex is a rare disease characterized by the presence of hamartomas in multiple organs such as skin, brain, kidney, lung, heart, and bone[1]. Hamartomas are benign tumors consisting of disorganized tissue and, although

most are benign, they can cause morbidity by affecting the function of the organ in which they are located[8].

Criteria for clinically diagnosing a patient with TSC were first established in 1998 and updated in 2012 and 2021[1,9,10]. According to the latest, a "definite TSC" diagnosis in a patient is reached when the patient presents two major features or one major and two minor features, while a patient is considered as "possible TSC" if their show either one major feature or two or more minor features (Table 1).

| Major criteria | Minor criteria |
|---|---|
| Hypomelanotic macules (≥3) | "Confetti" skin lesions |
| Angiofibroma (≥3) or fibrous cephalic plaque | Dental enamel pits (≥3) |
| Ungual fibromas (≥2) | Intraoral fibromas (≥2) |
| Shagreen patch | Retinal achromic patch |
| Multiple retinal hamartomas | Multiple renal cysts |
| Multiple cortical tubers and/or radial migration lines | Nonrenal hamartomas |
| Subependymal nodule (≥2) | Sclerotic bone lesions |
| Subependymal giant cell astrocytoma | |
| Cardiac rhabdomyoma | |
| Lymphangiomyomatosis (LAM) | |
| Angiomyolipomas (≥2) | |

**Table 1. Diagnostic criteria for patients with suspected TSC.** Definite TSC is considered in a patient with 2 major features or 1 major feature with 2 minor features. Possible TSC is considered in a patient with either 1 major feature or ≥2 minor features. A pathogenic variant in *TSC1* or *TSC2* is sufficient to confirm the diagnosis of TSC.
*As an exception, the presence of lymphangiomyomatosis (LAM) and angiomyolipomas without other features, does not meet criteria for a definite diagnosis. Adapted from Northrup et al. 2021[1].

The main factors of morbidity in TSC are cerebral, renal and pulmonary alterations[11]. The characteristic brain alterations are cortical tubers, radial migration lines, subependymal nodules and subependymal giant cell astrocytomas (SEGAs). All these lesions can be detected by MRI. These structural alterations cause focal seizures or epileptic spasms in for 63-78% of patients[12]. In addition, around half of

the patients present intellectual disability ranging from moderate to severe. Indeed, very often patients show TSC-associated neuropsychiatric disorders (TAND), which include behavioral, psychiatric, neuropsychological, and psychosocial difficulties, as well as intellectual disability[1,2,13]. Regarding renal involvement, angiolipomas are found in approximately 75% of cases and those of larger size (>4 cm) can result in hemorrhage and renal failure. This has been described as the most common cause of early death in TSC patients' thus surveillance should include abdominal MRI and renal function assessment[14]. About the pulmonary system, the classic lesion is lymphangiomyomatosis (LAM), which can lead to shortness of breath, recurrent pneumothoracices and even death. This symptom is particularly common in adult women and can be detected by a chest computed tomography (CT)[11,13].

Other organs affected in TSC but with less morbidity are skin, heart and, less frequently, eyes (with multiple retinal hamartomas), teeth (with dental enamel pits) and bones (with sclerotic lesions). Specifically, skin is one of the most typically affected organs, although these alterations do not usually lead to severe health problems. Patients could present a variety of features including hypomelanotic macules, facial angiofibromas, a fibrous cephalic plaque, ungual or periungual fibromas, and a Shagreen patch. The most common are the hypomelanotic macules, found in 90% of patients, and the facial angiofibromas, reported in about three fourths of them[1,2]. Interestingly, newborns with TSC present cardiac rhabdomyomas in at least 50% of cases but they rarely develop in adults. Such lesions do not usually cause symptoms in patients, but depending on the size and location, they may increase the risk of heart failure, so an echocardiogram is recommended in childhood.

Penetrance of the disease is considered to be complete, as to date, it has not been reported any patient with a pathogenic variant in *TSC1* or *TSC2* that was asymptomatic after a careful evaluation[15]. In contrast, the expressivity is highly variable between patients, even within individuals of the same family. This can be explained by two main factors; firstly, Knudson's "second hit" phenomenon and, secondly, by the presence of mosaic variants. The "second hit" theory describes that

for a tumor development, the cell with one affected allele must be mutated to lose the function of the other allele, initially functional. This makes the growth and proliferation of tumors highly variable. In addition, mosaic variants have been widely described as a cause of TSC, resulting in patients with milder phenotypes or confined to a specific tissue or body area[15,16].

## 1.3 Genetics of TSC

Tuberous sclerosis complex is a genetic disorder with autosomal dominant inheritance. At the molecular level, it is caused by pathogenic variants in the *TSC1* and *TSC2* genes[4,5].

On the one hand, the *TSC1* gene is located in the long arm of chromosome 9 and it consists of 53.3 kb divided into 23 exons and 22 introns. The reading frame starts in the middle of exon 3 and ends at the beginning of exon 23, making a coding region of 3495 base pairs. *TSC1* encodes for hamartin, a 130kDa protein consisting of 1164 amino acids with three main conserved domains. Specifically, the protein presents an N-terminal domain, a C-terminal domain, and a central coiled-coil domain encoded by exons 15 to 19, necessary for binding tuberin (Figure 1). On the other hand, the *TSC2* gene is located in the short arm of chromosome 16 and it comprises 41.5 kb consisting of 42 exons and 41 introns. The reading frame extends from exon 2 to the beginning of exon 42, for a total of 5424 coding base pairs. *TSC2* codes for tuberin, a 200kDa protein of 1807 amino acids that contains a conserved N-terminal domain (exons 2-12) responsible for binding hamartin, and a GTPase-activating protein (GAP) domain (exons 36-40) providing the protein's function (Figure 1)[17,18].

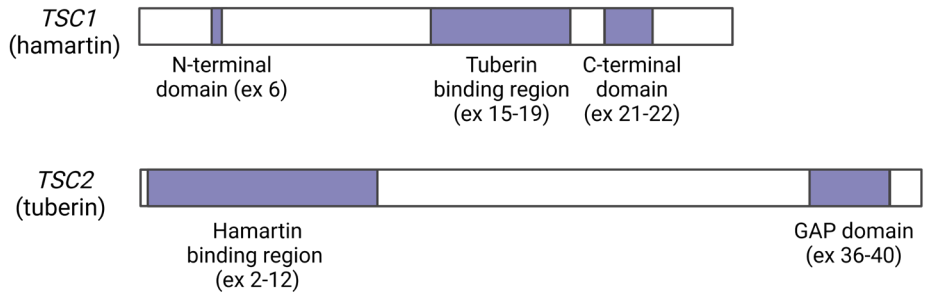



**Figure 1. Structural and functional domains of *TSC1* and *TSC2*.** *TSC1* contains an N-terminal domain, a C-terminal domain and a coiled-coil domain codified by exons 15 to 19 and necessary for the interaction with tuberin. *TSC2* contains a N-terminal domain (exons 2 to 12) essential for hamartin binding and a GTPase-activating protein (GAP) domain codified by exons 36 to 40 and responsible for the protein function. Adapted from Rosengren et al. 2020[18].

Both proteins bind together to form the TSC1-TSC2 complex, in which tuberin carries out the enzymatic function and hamartin stabilizes tuberin, avoiding its degradation. The main role of TSC complex is to inhibit Rheb (Ras homologue enriched in brain) protein, downregulating the mTORC1 (mammalian target of rapamycin complex 1) pathway and, therefore, inhibiting cell growth and proliferation. This is why *TSC1* and *TSC2* are considered tumor suppressor genes and the alteration of their function leads to increased cell proliferation and growth of benign tumors (Figure 2)[19–21].

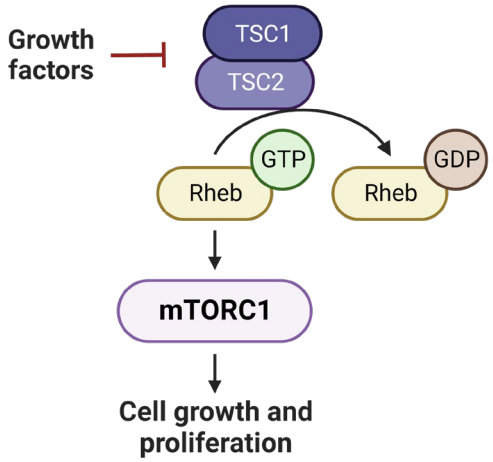



**Figure 2. Scheme of Tuberous Sclerosis Complex signaling.** TSC1-TSC2 complex inhibits Rheb (Ras homologue enriched in brain) through its GAP domain, thereby inhibiting mTORC1 and cell growth and proliferation.

By the end of 2022, around 3600 *TSC1* and *TSC2* variants had been reported as pathogenic or likely pathogenic in the Leiden Open Variation Database (LOVD)[22]. Approximately a quarter of them are in *TSC1* and the rest in *TSC2*, which agrees with the ratio of patients with pathogenic variants in *TSC1* and *TSC2*, 26% and 74%, respectively [23–25]. In both cases the variants are distributed along the entire gene and include single nucleotide variants, small deletions and insertions, and large deletions or rearrangements[15,22,26]. Regarding the origin of the variants, it has been reported that about two thirds of them are *de novo* and the rest are inherited from a symptomatic progenitor[6]. However, a third scenario is possible and that is a germline mosaicism in one of the progenitors. Indeed, mosaic variants are relatively common as the cause of TSC, even variants at very low frequency (allele frequency (AF) <1%) have been described[24].

## 1.4 TSC molecular diagnosis and therapies

As discussed above, to reach a clinical diagnosis of TSC a patient must meet Northup's clinical criteria. Nevertheless, the detection of a pathogenic variant in *TSC1* or *TSC2* is enough to confirm a molecular diagnosis of TSC, regardless of the clinical manifestations[15].

The routine diagnostic workflow usually includes sequencing of *TSC1* and *TSC2* exons, and copy number variant (CNVs) study in these genes. Sequencing can be carried out by Sanger sequencing or by NGS, the latter being increasingly common. In any case, both exons and exon-intron boundaries are usually sequenced in order to also detect splicing variants at canonical positions. About CNVs detection, the most commonly used techniques are multiplex ligation-dependent probe amplification (MLPA) or NGS with specific bioinformatics tools[27]. After applying these techniques, 10-25% of clinically diagnosed patients remain negative or NMI (no mutation identified)[25,28,29]. The most expanded hypothesis is that these cases are due to deep intronic or low-frequency mosaic variants, both of which are undetectable by these routine methods[24,30].

Once the diagnosis of TSC is established, either clinically or molecularly, disease-specific treatment can be administered. At present, therapy is based on treating the symptoms generated by the disease, generally hamartomas in different tissues. The main drugs used are mTOR inhibitors to reduce tumor growth and, if necessary, surgery to remove it. They may also require specific medication for seizures or TAND[1,15].

## 2. Spinal muscular atrophy (SMA)

Spinal muscular atrophy (SMA) is a severe neuromuscular disorder characterized by a progressive proximal muscles weakness and atrophy caused by degeneration of the alpha motor neurons in the spinal cord[31].

### 2.1 Epidemiology

SMA disease was first described in 1891 by the scientist Guido Werdnig, who reported two patients with muscle weakness developed around 10 months. Shortly after, in 1893, Johann Hoffmann added seven cases with a similar phenotype. Both already observed, in their patients, degeneration of the anterior horn cells from the spinal cord, distinctive in SMA disease. Consequently, the severe infantile form of SMA is traditionally referred to as Werdnig-Hoffmann's disease[32].

In the mid-1900s, almost simultaneously, the group of Wohlfart and the one of Kugelberg and Welander documented a mild ambulant form of Werdnig-Hoffmann disease. As a result, the mild form of spinal muscular atrophy is also called Wohlfart-Kugelberg-Welander disease[32–34]. Soon after, in the 1960s, Dubowitz reported 12 cases with an intermediate form. These patients were characterized by achieving the ability to sit but never become able to walk[35]. This led to the understanding of the three forms as a single disease with high variability and a continuous phenotypic spectrum.

It was not until 1995 that the SMA determining gene was identified by Lefebvre's group and called the survival motor neuron *(SMN)* gene, later renamed as *SMN1*

(OMIM #600354). They also describe the presence of an almost identical centromeric gene (later called *SMN2;* OMIM #601627) with similar structure than *SMN1* and only a few discrepancies, including two nucleotides in exons seven and eight. Studies at cDNA level already showed that *SMN2* gene undergoes alternative splicing of exon seven, producing a transcript lacking this exon[36].

Spinal muscular atrophy is the second most common recessive genetic disease of infancy and early childhood. The estimated pan-ethnic incidence is 1 in 11,000 live births and the worldwide carrier frequency is 1/54, ranging from 1/47 in the Caucasian population to 1/72 in the African-American population[37,38]. These high frequencies, together with the emergence of SMA therapies, explain why in recent years newborn screening (NBS) programs are being implemented in many countries to detect the disease as early as possible[39–41].

## 2.2 Clinical manifestations

Spinal muscular atrophy is characterized by progressive muscle atrophy, leading to weakness. This affectation is typically symmetrical, focused especially in the lower limbs and predominant in proximal muscles, although it also affects the axial, intercostal and bulbar muscles[31,32]. The muscle weakness that occurs in SMA patients is a consequence of the deficiency of the survival motor neuron (SMN) protein, encoded mainly by *SMN1* gene (see details in section 2.3). The human SMN protein has an essential role in the spliceosomal snRNP (small nuclear ribonucleoproteins) biogenesis and, although is ubiquitously expressed, alpha motor neurons in the spinal cord are the most affected cells to SMN deficit[42,43]. This leads to muscle denervation and, finally, muscular atrophy. To date, it is unclear why alpha motor neurons are more sensitive that the rest of cells and it is still an active field of investigation[44,45].

Both the severity of symptoms and age of onset are highly variable in SMA disease. This results in patients with neonatal symptoms and a life expectancy of few weeks, to patients who develop normal motor skills and start undergoing weakness in young or middle adulthood.

### *2.2.1 Clinical classification of SMA patients*

Historically, three SMA forms have been considered; Werdnig-Hoffmann, Dubowitz and Wohlfahrt-Kugelberg-Welander, depending on the phenotype severity. In 1992, at the International Consortium on Spinal Muscular Atrophy, a formalized classification in three SMA types was stablished[46]. The criteria for this categorization of patients were age of onset and maximum milestones achieved. Later on, this classification was expanded introducing type IV patients and adding subdivisions into types I, II and III (Figure 3)[47,48].

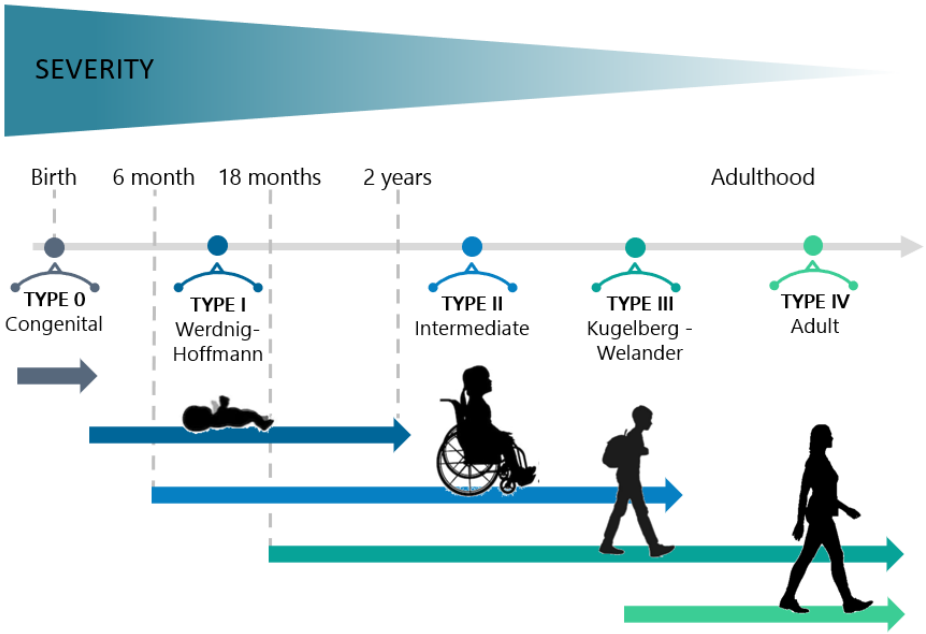



**Figure 3. Classification of SMA patients.** Patients are classified from more to less severity between type 0 (or Ia) and IV according to their age of onset and maximum milestones achieved.

SMA Type I or Werdnig-Hoffmann disease, the most severe form, appear in the first six months of life. These infants present "frog-leg posture" in supine position, indicating a generalized hypotonia. They are not able to roll over or to sit independently, being referred as "nonsitters". They also show bell-shaped thorax and diaphragmatic breathing. Early in their life, they develop feeding difficulties, with high risk of aspiration pneumonia. The respiratory failure is the most common

cause of death, usually before the age of two years without treatment. It could be subdivided in three subtypes depending on the onset of the symptoms. Infants with prenatal onset or in the first week of life are considerate Type IA (or type 0). It is the most severe presentation, with joint contractures, cardiopathy and early respiratory failure in the first weeks of life. If symptoms appear between the first week and three months of age, they are classified as type IB. They are characterized by never achieving cephalic control. Lastly, Type IC patients develop symptoms after three months of age, achieve cephalic control but are unable to sit independently (Table 2)[48–50].

| SMA Type | Onset | Milestones achieved |
|---|---|---|
| **0/Ia** | Prenatal | None |
| **Ib** | <3 months | Poor or none cephalic control |
| **Ic** | >3 months | Cephalic control |
| **IIa** | >6 months | Sitting position |
| **IIb** | >12 months | Sitting and bipedestation, but no independent gait |
| **IIIa** | >18 months | Walk independently, but lose it at an early age |
| **IIIb** | >36 months | Walk independently, until adolescence |
| **IV** | Second/third decade | Walk for most of their life |

**Table 2. SMA patients' subclassification according to their onset and achieved milestones.** Adapted from Serra-juhe et al. 2019[49].

The intermediate SMA, referred to as Type II or Dubowitz disease, presents its onset between six and 18 months of life. Patients with SMA type II are characterized by being able to sit down but never achieve the ability to walk independently, so they are confined to a wheelchair. These patients are called "sitters" and, depending on their motor milestones, we differentiate two subgroups. Patients with SMA Type IIA, considered weak sitters, debuted after six months of life and will lose the ability to stay seated. In contrast, the strong sitters, classified as SMA Type IIB, have the onset around 12 months, maintain the ability to sit independently and achieve bipedestation, but not independent gait. SMA Type II children show proximal weakness, especially in the lower limbs, hypotonia and areflexia. In the absence of treatment, they develop scoliosis and contractures, causing respiratory

complications that may require noninvasive respiratory support. Overall, life expectancy of these patients generally exceeds 25 years.

SMA Type III or Kugelberg–Welander disease is the milder form, in which patients are named "walkers". Children with this SMA form attain the ability to walk without support, although without treatment they will lose it eventually. Based on the age of onset, patients are subdivided in two subgroups. When first symptoms appear between 18 months and three years, patients are classified as Type IIIA. They tend to loss deambulation in childhood. In contrast, patients with Type IIIB present the disease onset after three years of age and retain the ability to walk into adolescence or even adulthood. At the onset, the symptoms are milder than in other forms of SMA. For example, the characteristic muscle weakness in the lower limbs is present, but reflexes are preserved. This makes the clinical diagnosis more complex and can be confused with other neuromuscular diseases such as Becker muscular dystrophy. It has been observed that life expectancy of SMA Type III patients is similar to that of the general population[51].

Lastly, there is the adult form of the disease, classified as Type IV. It is a very uncommon form of SMA in which patients show slowly progressive muscle weakness in the second or third decade of life. It is considered underdiagnosed due to its mild presentation since, in the less severe cases, the symptoms allow them to develop a virtually normal life, without major complications[52].

In addition, since the implementation of neonatal screening programs, a new type of SMA has to be considered: cases genetically confirmed in the presymptomatic or paucisymptomatic status[49].

## 2.3 Genetics of SMA

Spinal muscular atrophy is a genetic disorder with autosomal recessive inheritance. At the molecular level, it is caused by the loss of both copies of *SMN1* gene, in the 5q13 locus[36].

### 2.3.1 SMN locus

The structure of *SMN1* gene consists of nine exons, historically named from 1 to 8 and including exons 2a and 2b. It is located in a highly repetitive region that also contains *SMN2*, its nearly identical paralogous gene result of a segmental duplication. *SMN2* is located in centromeric position regarding *SMN1* and its orientation remains unclear, as it was originally described to be in the opposite orientation to *SMN1* (head-to-head), but further information suggests that they could be in the same direction[36,53,54]. Moreover, the fact that *SMN2* is exclusively present in humans indicates that the duplication of *SMN1* occurred recently in time.

Due to this duplication event, the region containing *SMN* genes is highly repetitive and dynamic. This leads to a high rate of genomic rearrangements such as deletions, duplications and gene conversions. Indeed, in the general population, it is relatively common to present more than two *SMN1* copies, while *SMN2* copy number is widely variable, being two the most common[55,56]. In addition, *SMN2-SMN1* hybrids genes have been widely reported, being the most detected structure the one consisting of exon 7 of *SMN2* and exon 8 of *SMN1* (see details in section 2.3.3)[57–59]. Finally, a partial *SMN* gene lacking exons 7 and 8 (*SMN1/2Δ*7-8) has been reported in patients and controls. Although its frequency varies greatly (0.3-21%) between different populations, its structure is conserved, presenting the breakpoint always in intron 6. This suggests that all *SMN1/2Δ*7-8 detected have a common ancestral origin, however, it is unclear whether it comes from *SMN1*, *SMN2* or whether both structures exist[60–62].

### 2.3.2 SMN1 gene: determinant of SMA

Spinal muscular atrophy is caused by the lack of functional *SMN1* gene. In about 95% of cases, the disease is due to the homozygous absence of exon 7 of *SMN1*, a fact consistent with the high genomic instability of the region as mentioned above. Of these, 90% show a complete deletion of *SMN1* while in the other 5% exon 8 of *SMN1* is detected, indicating that they present *SMN2-SMN1* hybrids genes as the cause of SMA[58,59].

Accordingly, only 5% of the positive cases present pathogenic variants in *SMN1* as the cause of the disease. Nearly all of them are compound heterozygous cases including one *SMN1* deletion and one pathogenic variant. Homozygous cases for a pathogenic variant are extremely rare and usually linked to familial consanguinity[52]. More than 80 pathogenic variants have been described in *SMN1* all over the gene but, especially in exons 3 and 6. The small deletion c.399_402delAGAG; p.(Arg133fs*15) is particularly common in Spanish patients while c.815A>G; p.(Tyr272Cys) variant is the most recurrent globally[57,59,63,64].

The *SMN1* gene encodes for the motor neuron survival (SMN) a 294 amino acid protein. The SMN protein presents different functional domains that allow interaction with several other proteins and itself to carry out its function in the assembly of the spliceosomal complex of proteins. The main domains are the Tudor domain, responsible for binding Sm proteins, and the conserved YG-box (a tyrosine-glycine-rich region) domain, required for SMN self-oligomerisation[42–44,65]. These domains are located in exons 3 and 6 respectively, corresponding with those exons with more described pathogenic variants[57,59,64].

### *2.3.3 SMN2 gene: modifier of SMA*

To explain how the same alteration (absence of functional *SMN1*) leads to such a variable expressivity among patients, it is necessary to focus on *SMN2* gene, main modifier of the SMA disease described so far[62,66,67].

As previously mentioned, the *SMN1* and *SMN2* paralogous genes are virtually identical, differing only in few positions called paralogous sequence variants (PSVs)[68]. One of these PSVs explains why *SMN2* does not have the same function as *SMN1*, and this is because the PSV c.840C>T, a silent transition in exon 7, alters the splicing pattern in most *SMN2* pre-mRNA transcripts. This leads to a SMN protein lacking exon 7 (SMN-Δ7) that is nonfunctional, unstable and rapidly degraded[69,70]. However, *SMN2* preserves part of its function, since it is able to generate functional full-length protein in 10-15% of the cases (Figure 4)[71–73]. This small amount of

protein is crucial for survival, having never described a case without both *SMN1* and *SMN2* [36,55].

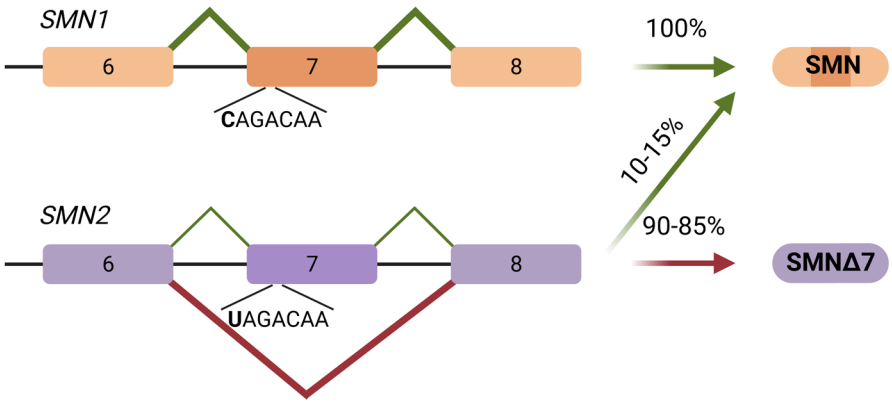



**Figure 4. Representation of alternative splicing in *SMN2*.** In the position c.840 of *SMN1* there is a cytosine (C), while in *SMN2* there is a thymine (T). This change in the sequence causes in the 90-85% of the times the skipping of exon 7, generating an incomplete protein. Only in the 10-15% of the times the splicing is done correctly and full-length protein SMN in generated. Adapted from Cartegni el al. 2002[74].

The partial function of the *SMN2* gene makes the copy number of *SMN2* (*SMN2*_CN) the main modifier of the disease, since the higher number of *SMN2*, the larger the complete protein produced and milder the SMA phenotype. Following this rule, it is possible to predict the SMA type of patients base on their *SMN2*_CN (Figure 5). However, this correlation is not absolute and discordant patients have been described in the literature, further classified as better-than-expected or worse-than-expected phenotypes according to their *SMN2*_CN[55,62].

Beyond *SMN2* copy number, there are two variants in the gene considered as positive modifiers of SMA phenotype. The first to be described was the variant c.859G>C; p.(Gly287Arg) located in exon 7 of *SMN2*. It creates a novel exonic splicing enhancer site that increase *SMN2* exon 7 inclusion and, therefore, one copy of *SMN2* with the variant will generate more functional SMN protein than a wild type *SMN2* [67,75]. The second modifier variant described was c.835-44A>G in intron 6 which, as in the previous case, enhances the inclusion of exon 7, producing more

full-length protein. Interestingly, this variant is one of the PSVs that defers between *SMN1* and *SMN2*, being the guanine the nucleotide typical from *SMN1* [76]. Other variants in *SMN2* have been proposed to be positive or negative disease modifiers, however, so far none have been validated with functional studies[62,77].

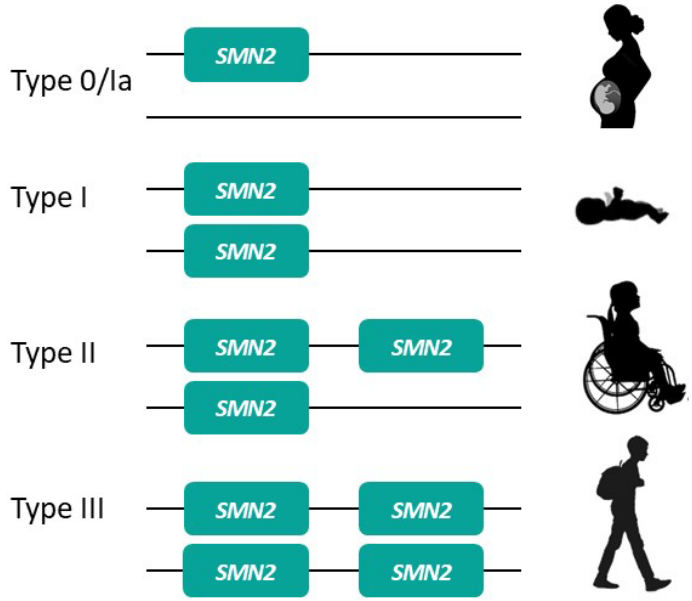



**Figure 5. Expected SMA type base on *SMN2*_CN.** Patients with 1 copy of *SMN2* are expected to present type 0 or Ia phenotype, with 2 *SMN2* copies they are predicted to develop type I SMA (non-sitter), with 3 *SMN2* copies patients will present a type 2 phenotype (sitters), and with 4 copies of *SMN2* they will show a type 3, being walkers at least part of their lives.

Lastly, hybrid genes consisting in exon 7 of *SMN2* and exon 8 of *SMN1* are present in about 5-10% of SMA patients[59,78]. Other hybrid structures have been reported but are less often detected due to technical limitations[58]. The involvement of *SMN2-SMN1* hybrid genes in SMA phenotype remains unknown, although they seem to have a positive effect[79].

## 2.4 SMA molecular diagnosis and therapies

Molecular confirmation of SMA and establishment of *SMN2* copy number is essential for the clinical management of patients, as this will be crucial for the selection of the most appropriate therapy.

### *2.4.1 Methods for SMA molecular diagnosis*

Using routine techniques, around 99% of SMA patients are molecular diagnosed. As discussed above, about 95% of SMA cases are due to *SMN1* exon 7 homozygous deletion, being 90% caused by *SMN1* complete deletions and 5% by *SMN2-SMN1* hybrid genes[58,59]. Accordingly, the first approach to the routine diagnosis of SMA is to determine the *SMN1* copy number (*SMN1*_CN). The principal techniques that can be used for this purpose are MLPA[80], quantitative PCR (qPCR)[81,82], digital droplet PCR (ddPCR)[83], and PCR plus capillary electrophoresis (PCR-CE)[84]. All of them have in common the use of the PSVs to make an *SMN1*-specific method, although each has its own limitations[85]. MLPA and PCR-CE are the most informative techniques, as they provide additional data besides *SMN1* exon 7 copy number. Among others, MLPA allows the detection of the most reported *SMN2-SMN1* hybrid gene, partial deletions and silent carrier variants (see below in this section). In contrast, PCR-CE could establish the copy number of exon 7 of *SMN1* and *SMN2* and detect some hybrid structures, based on probes in exons and introns 7. Additionally, PCR-CE can detect the modifier variant c.859C>G of *SMN2* and silent carrier variants[85].

When a patient presents a heterozygous deletion of *SMN1*, it is necessary to study pathogenic variants in the remaining allele. As the majority of described pathogenic variants are located in exons 3 and 6, the most common approach in routine laboratories is to sequence these two exons of *SMN* genes (*SMN1*-specific amplification is not available). When the patient with suspected SMA presents two copies of *SMN1*, alternative reasons for their symptoms should be explored[57,59,64].

A similar approach is applied to the study of SMA carriers, although some aspects must be considered. Most carriers are detectable by *SMN1*_CN study, as they present one wild type (wt) *SMN1* copy and a deletion of the gene on the other allele (1+0; Figure 6B). However, a small portion of carriers are undetectable by these methods and named as "silent carriers". This group comprises two types of carriers: 1) those with a wt *SMN1* and a *SMN1* pathogenic variant in the other allele (1+1 null; Figure 6C); and 2) those with two copies of *SMN1* in cis and a *SMN1* deletion in the other allele (2+0; Figure 6D) (Alias2018). In both cases the *SMN1*_CN result

would be of two copies, so they would not be detected as carriers by *SMN1*_CN study, although they present a null allele. Then, *SMN1* sequencing is performed, detecting (1+1 null) carriers, but (2+0) carriers would remain undetected. One way to detect them is to assess the presence of the variants c.*3+80T>G and c.*211_212del, which have been associated with the allele with two *SMN1* copies in cis and named "silent carrier variants"[84].

Once a diagnosis of SMA has been confirmed in a patient, it is important to study the copy number of *SMN2*, as it is considered to be the main disease modifier. The methods applied are the same as those used to determine *SMN1*_CN but the interpretation can be more challenging, especially in cases with three or more *SMN2* genes, since the sensitivity of the techniques is lower[86]. *SMN2*_CN accurate determination is extremely important as this will determine which therapy the patient will be eligible for (see details in section 2.4.2). Additionally, testing for the positive modifier variants c.859C>G and c.835-44A>G of *SMN2* is increasingly common in routine diagnostics, as it provides useful information on the patient's prognosis.

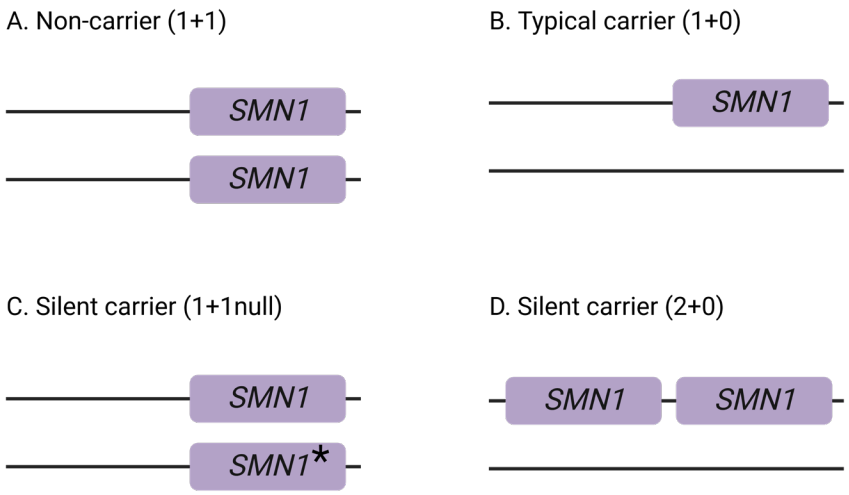



**Figure 6. Type of SMA carriers.** A) Non-carriers usually presented two *SMN1* copies in trans. B) Typical carriers presents one *SMN1* gene and a *SMN1* deletion in the other allele. C) Silent carriers (1+1 null) presented a wt *SMN1* and, in the other allele, a *SMN1* gene with a pathogenic variant (represented as an asterisk). D) Silent carriers (2+0) show an allele with two *SMN1* copies in cis and a *SMN1* deletion in the other. Adapted from Milligan et al. 2022[84].

### 2.4.2 New therapies for SMA

In the last decade, the therapeutic landscape in SMA disease has experienced a revolution due to the approval of three drugs for SMA. Prior to this, the management of SMA patients consisted of a multidisciplinary approach to improve disease-related symptoms, focusing on pulmonary, nutritional, rehabilitation and orthopedic, and palliative care. The implementation of the Consensus statement for standard of care in SMA, published in 2007, led to great improvements in the natural history of all the SMA types, including an increase on quality of life and survival[87–89].

After years of basic research and SMA animal models, several drugs for SMA were developed and, in the 2010s, clinical trials in patients were started. Nusinersen (Spinraza®) was the first treatment for SMA to be approved, in 2016 by the FDA (Food and Drug Administration) in United States, and in 2017 by the EMA (European Medicines Agency) in Europe. This drug consists on an antisense oligonucleotide that binds to *SMN2*, specifically to the intronic splicing silencer N1 (ISSN1) of intron 7 (Figure 7). Its binding displaces negative splicing factors and thereby increases the inclusion of *SMN2* exon 7 as well as the amount of full-lenght SMN protein produced for each copy of the endogenous *SMN2* gene. It is approved for use in patients with two or three *SMN2* copies and is administered intrathecally in a loading doses, in two months, followed by sustained dose every four months[90–92].

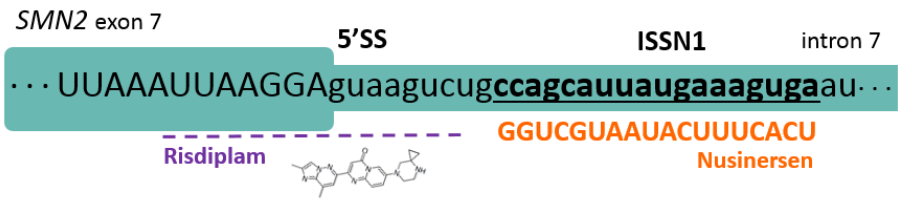



**Figure 7. Scheme of Nusinersen and Risdiplam binding regions.** Nusinersen (in orange) is an oligonucleotide complementary to the intronic splicing silencer N1 (ISSN1) of intron 7 of *SMN2*. Risdiplam (in purple) is a molecule that binds exonic splicing enhancer 2 (ESE2) and on the 5'-splite site (5'-ss) on exon 7 of *SMN2.*

Shortly thereafter, the AVXS-101 (ZolgenSMA®) was approved by the FDA in 2019, and by the EMA in 2020. It is a gene therapy consisting on a self-complementary adeno-associated virus 9 with the coding part of *SMN1*. A single intravenous administration provides a continuous production of exogenous SMN protein in the patient. Due to the recommended dosage of vector genomes (vg) per kg of body weight, the gene therapy is only approved in patients under two years. In addition, the presence of antibodies to the virus or a *SMN2* copy number different from two or three, contraindicates drug administration[90,93,94].

Lastly, the oral drug Risdiplam (Evrysdi®) was approved by the FDA in 2020 and by the EMA in 2022. It is a small molecule derivate from pyrimidines that binds to the exonic splicing enhancer 2 (ESE2) and on the 5'-splite site (5'-ss) on exon 7 of *SMN2* gene (Figure 7). Similarly to Nusinersen, modifies the splicing of the gene, increasing the inclusion of the exon 7 and, therefore, improving the production of full-length SMN protein by the endogenous *SMN2*. To date, it is approved to patients older than two years, and unlike the other two drugs, it is approved to patients with two, three or four copies of *SMN2* [90,95,96].

The emergence of SMA therapies, together with the implementation of the standard of care in SMA patients, have completely changed the SMA scenario in the last few years. Patients treated in the early stages of the disease extend their lifespan and significantly improved their motor milestones, while adult patients, treated at more advanced stages, achieve only stabilization of their symptoms[91,93,96,97]. This highlighted the importance of administering treatment as soon as possible, leading to the gradual implementation of new born screening worldwide.

In the current scenario, molecular diagnosis of patients and *SMN2*_CN study have become more relevant than ever. Both are crucial for the administration of therapies, as well as for assessing their effectiveness, since, despite being in the minority, cases of patients who do not respond to therapies (non-responders) have been described.

# 3. Next-generation sequencing (NGS)

The term next-generation sequencing (NGS) includes several technologies that allow massively parallel and deep sequencing of DNA. The emergence of these techniques revolutionized the field of genomics by dramatically reducing the cost and time of DNA sequencing[98].

## 3.1 Evolution of DNA sequencing

The structure of DNA was discovered in 1953, but it was not until 1977 that the first DNA sequencing techniques were developed[99–102]. The most relevant of these "first-generation sequencing" methods was Sanger sequencing, based on the addition of radioactively labeled chain-terminating dideoxynucleotides and fragment analysis to determine the DNA sequence. Sanger sequencing was optimized over the years, replacing radioactivity with fluorescence and automatizing part of the process, among others. These advances allowed the sequencing of the human genome for the first time in 2004, but it cost 13 years and $2.7 billion. These data highlighted the need for new techniques to further advance in genomic sequencing studies[103–105].

In the mid 2000's, the "next-generation sequencing" techniques appeared. One of the main advantages over Sanger sequencing was the possibility of sequencing different DNA molecules in parallel, increasing throughput and reducing costs. These "second-generation" techniques have in common the massive parallel sequencing of short DNA fragments (250-800 bp) and the three-steps workflow, consisting in library preparation, sequencing and bioinformatics analysis of the data. Several NGS technologies were developed, such as pyrosequencing (454 sequencing from Roche)[106] or ion semiconductor sequencing (Ion Torrent by Life Technologies)[107], but the technique that has become more relevant over time is the sequencing-by-synthesis (SBS) with fluorescent nucleotides and reversible terminators (by Illumina)[104,108].

## 3.2 NGS based on sequencing-by-synthesis

Illumina's technology is based on SBS with fluorescent reversible terminators. This method uses terminator nucleotides, similar those used in Sanger sequencing, so that elongation stops after the addition of each base. An image is then taken and the added nucleotides are modified allowing to incorporate one more base (more details above)[104,109].

However, prior to sequencing it is necessary to generate the libraries, that is, to prepare the target DNA to be sequenced in the NGS platform. For this purpose, it is necessary to define the target DNA, which in general can be one or several genes (gene panel), all exons of the genome (Whole Exome Sequencing, WES) or the entire genome (Whole Genome Sequencing, WGS). Two main approaches can be used for library preparation: PCR amplification (amplicon-based) or hybridization enrichment (Figure 8)[110].

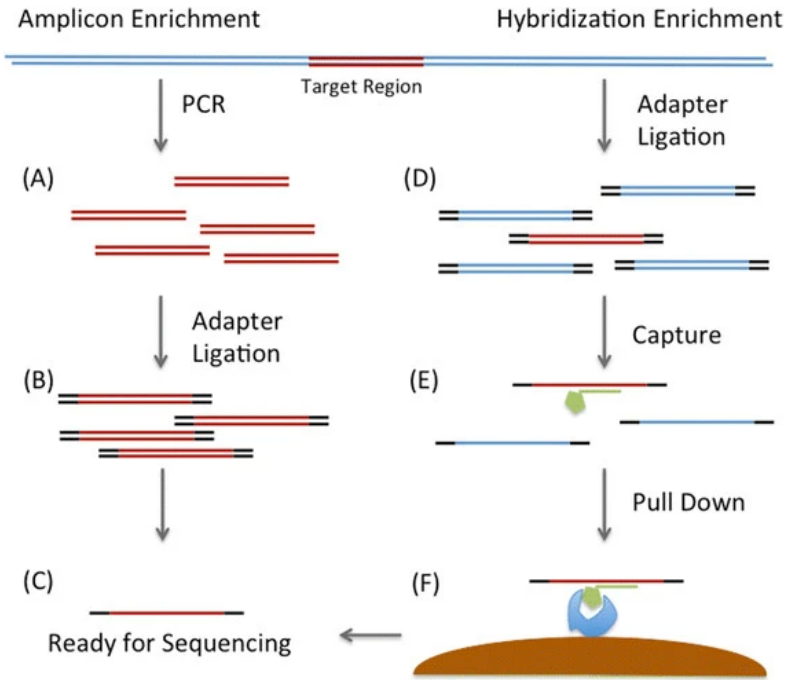



**Figure 8. Representation of library generation approaches.** On the left, is represented the amplicon-based method (A-B). On the right, hybridization capture procure is represented (D-F)[110].

In amplicon-based method, several primers are designed in order to amplify all the target regions by a multiple PCR reaction (Figure 8A). Then adaptor (necessary for sequencing) and index (necessary for sample identification) are added to PCR products, generating the library (Figure 8B). In the second case, genomic DNA (gDNA) is fragmented and adaptor and index are added, generating a "preliminary library" that contains all the gDNA (Figure 8D). The "preliminary library" is then hybridized to biotinylated probes complementary to the target DNA (Figure 8E). Subsequently, the probes are captured by streptavidin-coated magnetic beads, isolating the fragments containing target DNA and obtaining the libraries (Figure 8F). In both cases, extra amplifications steps are usually part of the protocol to increase the amount of DNA library generated[104,111].

After libraries have been generated, sequencing can be carried out. The first step is the DNA library amplification, which in Illumina's system is done by bridge amplification. The DNA library is hybridized to the flow cell, a glass slide that presents in its surface thousands of oligos complementary to the adapter. Then, the DNA fragments folds and the other adapter region hybridizes to another oligo from the surface. After amplification cycles, groups of molecules identical to each other are generated, called clusters. Thousands of clusters are generated in the flow cell (Figure 9.1). At this point, starts the sequencing-by-synthesis by the addition of fluorescently labeled nucleotides, each base with a different fluorophore. These nucleotides are modified to prevent elongation, so in each sequencing cycle only one nucleotide is added. This allows to stimulate the fluorophores and take an image to know which base has been incorporated in each cluster. Then the fluorophore is eliminated and the last nucleotide is modified to allow elongation and another sequencing cycle starts (Figure 9.2). After the sequencing process is complete, the indexes are read in a similar way. Indexes are small DNA sequences added to the target DNA during library generation and used to identify different DNA samples. Then DNA clusters folds again and the molecules are read in the complementary chain following the same procedure. The generation of clusters results in each initial molecule of the library being read multiple times, thus

achieving a very high coverage. All the images are collected and analyzed by bioinformatics tools to transform it into text and the sequences are aligned against the reference genome to variant identification[104,109,112].

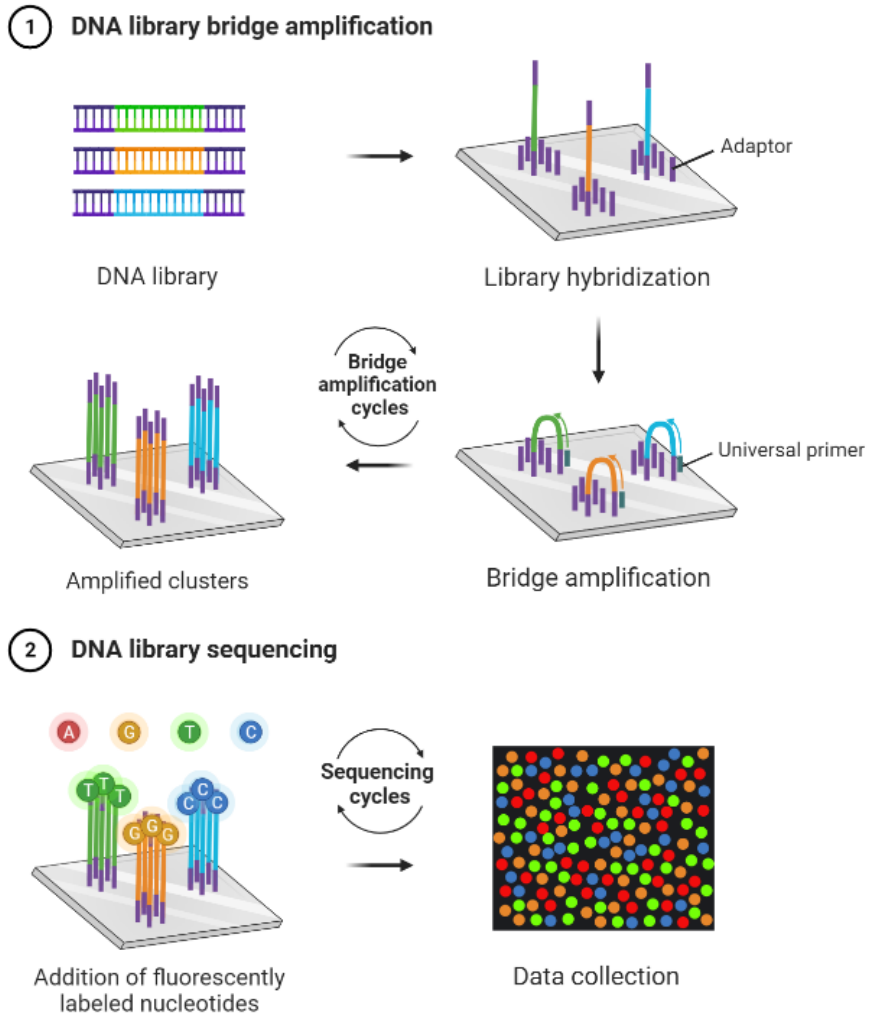



**Figure 9. Steps of bridge amplification and fluorophore-base sequencing.**

## 3.3 NGS impact on genetic diagnosis

Next-generation sequencing was initially developed with the objective of achieve large-scale sequencing, but its application soon proved to be much broader. Compared to Sanger sequencing, the high-throughput massive parallel sequencing enables the sequencing of several genes of multiple patients in the same

experiment. This resulted in a considerable decrease in both sequencing costs and time, thus establishing NGS, currently, as a standard method for genetic diagnostics. Depending on the case, NGS can be applied through gene panels (in which a group of genes, usually associated with the same phenotype, are studied), whole-exome sequencing or even whole-genome sequencing, each one being more expensive and challenging to analyze than the previous one. Thanks to the improvement of technology and bioinformatics tools in recent years, WES implementation in clinical practice is increasingly common. However, WES only covers about 2% of the entire genome, therefore, in inconclusive cases, WGS is applied but it is much more expensive and complex and remains mainly for research purposes[113–115].

In the field of tuberous sclerosis complex, the application of NGS is the goal standard, together with MLPA to determine CNVs. Nevertheless, in 10-25% of patients with clinical manifestations the molecular cause is not found, probably because the disease is caused by low allelic frequency mosaic variants or by deep intronic variants. Both will be undetected using standard NGS approaches, as normally coverage is not high enough to detect these mosaics and only exons and intron-exon boundaries are sequenced[24,30].

In routine molecular diagnosis of spinal muscular atrophy, the use of NGS is uncommon. This is probably due to the complexity of the region and the high homology of *SMN1* (causative gene) with *SMN2* (modifier gene). However, NGS is recently being used in an adapted way, such as the in-depth study of *SMN2* in SMA patients or the development of bioinformatics tools that allow the determination of *SMN1* and *SMN2* copy number based on WGS data[61,62].

# HYPOTHESIS

# Hypothesis

About a decade ago, the field of genetic diagnostics radically changed with the development of next-generation sequencing. NGS can be adapted for the study of rare diseases with special features, such as genes with deep intronic or low frequency mosaic candidate variants or genes in regions with high homology. Specifically, the thesis focuses on the rare diseases of TSC and SMA, with the following hypotheses:

- In TSC disease, the 10-25% of patients remain molecularly undiagnosed by routine methods (exon sequencing and CNVs study of *TSC1* and *TSC2*). The complete sequencing of the *TSC1* and *TSC2* genes by high-coverage NGS will allow the detection of the causal variants in these patients, which includes deep intronic and low AF variants.

- SMA patients present a very broad phenotypic spectrum. Routinely, *SMN2* copy number alone is used to predict the patient's phenotype but the correlation is not perfect. The study of the *SMN2* gene by NGS in SMA patients would allow the detection of new genetic markers of the disease or description of *SMN2* structures that could improve the existing genotype-phenotype correlation.

# AIMS

# Aims

Following our main hypothesis, the main objective of the thesis is to achieve further characterization of the genes causing TSC and SMA by an adapted NGS approach.

The specific objectives proposed in the field of TSC are:

1. Establish the optimal workflow and the most appropriate diagnostic tools for molecular diagnosis of TSC, in order to improve the diagnostic rate in the disease.
2. Identify the causal variant of TSC disease in clinically diagnosed patients by an adapted NGS approach, allowing detection of deep intronic and low AF variants in *TSC1* and *TSC2* genes.

The specific objectives proposed in the field of SMA are:

1. Develop an NGS-based methodology that allows the complete and deep analysis of *SMN2* genes in SMA patients.
2. Characterize the *SMN2* genes of SMA patients with the modifier variant c.859C>G of *SMN2*.
3. To study the variability of *SMN2* genes in SMA patients, such as variants or structures, to generate a data base that could be applied to study phenotype-genotype correlations.
4. To determine if the specific binding region of *SMN2* targeted therapies are conserved or show variability in SMA patients.

# RESULTS

## Summary of the results according to the aims of the thesis

The results obtained in this thesis are compiled in three main manuscripts (Chapters 1 to 3), which represent the major contributions of the thesis. Unpublished results related to a specific aim are included in the last part of results section (Chapter 4). In addition, the contribution of the PHD student to other articles published during the development of this thesis and not directly related to its aims are included as appendices (Appendix 1 to 4).

In the field of TSC:

1. Establish the optimal workflow and the most appropriate diagnostic tools for molecular diagnosis of TSC, in order to improve the diagnostic rate in the disease.
2. Identify the causal variant of TSC disease in clinically diagnosed patients by an adapted NGS approach, allowing detection of deep intronic and low AF variants in *TSC1* and *TSC2* genes.

**Chapter 1: An integral approach to the molecular diagnosis of tuberous sclerosis complex: the role of mosaicism and splicing variants**

**Laura Blasco-Pérez**, Leticia Iranzo-Nuez, Ricard López-Ortega, Desirée Martínez-Cruz, María Camprodon-Gómez, Anna Tenés, María Antolín, Eduardo F. Tizzano*, Elena García-Arumí*. Submitted to The Journal of Molecular Diagnostics; under review.

In the field of Spinal muscular atrophy:

1. Develop an NGS-based methodology that allows the complete and deep analysis of *SMN2* genes in SMA patients.

**Chapter 2**: **Beyond copy number: A new, rapid, and versatile method for sequencing the entire *SMN2* gene in SMA patients**

**Blasco-Pérez L**, Paramonov I, Leno J, Bernal S, Alias L, Fuentes-Prior P, Cuscó I, Tizzano EF. Hum Mutat. 2021 Jun;42(6):787-795. doi: 10.1002/humu.24200.

2. Characterize the *SMN2* genes of SMA patients with the modifier variant c.859C>G of *SMN2*.

**Chapter 3: Deep molecular characterization of milder spinal muscular atrophy patients carrying the c.859G>C variant in *SMN2***

**Blasco-Pérez L**, Costa-Roger M, Leno-Colorado J, Bernal S, Alias L, Codina-Solà M, Martínez-Cruz D, Castiglioni C, Bertini E, Travaglini L, Millán JM, Aller E, Sotoca J, Juntas R, Hoei-Hansen CE, Moreno-Escribano A, Guillén-Navarro E, Costa-Comellas L, Munell F, Boronat S, Rojas-García R, Povedano M, Cuscó I, Tizzano EF. Int J Mol Sci. 2022 Jul 27;23(15):8289. doi: 10.3390/ijms23158289.

3. To study the variability of *SMN2* genes in SMA patients, such as variants or structures, to generate a data base that could be applied to study phenotype-genotype correlations.
4. To determine if the specific binding region of *SMN2* targeted therapies are conserved or show variability in SMA patients.

**Chapter 4 (ongoing work): Towards implementation of a genomic data base of *SMN2* genes in SMA patients through complete sequencing of SMN locus.**

## Summary of results

This thesis aims to achieve further characterization of the genes causing TSC and SMA by an adapted NGS approach.

In TSC, 10-25% of patients remain undiagnosed after applying routinely techniques. **Chapter 1** presents the development of a workflow specially designed to allow detection of low-frequency and deep intronic variants. The workflow consists in high coverage NGS of the complete *TSC* genes together with MLPA for CNV study. Following this approach in our cohort enables the molecular diagnosis of 29 patients, including four previously classified as NMI.

On the other hand, molecular diagnosis in SMA is achieved in 99% of the suspected SMA patients. However, phenotype is highly variable between patients. The study of *SMN2* gene, the main modifier of the disease, helps to predict the expected phenotype. Specifically, the higher number of *SMN2* copies, the milder the SMA phenotype. But this correlation is not perfect and variant in *SMN2* have been described as modifiers of the disease. In **Chapter 2** we developed an NGS-based method that allows the deep characterization of the SMN region in SMA patients. The validation of the technique in 56 samples confirms its utility to asses *SMN1* and *SMN2* copy number, detection of *SMN2-SMN1* hybrid genes and SNV study of the region. Moreover, in **Chapter 3**, the method was successfully used to carefully study the *SMN2* genes of 11 SMA patients carrying the modifier variant c.859C>G of *SMN2*. This study allowed to identified a haplotype always associated with the modifier variant and a second haplotype possibly originated from the first one, pointing towards a common ancestral origin of all the alleles carrying the c.859C>G variant. Finally, **Chapter 4** presents the ongoing work, the application of the method in 313 new SMA samples so far. This has allowed us to update the list of PSVs between *SMN1* and *SMN2* to 15 positions, to detect hybrid structures in 41 patients and to study the genetic variability in SMA patients of the binding region of *SMN2* modulating drugs.

## Report of the thesis directors on the contribution of the PhD candidate

As directors of the doctoral thesis carried out by Laura Blasco Pérez, entitled "Adapted application of NGS to improve genetic characterization of tuberous sclerosis complex and spinal muscular atrophy patients", we certify her active participation in the design, experimental task, analysis of the results obtained and their discussion, formulation of conclusion and preparation of the manuscripts included in this thesis.

The specific contribution of the PhD candidate for each article are listed below, along with the impact factors (IF) of the journals. None of the included articles have been presented as part of other doctoral theses.

**Article 1: An integral approach to the molecular diagnosis of tuberous sclerosis complex: the role of mosaicism and splicing variants**

Laura Blasco-Pérez, Leticia Iranzo-Nuez, Ricard López-Ortega, Desirée Martínez-Cruz, María Camprodon-Gómez, Anna Tenés, María Antolín, Eduardo F. Tizzano*, Elena García-Arumí*.

Submitted to The Journal of Molecular Diagnostics; under review.

**Contribution:** conceptualization, methodology, formal analysis, investigation, data curation, writing - original draft, writing - review & editing.

**Article 2**: **Beyond copy number: A new, rapid, and versatile method for sequencing the entire *SMN2* gene in SMA patients**

Blasco-Pérez L, Paramonov I, Leno J, Bernal S, Alias L, Fuentes-Prior P, Cuscó I, Tizzano EF.

Hum Mutat. 2021 Jun;42(6):787-795. doi: 10.1002/humu.24200.

**IF= 4.27**; **Contribution:** conceptualization, methodology, formal analysis, investigation, data curation, writing - original draft, writing - review & editing.

**Article 3: Deep molecular characterization of milder spinal muscular atrophy patients carrying the c.859G>C variant in *SMN2***

Blasco-Pérez L, Costa-Roger M, Leno-Colorado J, Bernal S, Alias L, Codina-Solà M, Martínez-Cruz D, Castiglioni C, Bertini E, Travaglini L, Millán JM, Aller E, Sotoca J, Juntas R, Hoei-Hansen CE, Moreno-Escribano A, Guillén-Navarro E, Costa-Comellas L, Munell F, Boronat S, Rojas-García R, Povedano M, Cuscó I, Tizzano EF.

Int J Mol Sci. 2022 Jul 27;23(15):8289. doi: 10.3390/ijms23158289.

**IF=6.208**; **Contribution:** conceptualization, methodology, formal analysis, investigation, data curation, writing - original draft, writing - review & editing.

**Elena García Arumí** **Eduardo F. Tizzano Ferrari**

**Barcelona, March 18th 2023**

# CHAPTER 1

# An integral approach to the molecular diagnosis of tuberous sclerosis complex: the role of mosaicism and splicing variants

Laura Blasco-Pérez[1,2], Leticia Iranzo-Nuez[1,2], Ricard López-Ortega[3], Desirée Martínez-Cruz[1,2], María Camprodon-Gómez[4], Anna Tenés[1,2], María Antolín[1,2], Eduardo F. Tizzano[1,2,5]*, Elena García-Arumí[1,2,5,6]*.

Submitted to The Journal of Molecular Diagnostics; *under review*

## ABSTRACT

Tuberous sclerosis complex (TSC) is an autosomal dominant genetic disorder characterized by the presence of hamartomas in multiple organs. At the molecular level, the disease is caused by pathogenic variants in the *TSC1* and *TSC2* genes, and only 10-25% of clinically diagnosed patients remain negative after MLPA and exon sequencing of both genes. Here, to improve the molecular diagnosis of TSC, we developed an integral approach that includes MLPA and deep coverage NGS of the entire *TSC1* and *TSC2* genes, along with an adapted bioinformatic pipeline to detect variants at low allele frequencies (>1%). Using this workflow, the molecular cause was identified in 29 TSC patients out of 42, describing here, for the first time, 12 novel pathogenic variants in TSC genes. These variants included seven splicing variants, five of which were studied at the cDNA level determining their effect on splicing. In addition, eight of the 29 pathogenic variants were detected in mosaicism, including four patients with previous negative studies that presented extremely low mosaic variants (allele frequency <16%). We demonstrate that this integral approach allows the molecular diagnosis of TSC patients and improves the conventional one by adapting the technology to the detection of low-frequency mosaics.

1. Department of Clinical and Molecular Genetics, Vall d'Hebron University Hospital, Universitat Autònoma de Barcelona (UAB), 08035 Barcelona, Spain.

2. Medicine Genetics Group, Vall d'Hebron Research Institute (VHIR), Universitat Autònoma de Barcelona (UAB), 08035 Barcelona, Spain.
3. Laboratori Clínic ICS Lleida, Hospital Universitari Arnau de Vilanova, 25198 Lleida, Spain.
4. Unit of Rare Diseases and Hereditary Metabolic Disorders, Vall d'Hebron University Hospital, 08035 Barcelona, Spain.
5. Neuromuscular and Mitochondrial Disorders Group, Vall d'Hebron Research Institute (VHIR), Universitat Autònoma de Barcelona (UAB), 08035 Barcelona, Spain.
6. Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Instituto de Salud Carlos III, 08041 Barcelona, Spain.
* These authors share senior authorship.

**Funding information:** This work has been partially supported by Spanish Tuberous Sclerosis Association (AET).

**Disclosures:** none declared.

**Acknowledgements:** We thank Spanish Tuberous Sclerosis Association, for partially funding this work; Alejandro Moles-Fernández, for his assistance in splicing variants interpretation; and Mar Costa-Roger, for critical review of the manuscript.

# INTRODUCTION

Tuberous sclerosis complex (TSC, OMIM #191100 and #613254) is an autosomal dominant genetic disorder with an incidence of between 1:6000 and 1:10000 live births[1]. This rare disease is characterized by the presence of hamartomas in multiple organs such as brain, skin, kidney, heart, lung, and bone. The expressivity is highly variable between individuals, even within the same family, but penetrance is considered to be complete[2,3]. Criteria for clinically diagnosing a patient with suspected TSC were established in 2012 and updated recently in 2021. These criteria allow to classify patients according to their phenotype into "definite TSC" or "possible TSC" depending on the number of major and/or minor clinical features present in them[4].

At the molecular level, TSC is caused by pathogenic variants in the *TSC1* and *TSC2* genes[5,6]. *TSC1* encodes for the protein hamartin and *TSC2* for tuberin, both of which bind together to form the TSC complex[7]. Hamartin stabilizes the TSC complex while tuberin provides the enzymatic function due to its GTPase activating protein (GAP) domain[8]. The TSC complex downregulates the mTORC1 (mammalian target of rapamycin complex 1) pathway to inhibit cell growth and, therefore, pathogenic variants affecting these genes result in an increase in cell proliferation, leading to benign tumors[9,10].

By the end of 2022, around 3600 *TSC1* and *TSC2* variants had been reported as pathogenic or likely pathogenic in the Leiden Open Variation Database (LOVD), 950 in *TSC1* and 2650 in *TSC2* [11]. This is in accordance with the fact that ≈74% of molecularly diagnosed patients present a variant in *TSC2* and ≈26% in *TSC1* [3]. Only 10-25% of clinically diagnosed patients remain negative after being studied by conventional methods (MLPA and exon and exon-intron boundaries sequencing of *TSC1* and *TSC2*) and are referred to as no mutation identified (NMI) patients. In these cases, it is hypothesized that the causal variant could be a deep intronic variant or a low-frequency mosaic variant, which would not be detected by conventional techniques[12,13].

In this work, we present 42 patients clinically diagnosed as definitive TSC, in whom we performed a CNV study of *TSC1* and *TSC2* by MLPA and complete sequencing gene by NGS at high coverage to allow detection of mosaic and deep intronic variants.

## MATERIAL & METHODS

### *Patients*

From a cohort of 75 patients with clinical suspicion of TSC, we selected a total of 42 patients that met standard clinical criteria for the diagnosis of TSC, according to updated diagnostic criteria recommendations of 2021[4]. All had been molecularly studied at the Department of Clinical and Molecular Genetics of the Hospital Vall d'Hebron in Barcelona. Thirteen of these patients had prior genetic testing for TSC performed at other Centers with negative results and were included in this project as NMI patients to expand molecular studies.

DNA samples were extracted from peripheral blood, buccal swab, saliva, or skin using standard methods. RNA samples were extracted from peripheral blood using standard methods. Written informed consent was obtained from all participants or their parents/legal caregivers.

### *Molecular diagnosis*

The initial molecular diagnostic workflow consisted in studying all patients by MLPA and exonic NGS sequencing of *TSC1* and *TSC2* to detect deletions and exon variants, respectively. For this purpose, MLPA kits P124 and P046 from MRC-Holland were used and a custom NGS panel of amplicons was designed, using GeneRead technology (Qiagen, Hilden, Germany). This panel allows sequencing of *TSC1* and *TSC2* exons, exon-intron boundaries, and intronic pathogenic variants described at the time of design (June 2018). NGS results were analyzed using the GeneRead Panel Variant Calling as germline samples, which allows us to detect mosaicisms with a frequency >4%. Later, the workflow was improved by replacing the exonic amplicon panel with a capture panel, which contains the entire sequences of *TSC1* and *TSC2* as target. First, an in-house panel was designed using NimbleGen SeqCap EZ HyperCap technology (Roche), covering 75% of the *TSC1* and *TSC2* genes, and a customized pipeline was

used for its analysis, detecting mosaics >5%. Later, it was replaced by another in-house panel of similar characteristics, with Cell3 Target technology (Nonacus). The analysis was performed using the bioinformatic pipeline of Datagenomics platform (Health in Code), which allows the detection of mosaics at a frequency >1%. In both cases, *TSC1* and *TSC2* exons and exon-intron boundaries were completely covered and only part of deep intronic regions with highly repetitive sequences remained unstudied. Patients with negative results after applying the initial workflow were retested with the updated approach (Figure 1).

All variants detected by NGS were validated by standard Sanger sequencing and deletions detected by MLPA were confirmed by long-range PCR and Sanger sequencing to characterize the deletion junction. Primer sequences and PCR conditions used in this work are available upon request. The American College of Medical Genetics and Genomics (ACMG) guidelines were used to classify the candidate variants[14], and ACMG and Clinical Genome Resource (ClinGen) guidelines were used to classify the candidate intergenic CNVs[15]. Variant nomenclature refers to RefSeq NM_000368.5 for *TSC1* and RefSeq NM_000548.5 for *TSC2*.

### *Mosaicism validation*

Candidate variants found at low frequency (<20%), indicating a possible mosaic variant, were confirmed by amplicon-based deep sequencing (ADS)[16], ensuring a minimum coverage of 7000x of the variant. For this purpose, the region of interest in each case was amplified by standard PCR, the PCR product was fragmented with NEBNext® dsDNA Fragmentase® (New England Biolabs) obtaining 200 bp fragments and libraries were generated using NEBNext Ultra DNA library prep kit for Illumina and NEBNext® MultiplexOligos for Illumina® Dual Index Primers Set 1 (New England Biolabs). The resulting product was sequenced using a MiSeq instrument (Illumina).

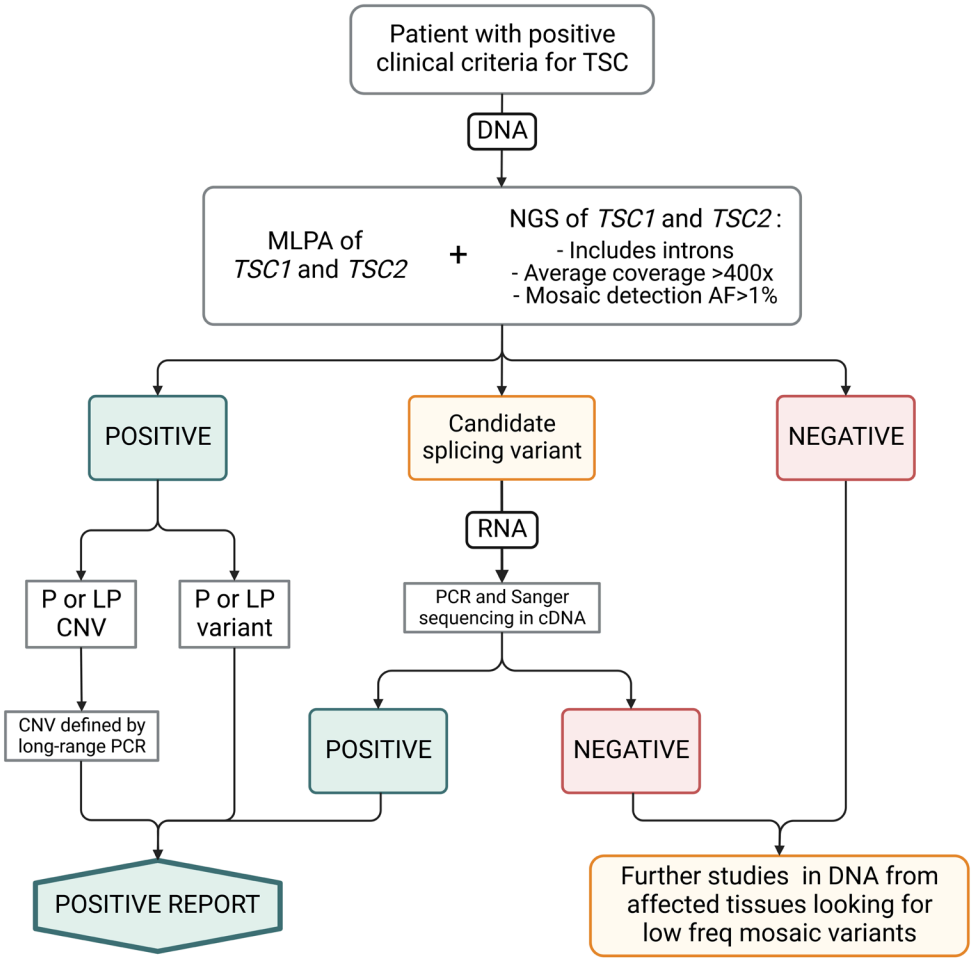



**Figure 1. Updated workflow for the molecular study of TSC patients.** Patients with positive clinical criteria of TSC were studied by MLPA and by complete NGS sequencing of *TSC* genes with an average coverage of >400x, detecting variants with an allele frequency (AF) >1%. If a candidate CNV was identified, long-range PCR was carried out to define the breakpoints. If a candidate splicing variant was identified, cDNA studies were carried out to determine its mRNA effect and thus classify the variant. Patients who remained negative after all these steps (NMI) were candidates for further studies in DNA from affected tissue. Abbreviations: P, pathogenic; LP, likely pathogenic; CNV, copy number variant.

### *mRNA studies*

To determine the effect of novel splice variants, mRNA studies were performed. Patients' mRNA was retrotranscribed using standard methods to obtain cDNA. Specific primers were designed to study each variant and the PCR products were analyzed through electrophoresis and Sanger sequencing. Bioinformatic splicing predictors used were Splice Site Finder (SSF-like)[17] and MaxEntScan (MES)[18] through Alamut™ Visual Plus (Sophia Genetics) and SpliceAI software[19].

## RESULTS

In our cohort of 42 unrelated patients with TSC positive clinical criteria, we identified the molecular cause in 29 patients, representing a total diagnostic rate of 69%. If patients were subdivided according to whether they had previous negative studies or not, the diagnostic rate was 31% (4/13) and 86% (25/29), respectively. As for the remaining cases, a potential causative variant was found in one case (P14) and 12 patients were classified as NMI (see Table 1).

In total, 27 different causative variants were detected (two variants present in two patients), 11 in *TSC1* (41%) and 16 in *TSC2* (59%). Of these, 12 had never been described or published (44%), and eight had been previously reported in LOVD or ClinVar but without associated literature, making a total of 20 unpublished variants. The TSC causative variants identified in our cohort were seven missense, one in-frame deletion and 19 loss-of-function (LoF) variants, including five nonsense, six splicing variants, five small indels, and three large deletions. In addition, a mosaic splicing variant was detected in patient 14 (*TSC2*:c.482-1G>C, 1% allele frequency (AF)) and was considered the possible cause of the disease. Interestingly, the four causative variants detected in previously studied patients (4/13) were in mosaic status, including one missense variant (P29), two small deletions (P12 and P17), and one large deletion (P21). All variants are detailed in Table 1.

**Table 1. Detailed information of candidate variants in TSC patients.** Patients are numbered according to the localization of their variants. All variants were named following HGVS recommendations. SNVs (single nucleotide variants) and intragenic CNVs were classified according to ACMG guidelines and the intergenic deletion of P13 was classified following ACMG and ClinGen guidelines. Variant nomenclature refers to RefSeq NM_000368.5 for *TSC1* and RefSeq NM_000548.5 for *TSC2*.

| Patient | Previous Studies | DNA Change | Protein Change | Allelic Frequency | Genetic Origin | ACMG variant classification and criteria | *in silico* or *in vitro* studies | First reported |
|---|---|---|---|---|---|---|---|---|
| P1 | no | *TSC1*:c.334_335delinsG | p.Leu112Valfs*6 | HTZ | *de novo* | P: PVS1, PM2, PM6, PP4 | - | This article |
| P2 | no | *TSC1*:c.388A>C | p.Thr130Pro | HTZ | *de novo* | LP: PM2, PM6, PP3, PP4 | - | This article |
| P3 | no | *TSC1*:c.473T>C | p.Phe158Ser | HTZ | *de novo* | LP: PS3, PM2, PM6, PP4 | Functional studies | Ref. [20] |
| P4 | no | *TSC1*:c.663+2T>G | p.? | HTZ | *de novo* | P: PVS1, PM2, PM6, PP4 | SpliceAI | This article |
| P5 | no | *TSC1*:c.838C>T | p.Gln280* | HTZ | n.a. | P: PVS1, PM2, PP4 | - | LOVD (TSC1_000267) |
| P6 | no | *TSC1*:c.994del | p.Ser332Profs*6 | 38% Sa; 42% NAS* | *de novo* | P: PVS1, PS2, PM2, PP4 | - | This article |
| P7 | no | *TSC1*:c.1258_1263+20dup | p.Glu422Valfs*29 | HTZ | *de novo* | P: PVS1, PM2, PM6, PP4 | cDNA studies | This article |
| P8 | no | *TSC1*:c.1798C>T | p.Gln600* | HTZ | n.a. | P: PVS1, PM2, PP4 | Functional studies | Ref. [21] |
| P9 | no | *TSC1*:c.2041+2T>G | p.Pro668Serfs*5 | HTZ | n.a. | P: PVS1, PM2, PP4 | cDNA studies | LOVD (TSC1_000758) |
| P10 | no | | | HTZ | n.a. | P: PVS1, PM2, PP4 | | |
| P11 | no | *TSC1*:c.2080C>T | p.Gln694* | HTZ | n.a. | P: PVS1, PM2, PP4 | - | LOVD (TSC1_000612) |
| P12 | yes | *TSC1*:c.2101_2107del | p.Gln701Serfs*21 | 4% PB | - | P: PVS1, PS2, PM2, PP4 | - | LOVD (TSC1_000696) |
| P13 | no | g.2091854_2100202del; *NTHL1* in4 to *TSC2* in2; 8349bp | No protein expressed | HTZ | *de novo* | P: Score of 1.45: 2C-1,5A | - | This article |
| P14 | yes | TSC2:c.482-1G>C | p.(Ala161Valfs*22) | 1% PB;1% OS | - | †P: PVS1, PS2, PM2, PP4 | cDNA studies | ClinVar (49826) |
| P15 | no | *TSC2*:c.976-14G>A | p.Ala326_Gln373del | HTZ | Father (TSC) | LP: PVS1_Strong, PM2, PP1, PP4 | cDNA studies | LOVD (TSC2_002150) |

(Continue in next page)

| Patient | Previous Studies | DNA Change | Protein Change | Allelic Frequency | Genetic Origin | ACMG variant classification and criteria | *in silico* or *in vitro* studies | First reported |
|---|---|---|---|---|---|---|---|---|
| P16 | no | *TSC2*:c.1119G>A | p.Ala326_Gln373del | 35% PB | - | LP: PVS1_Strong, PS2, PM2, PP4 | cDNA studies | LOVD (TSC2_002795) |
| P17 | yes | *TSC2*:c.1283_1285del | p.Ser428del | 10% OS; 16% PB | - | P: PS2, PS3, PM2, PP4 | Functional studies | Ref. [22] |
| P18 | no | *TSC2*:c.1397T>C | p.Leu466Pro | 19% PB | - | P: PS2, PM2, PM5, PP4, PP5 | - | LOVD (TSC2_001121) |
| P19 | no | *TSC2*:c.1478T>C | p.Leu493Pro | HTZ | *de novo* | P: PS3, PM2, PM5, PM6, PP4 | Functional studies | Ref. [23] |
| P20 | no | *TSC2*:c.1513C>T | p.Arg505* | HTZ | n.a. | P: PVS1, PM2, PP4 | Functional studies | Ref. [24,25] |
| P21 | yes | *TSC2:*c.1716+1745_2546-58del; g.2117381_2125742del; 8362bp | p.Thr573Leufs*45 | <50% PB | - | P: PVS1, PS2, PM2, PP4 | - | This article |
| P22 | no | *TSC2:*c.2356-156_2743-36del; g.2124045_2126456del; 2412bp | p.Arg786_Lys914del | HTZ | n.a. | LP: PVS1_Strong, PM2, PP4 | - | This article |
| P23 | no | *TSC2*:c.2838-122G>A | p.Ser946Argfs*6 | HTZ | n.a. | P: PSV1, PM2, PP4 | cDNA studies | Ref. [12] |
| P24 | no | | | HTZ | *de novo* | P: PSV1, PM2, PM6, PP4 | | |
| P25 | no | *TSC2*:c.4398delinsTT | p.Ala1467Cysfs*57 | 7% PB | - | P: PVS1, PS2, PM2, PP4 | - | This article |
| P26 | no | *TSC2*:c.4663-1_4667del | p.Ser1555Argfs*8 | HTZ | *de novo* | P: PVS1, PM2, PM6, PP4 | cDNA studies | This article |
| P27 | no | *TSC2*:c.4717G>T | p.Glu1573* | HTZ | n.a. | P: PSV1, PM2, PP4 | - | This article |
| P28 | no | *TSC2*:c.4953T>G | p.Asn1651Lys | HTZ | n.a. | LP: PM2, PM5, PM6, PP4 | - | This article |
| P29 | yes | *TSC2*:c.5024C>G | p.Pro1675Arg | 6% PB | - | P: PS2, PM2, PM5, PP3, PP4 | - | LOVD (TSC2_002575) |
| P30 | no | *TSC2*:c.5177A>C | p.His1726Pro | HTZ | n.a. | LP: PM2, PM6, PP3, PP4 | - | ClinVar (468144) |

*Analysis of the variant in DNA from peripheral blood (PB-DNA) was not possible as patient 6 had a previous bone marrow transplant not related to their TSC condition.
†This variant is classified as pathogenic following ACMG guidelines but not as disease-causing in P13 as it is present in mosaic at 1%AF. More information in discussion section.
Abbreviations: HTZ, heterozygous; NAS, Non-Affected Skin; Sa, Saliva; PB, Peripheral Blood; OS, Oral Swab. PV, Pathogenic Variant; LPV, Likely Pathogenic Variant; n.a.; not available

### *Mosaic variants*

In our cohort, mosaic alterations were established as the cause of TSC in eight of 29 patients (28%) and as the probable cause in a ninth patient (P14), in which more evidence is needed. The mosaic variants detected included two missense (6 and 19% AF respectively), two splicing variants (1 and 35% AF), one indel (7% AF), three small deletions (4, 10-16, and 38-42% AF) and one large deletion. The variant found at an AF of 38-42% (P6) was initially detected with an AF of 35% in DNA from non-affected skin (peripheral blood DNA was unavailable, see Table 1) by standard NGS (coverage >400x). It was later studied by ADS in DNA from non-affected skin (42%) and saliva (38%). In addition, the variant was confirmed to be *de novo*, indicating that it is likely a mosaic rather than a heterozygous variant. In cases with an AF <20%, the presence of the variant was confirmed by an independent approach and, when possible, a second DNA sample from other tissues were analyzed. Details regarding mosaic variants are shown in Table 1.

In particular, patient 21 presented a large deletion comprising introns 16 to 22 of *TSC2* (chr16:2117381-2125742), detected through MLPA. The signal decrease of these probes was compatible with a mosaic deletion as they were around 0.8 instead of 0.5 (which would indicate a heterozygous deletion) (Figure 2). This mosaic deletion was not detectable by CGH-Array nor NGS, therefore, the allele frequency could not be precisely determined; however, it was confirmed and delimited through long-range PCR and Sanger sequencing. Another interesting case was patient 14, with pathogenic variant c.482-1G>C at an AF of approximately 1% in DNA from peripheral blood and confirmed in DNA from an oral swab. Due to the variant's low AF, mRNA studies were performed by ADS since conventional Sanger sequencing lacks the sensitivity to study these cases (see details below).

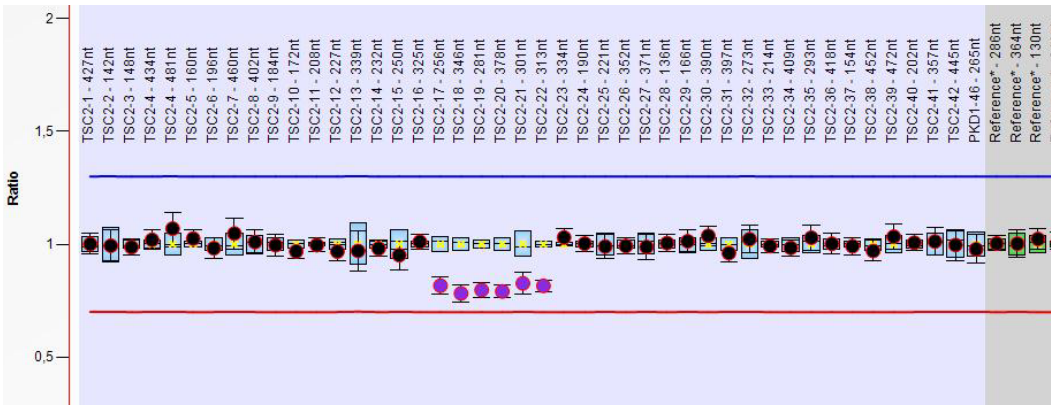



**Figure 2. MLPA result from patient 21.** MLPA analysis show a decrease in probe signals corresponding to exons 17 to 22. The fact that the probes are located approximately at a 0.8 ratio indicates a mosaic deletion rather than a heterozygous deletion, in which probes would be located at around a 0.5 ratio. Results were consistent in several repeated experiments.

### *mRNA studies*

Based on splicing predictors, a total of eight candidate variants to alter splicing were detected in 10 unrelated patients (three in *TSC1* and five in *TSC2* ), representing 34.5% of the causative variants detected (10/29). Of these, only the c.2838-122G>A (p.Ser946Argfs*6) variant in *TSC2* (P23 and P24) had been described and previous mRNA studies confirmed its pathogenicity[12]. The remaining variants were studied at the mRNA level, except for c.663+2T>G (p.?) of *TSC1* (P4) since it was not possible to obtain an RNA sample from the patient.

The study of the c.1258_1263+20dup variant in *TSC1* (P7), a 26-nucleotide duplication spanning the end of exon 12 and the beginning of intron 12, revealed that alternative splicing occurs using a new donor site generated by the duplication of exon 12 (Figure 3A). This leads to the inclusion of the 26 duplicated nucleotides in the mRNA, resulting in a frameshift that predicts a truncated protein (p.Glu422Valfs*29). Also, in *TSC1*, we studied the c.2041+2T>G variant located in intron 16 and found it in two unrelated patients (P9 and P10). In this case, the alteration produces skipping of exon 16, which also alters the reading frame and creates a premature stop codon predicting a truncated protein (p.Pro668Serfs*5) (Figure 3B).

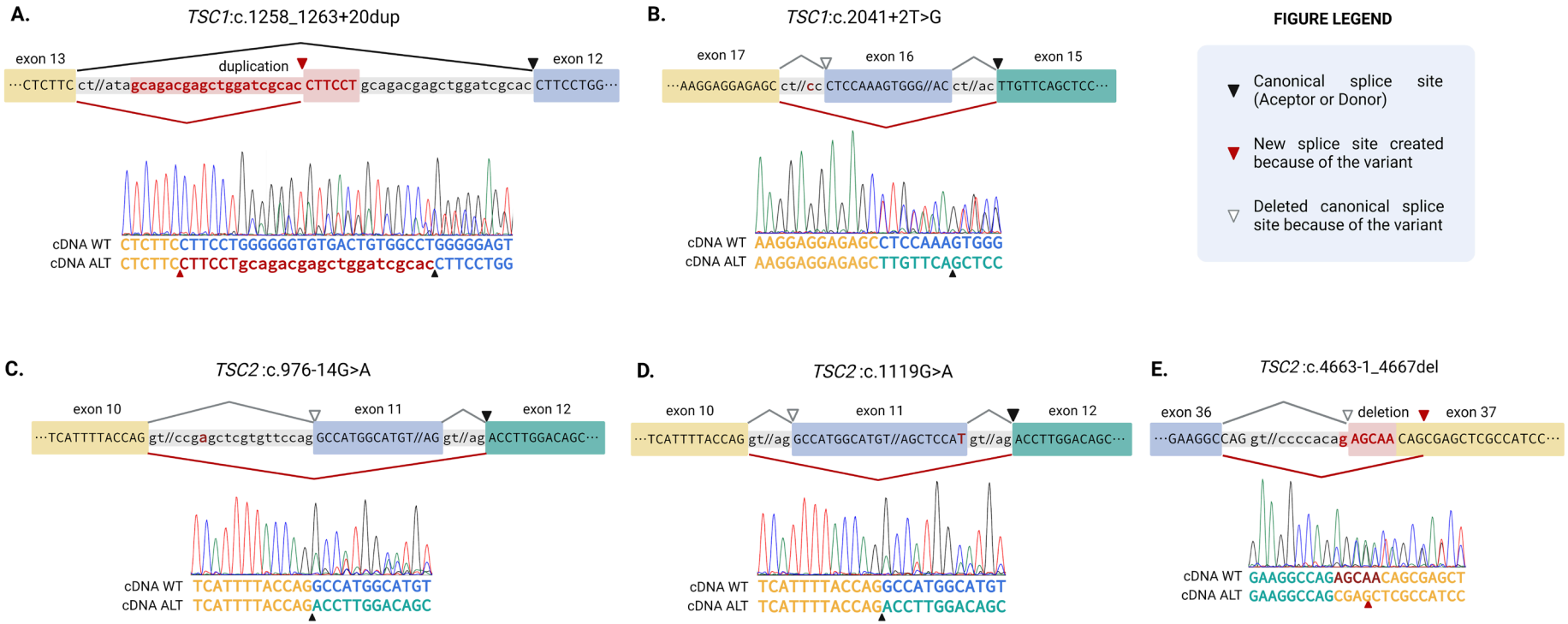



**Figure 3. Schematic representation of the observed effect of the studied splicing variants at the cDNA level.** In each case, the splicing variant is represented in red letters. Above, the gDNA sequence around the variant is presented, with exons colored and introns in gray. Duplication (A) and deletion (E) are also colored red. In addition, canonical and alternative splicing positions are indicated with an arrowhead. Below, cDNA Sanger sequencing results are shown, including the reference and alternative sequences. *TSC1* is represented in reverse strand and *TSC2* in forward, following gene orientation.

In *TSC2*, variant c.976-14G>A (P15) in intron 10 and mosaic variant c.1119G>A (P16) in exon 11 were studied. Interestingly, both variants produce the same effect in mRNA, exon skipping of exon 11 (p.Ala326_Gln373del) (Figure 3C and 3D); this leads to the loss of 48 amino acids belonging to the Hamartin binding region of the protein, necessary for conformation of the TSC complex. Another variant detected in *TSC2* was the small deletion c.4663-1_4667del (P26) located between intron 36 and exon 37. Due to this variant, the canonical splicing acceptor site is deleted and a new one is created three nucleotides downstream of the deletion (Figure 3E). This alternative splicing disrupts the reading frame, predicting a truncated protein (p.Ser1555Argfs*8).

A different technical approach was necessary to study the 1% mosaic variant c.482-1G>C in *TSC2* (P14). According to predictors, the variant resulted in the loss of the canonical splicing acceptor site and is expected to use an alternative site 17 nucleotides downstream of the canonical site (Figure 4A). This would produce an mRNA transcript lacking the first 17 nucleotides of exon 6, generating an alteration of the reading frame and a premature stop codon (p.Ala161Valfs*22). The study of the patient's mRNA by ADS revealed the expected alternative transcript at an extremely low AF (≈0.15%) (Figure 4A). This transcript was also observed in control samples and no other alternative transcripts were observed neither in the patient or controls. The complete sequencing of *TSC1* and *TSC2* in patient 14 allowed us to assure that, with the information available to date, they did not present any other candidate variant in the studied region.

Lastly, an alternative strategy was carried out to study the chr16:2091854-2100202 deletion (P13) at the mRNA level. This alteration involves the first two exons of *TSC2* and part of the *NTHL1* gene, upstream of *TSC2.* Therefore, we hypothesized that mRNA would not be generated from this allele as the deletion includes the promotor and the start codon of the gene. To confirm this, exon 15 was amplified since the patient presents an SNP (rs34012042) that distinguishes each allele. The results showed a loss of heterozygosity, meaning that only one allele was present, thus confirming our hypothesis (Figure 4B).

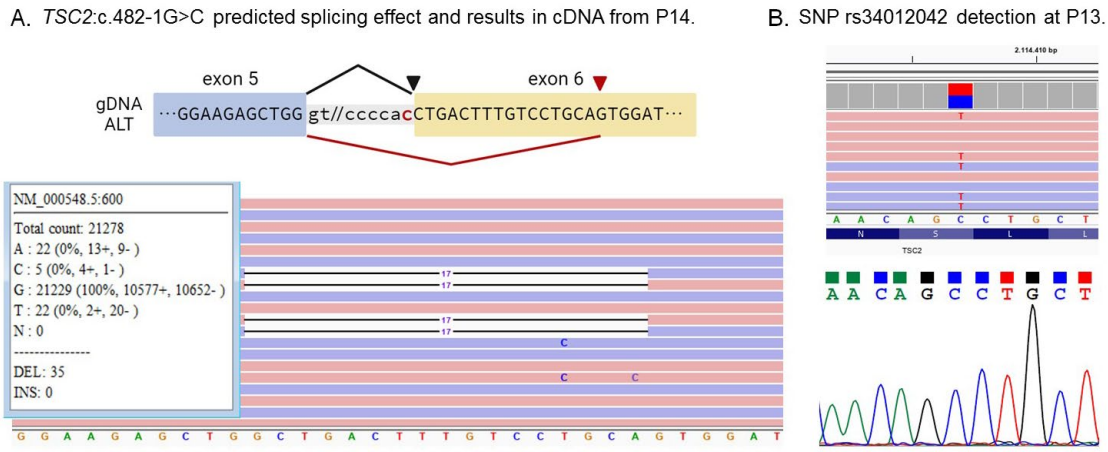



**Figure 4. Results of specific approaches to determine the effect at the cDNA level of variants *TSC2:*c.482-1G>C and g.2091854_2100202del.** A) Studies on P14 to establish the effect of *TSC2:* c.482-1G>C variant, detected at AF 1%, in cDNA. Above, schematic representation of the splicing alteration prediction. The canonical splicing donor site is marked with a black arrowhead and the alternative site, which is potentiated due to the variant, is indicated in red. Below, IGV image showing 35 reads out of 21313 (≈0.15%) compatible with splicing alteration. B) Studies on P13 to determine the loss of expression of the allele with the deletion g.2091854_2100202del (including *NTHL1* in4 to *TSC2* in2). Above, results from NGS in the gDNA of P13 show the presence of the SNP rs34012042 in a heterozygous state in exon 15 of *TSC2*. Below, Sanger results from cDNA, showing absence of the SNP rs34012042.

## DISCUSSION

In this work, we present the implementation of an integral approach to the molecular diagnosis of TSC in a heterogeneous cohort of clinically diagnosed patients. We designed a workflow that included MLPA analysis, to detect CNVs, and high-coverage complete sequencing of the *TSC1* and *TSC2* genes with an optimized bioinformatic analysis, to detect deep intronic and low-frequency mosaic variants.

After applying this workflow to our 42 patients, we identified the molecular cause in 29 cases and a likely causal variant in an additional patient. This represents a diagnostic rate of 69% (29/42), slightly lower than that previously reported of around 75-90%[13,26–28]. This could be explained because our cohort includes patients with prior negative molecular studies (MLPA and/or exonic sequencing of *TSC1* and

*TSC2*) classified as NMI (n=13). In fact, excluding these patients, our diagnostic rate would increase to 86%, consistent with that described in the literature.

Focusing on the positive cases, we detected 27 different variants as the molecular cause of TSC in 29 patients. In our cohort, variants in *TSC2* were more frequent than in *TSC1* (59% vs. 41%, respectively) which is in line with the literature[13,26]. Among these causal variants, more than two-thirds (70%) were loss-of-function changes. This fact is consistent with the extremely low tolerance of *TSC1* and *TSC2* to this type of variant, as both genes show the highest score for loss-of-function intolerance (pLI) of 1, reaching the maximum score[29]. Moreover, almost half of the variants detected had not been previously reported (12/27, 44%), similar to other studies[30,31].

It has been described that the most likely molecular causes of TSC in NMI patients are deep intronic and low-frequency mosaic variants[12,13]. Regarding allelic frequency, eight patients of the 29 molecularly diagnosed in our series (28%) had a mosaic variant, with an AF of between 4 and 42%, and a ninth patient presenting the likely causal variant as a 1% mosaic. In our study, *TSC* genes were sequenced at high coverage NGS (>350x) to allow the detection of low-frequency variants, and those with an AF <20% were also studied by ADS (minimum coverage 7000x) to confirm the results. The variant initially detected with an AF of 35% was considered likely to be in mosaic, and not in heterozygosis, as it was found *de novo* in the patient and confirmed by ADS in two different samples (38% in saliva and 42% in non-affected skin). It is important to note that mosaic variants, especially those at low frequency (<20%), would probably not have been detected using the exome (WES) or genome (WGS) techniques that are increasingly being used in routine genetic diagnosis. In fact, the four patients with previous molecular studies in which we identified the molecular cause presented mosaic variants, three of them with an AF <20%. The fourth case (P20) was the 8349 bp large mosaic deletion in *TSC2*, identified by MLPA but undetected by NGS or CGH-Array. In addition, a previous MLPA study performed in another center was reported as negative in this patient. These results make evident the complexity of detecting mosaic deletions,

highlighting the use and careful interpretation of MLPA as the reference technique to study CNVs in the routine diagnosis of TSC.

As regards deep intronic variants, we detected two cases with the known pathogenic variant c.2838-122G>A in *TSC2* but we did not find any novel candidate variant of this type in our cohort. Therefore, deep intronic variants explain the cause of the disease in 5% of our patients, data similar to those previously reported (1-12%)[12,13,32]. In contrast, we detected six novel splicing variants near exons as the cause of the disease in seven unrelated patients (Table 1). Five of these variants were studied at the cDNA level, and these studies revealed that three of them lead to a frameshift. The other two, produced the skipping of exon 11 of *TSC2,* resulting in an in-frame deletion of 48 amino acids that were localized in the functional domain of the protein (Figure 3). In all cases, mRNA studies allowed us to determine the specific alteration produced in the mRNA, since the SpliceAI predictor was able to correctly predict a splicing alteration but failed to predict the type of alteration that would occur in three of the variants (P7, P15, P16; data not shown). mRNA studies led as to know the exact alteration that each variant produces and made it possible to correctly classify the variants using the ACMG guidelines and confirm them as the cause of the disease.

Furthermore, variant c.482-1G>C in *TSC2* remains as the probable cause of the disease in patient 14, without further validation. This variant is classified as pathogenic according to ACMG guidelines as it affects the canonical splicing site, however, it was present at an AF of 1% so we would expect its contribution to be lower than if it was in the heterozygous state. The next step forward should be to study whether affected neural crest tissues show the variant at a higher AF and perform the same mRNA studies to analyze the presence of the alternative transcript in target cells. In fact, several patients with causative mosaic variants at <1% have been described in the literature and, specifically, one patient with a mosaic splicing variant (*TSC2*:c.4850-1G>A, AF of 0.71%)[13].

## CONCLUSION

To sum up, the application of our workflow (MLPA and deep NGS sequencing of complete *TSC1* and *TSC2*) allowed the molecular diagnosis of 86% of TSC patients, with about one third of patients showing mosaic variants and another third presenting splicing variants, 5% of them being deep intronic. These includes the diagnosis of four NMI patients due to the high coverage used, as all presented <20% mosaic variant. Based on the above, we consider that this strategy would improve the genetic diagnostic rate of TSC in relation to the one achieved with routine techniques (24%). The use of WES or WGS would not be the most convenient approach as, in general, lower coverage is achieved and, to allow the detection of mosaics at >1% AF, it is necessary a high coverage and an optimized bioinformatics analysis. In the case of CNVs, the most sensitive technique is MLPA as it allows the detection of mosaic deletions, which are very unlikely to be identified by NGS or CGH-array. Lastly, regarding splicing variants, it is important to note that the *in silico* predictors should only be a guide and the effect of splicing variants should be established by mRNA studies in order to properly classify the variants. Our results show the importance of implementing specific techniques to improve the molecular diagnosis of TSC and adapting the analysis of the results according to the high recurrence of low-frequency variants.

## BIBLIOGRAPHY


1. Osborne JP, Fryer A, Webb D. Epidemiology of tuberous sclerosis. *Ann N Y Acad Sci*. 1991;615(1):125-127. doi:10.1111/J.1749-6632.1991.TB37754.X

2. Northrup H, Wheless JW, Bertin TK, Lewis RA. Variability of expression in tuberous sclerosis. *J Med Genet*. 1993;30(1):41-43. doi:10.1136/JMG.30.1.41

3. Northrup H, Koenig MK, Pearson DA, Au KS. Tuberous Sclerosis Complex. *GeneReviews®*. Published online 1993. Accessed November 8, 2022. https://pubmed.ncbi.nlm.nih.gov/20301399/

4. Northrup H, Aronow ME, Bebin EM, et al. Updated International Tuberous Sclerosis Complex Diagnostic Criteria and Surveillance and Management Recommendations. *Pediatr Neurol*. 2021;123:50-66. doi:10.1016/j.pediatrneurol.2021.07.011

5. European Chromosome 16 Tuberous Sclerosis Consortium. Identification and characterization of the tuberous sclerosis gene on chromosome 16. *Cell*. 1993;75(7):1305-1315. doi:10.1016/0092-8674(93)90618-Z

6. Van Slegtenhorst M, De Hoogt R, Hermans C, et al. Identification of the tuberous sclerosis gene TSC1 on chromosome 9q34. *Science*. 1997;277(5327):805-808. doi:10.1126/SCIENCE.277.5327.805

7. Van Slegtenhorst M, Nellist M, Nagelkerken B, et al. Interaction between hamartin and tuberin, the TSC1 and TSC2 gene products. *Hum Mol Genet*. 1998;7(6):1053-1057. doi:10.1093/HMG/7.6.1053

8. Benvenuto G, Li S, Brown SJ, et al. The tuberous sclerosis-1 (TSC1) gene product hamartin suppresses cell growth and augments the expression of the TSC2 product tuberin by inhibiting its ubiquitination. *Oncogene*. 2000;19(54):6306-6316. doi:10.1038/SJ.ONC.1204009

9. Laplante M, Sabatini DM. mTOR signaling in growth control and disease. *Cell*. 2012;149(2):274-293. doi:10.1016/J.CELL.2012.03.017

10. Huang J, Manning BD. The TSC1-TSC2 complex: a molecular switchboard controlling cell growth. *Biochem J*. 2008;412(2):179-190. doi:10.1042/BJ20080281

11. Fokkema IFAC, Taschner PEM, Schaafsma GCP, Celli J, Laros JFJ, den Dunnen JT. LOVD v.2.0: the next generation in gene variant databases. *Hum Mutat*. 2011;32(5):557-563. doi:10.1002/HUMU.21438

12. Nellist M, Brouwer RWW, Kockx CEM, et al. Targeted Next Generation Sequencing reveals previously unidentified TSC1 and TSC2 and mutations. *BMCNellist, M, Brouwer, R W W, Kockx, C E M, van Veghel-Plandsoen, M, Withagen-Hermans, C, Prins-Bakker, L, Hoogeveen-Westerveld, M, Mrsic, A, van den Berg, M M P, Koopmans, A E, Wit, M C, Jansen, F E, Maat-Kievit, A J A, van*. 2015;16(1):1-11. doi:10.1186/s12881-015-0155-4

13. Tyburczy ME, Dies KA, Glass J, et al. Mosaic and Intronic Mutations in TSC1/TSC2 Explain the Majority of TSC Patients with No Mutation Identified by Conventional Testing. *PLoS Genet*. 2015;11(11). doi:10.1371/JOURNAL.PGEN.1005637

14. Richards S, Aziz N, Bale S, et al. Standards and Guidelines for the Interpretation of Sequence Variants: A Joint Consensus Recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17(5):405. doi:10.1038/GIM.2015.30

15. Riggs ER, Andersen EF, Cherry AM, et al. Technical standards for the interpretation and reporting of constitutional copy-number variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics (ACMG) and the Clinical Genome Resource (ClinGen). *Genet Med*. 2020;22(2):245-257. doi:10.1038/S41436-019-0686-8

16. Fernández-Álvarez P, Codina-Sola M, Valenzuela I, et al. A systematic study and literature review of parental somatic mosaicism of FBN1 pathogenic variants in Marfan syndrome. *J Med Genet*. 2022;59(6):605-612. doi:10.1136/JMEDGENET-2020-107604

17. Shapiro MB, Senapathy P. RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucleic Acids Res*. 1987;15(17):7155-7174. doi:10.1093/NAR/15.17.7155

18. Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J Comput Biol*. 2004;11(2-

3):377-394. doi:10.1089/1066527041410418

19. Jaganathan K, Kyriazopoulou Panagiotopoulou S, McRae JF, et al. Predicting Splicing from Primary Sequence with Deep Learning. *Cell*. 2019;176(3):535-548.e24. doi:10.1016/J.CELL.2018.12.015

20. Hoogeveen-Westerveld M, Ekong R, Povey S, et al. Functional assessment of TSC1 missense variants identified in individuals with tuberous sclerosis complex. *Hum Mutat*. 2012;33(3):476-479. doi:10.1002/humu.22007

21. Potter CJ, Huang H, Xu T. Drosophila Tsc1 functions with Tsc2 to antagonize insulin signaling in regulating cell growth, cell proliferation, and organ size. *Cell*. 2001;105(3):357-368. doi:10.1016/S0092-8674(01)00333-6

22. Rosengren T, Nanhoe S, de Almeida LGD, et al. Mutational analysis of TSC1 and TSC2 in Danish patients with tuberous sclerosis complex. *Sci Rep*. 2020;10(1):1-9. doi:10.1038/s41598-020-66588-4

23. Hoogeveen-Westerveld M, Ekong R, Povey S, et al. Functional Assessment of TSC2 Variants Identified in Individuals with Tuberous Sclerosis Complex. *Hum Mutat*. 2013;34(1):167-175. doi:10.1002/humu.22202

24. Wilson PJ, Ramesh V, Kristiansen A, et al. Novel mutations detected in the TSC2 gene from both sporadic and familial TSC patients. *Hum Mol Genet*. 1996;5(2):249-256. Accessed March 9, 2023. https://academic.oup.com/hmg/article/5/2/249/2568724

25. Jóźwiak S, Kwiatkowski D, Kotulska K, et al. Tuberin and hamartin expression is reduced in the majority of subependymal giant cell astrocytomas in tuberous sclerosis complex consistent with a two-hit model of pathogenesis. *J Child Neurol*. 2004;19(2):102-106. doi:10.1177/08830738040190020401

26. Au KS, Williams AT, Roach ES, et al. Genotype/phenotype correlation in 325 individuals referred for a diagnosis of tuberous sclerosis complex in the United States. *Genet Med*. 2007;9(2):88-100. doi:10.1097/GIM.0B013E31803068C7

27. Sancak O, Nellist M, Goedbloed M, et al. Mutational analysis of the TSC1 and TSC2 genes in a diagnostic setting: genotype--phenotype

correlations and comparison of diagnostic DNA techniques in Tuberous Sclerosis Complex. *Eur J Hum Genet*. 2005;13(6):731-741. doi:10.1038/SJ.EJHG.5201402

28. Dabora SL, Kwiatkowski DJ, Franz DN, et al. Mutational Analysis in a Cohort of 224 Tuberous Sclerosis Patients Indicates Increased Severity of TSC2, Compared with TSC1, Disease in Multiple Organs. *Am J Hum Genet*. 2001;68(1):64. doi:10.1086/316951

29. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020;581(7809):434-443. doi:10.1038/s41586-020-2308-7

30. Reyna-Fabián ME, Hernández-Martínez NL, Alcántara-Ortigoza MA, et al. First comprehensive TSC1/TSC2 mutational analysis in Mexican patients with Tuberous Sclerosis Complex reveals numerous novel pathogenic variants. *Sci Rep*. 2020;10(1):1-14. doi:10.1038/s41598-020-62759-5

31. Yu T, He Y, Li N, et al. Novel TSC1 and TSC2 gene mutations in Chinese patients with tuberous sclerosis complex. *Clin Neurol Neurosurg*. 2017;154:104-108. doi:10.1016/J.CLINEURO.2017.01.015

32. Togi S, Ura H, Hatanaka H, Niida Y. Genotype and Phenotype Landscape of 283 Japanese Patients with Tuberous Sclerosis Complex. *Int J Mol Sci*. 2022;23(19). doi:10.3390/IJMS231911175/S1

# CHAPTER 2

Received: 4 September 2020 | Revised: 13 January 2021 | Accepted: 11 March 2021

DOI: 10.1002/humu.24200

METHODS

Human Mutation HGVS HUMAN GENOME VARIATION SOCIETY WILEY

# Beyond copy number: A new, rapid, and versatile method for sequencing the entire *SMN2* gene in SMA patients

Laura Blasco-Pérez[1,2] | Ida Paramonov[1,2] | Jordi Leno[1,2] | Sara Bernal[3,4] | Laura Alias[3,4] | Pablo Fuentes-Prior[5] | Ivon Cuscó[1,2,4] | Eduardo F. Tizzano[1,2]

[1]Medicine Genetics Group, Vall d'Hebron Research Institute (VHIR), Barcelona, Spain

[2]Department of Clinical and Molecular Genetics, Hospital Vall d'Hebron, Barcelona, Spain

[3]Genetics Department, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain

[4]Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER)

[5]Molecular Bases of Disease, Biomedical Research Institute Sant Pau (IIB Sant Pau), Hospital de la Santa Creu i Sant Pau, Barcelona, Spain

**Correspondence**
Ivon Cuscó and Eduardo F. Tizzano, Department of Clinical and Molecular Genetics, Hospital Vall d'Hebron, 08035, Barcelona, Spain.
Email: icusco@vhebron.net and etizzano@vhebron.net

**Funding information**
Fundació Daniel Bravo Andreu, Grant/Award Number: 2016; Spanish Instituto de Salud Carlos III, Fondo de Investigaciones Sanitarias and cofunded with ERDF, Grant/Award Number: FIS PI18/000687; Galicia AME, Grant/Award Number: 2019; Biogen, Grant/Award Number: ESP-SMG-17-11256; SMA Europe, Grant/Award Number: 2014-2016

## Abstract

Spinal muscular atrophy (SMA) is caused by bi-allelic loss or pathogenic variants in the *SMN1* gene. *SMN2*, the highly homologous copy of *SMN1*, is considered the major phenotypic modifier of the disease. Determination of *SMN2* copy number is essential to establish robust genotype–phenotype correlations and predict disease evolution, to stratify patients for clinical trials, as well as to define those eligible for treatment. Discordant genotype–phenotype correlations are not uncommon in SMA, some of which are due to intragenic *SMN2* variants that may influence the amount of complete *SMN* transcripts and, therefore, of full-length SMN protein. Detection of these variants is crucial to predict SMA phenotypes in the present scenario of therapeutic advances and with the perspective of SMA neonatal screening and early diagnosis to start treatments. Here, we present a novel, affordable, and versatile method for complete sequencing of the *SMN2* gene based on long-range polymerase chain reaction and next-generation sequencing. The method was validated by analyzing samples from 53 SMA patients who lack *SMN1*, allowing to characterize paralogous, rare variants, and single-nucleotide polymorphisms of *SMN2* as well as *SMN2*–*SMN1* hybrid genes. The method identifies partial deletions and can be adapted to determine rare pathogenic variants in patients with at least one *SMN1* copy.

**KEYWORDS**
next-generation sequencing, paralogous variants, phenotype–genotype correlations, *SMN2* copies, spinal muscular atrophy

## 1 | INTRODUCTION

Spinal muscular atrophy (SMA) is the second most common recessive genetic disease of infancy and early childhood, with an incidence of 1 in 5000–10,000 live births and a worldwide carrier frequency of 1:51 (Sugarman et al., 2012). SMA patients are classified into different clinical groups based on the age of onset, clinical severity, and achieved motor milestones. In the most severe form, type I SMA, patients are never able to sit and generally die of respiratory failure before the age of 2 years. Intermediate type II SMA patients are able to sit but never walk, thus

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.
© 2021 The Authors. *Human Mutation* published by Wiley Periodicals LLC.

being confined to wheel-chair. Type III patients walk unassisted but may lose this ability during infancy or adolescence (Wang et al., 2007; Zerres & Rudnik-Schöneborn, 1995).

Bi-allelic absence or pathogenic variants of the *Survival of Motor Neuron 1* (*SMN1*) gene cause SMA (Lefebvre et al., 1995). A centromeric and nearly identical paralog, *SMN2*, encodes in principle the same protein as *SMN1* (Lefebvre et al., 1995; Monani et al., 1999; Rochette et al., 2001). However, a silent transition within exon 7 of the *SMN2* gene causes exon skipping and results in a truncated, nonfunctional variant (SMN-Δ7) (Lorson et al., 1999). It has been estimated that each *SMN2* copy can produce only around 10% to 15% of functional SMN protein, depending on the cells and tissues studied (Boza-Morán et al., 2015; Soler-Botija et al., 2005; Wirth et al., 2013).

The number of *SMN2* copies and the presence of intragenic *SMN2* variants are known modifiers of SMA disease severity (Bernal et al., 2010; Prior et al., 2009; Ruhno et al., 2019). Indeed, numerous studies show that the higher the *SMN2* copy number, producing larger amount of full-length SMN protein, the milder the associated SMA phenotype and vice versa. However, this inverse correlation is not absolute (Calucho et al., 2018). Whereas the determination of *SMN2* copy number is widely implemented to study SMA patients, the actual structures and genomic sequences of *SMN2* copies are usually not included in the characterization of SMA patients.

The current scenario of SMA therapy is rapidly evolving due to the approval in the last years of nusinersen/Spinraza, an antisense-tailored therapy (Finkel et al., 2017), AVXS101/Zolgensma, an adeno-associated viral-based gene therapy (Mendell et al., 2017), as well as the recent approval of the first oral drug to treat SMA, risdiplam/Evrysdi (http://www.fda.gov). However, these disease-modifying therapies are expensive treatments, and their efficacy needs to be periodically assessed. Although responses to treatment vary in SMA patients, it is not yet known whether specific features of *SMN2* are correlated with these responses (Cuscó et al., 2020). Thus, it becomes crucial to investigate genomic *SMN2* data to better characterize SMA patients and accurately predict disease evolution.

Here, we report a novel method for sequencing the whole *SMN2* gene based on long-range polymerase chain reaction (PCR) and next-generation sequencing (NGS). The method allows determining all variants described so far as disease modifiers in SMA patients without *SMN1* as well as to identify new variants and structural changes. Furthermore, the technique can be adapted to determine rare pathogenic variants in heterozygous patients with at least one *SMN1* copy. Inclusion of this technique in the routine diagnosis of SMA patients is expected to improve individual genotype–phenotype correlations and, therefore, to help predict more accurately the evolution of the disease.

## 2 | MATERIALS AND METHODS

### 2.1 | Patients

We studied 53 genetically confirmed SMA patients with homozygous absence of *SMN1*. The vast majority of studied patients had three *SMN2* gene copies ($n = 51$), but the cohort included one patient each with two and four *SMN2* copies, respectively. In addition, samples from three patients who carry one *SMN1* copy were studied to assess the versatility of the method to detect *SMN1*-specific pathogenic variants. These patients were: patient SMA54, who had a heterozygous deletion of *SMN1*, the pathogenic variant c.399_402del (p.(Glu134Serfs*14)) in the other *SMN1* allele, and three *SMN2* copies; patient SMA55, who had a heterozygous deletion of *SMN1*, the pathogenic variant c.815A>G (p.(Tyr272Cys)) in the other *SMN1* allele, one complete *SMN2* copy, and two partial *SMN* copies comprising exons 1 to 6 (also known as *SMN1/2*Δ7-8 deletion, Arkblad et al., 2006), and SMA55F, father of SMA55 harboring 2 *SMN1* copies (one with the variant c.815A>G), and two *SMN2* copies.

Genetic confirmation of SMA by bi-allelic defects in *SMN1* (Alías et al., 2009) as well as *SMN2* copy number determination by multiplex ligation-dependent probe amplification (MLPA) were carried out as previously described (Alías et al., 2011). Patients were classified as I, II, or III according to their severity and motor milestone achievements. One patient presenting with type 0 (congenital) SMA was also studied. All patients were unrelated with the exception of the SMA55/55F pair mentioned above (a child with type 0 SMA and his father) and two pairs of siblings, SMA17/SMA18 (both type II) and SMA27/SMA51 (types II and III, respectively). DNA samples were obtained from peripheral blood. All participants or their legal guardians signed written informed consent. The study was approved by the Ethical Committee of our Hospital (PR(AG)229/2018).

### 2.2 | PCR design and library preparation

We studied the complete genomic *SMN2* sequence, including promoter, 5′-UTR and 3′-UTR regions. To this end, we designed three overlapping PCRs (~12-kb each) to amplify a target region of approximately 31.5 Kb (chr5:69,342,511-69,374,064). These long-range PCRs were successfully set up using TaKaRa LA Taq® DNA polymerase (#RR002A; Takara Bio). Given the high homology between *SMN1* and *SMN2* sequences, primers were not specific to *SMN2*, but the 53 patients analyzed have zero *SMN1* copies, allowing in principle an *SMN2* exclusive analysis. In addition, samples from three patients with at least one *SMN1* copy were analyzed. Primer sequences and PCR conditions are given in Table S1.

After amplification, the concentration of the three PCR products was measured using Qubit 2.0 Fluorometer (Thermo Fisher Scientific) and mixed equimolarly. One thousand nanogram of the obtained mixture was fragmented with NEBNext® dsDNA Fragmentase® (New England Biolabs) to generate DNA fragments of ~200 bp. Then, the NEBNext Ultra DNA library prep kit for Illumina and NEBNext® MultiplexOligos for Illumina® Dual Index Primers Set 1 (New England Biolabs) were used to generate the libraries. The necessary purifications and size selections were performed using AMPure XP beads (Beckman Coulter).

The quality and size of the libraries were assessed using QIAxcel (Qiagen) and were quantified with Qubit. Finally, the

libraries of all patients were equimolarly mixed and sequenced using a 500-cycle MiSeq reagent kit v2 with a paired-end run of 2 × 251 bp reads in a MiSeq instrument (Illumina). The number of patients included in each run was calculated to ensure a minimum coverage of ×400. All procedures were performed following the manufacturer's instructions.

In samples harboring at least one copy of *SMN1*, pathogenic variants were ascribed to *SMN1* using a long-range PCR and Sanger sequencing of *SMN1* as previously described (Kubo et al., 2015), with slight modifications.

### 2.3 | Bioinformatics analysis

The data analysis pipeline included the quality trimming of Illumina sequences using Trimmomatic (Bolger et al., 2014). Sequences were mapped to an artificial genome reference that contains only the *SMN2* coordinates, based on the reference genome (UCSC hg 19 version, build 37.1). This strategy avoids arbitrarily reporting only one of the possible alignments produced by most mapping algorithms and therefore avoids dispersion of the read-depth signal from all *SMN* genes. This approach increases the power to detect small changes of coverage and variants with lower AB ratios, which are characteristic of heterozygous variants in multicopy regions. Results show a single location per each read that corresponds to any of the *SMN* genes. Mapping was performed using burrows-wheeler aligner (BWA)-align and BWA-sample with default parameters. Variant calling was performed with Genome Analysis Toolkit (GATK) Unified Genotyper and Haplotype Caller (McKenna et al., 2010), and variant annotation with ANNOVAR (Codina-Solà et al., 2016; Wang et al., 2010).

### 2.4 | Genetic variant types

We discriminated the genetic variants identified in three categories: (1) paralogous sequence variants (PSVs), which are positions differing between duplicated genes (in this context, variants differing between *SMN1* and *SMN2*); (2) rare single-nucleotide variants (SNVs), which are those present in less than 1% of the population or never described; and (3) single-nucleotide polymorphisms (SNPs), which are variations in a single position present in over 1% of the population. Determining whether the variants found are SNPs, SNVs, or PSVs allowed us to evaluate the specificity of the method in the detection of *SMN1* versus *SMN2* and to identify hybrid structures in the samples. This classification is not related to a possible effect of variants on *SMN2* expression and phenotype, a topic out of the scope of this article.

PSV ratios were computed as quotients of the number of reads belonging to a given functional copy and the total coverage.

Mutation nomenclature refers to GenBank NC_000005.9, RefSeq NM_000344.3 for *SMN1*, and RefSeq NM_017411.4 for *SMN2*.

## 3 | RESULTS

In all analyzed samples, a mean depth of coverage of 1720× (435×–7478×) was obtained. This high coverage allows to accurately determine the allelic frequency of the detected variants in each sample (AB ratio). Given that the copy number of *SMN2* genes had been already studied by MLPA, we could use the AB ratio to calculate the number of copies in which the detected variant was present.

To assess the specificity of our method to identify genetic variants in the 5q13 complex region, we first analyzed PSVs that are specific for the *SMN2* gene. To this end, a list of *SMN1* and *SMN2* PSVs was created based initially on a literature review (Monani et al., 1999), and expanded with the information obtained from a basic local alignment search tool-like alignment tool (BLAT) between the two genes using the reference genome Hg19. The detailed information obtained from sequencing the *SMN2* gene in samples from 53 patients was used to determine and eventually confirm which positions should be considered real PSVs or SNPs or SNVs. Based on previous reports and BLAT data, we initially considered 22 PSVs (Table 1). Four of these variants exhibited high variability between samples, indicating that they should not be considered genuine PSVs (by definition). Two other variants, including the candidate position g.69371981A/C previously described as a PSV (Monani et al., 1999) and g.69367553G/A were not found neither in *SMN1* nor *SMN2* genes from our samples and were considered as very infrequent variants. The remaining 16 changes are genuine PSVs between *SMN1* and *SMN2* genes, being c.835-1606C/T in intron 6 considered here for the first time as a PSV (Table 1). These PSVs were present in all the reads and in more than 95% of the samples, confirming that all these positions were specific for *SMN2*. However, in two patients (SMA04 and SMA39) we observed some discrepancies in the PSV ratio compatible with the presence of hybrid genes. Both patients presented six *SMN1* PSVs all located in intron 6 (Chr5: 69370451- 69370895). The PSV ratio indicates that in SMA04, two of the three *SMN* copies are *SMN2–SMN1* hybrids, while in SMA39, only one of the three copies is an *SMN2–SMN1* hybrid (Figure 1).

We also detected variants previously described as beneficial for *SMN2* function (Bernal et al., 2010; Prior et al., 2009; Ruhno et al., 2019; Wu et al., 2017). Using the AB ratio data, we were able to confirm the presence of the c.859G>C (p.(Gly287Arg)) variant in patient SMA52 with two *SMN2* copies in the heterozygous state and of c.835-44A>G (g.69372304A>G, commonly known as A-44G) in one of the three *SMN2* copies of patient SMA21 (Figure 2 and Table S2). Furthermore, in two sisters with different phenotypes (SMA27, type II and SMA51, type III) and three *SMN2* copies, we found the PSV conversion c.835-1897C>T (g.69370451C>T) in one of their three alleles. In our series, we did not identify any of the other more recently published variants considered as candidates to be modifiers (Ruhno et al., 2019; Wadman et al., 2020).

Finally, to verify whether our technique could also be extended to identify *SMN1* variants, we studied samples from three SMA patients harboring at least one *SMN1* copy. The PSV genotype in all three cases was in agreement with the copy numbers previously determined by MLPA. SMA54 (1_*SMN1*/3_*SMN2* MLPA genotype) showed the *SMN1* PSVs at a frequency of ~25% (one of the four *SMN* copies of the

**TABLE 1** The 22 candidate positions for paralogous sequence variants (PSVs) between *SMN1 and SMN2* are shown

| *SMN1* position | *SMN2* position | Gene location | Ref *SMN1* | Ref *SMN2* | Categorization | Conversion nomenclature *SMN2>SMN1* | Source of information |
|---|---|---|---|---|---|---|---|
| 70231509 | 69356085 | Intron 1 | G | A | SNP[a] | c.82-3157A>G | BLAT |
| 70240028 | 69364605 | Intron 4 | G | A | SNP[a] | c.628-457A>G | BLAT |
| 70242435 | 69367010 | Intron 6 | T | C | SNP[a] | c.834+432C>T | BLAT |
| 70242978 | 69367553 | Intron 6 | A | G | SNV[b] | c.834+975G>A | BLAT |
| 70244142 | 69368717 | Intron 6 | A | G | SNP[a] | c.834+2139G>A | Monani et al. /BLAT |
| **70245876** | **69370451** | **Intron 6** | **T** | **C** | **PSV** | **c.835-1897C>T** | **Monani et al. /BLAT** |
| **70246016** | **69370591** | **Intron 6** | **G** | **A** | **PSV** | **c.835-1757A>G** | **Monani et al. /BLAT** |
| **70246019** | **69370594** | **Intron 6** | **T** | **C** | **PSV** | **c.835-1754C>T** | **Monani et al. /BLAT** |
| **70246156** | **69370731** | **Intron 6** | **G** | **A** | **PSV** | **c.835-1617A>G** | **Monani et al. /BLAT** |
| **70246167** | **69370742** | **Intron 6** | **T** | **C** | **PSV** | **c.835-1606C>T** | **BLAT** |
| **70246320** | **69370895** | **Intron 6** | **G** | **A** | **PSV** | **c.835-1453A>G** | **Monani et al. /BLAT** |
| **70246793** | **69371368** | **Intron 6** | **G** | **A** | **PSV[c]** | **c.835-980A>G** | **Monani et al. /BLAT** |
| **70246872** | **69371448** | **Intron 6** | **-** | **AGGCA** | **PSV[c]** | **c.835-900_835-896del** | **Monani et al. /BLAT** |
| **70246919** | **69371499** | **Intron 6** | **A** | **C** | **PSV[c]** | **c.835-849C>A** | **Monani et al. /BLAT** |
| **70247219** | **69371799** | **Intron 6** | **G** | **A** | **PSV[c]** | **c.835-549A>G** | **Monani et al. /BLAT** |
| **70247290** | **69371870** | **Intron 6** | **T** | **C** | **PSV** | **c.835-478C>T** | **Monani et al. /BLAT** |
| 70247401 | 69371981 | Intron 6 | C | A | SNV[b,c] | c.835-367A>C | Monani et al. /BLAT |
| **70247724** | **69372304** | **Intron 6** | **G** | **A** | **PSV[c]** | **c.835-44A>G** | **Monani et al. /BLAT** |
| **70247773** | **69372353** | **Exon 7** | **C** | **T** | **PSV[c]** | **c.840T>C** | **Monani et al. /BLAT** |
| **70247921** | **69372501** | **Intron 7** | **A** | **G** | **PSV[c]** | **c.*3+100G>A** | **Monani et al. /BLAT** |
| **70248036** | **69372616** | **Intron 7** | **A** | **G** | **PSV[c]** | **c.*3+215G>A** | **Monani et al. /BLAT** |
| **70248501** | **69373081** | **Exon 8** | **G** | **A** | **PSV[c]** | **c.*239A>G** | **Monani et al. /BLAT** |

*Note*: These positions were obtained from a previous bibliographic compilation (Monani et al., 1999) and were complemented with a BLAT between the two genes, as deposited in the reference genome Hg19. Repetitive regions (polyA, polyT, and polyGT) were discarded. The candidate positions were genotyped in the patients studied in this study ($n = 53$) and in samples with at least one SMN1 ($n = 3$) to check for consistency. From the 22 candidates, six were discarded as PSVs, four of them are listed as SNPs instead, and the remaining two as rare SNV. Therefore, a total of 16 nucleotides (in bold in the Table) differentiate *SMN1* and *SMN2* genes, 10 of which had been previously described and validated, 5 had been described but not validated, and 1 is considered here as a PSV for the first time.

Abbreviations: BLAT, BLAST-like alignment tool; SNP, single-nucleotide polymorphism; SNV, single-nucleotide variant.

[a]These positions show high variability between samples and were therefore classified as SNPs (including position c.835-367C/A previously validated by Monani et al., 1999).

[b]In these positions, the same nucleotide has always been found in both SMN1 and SMN2. Thus, G>A and A>C exchanges appears to be very rare SNV found in the reference genome.

[c]These 11 positions were previously classified as bona fide PSVs, after being tested in a control population of 15 individuals (Monani et al., 1999).

patient). The *SMN1* PSV frequency in SMA55 and SMA55F was ~50% in line with their genotypes (1_*SMN1*/1_*SMN2* and 2_*SMN1*/2_*SMN2*, respectively). A similar consistency was obtained with *SMN1* pathogenic variants. Thus, c.399_402del (p.(Glu134Serfs*14)) in SMA54 showed a frequency of ~25%, while c.815A>G (p.(Tyr272Cys)) showed a frequency of ~50% in SMA55 and of 25% in SMA55F. *SMN1*-specific Sanger sequencing (Kubo et al., 2015) revealed that patients SMA54 and SMA55 present this variant in hemizygous state, whereas in SMA55F it was detected in heterozygous state (Figure 3).

In addition, in patient SMA55, it was also possible to corroborate the presence of the common partial *SMN1*/2Δ7-8 deletion of the 3′ region previously detected by MLPA (Figure 4). Indeed, analysis of the AB ratios of all SNVs revealed a discrepancy between the 5′ and 3′ *SMN* region, showing ratios compatibles with four copies (25%–75%) in the 5′ region (from promoter to exon 6), but compatible with two copies (50%–100%) in the 3′ region confirming the presence of the two alleles with the *SMN1*/2Δ7-8 deletion (further explanation in Figure 4).

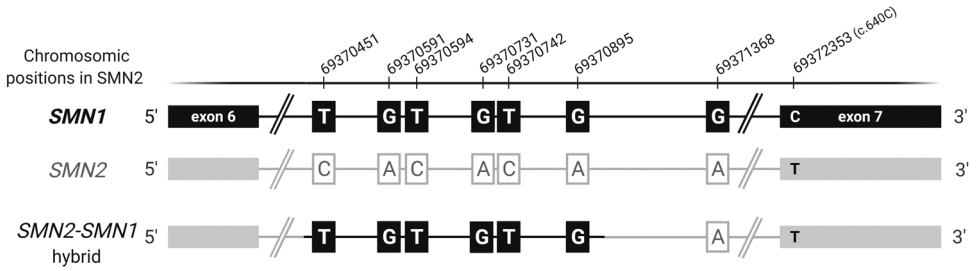



**FIGURE 1** Structure of the *SMN2–SMN1* hybrid detected in patients SMA4 and SMA39. PSVs were genotyped in all patients. In SMA4 and SMA39, six *SMN1* PSVs located in intron 6 (Chr5: 69370451-69370895) were detected, which indicates the presence of hybrid genes. The AB ratio indicates that in SMA04, two of the three copies are *SMN2–SMN1* hybrids (*SMN1* PSVs in 66%) while in SMA39, only one of the three copies is an *SMN2–SMN1* hybrid (*SMN1* PSVs in 33%). PSV, paralogous sequence variant

## 4 | DISCUSSION

We have developed a novel, practical method for the genomic analysis of both *SMN1* and *SMN2* regions. The technique focuses on the genomic characterization of each *SMN2* copy in SMA patients regardless of the gene copy number and can also be applied to detect pathogenic variants in heterozygous SMA patients with at least one *SMN1* copy.

Currently published methods to study the complete sequence of the *SMN2* gene include whole genome sequencing (WGS) and multiplexed direct genomic selection (MDiGS) sequencing (Chen et al., 2020; Ruhno et al., 2019). WGS is a rather expensive and laborious technique and usually needs a complex bioinformatics analysis. In MDiGS, whole DNA is prepared in libraries, target regions are captured by bacterial artificial chromosomes (BAC) probes of *SMN* (not specific for *SMN1* or *SMN2*), as well as of *CFTR* and *PLS3*, for quantitative comparison. Thus, the hybridization step is rather large and complex. By contrast, the new method described here based on long-range PCR and NGS can be easily implemented in any genetics laboratory performing NGS applications.

The main advantages of our method, specifically designed to deal with the complexity of the SMA region, include accessible cost, relative simplicity, and speedy results (usually obtained in around 3 days). This practical protocol can easily genotype both PSVs and rare variants already described in SMA patients, but it might also help to identify new variants and *SMN2–SMN1* hybrid genes as part of more investigative and personalized approaches. The high coverage obtained at the nucleotide level with the new method allows the determination of the number of copies in which a specific variant is present, which, in turn, is essential to characterize the genomic architecture of each *SMN2* copy. Indeed, although the method does not directly quantify *SMN2* copy numbers, calculation of the number of copies and the presence of partial *SMN* genes is straightforward using the allelic frequencies of the variants and might be used to, for example, confirm results reported by other methods, such as MLPA.

Employing this method to samples from SMA patients, we elaborated an updated list of 16 PSVs between *SMN1* and *SMN2*, including one previously never described. PSV genotyping would help to detect hybrid genes and to discover new, potentially relevant conversions that

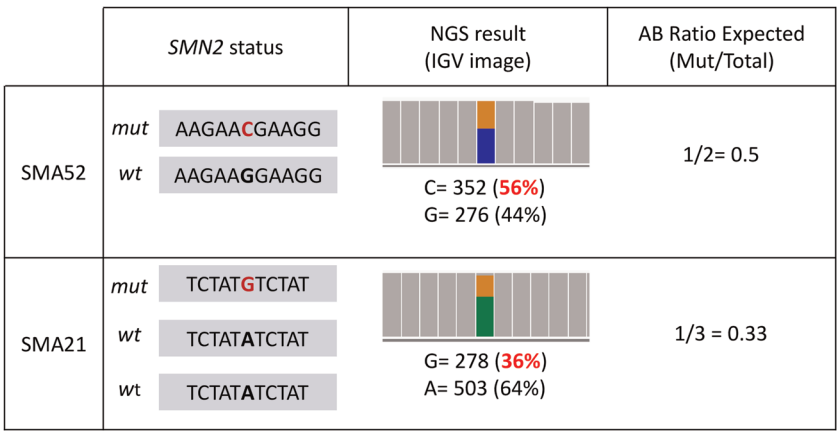



**FIGURE 2** Utility of AB ratios to calculate the number of the copies in which variants are present. Patient SMA52 has two *SMN2* copies (determined by MLPA) and the variant NM_017411.4:c.859G>C (p.(Gly287Arg)) was detected with a frequency of 56%, in agreement with the AB ratio expected for the variant in one over two alleles. Patient SMA21 has three *SMN2* copies (determined by MLPA), and the variant c.835-44A>G (NC_000005.9: g.69372304A>G) was detected with a frequency of 36%, in agreement with the AB ratio expected for the variant in one over three alleles. MLPA, multiplex ligation-dependent probe amplification

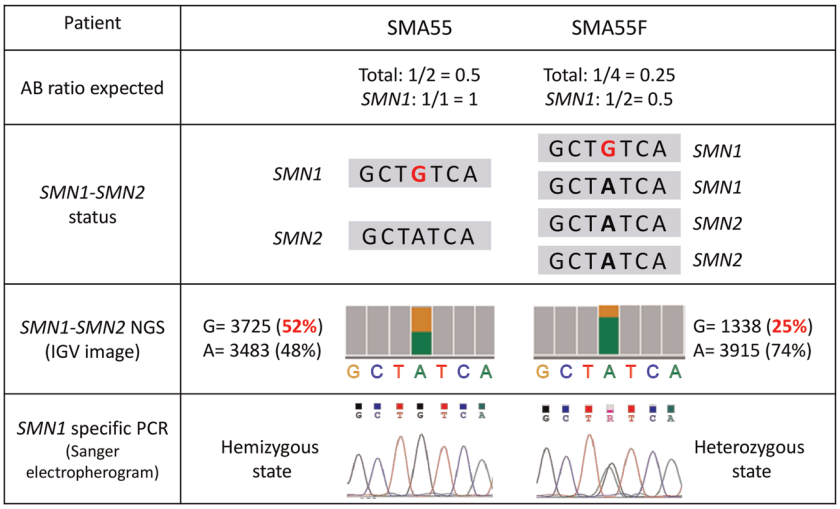



**FIGURE 3** Detection of the pathogenic variant NM_017411.4:c.815A>G in samples SMA55 and SMA55F. Patient SMA55 has one *SMN1* copy (with the variant c.815A>G) and one *SMN2* copy, while his father (SMA55F) has two *SMN1* (one copy with the variant c.815A>G) and two *SMN2* copies (determined by MLPA and Sanger). The pathogenic variant c.815A>G was detected in SMA55 and SMA55F through NGS with a frequency of 52% and 25%, respectively. The *SMN1*-specific PCR performed confirms that the pathogenic variant c.815A>G is present in *SMN1* since we observed the variant in hemizygous status in SMA55 and in heterozygous status in SMA55F. MLPA, multiplex ligation-dependent probe amplification; NGS, next-generation sequencing; PCR, polymerase chain reaction

might act as disease modifiers (Ruhno et al., 2019; Wu et al., 2017). Our method is highly specific given that it allows detection of positive modifiers, such as c.859G>C, c.835-44A>G (g.69372304A>G), and c.835-1897C>T (g.69370451C>T). The method also determines the number of *SMN2* copies carrying modifier variants through analysis of allelic frequencies (Figure 2).

Based on the analysis of the entire genomic *SMN2* region and according to the PSV genotyping, we also identified specific *SMN2–SMN1* hybrid structures previously undetected by MLPA. The presence of hybrid *SMN* genes has been previously described based mainly on the analysis of exons 7 and 8 (Cuscó et al., 2001; Hahnen et al., 1996). This phenomenon occurs because the complex 5q13 region contains segmental duplications prone to nonallelic homologous recombinations, deletions, duplications, and gene conversion events. Usually, hybrid genes are detected because of the homozygous absence of *SMN1* exon 7 coupled to the presence of *SMN1* exon 8. Thus, hybrids upstream exon 7 cannot be detected by common methods of diagnosis, including MLPA. In contrast, with our long PCR-based approach, we were able to detect

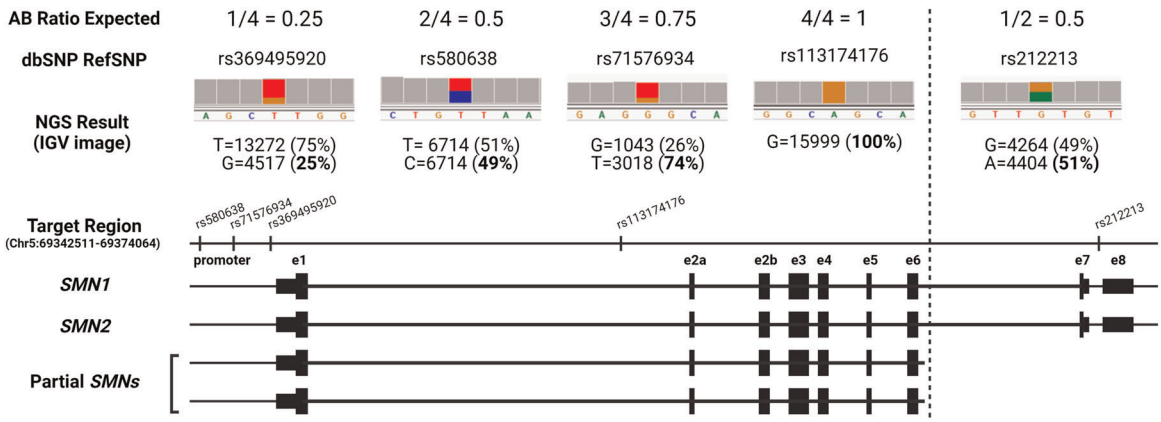



**FIGURE 4** Description of the utility of AB ratios to determine the presence of two partial *SMN* genes (*SMN1/2Δ7/8*). The patient (SMA55) has one *SMN1* (with a pathogenic variant in exon 6*), one *SMN2*, and two partial *SMN*s from promoter to exon 6 (determined by MLPA). In the 5′ region (promoter-ex6), the patient has a total of four *SMN* copies; consequently, we detected different SNPs with an allelic frequency of 25%, 50%, 75%, and 100%. By contrast, in the 3′ region (in6-ex8), the patient has two *SMN* copies, and only variants with an allelic frequency of 50% are detected. Note that PSVs are located in the 3′ region; therefore, it is not possible to determine whether these partial genes are derived from *SMN1* or *SMN2*. MLPA, multiplex ligation-dependent probe amplification; PSV, paralogous sequence variant; SNP, single-nucleotide polymorphisms

novel hybrid genes consisting of a *SMN2* gene with a fragment of intron 6 derived from *SMN1* (Figure 1). Although additional studies are needed to characterize the function of these two hybrid genes, their detection might provide clues about possible functional differences between *SMN2* genes.

The method described here also allows a complete genomic *SMN1* analysis. In fact, we were able to detect the pathogenic variants c.399_402del (p.(Glu134Serfs*14)) and c.815A>G (p.(Tyr272Cys)) in patients and carriers, and to determine their frequencies (e.g., homozygous vs. heterozygous cases). Further confirmation that a certain variant is located in *SMN1* can be achieved with a specific PCR of the gene (Kubo et al., 2015). Given that in our method, we analyze the whole *SMN1*, including promoter, 5′ and 3′ regions as well as all introns, it emerges as potentially useful to study complex SMA cases in which only one allele alteration (deletion or point mutation) has been detected using conventional techniques (Alías et al., 2009).

Finally, we were able to confirm the presence of partial *SMN* copies (*SMN1/2*Δ7-8) using the AB ratios. The *SMN1/2*Δ7-8 deletion has been widely described in the literature (Arkblad et al., 2006; Calucho et al., 2018; Chen et al., 2020; Ruhno et al., 2019; Vijzelaar et al., 2019) as a variant relatively common in general population (Europeans 15.7%, Vijzelaar et al., 2019). In almost all cases, the breakpoints are consistent (3643 bp before exon 7 and 1587 bp after exon 8 according to Ruhno et al., 2019), so it is considered that this deletion is the result of a single deletion event (Vijzelaar et al., 2019).

The number of *SMN2* copies and specific variants of the gene have been established as the main disease modifiers of the SMA phenotype. However, to date, studies in discordant siblings do not support the hypothesis that the intra-familiar variability is due to variants in the *SMN2* locus (Calucho et al., 2018; Cuscó et al., 2006; Ruhno et al., 2019) as is the case of our sisters sharing the modifier variant c.835-1897C>T (see Table S2). Although *SMN2* is the target for splicing modifiers in the current therapeutic scenario, thorough sequencing of the gene is almost never performed in genetic diagnostic laboratories. Application of our method and thus the availability of detailed *SMN2* sequences of SMA patients would help to solve discrepancies in genotype–phenotype correlations, as well as deepen the study of intra-familiar variability, in a prospective manner, for the analysis of presymptomatic cases detected in newborn screening. Indeed, being the method so specific and fast, it is envisaged that it could be accommodated within the time frame for therapeutic decisions in SMA newborns. It is also essential to unveil possible linkages between specific *SMN2* variants, factors involved in *SMN2* splicing, and responses to therapies. SMA treatments are very expensive, and proof of their efficacy is periodically assessed in SMA patients. In particular, nusinersen is an 18-mer oligonucleotide that binds the ISNN1 region of the intron 7, and to date, none of the modifier variants described is located in this region (Ruhno et al., 2019; Wadman et al., 2020; this study). However, in addition to the known modifier variants, other features of their *SMN2* genes may be correlated with the level of responsiveness and effectiveness, an issue that warrants further investigation. Thus, the discovery and validation of positive and negative *SMN2* variants in each patient remain a crucial issue in SMA diagnosis and research. An additional benefit of implementing the new method besides the characterization of *SMN2* sequences in patients with a homozygous deletion of *SMN1* is its application to study *SMN1* in SMA patients retaining at least one *SMN1* copy. Therefore, the versatile method described here is a useful tool to approach *SMN1* and *SMN2* deep sequencing, which can be easily implemented in most SMA diagnostic laboratories.

## ACKNOWLEDGMENTS

This study was supported by Grants from Biogen ESP-SMG-17-11256 and Galicia AME (to E. F. T. supporting L. B.-P.), "Fundación Daniel Bravo Andreu" (to E. F. T. and P. F.-P.), SMA Europe (to E.F.T and P.F.-P.), and Spanish Instituto de Salud Carlos III, Fondo de Investigaciones Sanitarias and cofunded with ERDF funds (Grant No. FIS PI18/000687) (to E. F. T. and L. A.).

## CONFLICT OF INTERESTS

A patent is in preparation for this methodology and pipeline (I.C. and E. F. T.). E.F.T has received grant support to conduct clinical trials on SMA from Ionis/Biogen and serves as a consultant to AveXis, Novartis, Biogen, Biologix, Cytokinetics, Roche. The remaining authors declare that there are no conflict of interests.

## AUTHOR CONTRIBUTIONS

The enclosed manuscript has been seen and approved by all the authors, and they have taken care to ensure the integrity of the work.

## WEB RESOURCES

The following web sources have been used to carry out this study:

http://www.fda.gov,
https://genome-euro.ucsc.edu/,
http://www.hgmd.cf.ac.uk/ac/index.php,
https://mutalyzer.nl/,
https://gnomad.broadinstitute.org/,
https://arxiv.org/abs/1303.3997v2.

## DATA AVAILABILITY STATEMENT

All data and scripts used to generate the analyses of this paper are available upon request unless that the type of request compromises ethical standards or legal requirements. All variants reported in this article are now listed in the ClinVar database (https://www.ncbi.nlm.nih.gov/clinvar/).

## ORCID

*Laura Blasco-Pérez* https://orcid.org/0000-0001-6483-5742
*Jordi Leno* https://orcid.org/0000-0001-6049-256X
*Laura Alias* https://orcid.org/0000-0003-2193-6185
*Pablo Fuentes-Prior* http://orcid.org/0000-0002-6618-3204
*Ivon Cuscó* https://orcid.org/0000-0003-2104-9332
*Eduardo F. Tizzano* http://orcid.org/0000-0002-7116-6310

## REFERENCES


Alías, L., Bernal, S., Barceló, M. J., Also-Rallo, E., Martínez-Hernández, R., Rodríguez-Alvarez, F. J., Hernández-Chico, C., Baiget, M., & Tizzano, E. F. (2011). Accuracy of marker analysis, quantitative real-time polymerase chain reaction, and multiple ligation-dependent probe amplification to determine SMN2 copy number in patients with spinal muscular atrophy. *Genetic Testing and Molecular Biomarkers*, *15*(9), 587–594. https://doi.org/10.1089/gtmb.2010.0253

Alías, L., Bernal, S., Fuentes-Prior, P., Barceló, M. J., Also, E., Martínez-Hernández, R., Rodríguez-Alvarez, F. J., Martín, Y., Aller, E., Grau, E., Peciña, A., Antiñolo, G., Galán, E., Rosa, A. L., Fernández-Burriel, M., Borrego, S., Millán, J. M., Hernández-Chico, C., Baiget, M., & Tizzano, E. F. (2009). Mutation update of spinal muscular atrophy in Spain: Molecular characterization of 745 unrelated patients and identification of four novel mutations in the SMN1 gene. *Human Genetics*, *125*(1), 29–39. https://doi.org/10.1007/s00439-008-0598-1

Arkblad, E. L., Darin, N., Berg, K., Kimber, E., Brandberg, G., Lindberg, C., Holmberg, E., Tulinius, M., & Nordling, M. (2006). Multiplex ligation-dependent probe amplification improves diagnostics in spinal muscular atrophy. *Neuromuscular Disorders*, *16*(12), 830–838. https://doi.org/10.1016/j.nmd.2006.08.011

Bernal, S., Alías, L., Barceló, M. J., Also-Rallo, E., Martínez-Hernández, R., Gámez, J., Guillén-Navarro, E., Rosell, J., Hernando, I., Rodriguez-Alvarez, F. J., Borrego, S., Millán, J. M., Hernández-Chico, C., Baiget, M., Fuentes-Prior, P., & Tizzano, E. F. (2010). The c.859>C variant in the SMN2 gene is associated with types II and III SMA and originates from a common ancestor. *Journal of Medical Genetics*, *47*(9), 640–642.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Boza-Morán, M. G., Martínez-Hernández, R., Bernal, S., Wanisch, K., Also-Rallo, E., Le Heron, A., Alías, L., Denis, C., Girard, M., Yee, J.-K., Tizzano, E. F., & Yáñez-Muñoz, R. J. (2015). Decay in survival motor neuron and plastin 3 levels during differentiation of iPSC-derived human motor neurons. *Scientific Reports*, *5*, 11696.

Calucho, M., Bernal, S., Alías, L., March, F., Venceslá, A., Rodríguez-Álvarez, F. J., Aller, E., Fernández, R. M., Borrego, S., Millán, J. M., Hernández-Chico, C., Cuscó, I., Fuentes-Prior, P., & Tizzano, E. F. (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*(3), 208–215. https://doi.org/10.1016/j.nmd.2018.01.003

Chen, X., Sanchis-Juan, A., French, C. E., Connell, A. J., Delon, I., Kingsbury, Z., Chawla, A., Halpern, A. L., Taft, R. J., Bentley, D. R., Butchbach, M. E. R., Raymond, F. L., & Eberle, M. A. (2020). Spinal muscular atrophy diagnosis and carrier screening from genome sequencing data. *Genetics in Medicine*, *22*(5), 945–953. https://doi.org/10.1038/s41436-020-0754-0

Codina-Solà, M., Flores, R., Homs, A., Aguado, C., Pérez-Jurado, L., & Cuscó, I. (2016). Genomic and genetic variation at complex segmental duplications in Autism Spectrum disorders [Conference presentation abstract]. *European Journal of Human Genetics*, *24*, 181.

Cuscó, I., Barceló, M. J., Del Rio, E., Martín, Y., Hernández-Chico, C., Bussaglia, E., Baiget, M., & Tizzano, E. F. (2001). Characterisation of SMN hybrid genes in Spanish SMA patients: De novo, homozygous and compound heterozygous cases. *Human Genetics*, *108*(3), 222–229. https://doi.org/10.1007/s004390000452

Cuscó, I., Barceló, M. J., Rojas-García, R., Illa, I., Gámez, J., Cervera, C., Pou, A., Izquierdo, G., Baiget, M., & Tizzano, E. F. (2006). SMN2 copy number predicts acute or chronic spinal muscular atrophy but does not account for intrafamilial variability in siblings. *Journal of Neurology*, *253*(1), 21–25. https://doi.org/10.1007/s00415-005-0912-y

Cuscó, I., Bernal, S., Blasco-Pérez, L., Calucho, M., Alias, L., Fuentes-Prior, P., & Tizzano, E. F. (2020). Practical guidelines to manage discordant situations of SMN2 copy number in patients with spinal muscular atrophy. *Neurology: Genetics*, *6*(6), e530. https://doi.org/10.1212/nxg.0000000000000530

Finkel, R. S., Mercuri, E., Darras, B. T., Connolly, A. M., Kuntz, N. L., Kirschner, J., Chiriboga, C. A., Saito, K., Servais, L., Tizzano, E., Topaloglu, H., Tulinius, M., Montes, J., Glanzman, A. M., Bishop, K., Zhong, Z. J., Gheuens, S., Bennett, C. F., Schneider, E., … De Vivo, D. C. (2017). Nusinersen versus sham control in infantile-onset spinal muscular atrophy. *New England Journal of Medicine*, *377*(18), 1723–1732. https://doi.org/10.1056/NEJMoa1702752

Hahnen, E., Schönling, J., Rudnik-Schöneborn, S., Zerres, K., & Wirth, B. (1996). Hybrid survival motor neuron genes in patients with autosomal recessive spinal muscular atrophy: New insights into molecular mechanisms responsible for the disease. *American Journal of Human Genetics*, *59*(5), 1057–1065.

Kubo, Y., Nishio, H., & Saito, K. (2015). A new method for SMN1 and hybrid SMN gene analysis in spinal muscular atrophy using long-range PCR followed by sequencing. *Journal of Human Genetics*, *60*(5), 233–239. https://doi.org/10.1038/jhg.2015.16

Lefebvre, S., Bürglen, L., Reboullet, S., Clermont, O., Burlet, P., Viollet, L., Benichou, B., Cruaud, C., Millasseau, P., Zeviani, M., Le Paslier, D., Frézal, J., Cohen, D., Weissenbach, J., Munnich, A., & Melki, J. (1995). Identification and characterization of a spinal muscular atrophy-determining gene. *Cell*, *80*(1), 155–165.

Lorson, C. L., Hahnen, E., Androphy, E. J., & Wirth, B. (1999). A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(11), 6307–6311.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & DePristo, M. A. (2010). The genome analysis toolkit: A map reduce framework for analyzing next-generation DNA sequencing data. Genome Res, 20(9), 1297–303. https://doi.org/10.1101/gr.107524.110.20

Mendell, J. R., Al-Zaidy, S., Shell, R., Arnold, W. D., Rodino-Klapac, L. R., Prior, T. W., Lowes, L., Alfano, L., Berry, K., Church, K., Kissel, J. T., Nagendran, S., L'Italien, J., Sproule, D. M., Wells, C., Cardenas, J. A., Heitzer, M. D., Kaspar, A., Corcoran, S., … Kaspar, B. K. (2017). Single-dose gene-replacement therapy for spinal muscular atrophy. *New England Journal of Medicine*, *377*(18), 1713–1722. https://doi.org/10.1056/NEJMoa1706198

Monani, U. R., Lorson, C. L., Parsons, D. W., Prior, T. W., Androphy, E. J., Burghes, A. H. M., & McPherson, J. D. (1999). A single nucleotide difference that alters splicing patterns distinguishes the SMA gene SMN1 from the copy gene SMN2. *Human Molecular Genetics*, *8*(7), 1177–1183.

Prior, T. W., Krainer, A. R., Hua, Y., Swoboda, K. J., Snyder, P. C., Bridgeman, S. J., Burghes, A. H. M., & Kissel, J. T. (2009). A positive modifier of spinal muscular atrophy in the SMN2 gene. *American Journal of Human Genetics*, *85*(3), 408–413.

Rochette, C. F., Gilbert, N., & Simard, L. R. (2001). SMN gene duplications and the emergence of the SMN2 gene ocurred in distinct hominids: SMN2 is unique to Homo sapiens. *Human Genetics*, *108*(3), 255–266.

Ruhno, C., McGovern, V. L., Avenarius, M. R., Snyder, P. J., Prior, T. W., Nery, F. C., Muhtaseb, A., Roggenbuck, J. S., Kissel, J. T., Sansone, V. A., Siranosian, J. J., Johnstone, A. J., Nwe, P. H., Zhang, R. Z., Swoboda, K. J., & Burghes, A. H. M. (2019). Complete sequencing of the SMN2 gene in SMA patients detects SMN gene deletion junctions and variants in SMN2 that modify the SMA phenotype. *Human Genetics*, *138*, 241–256. https://doi.org/10.1007/s00439-019-01983-0

Soler-Botija, C., Cuscó, I., Caselles, L., López, E., Baiget, M., & Tizzano, E. F. (2005). Implication of fetal SMN2 expression in type I SMA pathogenesis: Protection or pathological gain of function? *Journal of Neuropathology and Experimental Neurology*, *64*(3), 215–223.

Sugarman, E. A., Nagan, N., Zhu, H., Akmaev, V. R., Zhou, Z., Rohlfs, E. M., Flynn, K., Hendrickson, B. C., Scholl, T., Sirko-Osadsa, D. A., & Allitto, B. A. (2012). Pan-ethnic carrier screening and prenatal diagnosis for spinal muscular atrophy: Clinical laboratory analysis of >72 400 specimens. *European Journal of Human Genetics*, *20*(1), 27–32. https://doi.org/10.1038/ejhg.2011.134

Vijzelaar, R., Snetselaar, R., Clausen, M., Mason, A. G., Rinsma, M., Zegers, M., Molleman, N., Boschloo, R., Yilmaz, R., Kuilboer, R., Lens, S., Sulchan, S., & Schouten, J. (2019). The frequency of SMN gene variants lacking exon 7 and 8 is highly population dependent. *PLOS One*, *14*(7), 1–12. https://doi.org/10.1371/journal.pone.0220211

Wadman, R. I., Jansen, M. D., Stam, M., Wijngaarde, C. A., Curial, C. A. D., Medic, J., Sodaar, P., Schouten, J., Vijzelaar, R., Lemmink, H. H., van den Berg, L. H., Groen, E. J. N., & van der Pol, W. L. (2020). Intragenic and structural variation in the SMN locus and clinical variability in spinal muscular atrophy. *Brain Communications*, *2*(2), 1–13. https://doi.org/10.1093/braincomms/fcaa075

Wang, C. H., Finkel, R. S., Bertini, E. S., Schroth, M., Simonds, A., Wong, B., Aloysius, A., Morrison, L., Main, M., Crawford, T. O., & Trela, A. (2007). Consensus statement for standard of care in spinal muscular atrophy. *Journal of Child Neurology*, *22*(8), 1027–1049.

Wang, K., Li, M., & Hakonarson, H. (2010). ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Research*, *38*(16), 1–7. https://doi.org/10.1093/nar/gkq603

Wirth, B., Garbes, L., & Riessland, M. (2013). How genetic modifiers influence the phenotype of spinal muscular atrophy and suggest future therapeutic approaches. *Current Opinion in Genetics and Development*, *23*(3), 330–338.

Wu, X., Wang, S. H., Sun, J., Krainer, A. R., Hua, Y., & Prior, T. W. (2017). A-44G transition in SMN2 intron 6 protects patients with spinal muscular atrophy. *Human Molecular Genetics*, *26*(14), 2768–2780. https://doi.org/10.1093/HMG/DDX166

Zerres, K., & Rudnik-Schöneborn, S. (1995). Natural history in proximal spinal muscular atrophy. Clinical analysis of 445 patients and suggestions for a modification of existing classifications. *Archives of Neurology*, *52*(5), 518–523.

## SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

**How to cite this article:** Blasco-Pérez, L., Paramonov, I., Leno, J., Bernal, S., Alias, L., Fuentes-Prior, P., Cuscó, I., & Tizzano, E. F. (2021). Beyond copy number: A new, rapid, and versatile method for sequencing the entire *SMN2* gene in SMA patients. *Human Mutation*, 42, 787–795. https://doi.org/10.1002/humu.24200

**Supplementary Table 1. A)** Detail of the sequence and target region of the primers used to amplify the SMN2 region. **B)** Detail of the conditions used to amplify the target region.

**A. Primers sequences**

| Primer Name | Primer Sequence | Target Region (Hg19) |
|---|---|---|
| SMN2.frag1 F | TTAGCCACCACCCCGACATTATTTGAAA | chr5:70,217,918-70,229,238 |
| SMN2.frag1 R | AAGTTAAAGACGCTCCTATGACACAGCC | |
| SMN2.frag2 F | ACCCGGCCAACTTACCCATCTTTTAATT | chr5:70,227,343-70,239,599 |
| SMN2.frag2 R | TCCATGGAAGAAATCACCGCTTTACACA | |
| SMN2.frag3 F | GCAAAATCTGTCCGATCTACTTTCCCCA | chr5:70,238,316-70,249,484 |
| SMN2.frag3 R | CTCTTCCAGTTGCCCTCTTCTTTGACAA | |

**B. Long-Range PCR Conditions**

| Step | Temperature | Time | Cycles |
|---|---|---|---|
| Initial Denaturation | 94ºC | 2 min | 1 |
| Denaturation | 94ºC | 15 sec | 10 |
| Annealing | 60ºC | 30 sec | |
| Elongation | 68ºC | 8 min | |
| Denaturation | 94ºC | 15 sec | 20 |
| Annealing | 60ºC | 30 sec | |
| Elongation | 68ºC | 8 min (+20sec per cycle) | |
| Final Elongation | 72ºC | 20 min | 1 |

**Supplementary Table 2.** Phenotype of patients with variants previously described as modifier variants. The SMA type was defined according to Cuscó et al., 2020. RefSeq NG_008728.1; NM_017411.4.

| Patient | SMA Type | *SMN2* copy number | Modifier variant in *SMN2* | Ratio of the modifier variant |
|---|---|---|---|---|
| SMA52 | IIIa | 2 | c.859G>C | 1/2 |
| SMA21 | IIIb | 3 | c.835-44A>G | 1/3 |
| SMA27* | IIb | 3 | c.835-1897C>T | 1/3 |
| SMA51* | IIIa | 3 | c.835-1897C>T | 1/3 |

*siblings

# CHAPTER 3

International Journal of
*Molecular Sciences*

MDPI

*Article*

# Deep Molecular Characterization of Milder Spinal Muscular Atrophy Patients Carrying the c.859G>C Variant in *SMN2*

**Laura Blasco-Pérez [1,2], Mar Costa-Roger [1,2], Jordi Leno-Colorado [1,2], Sara Bernal [3,4], Laura Alias [3,4], Marta Codina-Solà [1,2], Desirée Martínez-Cruz [1,2], Claudia Castiglioni [5], Enrico Bertini [6], Lorena Travaglini [6], José M. Millán [4,7], Elena Aller [4,7], Javier Sotoca [8], Raúl Juntas [8], Christina Engel Hoei-Hansen [9,10], Antonio Moreno-Escribano [11], Encarna Guillén-Navarro [4,11], Laura Costa-Comellas [12], Francina Munell [12], Susana Boronat [13], Ricardo Rojas-García [4,14], Mónica Povedano [15], Ivon Cuscó [1,2,4] and Eduardo F. Tizzano [1,2,*]**

1 Medicine Genetics Group, Vall d'Hebron Research Institute (VHIR), Universitat Autònoma de Barcelona (UAB), 08035 Barcelona, Spain; lblasco@vhebron.net (L.B.-P.); mar.costa@vhebron.net (M.C.-R.); jleno@vhebron.net (J.L.-C.); mcodina@vhebron.net (M.C.-S.); desiree.martinez@vhebron.net (D.M.-C.); icusco@santpau.cat (I.C.)
2 Department of Clinical and Molecular Genetics, Hospital Universitari Vall d'Hebron, Universitat Autònoma de Barcelona (UAB), 08035 Barcelona, Spain
3 Genetics Department and Sant Pau Biomedical Research Institute, Hospital de la Santa Creu i Sant Pau, 08025 Barcelona, Spain; sbernal@santpau.cat (S.B.); lalias@santpau.cat (L.A.)
4 Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), 28029 Madrid, Spain; jose_millan@iislafe.es (J.M.M.); aller_ele@gva.es (E.A.); guillen.encarna@gmail.com (E.G.-N.); rrojas@santpau.cat (R.R.-G.)
5 Departamento de Neurología Pediátrica, Clínica Las Condes, 7591047 Santiago de Chile, Chile; castiglionic@gmail.com
6 Unit of Neuromuscular and Neurodegenerative Disease, Ospedale Pediatrico Bambino Gesu, IRCCS, 00165 Rome, Italy; ebertini@gmail.com (E.B.); lorena.travaglini@opbg.net (L.T.)
7 Unidad de Genética, Hospital La Fe and IIS La Fe, 46026 Valencia, Spain
8 Neuromuscular Diseases Unit, Neurology Department, Hospital Universitari Vall d'Hebron, 08035 Barcelona, Spain; jsotoca@vhebron.net (J.S.); rjuntas@vhebron.net (R.J.)
9 Department of Paediatrics, Copenhagen University Hospital, 2100 Copenhagen, Denmark; christina.hoei-hansen@regionh.dk
10 Department of Clinical Medicine, University of Copenhagen, 1165 Copenhagen, Denmark
11 Neurology Service and Medical Genetics Section, Hospital Clínico Universitario Virgen de la Arrixaca, IMIB-Arrixaca, Universidad de Murcia, 30120 Murcia, Spain; antonio_moreno9@yahoo.es
12 Pediatric Neurology Section, Vall d'Hebron Research Institute (VHIR), Hospital Universitari Vall d'Hebron, Universitat Autònoma de Barcelona (UAB), 08035 Barcelona, Spain; laura.costa@vhebron.net (L.C.-C.); fmunell@vhebron.net (F.M.)
13 Pediatrics Department, Hospital de la Santa Creu i Sant Pau, 08025 Barcelona, Spain; sboronat@santpau.cat
14 MND Clinic, Department of Neurology, Hospital de la Santa Creu i Sant Pau, Universitat Autònoma de Barcelona, 08025 Barcelona, Spain
15 Unidad Funcional de Enfermedad de Motoneurona, Servicio de Neurología, Hospital Universitario de Bellvitge, 08907 Barcelona, Spain; 30058mpp@gmail.com
* Correspondence: etizzano@vhebron.net; Tel.: +34-93274600 (ext. 6945)

check for updates

**Citation:** Blasco-Pérez, L.; Costa-Roger, M.; Leno-Colorado, J.; Bernal, S.; Alias, L.; Codina-Solà, M.; Martínez-Cruz, D.; Castiglioni, C.; Bertini, E.; Travaglini, L.; et al. Deep Molecular Characterization of Milder Spinal Muscular Atrophy Patients Carrying the c.859G>C Variant in *SMN2*. *Int. J. Mol. Sci.* **2022**, *23*, 8289. https://doi.org/10.3390/ijms23158289

Academic Editor: Giuseppe Novelli

Received: 22 June 2022
Accepted: 22 July 2022
Published: 27 July 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

**Abstract:** Spinal muscular atrophy (SMA) is a severe neuromuscular disorder caused by biallelic loss or pathogenic variants in the *SMN1* gene. Copy number and modifier intragenic variants in *SMN2*, an almost identical paralog gene of *SMN1*, are known to influence the amount of complete SMN proteins. Therefore, *SMN2* is considered the main phenotypic modifier of SMA, although genotype–phenotype correlation is not absolute. We present eleven unrelated SMA patients with milder phenotypes carrying the c.859G>C-positive modifier variant in *SMN2*. All were studied by a specific NGS method to allow a deep characterization of the entire *SMN* region. Analysis of two homozygous cases for the variant allowed us to identify a specific haplotype, *Smn2-859C.1*, in association with c.859G>C. Two other cases with the c.859G>C variant in their two *SMN2* copies showed a second haplotype, *Smn2-859C.2*, in cis with *Smn2-859C.1*, assembling a more complex allele. We also identified a previously unreported variant in intron 2a exclusively linked to the *Smn2-859C.1* haplotype (c.154-1141G>A), further suggesting that this region has been ancestrally conserved. The deep molecular characterization of *SMN2* in our cohort highlights the importance of testing c.859G>C,

as well as accurately assessing the *SMN2* region in SMA patients to gain insight into the complex genotype–phenotype correlations and improve prognostic outcomes.

**Keywords:** spinal muscular atrophy; *SMN2* copies; phenotype–genotype correlations; positive modifiers; next-generation sequencing

## 1. Introduction

Spinal muscular atrophy (SMA) is a neuromuscular disorder characterized by the degeneration and loss of alpha motor neurons in the spinal cord anterior horns, leading to progressive atrophy of proximal muscles, weakness, respiratory failure, and even death. It is the second most common recessive genetic disease of infancy and early childhood with an incidence around 1:11,000 live births and a carrier frequency of 1:51 worldwide [1,2].

SMA patients are mainly classified into five clinical groups on the basis of age of onset, achieved motor milestones, and clinical severity. Type 0 or congenital, the most severe, appears prenatally, and the patient's life expectancy is very short, usually a few weeks or months. Patients with type I, or Werdnig–Hoffmann disease (onset within the first six months of life), are never able to sit unsupported and generally do not survive beyond the age of two years. In the intermediate SMA type II (onset between 6 and 18 months of life), children acquire the ability to sit unsupported, but they never walk unaided and usually reach adolescence. Type III patients (Kugelberg–Welander disease) walk independently for a long time but eventually become wheelchair-bound. They can be further subdivided into type IIIa and IIIb depending on the age of disease onset (before or after three years of age). Finally, patients with SMA type IV present an adult onset and milder disease course [3–5]. It is important to bear in mind that current SMA therapies can modify the trajectory of SMA patients; therefore, this classification is mainly applied on clinical data prior to treatment [6,7].

At the molecular level, SMA is caused by the loss or mutation of both copies of the *survival of motor neuron 1 (SMN1)* gene, which encodes the survival motor neuron protein (SMN). In most cases, the disease is due to the homozygous absence of *SMN1* (95%), although pathogenic point variants have also been described [8–10].

Adjacent to *SMN1*, in a more centromeric position, lies *SMN2*, an almost identical paralog gene generated by a segmental duplication [11]. The fact that *SMN2* is present in humans and not in any other species suggests that the duplication of *SMN1* occurred recently in time. Consequently, the homology between both genes is extremely high, differing only in 16 positions called paralogous sequence variants (PSVs) [12,13]. This makes the region highly unstable, which leads to genomic instability predisposing to gene deletions, duplications, and conversions between both genes. Indeed, *SMN1* and *SMN2* genes can be present in multiple copies in the general population, both in cis and trans configuration [14].

Theoretically, the *SMN2* gene encodes the same protein as *SMN1,* but one of the PSVs, a silent transition in exon 7, alters the splicing pattern in most *SMN2* pre-mRNA transcripts. This causes the skipping of exon 7, resulting in a non-functional protein (SMN-Δ7) instead of the full-length protein [13]. As SMN-Δ7 is highly unstable and rapidly degraded, it is unable to compensate the absence or deficiency of *SMN1* in SMA patients [15]. It has been reported that each copy of *SMN2* can only produce about 10–15% of functional SMN proteins [16–18], being the number of *SMN2* copies the main modifier of SMA disease described to date.

Concretely, an inverse correlation between the number of *SMN2* copies and the severity of the phenotype has been widely reported, given that the higher the number of *SMN2* copies producing SMN functional protein, the milder the SMA phenotype [19–21]. Nevertheless, this correlation is not absolute, since discordant patients have been described in the literature, further classified as better-than-expected or worse-than-expected phenotypes

according to their *SMN2* copy number [19,21]. It is known that the presence of the c.859G>C and c.835-44A>G (-44G) variants, located in exon 7 and intron 6 of *SMN2*, respectively, explains some of the better-than-expected discordant phenotypes. These SNVs, considered positive modifiers of SMA disease, increase the inclusion of exon 7 and therefore generate greater amounts of functional SMN protein [21–25].

The full characterization of *SMN2*, including dosage and structure, will be more relevant in the current scenario where new therapies for SMA are being implemented. It is well known that *SMN2* dosage is the main modifier of SMA, but it seems that this could be just the tip of the iceberg of a much more complicated framework. Indeed, all differences between *SMN1* and *SMN2* can be revealed by specific NGS studies [12]. It is also possible that these findings may relate to phenotype variability or to *SMN2*-specific treatment response [20].

In this work, we performed an in-depth characterization of the *SMN* region in eleven SMA patients carrying the c.859G>C modifier variant in the *SMN2* gene ($SMN2^{859C}$) and presenting a milder phenotype. By defining the genetic background of $SMN2^{859C}$, we discovered the existence of a common haplotype alongside the *SMN2* gene in linkage disequilibrium with the variant and a second less common haplotype harboring two $SMN2^{859C}$ copies in cis.

## 2. Results

### 2.1. Clinical and Molecular Characterization of Patients

All SMA patients described in this study (ten males and one female) presented a biallelic absence of *SMN1* as the determinant of SMA and shared the presence of at least one copy of $SMN2^{859C}$. Seven of these individuals carried two *SMN2* copies, including five with the c.859G>C modifier variant in their two *SMN2* genes (patients 1 to 5) and two with the variant in only one *SMN2* (patients 6 and 7). The other four patients presented three *SMN2* copies, and the variant was only present in one of their *SMN2* alleles (patients 8 to 11). A summary of the clinical and molecular data of the patients is shown in Table 1. Our cohort comprised SMA patients of Spanish, Italian, Danish, and Chilean origins, and the majority were classified as SMA type IIIb (8/11) and the remaining patients as SMA type IIIa (2/11). The remaining case (Patient 6 in Table 1) was classified as type II based on his age of onset, which was prior to 18 months. Currently, at three years of age, he has not yet achieved independent ambulation.

**Table 1.** Clinical and molecular data of patients. Information regarding general characteristics of patients, SMA phenotype, and *SMN1/2* genotypes.

| Patient | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Origin** | Spanish | Spanish | Chilean | Italian | Spanish | Spanish | Danish | Spanish | Spanish | Spanish | Spanish |
| **Gender** | Male | Male | Male | Male | Male | Male | Female | Male | Male | Male | Male |
| **Age (years)** | 49 | 35 | 21 | 17 | 28 | 3 | 8 | 71 | 54 | 57 | 47 |
| **Consanguinity** | Yes | Yes | No | No | No | No | No | No | No | No | No |
| **Reported** | Bernal et al., 2010 (P2) [23] | Bernal et al., 2010 (P3) [23] | This work | This work | This work | This work | Blasco-Pérez et al., 2021 [12] | Bernal et al., 2010 (P4) [23] | This work | This work | Bernal et al., 2010 (P5) [23] |
| **SMA type** | IIIb | IIIb | IIIb | IIIb | IIIb | II | IIIa | IIIb | IIIb | IIIa | IIIb |
| **Age of onset** | 10 years | 4 years | 12 years | 13 years | 15 years | 12 months | 18 months | 14 years | 9-10 years | 24 months | 13 years |
| **Walked unaided** | Yes | Yes | Yes | Yes | Yes | Not yet accomplished | Yes | Yes | Yes | Yes | Yes |
| **Wheelchair bound age** | 41 years | 23 years | No | No | No | Not applicable | No | 49 years | No | 22 years | 37 years |
| ***SMN1* copies** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| ***SMN2* copies** | 2 | 2 | 2 | 2 | 2 * | 2 | 2 | 3 | 3 | 3 | 3 |
| **Presence of c.859G>C** | 2/2 | 2/2 | 2/2 | 2/2 | 2/2 | 1/2 | 1/2 | 1/3 | 1/3 | 1/3 | 1/3 |

* In addition, presents a partial *SMN* gene comprising only exons 1 to 6 (*SMN1/2*Δ7-8) (see Figure 1).

### 2.2. Haplotype Characterization by Deep Sequencing of SMN2 Genes

NGS data confirmed the biallelic absence of the entire *SMN1* gene in all patients, since specific nucleotides of *SMN2* were found in homozygous state (AB ratio of 100%) in all PSV positions. Similarly, the NGS results corroborated the *SMN2* copy number previously assigned by MLPA via the AB ratio analysis of all the different variants detected in the SMN region of each patient. In patients with two *SMN2* copies (except for patient 5), all variants were detected with an approximate allele frequency of 50% or 100%, whereas in patients with three *SMN2* copies, variants were found at a frequency of around 33%, 66%, or 100% (data not shown, available upon request). Patient 5 was a special case where variants were observed at a frequency of around 33–66–100% in the 5′ region and around 50–100% frequency in the 3′ region. This phenomenon was due to the presence of two complete *SMN2* genes and a partial *SMN* gene comprising exons 1 to 6 (*SMN1/2*Δ7-8) (see Table 2). In addition, the AB analysis of all patients confirmed the copy number of the c.859G>C modifier in each case.

Overall, our 11 patients represented 16 alleles with the c.859G>C variant, including five cases with two $SMN2^{859C}$ and the remainder with just one allele with the variant (Table 2).

#### 2.2.1. Establishment of Two Haplotypes Associated with the c.859G>C Modifier Variant

We initially performed an in-depth analysis of the complete SMN2 region in patients 1 and 2, who carried two $SMN2^{859C}$ genes and had consanguineous parents. The studies revealed that both patients were completely homozygous for the entire studied region and identical between them. Thus, we were able to determine the specific SMN2 sequence associated with the c.859G>C modifier in their alleles, establishing a haplotype called Smn2-859C.1 (Table 2). Similarly, sequencing results in patient 3 revealed an almost identical sequence to Smn2-859C.1 in his two SMN2 genes, with the exception of one rare variant (69356349-A-G) with an allele frequency of ~50%. In contrast, patients 4 and 5, who also presented two $SMN2^{859C}$ copies, showed several variants in only one of their $SMN2^{859C}$ along the studied region. Nonetheless, it was possible to infer that one of their SMN2 genes matched the sequence of the Smn2-859C.1 haplotype. Interestingly, in both patients, it was possible to assume a second haplotype associated with the c.859G>C variant that we defined as Smn2-859C.2 (Table 2). Applying this preliminary information, the Smn2-859C.1 haplotype was also inferred in one of the SMN2 copies of the remaining patients (patients 6 to 11), with few discrepant positions in patients 7, 10, and 11 (see Table 2).

To explore deeper into the structure of the *SMN2* genes, co-segregation studies from patients with two $SMN2^{859C}$ were carried out through MLPA together with NGS or allele-specific PCR. These investigations showed that patient 2 carried his two *Smn2-859C.1* haplotypes in trans, inheriting one from each progenitor (Figure 1B). Patient 1′s co-segregation was incomplete, as a sample from his father was not available, but this family was consanguineous, and the mother only presented one *Smn2-859C.1* haplotype. Therefore, we could assume that his father also presented one *Smn2-859C.1* haplotype, and he should harbor both *Smn2-859C.1* haplotypes in trans (Figure 1A). In contrast, the co-segregation study in patient 3 revealed that both *Smn2-859C.1* haplotypes were in cis, forming a complex allele inherited from the mother (Figure 1C). Co-segregation in patients 4 and 5 indicated that the two $SMN2^{859C}$ genes (*Smn2-859C.1* and *Smn2-859C.2* haplotypes) were located in cis. Specifically, patient 4 inherited this complex allele from his father and a null allele (without *SMN1* and *SMN2*) from his mother (Figure 1D), while patient 5 inherited the complex allele from his mother and the other allele with a partial non-functional *SMN1/2*Δ7-8 gene from his father (Figure 1E).

**Table 2.** Haplotype characterization of SMN2 genes. Detail of the 30 positions comprising the *Smn2-859C.1* (green) and *Smn2-859C.2* (blue) haplotypes in our patients. Punctual discrepancies are represented in red. The novel variant c.154-1141G>A (69360651-G-A, hg19/GRCh37), exclusively associated with the *Smn2-859C.1* haplotype, is indicated in green in the first column. The c.859G>C modifier variant is marked in red. The remaining alleles of each patient not carrying the c.859G>C are represented in grey.

| | | **Patient** | | **1** | | **2** | | **3** | | **4** | | **5** | | | **6** | | **7** | | **8** | | | **9** | | | **10** | | | **11** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Location | hg19 | *Smn2-859C.1* | *Smn2-859C.2* | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 2 | - | 1 | - | 1 | - | 1 | - | - | 1 | - | - | 1 | - | - | 1 | - | - |
| intron 1 | 69342881-T-C | C | C | C | C | C | C | C | C | C | C | C | C | C | C | C | C | C | C | C | C | C | C/T | C/T | C | C | C | C | C/T | C/T |
| | 69343230-C-T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | T |
| | 69343570-G-T | T | T | T | T | T | T | T | T | T | T | T | T | T | T | G | T | T | T | T | T | T | T | T | T | T | T | T | T/G | T/G |
| | 69347007-C-T | T | | T | T | T | T | T | T | T | C | T | C | C | T | C | C | C | T | C | C | T | C | C | T | C | C | T | C | C |
| | 69349821-T-C | C | | C | C | C | C | C | C | C | T | C | T | T | C | T | C | T | C | T | T | C | T | T | C | T | T | C | T | T |
| | 69350284-A– | - | - | - | - | - | - | - | - | - | - | - | - | - | - | A | - | - | - | - | - | - | - | - | - | - | - | A | A | A |
| | 69351711-A-G | | G | A | A | A | A | A | A | A | G | A | G | A | A | A | A | A | A | A | A | A | A | A | A | A | A | A | A | A |
| | 69353192-G-A | | A | G | G | G | G | G | G | G | A | G | A | A | G | G | G | G | G | G | G | G | G | G | G | G | G | G | G | G |
| | 69354973-A-G | G | | G | G | G | G | G | G | G | A | G | A | G | G | A | G | G | G | G | G | G | G | G | G | G | G | G | G/A | G/A |
| | 69356085-A-G | G | G | G | G | G | G | G | G | G | G | G | G | G | G | G | G | G | G | G | G | G | G | G | G | G | G | G | G | G |
| | 69356349-A-G | | G | A | A | A | A | A | G | A | G | A | G | A | A | A | A | A | A | A | A | A | A | A | A | A | A | A | A | A |
| | 69357245-C-G | | G | C | C | C | C | C | C | C | G | C | G | G | C | C | C | G | C | G | G | C | G | G | C | G | G | C | C/G | C/G |
| | 69357509-G-A | A | | A | A | A | A | A | A | A | G | A | G | G | A | G | A | G | A | G | G | A | G | G | A | G | G | A | G | G |
| | 69358318-A-G | | G | A | A | A | A | A | A | A | G | A | G | A | A | A | A | A | A | A | A | A | A | A | A | A | A | A | A | A |
| | 69358605-A-G | G | G | G | G | G | G | G | G | G | G | G | G | G | G | A | G | G | G | G | G | G | G | G | G | G | G | G | G/A | G/A |
| | 69359034-C-T | T | | T | T | T | T | T | T | T | C | T | C | C | T | C | T | C | T | C | C | T | C | C | T | C | C | T | C | C |
| intron 2a | 69360020-G-T | T | | T | T | T | T | T | T | T | G | T | G | G | T | G | T | G | T | G | G | T | G | G | T | G | G | G | G | G |
| | 69360651-G-A | A | | A | A | A | A | A | A | A | G | A | G | G | A | G | A | G | A | G | G | A | G | G | A | G | G | A | G | G |
| intron 2b | 69362410-T-C | C | C | C | C | C | C | C | C | C | C | C | C | T | C | T | C | T | C | T | T | C | T | T | C | T | T | C | T | T |
| exon 3 | 69362949-A-G | G | G | G | G | G | G | G | G | G | G | G | G | G | G | A | G | G | G | G | G | G | G | G | G | G | G | G | G/A | G/A |
| intron 4 | 69363717-C-T | T | T | T | T | T | T | T | T | T | T | T | T | | T | C | T | C | T | C | C | T | C | C | C | C | C | T | C | C |
| | 69363866-A-G | | G | A | A | A | A | A | A | A | G | A | G | | A | A | A | A | A | A | A | A | A | A | A | A | A | A | A | A |
| | 69364605-A-G | G | G | G | G | G | G | G | G | G | G | G | G | | G | A | G | G | G | G | G | G | G | G | G | G | G | G | G/A | G/A |
| intron 5 | 69365216-G-C | C | | C | C | C | C | C | C | C | G | C | G | | C | G | C | G | C | G | G | C | G | G | C | G | G | C | G | G |
| intron 6 | 69368084-A-G | G | G | G | G | G | G | G | G | G | G | G | G | | G | A | G | G | G | G/A | G/A | G | G | G | G | G | G | G | G/A | G/A |
| | 69368329-G-A | A | A | A | A | A | A | A | A | A | A | A | A | | A | G | A | A | A | G/A | G/A | A | A | A | A | A | A | A | G/A | G/A |
| | 69371981-C-A | A | A | A | A | A | A | A | A | A | A | A | A | | A | C | A | A | A | A | A | A | A | A | A | A | A | A | A/C | A/C |
| exon 7 | 69372372-G-C | C | C | C | C | C | C | C | C | C | C | C | C | | C | G | C | G | C | G | G | C | G | G | C | G | G | C | G | G |
| downstream | 69373667-A-G | G | G | G | G | G | G | G | G | G | G | G | G | | G | A | G | G | G | G | G | G | G | G | G | G | G | G | G | G |
| | 69373682-C-G | G | G | G | G | G | G | G | G | G | G | G | G | | G | C | G | C | G | C | C | G | C | C | G | C | C | G | G/C | G/C |

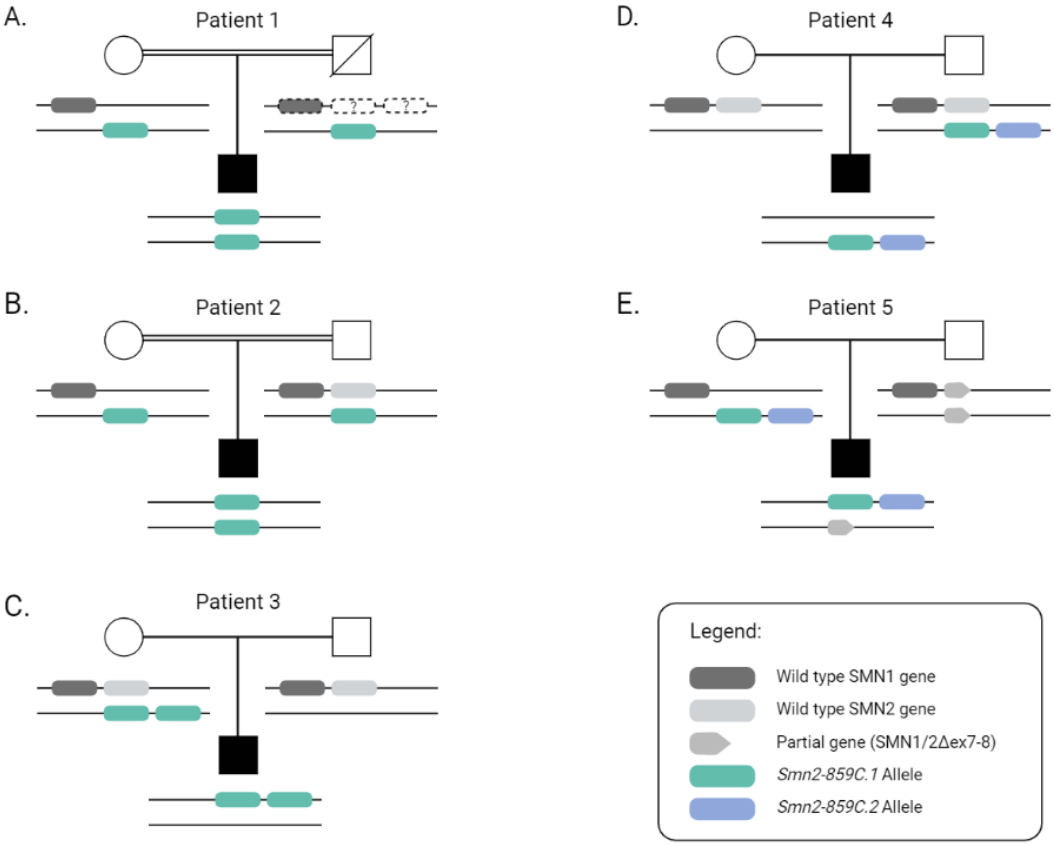



**Figure 1.** Pedigree representation of cases with two *SMN2*$^{859C}$ copies. According to segregation studies, a cis or trans configuration was defined in each patient. (**A**) In patient 1, the father's sample was not available, and *SMN2* configuration was inferred based on the results from the mother and patient. Given that the father did not present symptoms, we can assume that he carries at least one *SMN1* gene. In addition, being a consanguineous family, we assumed that *Smn2-859.C1* was transmitted by both parents (untested inferred alleles are represented by a dashed line). (**B**) In patient 2, the *Smn2-859.C1* haplotype was inherited from both parents, in agreement with the consanguinity in the family. (**C**) Patient 3 had two copies of *SMN2* with *Smn2-859.C1* in cis, inherited from his mother. (**D**) Patient 4 also had two copies of *SMN2* in cis, one with *Smn2-859.C1* and the other with *Smn2-859.C2* haplotype, forming a complex allele inherited from the father. (**E**) Patient 5 inherited the complex allele from his mother and the other allele with a partial non-functional *SMN1*/2Δ7-8 gene from his father.

All together, these results indicated that the *Smn2-859C.1* haplotype was consistent in our cohort, since all patients presented it in association with the c.859G>C variant, either as a single allele or as part of a more complex allele formed by the *Smn2-859C.1* and *Smn2-859C.2* haplotypes in cis. Based on our 11 SMA patients, we have not observed any clinical difference between *Smn2-859C.1* and *Smn2-859C.2* haplotypes, although we only found two cases carrying the *Smn2-859C.2* haplotype.

#### 2.2.2. Difference between Haplotypes and Detection of a Novel Variant Exclusively Associated with the *Smn2-859C.1* Haplotype

Analyzing the sequence of both haplotypes, *Smn2-859C.1* consists of 24 variants while *Smn2-859C.2* comprises 22 variants, sharing 16 of these positions and differing in the other 14. In particular, the sequence near the c.859G>C variant is shared between *Smn2-859C.1* and *Smn2-859C.2* haplotypes and spans at least 8848 bp (chr5:69365217-69374064). These haplotypes were not found in a total of 338 SMA patients without the c.859G>C variant, although some of the variants contained in the haplotypes are present in this larger cohort.

Interestingly, we noticed the presence of a novel variant, c.154-1141G>A (69360651-G-A, hg19/GRCh37), located in intron 2a (Table 2). This variant was detected in all patients with the *Smn2-859C.1* haplotype but absent in the *Smn2-859C.2* haplotype. Moreover, this variant was not detected in the 338 SMA patients without the c.859G>C variant. The c.154-1141G>A change has not been reported in the general population according to gnomAD, ISB Kaviar3, and Bravo (as of 18 July 2022) [26–28]. In silico analysis of this deep intronic variant using the software SpliceAI [29], Alamut Visual Software version 2.11 (SOPHiA GENETICS), and ESRseq [30] did not predict an effect on the splicing process.

## 3. Discussion

Here, we present 11 patients with a clinical and molecular diagnosis of SMA caused by the biallelic absence of *SMN1* and with a milder phenotype explained by at least one *SMN2*$^{859C}$ gene, given that 10 out of 11 patients were walkers. We identified a specific sequence, named *Smn2-859C.1,* present in all patients from our cohort in linkage disequilibrium with the c.859G>C variant. In addition, two cases showed a more complex allele, assembled by *Smn2-859C.1* and *Smn2-859C.2* in cis.

In order to study the genetic origin of the c.859G>C variant, we expanded our cohort of Spanish cases with patients from Denmark, Italy, and Chile. We applied NGS methodologies exclusively focused on the SMN region to determine the exact sequence of *SMN2* associated with the c.859G>C variant in each patient [12]. By studying the patients with two *SMN2*$^{859C}$, we were able to determine two haplotypes associated with the variant, *Smn2-859C.1* and *Smn2-859C.2*. The *Smn2-859C.1* haplotype, with minor modifications, was present in all 11 patients (14/16 *SMN2*$^{859C}$ alleles), either in cis or trans configuration, while the *Smn2-859C.2* haplotype was only found in two patients (2/16 *SMN2*$^{859C}$ alleles), always in cis configuration with the *Smn2-859C.1* haplotype (Figure 2). Notably, no patient was found to harbor the c.859G>C variant in association with any other haplotype, regardless of their ethnic lineage, which points towards a common ancestral origin in all cases.

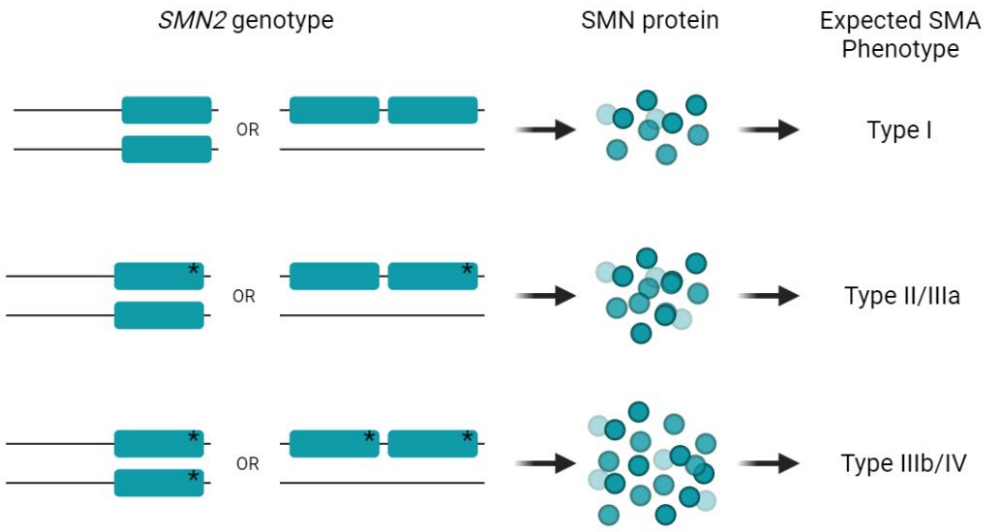



**Figure 2.** Expected SMA phenotype in cases with two *SMN2* copies according to the presence of c.859G>C. An additive effect on SMA phenotype is observed depending on whether the c.859G>C variant is found in one or both *SMN2* copies. *SMN2* gene is represented as a rectangle, and the presence of the c.859G>C variant in exon 7 is indicated by an asterisk. Not all *SMN2* genotypes represented in this figure were detected in this study (see Figure 1 for more details).

The c.859G>C variant has been previously reported to increase the inclusion of *SMN2* exon 7 by 20%, which leads to the generation of higher amounts of functional proteins than the wild-type *SMN2* gene [22,25]. Patients carrying this variant developed milder SMA phenotypes compared with those with the same *SMN2* copy number but without the

variant [23]. In our case, the deep characterization of the entire SMN region supports that *SMN2*$^{859C}$ is, at first sight, primarily responsible for the milder phenotype in our patients.

To date, together with our six newly described patients, a total of 44 patients carrying c.859G>C have been reported worldwide, including a patient recently detected by newborn screening [21–23,25,31–34]. In general population databases, the c.859G>C variant is reported at a frequency of approximately 0.3% with 132 homozygotes detected [26]. However, it is possible that the data are not accurate given the high homology between *SMN1* and *SMN2* and their copy number variability, which poses a challenge in the analysis and proper annotation of the SMN region with non-specific NGS techniques, such as exome or genome. Nevertheless, it is possible to estimate the frequency of this variant in the SMA population based on previous studies. According to the data of Calucho et al. (2018) [19], the allelic frequency of the c.859G>C variant is 1.04% (13/1250 alleles) in a series of 625 Spanish SMA patients. In fact, in this cohort, approximately 25% of better-than-expected cases with two *SMN2* copies carried the variant [19]. Although c.859G>C appears to be relatively uncommon, at present it is not routinely tested in SMA patients, deserving more studies to clearly establish its incidence.

Concerning clinical classification, patients with two *SMN2* genes usually debut in the first six months of life and are classified as SMA type I [19]. In our series, patients with two *SMN2* copies and the positive modifier presented at least type II or type III disease (Table 1). Furthermore, an additive effect was observed since patients with the c.859G>C change in both *SMN2* genes had a better phenotype than patients carrying the variant only in one *SMN2*, confirming previous observations [23] (Figure 2). For instance, patient 3 (with two *SMN2*$^{859C}$ copies) developed the first SMA symptoms at 12 years of age, being classified as type IIIb, whereas patient 7 (with the variant in one of his *SMN2* genes) had manifestations at 18 months of life with a clinical diagnosis of type IIIa. Regarding cases with three *SMN2* copies, all patients presented the c.859G>C variant in only one of their alleles, developing a type III phenotype. Interestingly, we did not find any patient with three *SMN2* copies and the variant in more than one allele and, in fact, no patient with this genotype has been described in the literature either. This could be due to the fact that patients with three *SMN2* copies showing SMA type II or III are not currently tested for the variant. Another reason could be that cases with this genotype perhaps do not manifest clear disease symptoms due to the higher production of SMN protein and therefore may never be diagnosed. Similarly, it has been previously speculated that some individuals with zero *SMN1* and four or five *SMN2* copies may present minimal symptoms or be asymptomatic throughout their lives, remaining undetected [35]. This corroborates the importance of implementing detection of the c.859G>C-positive modifier as part of the genetic diagnosis routine in SMA.

At this level of analysis and based on the clinical information available for each patient, we did not observe categorical phenotypic differences between the *Smn2-859C.1* or *Smn2-859C.2* haplotypes, nor the cis or trans configuration of the *Smn2-859C.1* haplotype, since all cases with two *SMN2*$^{859C}$ copies presented a milder phenotype (IIIb). Interestingly, patient 2, with the exact same sequence and configuration as patient 1, also developed type IIIb SMA, but his onset was noted earlier in comparison with patient 1 and the remaining cases with two *SMN2*$^{859C}$. At present, we are unable to explain this minor disparity considering all the studies performed in *SMN2*. Thus, this fact suggests disease onset could also be conditioned by as yet unknown factors, other than *SMN2* structure.

As mentioned above, the *Smn2-859C.2* haplotype was detected in cis configuration with respect to *Smn2-859C.1*, assembling a complex allele containing two different *SMN2*$^{859C}$ genes (Figure 1). These two haplotypes differ in several positions, but an identical block of at least 8848 bp around c.859G>C is present in both (Table 2 and Figure 3B). This observation, together with the fact that we also detected an allele formed by two *Smn2-859C.1* haplotypes in cis, points towards a possible origin of the complex allele through homologous recombination, implicating a double cross-over event [36]. In this event, two alleles would be involved (Figure 3A): allele A, consisting of two *SMN2* genes with the *Smn2-859C.1* haplotype, and allele B, formed by at least one *SMN2* with an unknown

haplotype containing part of *Smn2-859C.2* but without the c.859G>C variant. In the double homologous recombination process, allele A would maintain both c.859G>C variants as well as gain the part of the *Smn2-859C.2* haplotype from allele B, generating the complex allele that we detected in our patients (Figure 3B).

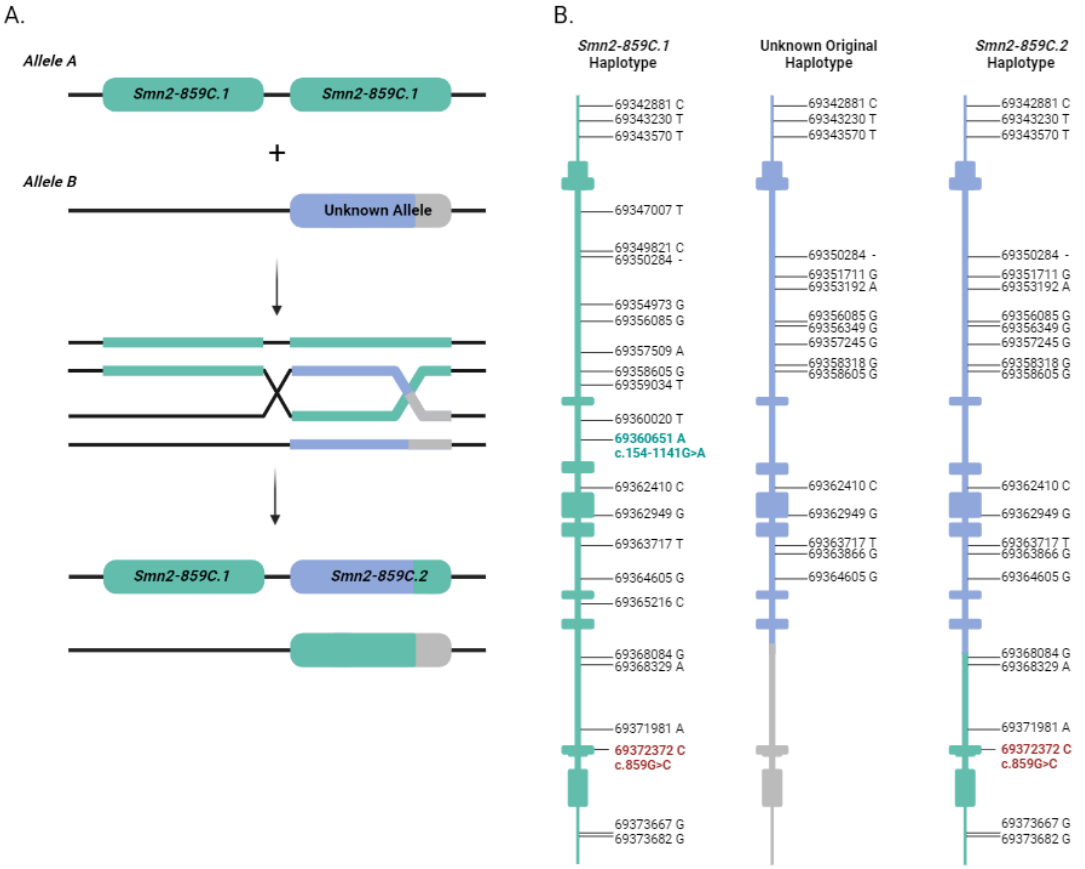



**Figure 3.** Proposed mechanism of origin and structure of *Smn2-859C.1* and *Smn2-859C.2* haplotypes. (**A**) Hypothetic origin of the *Smn2-859C.2* haplotype through homologous recombination involving a double cross-over event between allele A (with two *Smn2-859C.1* haplotypes represented in green) and allele B (containing the 5′ region of the *Smn2-859C.2* haplotype, in blue, and the 3′ end with an unknown sequence, in grey). (**B**) *SMN2* structure details (representing exons and introns from top to bottom) and location of variants of the *Smn2-859C.1* and *Smn2-859C.2* haplotypes, as well as the unknown original haplotype that presumably originated the *Smn2-859C.2* haplotype. The c.154-1141G>A (69360651-G-A, hg19/GRCh37) variant is indicated in green, whereas the c.859G>C modifier is shown in red (further explanation in the text and Table 2).

Finally, it should be noted that the *Smn2-859C.1* haplotype contains the novel variant c.154-1141G>A, located in intron 2a. According to our results, this variant is in linkage disequilibrium with the c.859G>C modifier given that, in our larger cohort of 349 SMA patients, it was only detected in those carrying the c.859G>C variant, and it was not found in population databases. This observation suggests that the sequence between this variant and the c.859G>C modifier has been ancestrally conserved. In silico splicing tools did not predict any specific effect of this deep intronic variant. However, we could not rule out some influence of this change, given the limitations of splicing predictors; thus, it deserves further investigation.

Our NGS approach to characterize these patients revealed new information that could be relevant for the different functions and/or alterations of *SMN2*. It is important to consider whether the function and expression of *SMN2* is not only modified depending on the cis or trans configuration of $SMN2^{859C}$ but also on the presence of the *Smn2-859C.1* or

*Smn2-859C.2* haplotype. Long regulators, cis- or trans-acting elements, may distinctively influence its function and/or expression according to the topography of the region.

## 4. Materials and Methods

### 4.1. Study Participants

We studied eleven unrelated SMA patients from different international centers with the presence of at least one *SMN2*$^{859C}$ gene. Patients were classified into SMA type according to age of onset, clinical severity, and achieved motor milestones, prior to receiving any modifying therapies. Criteria for correlating phenotype with *SMN2* dosage were type I (non-sitters) with two *SMN2* copies, type II (sitters) with three *SMN2* copies, and type III (walkers) patients with three–four *SMN2* copies [19]. Based on this model, our patients with two *SMN2* copies were considered discordant, as none presented a type I SMA phenotype (Table 1).

All patients were selected from a larger cohort of 349 SMA patients, undergoing an NGS study of the *SMN* region [12], based on the presence of the c.859G>C variant. Four of the patients were previously described as carriers of this variant (patients 1, 8, and 11 [23] and patient 7 [12]).

DNA samples were extracted from peripheral blood using standard methods. Ethics approval was granted by the Clinical Research Ethics Committee of Hospital Vall d'Hebron (Comité de Ética de Investigación con Medicamentos del Hospital Universitari Vall d'Hebron (PR(AG)229/2018)). Written informed consent was obtained from all participants or their parents/legal caregivers.

### 4.2. SMN2 Genotyping and Haplotype Characterization

All patients were genetically confirmed as SMA cases via previously described methods that also included testing *SMN2* modifier variants [10,23,37]. A detailed molecular characterization of *SMN2* was carried out in all patients by a specific NGS sequencing method [12].

In addition, to detect the presence of the c.859G>C variant in some progenitor samples, two specific PCRs were designed to amplify exons 7 and 8 of genes *SMN1* and *SMN2*. The allele-specific PCR technique [38] was used to amplify both genes separately to ascertain in which gene the variant was present. Standard Sanger sequencing was performed with the PCR products, allowing us to detect the c.859G>C variant. These primers are also designed to study the c.835-44A>G variant. Primer sequences and PCR conditions are provided in Table S1.

## 5. Conclusions

This series of patients with milder phenotypes demonstrates the relevance of testing the c.859G>C variant in all SMA patients, with special consideration in cases with two or three *SMN2* copies in the context of neonatal screening. Indeed, the presence of this rare variant in an asymptomatic neonate may help to predict a better phenotype by natural history per se, regardless of the therapeutic option chosen. This is crucial in order to evaluate the effects of the approved therapies to unmask long-term benefits in treated patients. Given that not all discordant cases can be explained by this positive variant, it is necessary to further analyze the *SMN2* region by NGS to detect other reported candidate variants [24] and the presence of hybrid *SMN1-SMN2* structures [20], as well as to unravel novel phenotypic modifier variants. In the current therapeutic context, genetic studies in patients confirmed with biallelic *SMN1* absence or pathogenic variants should consider not only testing for *SMN2* copies but also investigating *SMN2* variants and structures as part of the integral characterization of patients receiving expensive and sometimes lifelong therapies.

**Supplementary Materials:** The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/ijms23158289/s1.

**Author Contributions:** Conceptualization, L.B.-P., M.C.-R., I.C. and E.F.T.; methodology, L.B.-P., M.C.-R., S.B. (Sara Bernal), L.A. and D.M.-C.; formal analysis, J.L.-C., L.B.-P., M.C.-R., I.C. and E.F.T.; resources, C.C., E.B., L.T., J.M.M., E.A., J.S., R.J., C.E.H.-H., A.M.-E., E.G.-N., L.C.-C., F.M., S.B. (Susana Boronat), R.R.-G. and M.P.; data curation, L.B.-P., M.C.-R., M.C.-S., I.C. and E.F.T.; writing—original draft preparation, L.B.-P., M.C.-R., I.C. and E.F.T.; writing—review and editing, L.B.-P., M.C.-R., I.C., S.B. (Sara Bernal), M.C.-S. and E.F.T.; supervision, E.F.T. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by grants from Biogen (ESP-SMG-17-11256), Roche, GaliciAME and Spanish Instituto de Salud Carlos III, Fondo de Investigaciones Sanitarias and co-funded with ERDF funds (grant no. FIS PI18/000687). A grant from Horizon 2020 IMI2 Screen4Care is acknowledged by E.B., and L.T., E.F.T., R.J., J.S., L.C.-C., F.M., E.B., and L.T. are members of the ERN NMD Network for Rare Diseases. E.F.T. is a member of the ERN ITHACA Network for Rare Diseases.

**Institutional Review Board Statement:** The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of Hospital Universitari Vall d'Hebron, (PR(AG)229/2018).

**Informed Consent Statement:** Written informed consent was obtained from all participants or their parents/legal caregivers.

**Data Availability Statement:** All data and scripts used to generate the analyses of this paper are available upon request, unless the type of request compromises ethical standards or legal requirements.

**Acknowledgments:** We thank the patients and their families for participating in this study.

**Conflicts of Interest:** The authors declare no conflict of interest in relation to this work.

## References

1. Cusin, V.; Clermont, O.; Gérard, B.; Chantereau, D.; Elion, J. Prevalence of SMN1 deletion and duplication in carrier and normal populations: Implication for genetic counselling. *J. Med. Genet.* **2003**. [CrossRef] [PubMed]
2. Sugarman, E.A.; Nagan, N.; Zhu, H.; Akmaev, V.R.; Zhou, Z.; Rohlfs, E.M.; Flynn, K.; Hendrickson, B.C.; Scholl, T.; Sirko-Osadsa, D.A.; et al. Pan-ethnic carrier screening and prenatal diagnosis for spinal muscular atrophy: Clinical laboratory analysis of >72 400 specimens. *Eur. J. Hum. Genet.* **2012**, *20*, 27–32. [CrossRef] [PubMed]
3. Zerres, K.; Rudnik-Schöneborn, S. Natural history in proximal spinal muscular atrophy. Clinical analysis of 445 patients and suggestions for a modification of existing classifications. *Arch Neurol.* **1995**, *52*, 518–523. [CrossRef] [PubMed]
4. Zerres, K.; Wirth, B.; Rudnik-Schöneborn, S. Spinal muscular atrophy clinical and genetic correlations. *Neuromuscul. Disord.* **1997**, *7*, 202–207. [CrossRef]
5. Wang, C.H.; Finkel, R.S.; Bertini, E.S.; Schroth, M.; Simonds, A.; Wong, B.; Aloysius, A.; Morrison, L.; Main, M.; Crawford, T.O.; et al. Consensus statement for standard of care in spinal muscular atrophy. *J. Child Neurol.* **2007**, *22*, 1027–1049. [CrossRef] [PubMed]
6. Tizzano, E.F.; Finkel, R.S. Spinal muscular atrophy: A changing phenotype beyond the clinical trials. *Neuromuscul. Disord.* **2017**, *27*, 883–889. [CrossRef] [PubMed]
7. Schorling, D.C.; Pechmann, A.; Kirschner, J. Advances in Treatment of Spinal Muscular Atrophy – New Phenotypes, New Challenges, New Implications for Care. *J. Neuromuscul. Dis.* **2020**, *7*, 1–13. [CrossRef] [PubMed]
8. Lefebvre, S.; Bürglen, L.; Reboullet, S.; Clermont, O.; Burlet, P.; Viollet, L.; Benichou, B.; Cruaud, C.; Millasseau, P.; Zeviani, M.; et al. Identification and characterization of a spinal muscular atrophy-determining gene. *Cell* **1995**, *80*, 155–165. [CrossRef]
9. Burghes, A.H.M.; McGovern, V.L. Genetics of spinal muscular atrophy. In *Molecular and Cellular Therapies for Motor Neuron Diseases*; Elsevier: Amsterdam, The Netherlands, 2017; pp. 121–139.
10. Alías, L.; Bernal, S.; Fuentes-Prior, P.; Barceló, M.J.; Also, E.; Martínez-Hernández, R.; Rodríguez-Alvarez, F.J.; Martín, Y.; Aller, E.; Grau, E.; et al. Mutation update of spinal muscular atrophy in Spain: Molecular characterization of 745 unrelated patients and identification of four novel mutations in the SMN1 gene. *Hum. Genet.* **2009**, *125*, 29–39. [CrossRef]
11. Rochette, C.F.; Gilbert, N.; Simard, L.R. SMN gene duplications and the emergence of the SMN2 gene ocurred in distinct hominids: SMN2 is unique to Homo sapiens. *Hum. Genet.* **2001**, *108*, 255–266. [CrossRef]
12. Blasco-Pérez, L.; Paramonov, I.; Leno, J.; Bernal, S.; Alias, L.; Fuentes-Prior, P.; Cuscó, I.; Tizzano, E.F. Beyond copy number: A new, rapid, and versatile method for sequencing the entire SMN2 gene in SMA patients. *Hum. Mutat.* **2021**, *42*, 787–795. [CrossRef] [PubMed]
13. Monani, U.R.; Lorson, C.L.; Parsons, D.W.; Prior, T.W.; Androphy, E.J.; Burghes, A.H.; McPherson, J.D. A single nucleotide difference that alters splicing patterns distinguishes the SMA gene SMN1 from the copy gene SMN2. *Hum. Mol. Genet.* **1999**, *8*, 1177–1183. [CrossRef] [PubMed]

14. Prior, T.W.; Nagan, N.; Sugarman, E.A.; Batish, S.D.; Braastad, C. Technical standards and guidelines for spinal muscular atrophy testing. *Genet. Med.* **2011**, *13*, 686–694. [CrossRef] [PubMed]
15. Vitte, J.; Fassier, C.; Tiziano, F.D.; Dalard, C.; Soave, S.; Roblot, N.; Brahe, C.; Saugier-Veber, P.; Bonnefont, J.P.; Melki, J. Refined characterization of the expression and stability of the SMN gene products. *Am. J. Pathol.* **2007**, *171*, 1269–1280. [CrossRef] [PubMed]
16. Soler-Botija, C.; Cuscó, I.; Caselles, L.; López, E.; Baiget, M.; Tizzano, E.F. Implication of fetal SMN2 expression in type I SMA pathogenesis: Protection or pathological gain of function? *J. Neuropathol. Exp. Neurol.* **2005**, *64*, 215–223. [CrossRef] [PubMed]
17. Boza-Morán, M.G.; Martínez-Hernández, R.; Bernal, S.; Wanisch, K.; Also-Rallo, E.; Le Heron, A.; Alías, L.; Denis, C.; Girard, M.; Yee, J.-K.; et al. Decay in survival motor neuron and plastin 3 levels during differentiation of iPSC-derived human motor neurons. *Sci. Rep.* **2015**, *5*, 11696. [CrossRef] [PubMed]
18. Wirth, B.; Garbes, L.; Riessland, M. How genetic modifiers influence the phenotype of spinal muscular atrophy and suggest future therapeutic approaches. *Curr. Opin. Genet. Dev.* **2013**, *23*, 330–338. [CrossRef]
19. Calucho, M.; Bernal, S.; Alías, L.; March, F.; Venceslá, A.; Rodríguez-Álvarez, F.J.; Aller, E.; Fernández, R.M.; Borrego, S.; Millán, J.M.; et al. Correlation between SMA type and SMN2 copy number revisited: An analysis of 625 unrelated Spanish patients and a compilation of 2834 reported cases. *Neuromuscul. Disord.* **2018**, *28*, 208–215. [CrossRef]
20. Costa-Roger, M.; Blasco-Pérez, L.; Cuscó, I.; Tizzano, E.F. The Importance of Digging into the Genetics of SMN Genes in the Therapeutic Scenario of Spinal Muscular Atrophy. *Int. J. Mol. Sci.* **2021**, *22*, 9029. [CrossRef]
21. Ruhno, C.; McGovern, V.L.; Avenarius, M.R.; Snyder, P.J.; Prior, T.W.; Nery, F.C.; Muhtaseb, A.; Roggenbuck, J.S.; Kissel, J.T.; Sansone, V.A.; et al. Complete sequencing of the SMN2 gene in SMA patients detects SMN gene deletion junctions and variants in SMN2 that modify the SMA phenotype. *Hum. Genet.* **2019**, *138*, 241–256. [CrossRef]
22. Prior, T.W.; Krainer, A.R.; Hua, Y.; Swoboda, K.J.; Snyder, P.C.; Bridgeman, S.J.; Burghes, A.H.M.; Kissel, J.T. A positive modifier of spinal muscular atrophy in the SMN2 gene. *Am. J. Hum. Genet.* **2009**, *85*, 408–413. [CrossRef] [PubMed]
23. Bernal, S.; Alías, L.; Barceló, M.J.; Also-Rallo, E.; Martínez-Hernández, R.; Gámez, J.; Guillén-Navarro, E.; Rosell, J.; Hernando, I.; Rodríguez-Alvarez, F.J.; et al. The c.859G>C variant in the SMN2 gene is associated with types II and III SMA and originates from a common ancestor. *J. Med. Genet.* **2010**, *47*, 640–642. [CrossRef] [PubMed]
24. Wu, X.; Wang, S.H.; Sun, J.; Krainer, A.R.; Hua, Y.; Prior, T.W. A-44G transition in SMN2 intron 6 protects patients with spinal muscular atrophy. *Hum. Mol. Genet.* **2017**, *26*, 2768–2780. [CrossRef] [PubMed]
25. Vezain, M.; Saugier-Veber, P.; Goina, E.; Touraine, R.; Manel, V.; Toutain, A.; Fehrenbach, S.; Frébourg, T.; Pagani, F.; Tosi, M.; et al. A rare SMN2 variant in a previously unrecognized composite splicing regulatory element induces exon 7 inclusion and reduces the clinical severity of spinal muscular atrophy. *Hum. Mutat.* **2010**, *31*, E1110–E1125. [CrossRef] [PubMed]
26. Karczewski, K.J.; Francioli, L.C.; Tiao, G.; Cummings, B.B.; Alföldi, J.; Wang, Q.; Collins, R.L.; Laricchia, K.M.; Ganna, A.; Birnbaum, D.P.; et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **2020**, *581*, 434–443. [CrossRef]
27. Glusman, G.; Caballero, J.; Mauldin, D.E.; Hood, L.; Roach, J.C. Kaviar: An accessible system for testing SNV novelty. *Bioinformatics* **2011**, *27*, 3216–3217. [CrossRef]
28. NHLBI and University of Michigan. The NHLBI TransOmics for Precision Medicine (TOPMed) Whole Genome Sequencing Program. BRAVO Variant Browser (University of Michigan) 2018. Available online: https://bravo.sph.umich.edu/freeze5/hg38/ (accessed on 20 June 2022).
29. Jaganathan, K.; Panagiotopoulou, S.K.; McRae, J.F.; Darbandi, S.F.; Knowles, D.; Li, Y.I.; Kosmicki, J.A.; Arbelaez, J.; Cui, W.; Schwartz, G.B.; et al. Predicting Splicing from Primary Sequence with Deep Learning. *Cell* **2019**, *176*, 535–548. [CrossRef]
30. Ke, S.; Shang, S.; Kalachikov, S.M.; Morozova, I.; Yu, L.; Russo, J.J.; Ju, J.; Chasin, L.A. Quantitative evaluation of all hexamers as exonic splicing elements. *Genome Res.* **2011**, *21*, 1360–1374. [CrossRef]
31. Souza, P.V.S.; Pinto, W.B.V.R.; Ricarte, A.; Badia, B.M.L.; Seneor, D.D.; Teixeira, D.T.; Caetano, L.; Gonçalves, E.A.; Chieia, M.A.T.; Farias, I.B.; et al. Clinical and radiological profile of patients with spinal muscular atrophy type 4. *Eur. J. Neurol.* **2021**, *28*, 609–619. [CrossRef]
32. Wadman, R.I.; Jansen, M.D.; Stam, M.; Wijngaarde, C.A.; Curial, C.A.D.; Medic, J.; Sodaar, P.; Schouten, J.; Vijzelaar, R.; Lemmink, H.H.; et al. Intragenic and structural variation in the SMN locus and clinical variability in spinal muscular atrophy. *Brain Commun.* **2020**, *2*, fcaa075. [CrossRef]
33. Bowen, B.M.; Truty, R.; Aradhya, S.; Bristow, S.L.; Johnson, B.A.; Morales, A.; Tan, C.A.; Westbrook, M.J.; Winder, T.L.; Chavez, J.C. SMA Identified: Clinical and Molecular Findings From a Sponsored Testing Program for Spinal Muscular Atrophy in More Than 2,000 Individuals. *Front. Neurol.* **2021**, *12*, 663911. [CrossRef] [PubMed]
34. Pane, M.; Donati, M.A.; Cutrona, C.; De Sanctis, R.; Pirinu, M.; Coratti, G.; Ricci, M.; Palermo, C.; Berti, B.; Leone, D.; et al. Neurological assessment of newborns with spinal muscular atrophy identified through neonatal screening. *Eur. J. Pediatr.* **2022**, *181*, 2821–2829. [CrossRef] [PubMed]
35. Cuscó, I.; Bernal, S.; Blasco-Pérez, L.; Calucho, M.; Alias, L.; Fuentes-Prior, P.; Tizzano, E.F. Practical guidelines to manage discordant situations of SMN2 copy number in patients with spinal muscular atrophy. *Neurol. Genet.* **2020**, *6*, e530. [CrossRef] [PubMed]
36. Griffiths, W.M.; Miller, A.J.F.; Suzuki, J.H.; Lewontin, D.T.; Gelbart, R. Chapter 14—Mutation, repair, and recombination. In *An introduction to Genetic Analysis*; W.H. Freemand and Company: New York, NY, USA, 2000.

37. Alías, L.; Bernal, S.; Barceló, M.J.; Also-Rallo, E.; Martínez-Hernández, R.; Rodríguez-Alvarez, F.J.; Hernández-Chico, C.; Baiget, M.; Tizzano, E.F. Accuracy of marker analysis, quantitative real-time polymerase chain reaction, and multiple ligation-dependent probe amplification to determine SMN2 copy number in patients with spinal muscular atrophy. *Genet. Test. Mol. Biomark.* **2011**, *15*, 587–594. [CrossRef] [PubMed]
38. Liu, J.; Huang, S.; Sun, M.; Liu, S.; Liu, Y.; Wang, W.; Zhang, X.; Wang, H.; Hua, W. An improved allele-specific PCR primer design method for SNP marker analysis and its application. *Plant Methods* **2012**, *8*, 34. [CrossRef]

# CHAPTER 4

# Work in progress: Towards implementation of a genomic data base of *SMN2* genes in SMA patients through complete sequencing of SMN locus

The scenario of spinal muscular atrophy has changed radically in recent years with the approval of therapies to treat the disease and the implementation of NBS in many countries[1–6]. In this context, the study of the *SMN2* copy number in patients is essential as a predictor of the SMA phenotype, since it is considered the main modifier of the disease. However, patients with the same *SMN2*_CN can develop different phenotypes and SMA subtypes, indicating that the correlation between *SMN2*_CN and patient phenotype is not absolute and that not all *SMN2* copies are equivalent. Some of these discrepancies may be explained by *SMN2* modifier variants, such as c.859G>C and c.835-44A>G, but these variants are very infrequent implying that the genomic data of *SMN2* should be investigated[7–9]. This may clarify the involvement of *SMN2* structures and variants in the phenotype and the evolution of the disease.

Following the objectives of this thesis, and with the availability of an NGS method that allows deep characterization of *SMN* genes (Chapter 2), the next step was to implement the genetic study of the *SMN2* gene in SMA patients to generate a database of variants and structures of their *SMN2* genes.

***Patients' collection & sample processing***

Initially, with the aim to study 250 samples, a total of 418 new samples have been collected from 21 different centers, both national and international. The collection of almost twice expected number of samples has extended the processing and analysis time. At present, *SMN1*_CN and *SMN2*_CN has been performed by MLPA probe mix P021 on all samples, following the manufacturer instructions. This kit allows quantitative analysis of all *SMN1* and *SMN2* exons. Furthermore, complete sequencing of *SMN2* has been carried out in 336 of the 418 patients, following the procedure previously described (see Chapter 2). In the remaining 82 samples, the NGS studies are in progress.

***Clinical classification of SMA patients***

All patients' phenotypes of this cohort were established according to the information provided by the clinician of the referral center by means of a questionnaire based on the motor milestones achieved prior to treatment and discussion of the doubtful cases.

Almost half of our patients were referred as type II SMA (45%), about a quarter of them were type I phenotype (26.5%) and another quarter type III (28%). Only one patient was referred as type IV phenotype and three detected by NBS (presymptomatic).

***Molecular cause of spinal muscular atrophy***

MLPA analysis allow to stablish *SMN1*_CN in all the 418 patients. In 401 patients (95.9%) the molecular cause of SMA was the homozygous deletion of *SMN1,* while 15 patients (3.6%) presented one *SMN1* copy and two patients showed two *SMN1* genes (0.5%). In these 17 patients, the most common pathogenic variant detected was *SMN1*:c.399_402del, also known as Spanish Mutation[10]. It was found in seven patients of Spanish origin, including one with the variant in homozygous state. The second was *SMN1*:c.815A>G, present in four patients of Spanish and Argentinean origin. All the *SMN1* variants detected are listed in Table 1.

***SMN2 copy number determination by MLPA***

Through MLPA study it was possible to determine the *SMN2*_CN in 412 out of the 418 patients, as six samples showed poor quality, which prevented the analysis. Results show that more than half of the patients presented three *SMN2* copies (56%), followed by cases with two *SMN2* copies (25%) and four *SMN2* copies (16%) (more details in Table 2).

| SMA Patient | *SMN1*_ CN | Molecular Cause | Localization |
|---|---|---|---|
| SMA055 | 1 | *SMN1*:c.399_402del | exon 3 |
| SMA056 | 1 | *SMN1*:c.815A>G | exon 6 |
| SMA099 | 1 | *SMN1*:c.188C>A | exon 2b |
| SMA100 | 1 | *SMN1*:c.188C>A | exon 2b |
| SMA103 | 2 | *SMN1*:c.399_402del (HMZ) | exon 3 |
| SMA168 | 1 | *SMN1*:c.399_402del | exon 3 |
| SMA179 | 1 | *SMN1*:c.815A>G | exon 6 |
| SMA184 | 1 | *SMN1*:c.399_402del | exon 3 |
| SMA191 | 1 | *SMN1*:c.815A>G | exon 6 |
| SMA193 | 2 | *SMN1*:c.399_402del | exon 3 |
| | | *SMN1*:c.788T>G | exon 6 |
| SMA215 | 1 | *SMN1*:c.399_402del | exon 3 |
| SMA224 | 1 | *SMN1*:c.399_402del | exon 3 |
| SMA275 | 1 | *SMN1*:c.347T>C | exon 3 |
| SMA331 | 1 | *SMN1*:c.332C>G | exon 3 |
| SMA360 | 1 | *SMN1*:c.245_246del | exon 2b |
| SMA432 | 1 | *SMN1*:c.815A>G | exon 6 |
| SMA610 | 1 | *SMN1*:c.43C>T | exon 1 |

**Table 1. List of all the pathogenic variants of *SMN1* detected in cohort.** Patients 103 is homozygous (HMZ) for the variant *SMN1*:c.399_402del while patient 193 showed two different variants: *SMN1*:c.399_402del and *SMN1*:c.788T>G.

| *SMN2*_ CN | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Nº of patients | 6 | 103 | 231 | 67 | 4 | 1 |

**Table 2. *SMN2* copy number of SMA patients based on MLPA determination.** The *SMN2*_CN was analyzed taking into account only exon 7 determination

In order to be able to compare our results with the ones obtained in the referral centers, *SMN2*_CN has been calculated taking into account only *SMN2* exon 7 (*SMN2-SMN1* hybrid genes are considered as wild type *SMN2)*. We detected a total of 33 discrepancies, corresponding to a discrepancy index of 8%. Specifically, 22 were

underestimations, whereas 11 were overestimations. Specifically, MLPA study has detected *SMN2-SMN1* hybrid structures in 43 patients.

***SMN2 characterization by NGS***

To date, 336 patients have been studied by NGS, obtaining reliable results in 313 of them (23 samples presented poor DNA quality). A preliminary analysis has been carried out in which PSVs, modifier variants and consensus regions for Nusinersen and Risdiplam drugs have been studied.

First, the analysis of the AB ratio of PSVs, SNPs and rare variants in each patient (following the method described in chapter 2) allowed us to determine *SMN1*_CN and *SMN2*_CN, thus confirming in all cases the copy number assigned by our MLPA study. Interestingly, when expanding the NGS analysis to 368 SMA samples (including 55 patients studied in chapter 2), we observed that PSV 8 (chr5:70246872:-/chr5:69371448:AGGCA) presented a high variability between patients, being detected in almost half of the samples. Therefore, PSV 8 should not be considered a PSV but a polymorphism. The updated scheme of PSV consider only 15 nucleotides differing between the two *SMN* genes (Figure 1).

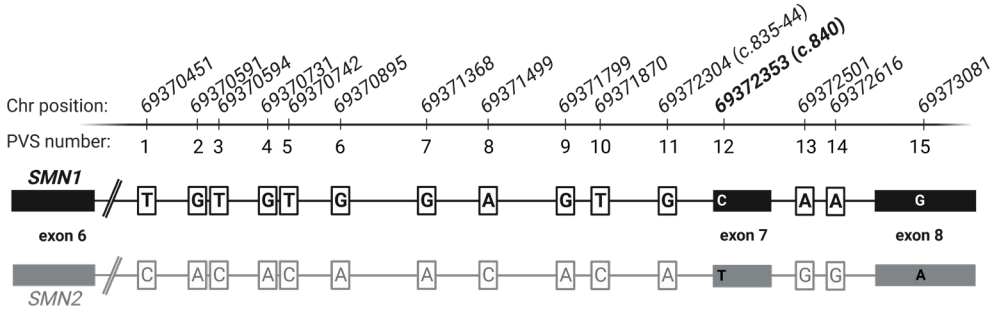



**Figure 1. Updated scheme representing the 15 PSVs that differentiates *SMN* genes.** *SMN1* gene is represented in black, while *SMN2* gene is in grey. The positions of the 15 PSVs are given in *SMN2* gene.

Based on the analysis of these 15 PSVs, *SMN2-SMN1* hybrid structures have been detected in 43 patients. Only 31 were identified by MLPA and of them, nine are the classical reported hybrids (exons 1 to 7 of *SMN2* and exon 8 of *SMN1)* (Figure 2A), while the others presented different complex structures. MLPA only studies PSVs

located in exons 7 (PSV 12) and 8 (PSV 15) of both *SMN* genes, therefore, these 22 structures, although detectable by MLPA, were not properly characterized as they present different PSVs from *SMN1*, in addition to the PSV 15 (Figure 2B). The remaining 12 patients presented hybrids structures that MLPA was not able to detect, as they present both PSVs of exons 7 and 8 from *SMN2* (Figure 2C). It is worth mentioning the identification of two hybrid structures in three patients that include the PSV c.835-44A>G, reported as positive modifier variant[9]. In addition, five patients showed only hybrid structures but no wild type *SMN2*.

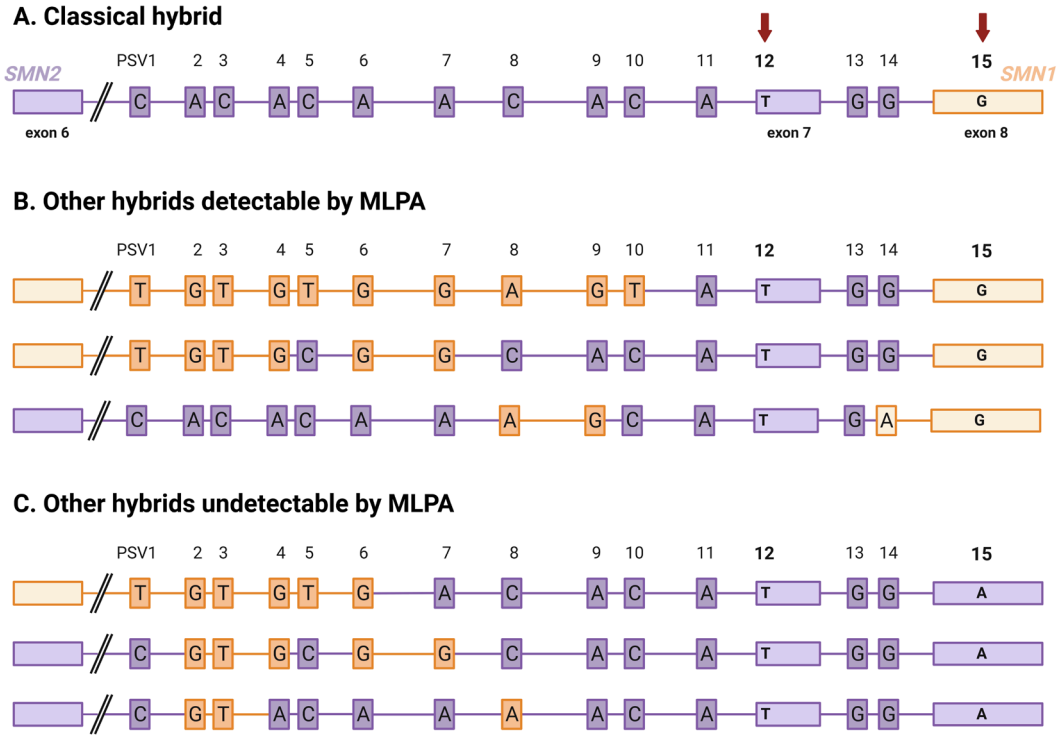



**Figure 2. Examples of hybrid structures detected.** *SMN2* gene is colored in purple and *SMN1* in orange. Red arrows point to PSVs 12 (exon 7) and 15 (exon 8), the only ones studied by MLPA. A) Classical hybrid consisting in all PSVs from *SMN2* except PSV 15 of exon 8, which correspond to *SMN1*. B) Examples of other hybrid structures detected in our cohort which differs from the classical hybrid but are detectable by MLPA as PSV 12 is from *SMN2* and PSV 15 is from *SMN1*. C) Examples of other hybrid structures different from the classical and undetectable by MLPA, as PSV 12 and 15 correspond to *SMN2*. These hybrid genes would be classified as *SMN2* genes by MLPA analysis.

Lastly, *SMN2* whole sequencing by NGS allowed the investigation of the target region of Nusinersen and Risdiplam, located in intron 7 and exon 7 of *SMN2*, respectively. Analysis of our results in 313 patients showed that there were no variants in the drug binding regions mentioned.

Once the rest of the samples have been processed, an in-depth analysis will be carried out to study the association between specific variants, structures of *SMN2,* and the phenotypes of the patients.

## REFERENCES


1. De Vivo DC, Bertini E, Swoboda KJ, et al. Nusinersen initiated in infants during the presymptomatic stage of spinal muscular atrophy: Interim efficacy and safety results from the Phase 2 NURTURE study. *Neuromuscul Disord*. 2019;29(11):842. doi:10.1016/J.NMD.2019.09.007
2. Lowes LP, Alfano LN, Arnold WD, et al. Impact of Age and Motor Function in a Phase 1/2A Study of Infants With SMA Type 1 Receiving Single-Dose Gene Replacement Therapy. *Pediatr Neurol*. 2019;98:39-45. doi:10.1016/J.PEDIATRNEUROL.2019.05.005
3. Baranello G, Darras BT, Day JW, et al. Risdiplam in Type 1 Spinal Muscular Atrophy. *N Engl J Med*. 2021;384(10):915-923. doi:10.1056/nejmoa2009965
4. Abiusi E, Vaisfeld A, Fiori S, et al. Experience of a 2-year spinal muscular atrophy NBS pilot study in Italy: towards specific guidelines and standard operating procedures for the molecular diagnosis. *J Med Genet*. Published online 2022. doi:10.1136/JMG-2022-108873
5. Boemer F, Caberg JH, Beckers P, et al. Three years pilot of spinal muscular atrophy newborn screening turned into official program in Southern Belgium. *Sci Rep*. 2021;11(1). doi:10.1038/S41598-021-99496-2
6. Pane M, Donati MA, Cutrona C, et al. Neurological assessment of newborns with spinal muscular atrophy identified through neonatal screening. *Eur J Pediatr*. 2022;181(7):2821. doi:10.1007/S00431-022-04470-3

7. Prior TW, Krainer AR, Hua Y, et al. A positive modifier of spinal muscular atrophy in the SMN2 gene. *Am J Hum Genet*. 2009;85(3):408-413.

8. Bernal S, Alías L, Barceló M, et al. The c.859G>C variant in the SMN2 gene is associated with types II and III SMA and originates from a common ancestor. *J Med Genet*. 2010;47(9):640-642. doi:10.1136/jmg.2010.079004

9. Wu X, Wang SH, Sun J, Krainer AR, Hua Y, Prior TW. A-44G transition in SMN2 intron 6 protects patients with spinal muscular atrophy. *Hum Mol Genet*. 2017;26(14):2768-2780. doi:10.1093/HMG/DDX166

10. Alías L, Bernal S, Fuentes-Prior P, et al. Mutation update of spinal muscular atrophy in Spain: Molecular characterization of 745 unrelated patients and identification of four novel mutations in the SMN1 gene. *Hum Genet*. 2009;125(1):29-39. doi:10.1007/s00439-008-0598-1

# DISCUSSION

The development of NGS completely transformed the field of genetics. Until then, the sequencing methods (Sanger) only allow to study individually genes in a small. In contrast, NGS allows massively parallel sequencing of large DNA regions in multiple patients in the same experiment. Therefore, it greatly reduces the economic cost of sequencing and processing time, which has made NGS a widely used technique for routine diagnosis of genetic diseases. In recent years, the trend in genetic diagnostics has been towards the use of large panels, even towards whole exome sequencing (WES) implementation. WES is particularly useful in cases where the clinical presentation of the patient is not specific enough to have a clear suspicion of a concrete genetic disease.

In this work, we propose the adaptation of NGS to deep into the challenging genetic characteristics presented in TSC and SMA. In TSC, we have explored the implication of low frequency and splicing variants as the possible cause of the disease, while in SMA we have developed an NGS-based method for the in-depth study of *SMN2* copies, variants and structures in relation with the SMA phenotypes.

## The role of mosaic and splicing variants in tuberous sclerosis complex

It is well established that the molecular cause of tuberous sclerosis complex is pathogenic variants in *TSC1* and *TSC2* genes[4,5]. Even so, the causal variant of the disease is still unknown in 10-25% of the patients studied by the standard techniques[25,28,29]. Based on the presence of undiagnosed patients, in the past, it was hypothesized the existence of a third gene responsible for the disease *(TSC3).* However, after years of research it was concluded that probably this third gene did not exist and that the most likely cause for NMI patients were mosaic and deep intronic variants[24,30,116].

In accordance with this information, we developed a workflow for molecular diagnosis of TSC, consisting of deep study of the *TSC1* and *TSC2* genes. MLPA was applied for the identification of CNVs, and high-coverage NGS of the complete genes together with an optimized bioinformatic analysis for the detection of both

deep intronic and low frequency mosaic variants. This strategy was followed in 42 patients with known positive clinical criteria as defined by the updated guidelines[1], 13 of whom had previous negative studies. The implementation of this approach resulted in an overall diagnostic rate of 69% (29/42), slightly lower than the reported (75-90%)[25,28,29]. This is probably due to the presence in our cohort of 13 patients with prior negative molecular studies (MLPA and/or exonic sequencing of *TSC1* and *TSC2*). Indeed, if these patients were excluded, our diagnostic rate would increase to 86% (25/29), corresponding to that reported in the literature.

Regarding the pathogenic variants identified, our results are mainly consistent with those previously described. First, causal variants were detected more frequently in *TSC2* than in *TSC1* (16 vs. 11). It has been described that variants in *TSC1* are associated with less severe phenotypes of the disease, therefore, it could be expected that milder cases would remain virtually asymptomatic and never be detected, underdiagnosing TSC patients with causal variants in *TSC1* [117,118]. Second, 70% of the pathogenic variants detected (19/27) were loss-of-function. This is in line with the predictors, as both genes have the highest score for loss-of-function intolerance (pLI=1), meaning that variants of this type are expected to be pathogenic in these genes[119]. In contrast, only one of the 29 diagnosed patients (3%) inherited the variant from a symptomatic progenitor, whereas the literature reports that about one third of TSC cases are inherited[6]. This discrepancy could be explained by improvements in genetic diagnosis, as patients with genetic disorders nowadays have more reproductive options to avoid the transmission of the disease to their offspring, including prenatal testing or even preimplantation genetic diagnosis (PGD)[120].

The applied workflow was intended to improve two issues with respect to the techniques routinely applied in TCS: detection of mosaic variants by means of high coverage NGS and detection of deep intronic variants by complete sequencing of *TSC* genes. Concerning the detection of mosaics, this workflow allowed the identification of eight mosaic variants, five of them with AF <20%, which would probably have remained undetected by routine techniques. Indeed, molecular

diagnosis was achieved in four patients previously classified as negative after applying standard techniques in other centers. Three of them presented pathogenic variants at low allele frequencies (<20%) detected by NGS, and the fourth patient showed a large mosaic deletion identified by MLPA and confirmed by long-range PCR. Interestingly, this deletion was not detectable using NGS or CGH-array, which would establish MLPA as the gold standard for CNV mosaic detection in TSC. In relation to deep intronic variants, whole sequencing of *TSC1* and *TSC2* identified the previously reported variant *TSC2*:c.2838-122G>A in two patients, but did not reveal any novel deep intronic candidate variants. Therefore, this type of variants explains the disease in 5% of our patients, which is consistent with the literature describing 1-12% of patients with deep intronic causative variants[24,30,121].

Based on our experience, the application of this workflow allowed the diagnosis of patients previously classified as NMI, being especially useful to detect low AF variants. This strategy also allows the identification of deep intronic variants, even though the yield was lower as they are less frequent among TSC patients. Consequently, the best strategy for TSC genetic diagnosis will be the use of high-coverage NGS of the complete *TSC* genes, combined with MLPA. Finally, patients who remain undiagnosed after being studied in depth probably present mosaic variants at very low frequency (<1%) or confined to a specific tissue. The recommended approach in these NMI patients is the application of this workflow in DNA from affected tissue, such as angiofibroma. Implementing this last strategy in all TSC patients would increase the diagnosis rate. However, it seems unnecessary to expose all patients to an invasive intervention, such as biopsy of affected tissue, and only considered this procedure in the remaining 15% of NMI patients.

## Towards a deep *SMN2* characterization in SMA patients

Spinal muscular atrophy affects one in 11,000 live births worldwide, but due to the high phenotypic variability of the disease, the diagnosis of SMA is not sufficient to predict the clinical course of the patient[38]. Obtaining information about the expected phenotype is crucial, since symptoms can range from severe involvement

with perinatal onset to mild involvement with onset in adolescence or even adulthood[47,48]. To date, the main modifier described for the disease is the number of copies of the *SMN2* gene, with the association that the higher the number of copies, the milder the phenotype. This genotype-phenotype correlation makes *SMN2*_CN determination of SMA patients an essential part of the disease prognosis. However, this correlation is not perfect and some patients present a discordant phenotype, either better-than-expected or worse-than-expected (based on their *SMN2*_CN)[55,62].

In this context, and with the aim to improve the existing genotype-phenotype correlation, we have developed an NGS method to explore the genetic variability of the *SMN2* gene in SMA patients. Its application is particularly relevant with the increasing diagnosis of presymptomatic/paucisymptomatic patients due to the implementation of NBS programs and the need to predict their evolution for treatment decisions[39–41].

Our method consists of three overlapping long-range PCRs including the promoter, 5' and 3' regions, exons and introns of the *SMN* genes, followed by high coverage NGS sequencing. To date, we have applied it to 368 SMA samples with a variable *SMN2*_CN between zero and six. Of these, 349 presented homozygous deletion of *SMN1* as the cause of SMA. Based on bibliographic information, the reference genome (Hg19) and our own NGS results, we established a revised list of 15 PSVs, corresponding to 15 nucleotides, as the only differences between *SMN1* and *SMN2* genes[68]. Nowadays, most SMA diagnostic techniques are based only on exon 7 and 8 PSVs, described in 1999[80,84]. The definition of all the PSVs in *SMN* genes enables the molecular differentiation of both genes through NGS, allowing the establishment of *SMN1*_CN and *SMN2*_CN and the detection of *SMN2-SMN1* hybrid genes, as well as the possibility to study *SMN1* or *SMN2* separately.

Although the method was originally designed to study sequence variability of *SMN2*, it can also be applied to determine *SMN1*_CN and *SMN2*_CN by analyzing the AB ratios of PSVs, as the results obtained by NGS correlate well with those of MLPA. Interestingly, comparing the information of *SMN2*_CN supplied by the

referral centers, we have observed a discrepancy in 8% of the patients respect to our MLPA results that were also confirmed by NGS analysis. This highlight the complexity to establish an accurate determination of *SMN2*_CN in some cases[122].

Furthermore, PSVs analysis allows the identification and characterization of *SMN2-SMN1* hybrid structures. Traditionally, *SMN2-SMN1* hybrid genes have been considered when there is the presence of exon 7 from *SMN2* and the exon 8 from *SMN1*. This structure could be detected through routine approaches based on PSVs 12 and 15 from exons 7 and 8 respectively (such as MLPA or qPCR), and for this reason it is the more reported hybrid, although other structures have been also described[77,123]. Specifically, a total of 45 patients with hybrid genes were detected in our cohort (45/368, 12.2%). This represents a higher proportion than the 5-10% described in the literature[59,78], a fact than can be explained by the wide information obtained by our method. The analysis of the 15 PSVs reveals 14 patients with hybrids structures undetectable by MLPA, as they are formed by a *SMN2* gene with a fragment of intron 6 from *SMN1*, not involving exons 7 and 8 (Figure 2; Chapter4). Then, the detection of these atypical hybrids could partially explain the higher percentage of hybrids identified in our cohort.

It should be noted that, besides detecting more hybrids, NGS provides a detailed characterization of them. In particular, from our 45 patients with hybrid genes, only nine showed the traditional reported hybrid (exons 1 to 7 of *SMN2* and exon 8 of *SMN1)* and, among them, a total of 21 different structures were identified. Since each hybrid structure consists in different PSVs composition, it is to be expected that not all of them will be equally functional. Accordingly, hybrid genes must be studied individually, and not as a whole, to help elucidate the contribution of each hybrid structure to SMN-full length production, and therefore, to the SMA phenotype.

Regarding SNVs analysis, this NGS method could be used to identify the modifier variants c.859C>G and c.835-44A>G of *SMN2.* It has been widely described that both variants increase the inclusion of exon 7 of *SMN2*, thereby generating more full-length SMN protein and improving SMA phenotype in patients.

Moreover, c.859C>G variant was identified in 11 patients, our method allowed to determine the number of *SMN2* copies with the variant, identifying two cases with the variant present in their two *SMN2* copies. All patients had a better-than-expected phenotype according to their *SMN2*_CN but those patients with the modifier variant in two of their *SMN2* show an even better phenotype than those with the modifier variant in only one copy, confirming the additive character of the c.859C>G variant. (Table 1, Chapter 3). On the other hand, the complete analysis of the SMN region in these patients led us to identify a common haplotype named *Smn2-859C.1* that was consistent in all the patients with the modifier variant. In addition, a second haplotype, *Smn2-859C.2,* was detected in two patients as part of a complex allele formed by the *Smn2-859C.1* and *Smn2-859C.2* haplotypes in cis configuration (Figure 1, Chapter 3). In-depth study of the two haplotypes shows that they share a region of at least 8848 bp located around the variant, which points towards a homologous recombination event as the origin of the complex allele (Figure 3, Chapter 3). This theory is supported by the presence of an allele with two *Smn2-859C.1* haplotypes in cis in patient 3, necessary for this recombination to occur (Figure 1, Chapter 3). These finding are especially noteworthy since the 11 patients studied have different origins (including Spanish, Chilean, Italian and Danish) and, thus, suggests that the appearance of the c.859C>G variant was a single event in evolution and that all patients who present it have a common genetic ancestral origin.

Based on the available information, no clear phenotypic differences have been observed between patients with *Smn2-859C.1* haplotype in homozygosis and patients with the complex allele (*Smn2-859C.1* + *Smn2-859C.2* ). Similarly, patients with two *Smn2-859C.1* haplotypes in cis present similar phenotypes than the ones presenting in trans configuration. Even though, further investigation is needed to assert that there is no difference between c.859C>G associated haplotypes or configurations, as our cohort consist of only 11 patients, a limited number to perform association studies.

In our cohort, we detected the c.835-44A>G variant in four patients (including a two siblings). In three of them, the variant was detected as part of a *SMN2-SMN1* hybrid gene, while in one patient it was detected alone. All four patients had a better-than-expected phenotype based on their *SMN2*_CN, which is consistent with the literature[62,76]. In addition, the study of SNVs could be extended to *SMN1* variants as well. This method would be especially useful for studying the complete *SMN1* in patients with suspicion of SMA with one *SMN1* copy. In our cohort, 17 patients presented one or two *SMN1* genes, and in all of them NGS analysis allowed to detect the pathogenic variant.

As part of the *SMN2* analysis, we studied the binding regions of Nusinersen and Risdiplam, two of the more frequent treatments used in SMA patients[90–92,95,96]. Analysis of our results in 368 patients showed that there were no variants in the region. These results confirmed both binding sequences are highly conserved.

To sum up, we developed a NGS based method that was specially adapted to the complexity of the SMA region and it is a relatively simple and affordable. This method could be proposed as a second-step in SMA diagnosis after the establishment of *SMN1* and *SMN2* copy number by a quantitative method. Its application will allow, first, copy number confirmation; second, detection of *SMN2* modifier variants c.835-44A>G and c.859C>G, including its zygosity; and third, identification of hybrid structures and other candidate variants that could modify *SMN2* function. Collecting all this information from each SMA patient would greatly improve the knowledge of *SMN2* variability and could eventually improve the current genotype-phenotype correlation of the disease.

## Concluding remarks

The adapted application of the NGS has allowed an improved genetic characterization of TSC and SMA diseases. In the TSC field, the use of high-coverage NGS sequencing of the complete *TSC* genes enable the molecular diagnosis in 29 patients, including four previous NMI patients, demonstrating the importance of studying low-frequency and deep intronic variants.

In the SMA field, the development of an NGS-based method specific to the SMN region allowed an in-depth genetic characterization of SMA patients. Traditionally, in the *SMN2* gene, only copy number is studied. This method makes possible to further explore *SMN2* variants and structures, which, to date, have an unknown implication. Larger studies in well clinically characterized cohorts may demonstrate their potential utility in order to achieve a better genotype-phenotype correlation and an improved use of *SMN2* as a prognostic factor in SMA NBS setting.

# CONCLUSIONS

## Conclusions

- Further application of the NGS technique allowed specific development of protocols for in-depth genetic studies of TSC and SMA diseases.

The specific conclusions in the TSC field are:

- The workflow designed for TSC diagnosis enables the identification of low-frequency (AF>1%) and deep intronic variants by high coverage NGS of the complete *TSC* genes, as well as the identification of mosaic CNVs by MLPA.
- In our series, we identify the causal variant in the 69% of patients, 11 in *TSC1* and 16 in *TSC2*. We detected seven missense, one in-frame deletion and 19 loss of function variants. Five of the 29 patients diagnosed presented low-frequency variants (<20%) and two other patients showed deep intronic variants.
- Low-frequency variants constitute a frequent cause of TSC in NMI patients.

The specific conclusions in the SMA field are:

- The developed NGS based technique allows the detailed characterization of the SMN genomic region in SMA patients.
- We were able to establish an updated list of 15 PSVs that differentiates *SMN1* and *SMN2* genes, allowing *SMN1*_CN and *SMN2*_CN determination as well as *SMN2-SMN1* hybrid characterization.
- In our cohort, we determined the structure of *SMN2* genes of 11 patients, from different geographical origins, carrying the c.859C>G variant. Two haplotypes (*Smn-859.C1* and *Smn-859.C2*) were identified in association with the modifier variant. Whilst *Smn-859.C1* is present

in all cases, *Smn-859.C2* is only present in two patients, as part of a complex allele formed by both *Smn-859.C1* and *Smn-859.C2* in cis.

- Haplotypes *Smn-859.C1* and *Smn-859.C2* share a common region of at least 8848bp, pointing towards a common ancestral origin of all the alleles with the c.859C>G variant.
- We detected four patients with the c.835-44A>G modifier variant. This variant has been detected either as a single change or as part of a hybrid structure with other PSVs.
- We identified 45 patients with hybrid genes (12%), representing a total of 21 different structures. In 31 patients the hybrid was also detected by MLPA while the remaining 14 were only identified by NGS. Hybrid structures seem to be a relevant source of genetic variability in our cohort and their role in patients' phenotype should be investigated.
- The study of the specific binding regions of *SMN2* targeted therapies (Nursinesen and Risdiplam) did not show different sequences among all 368 patients with NGS data, suggesting that these regions are highly conserved in the SMA patients.

# REFERENCES

# REFERENCES


1. Northrup H, Aronow ME, Bebin EM, et al. Updated International Tuberous Sclerosis Complex Diagnostic Criteria and Surveillance and Management Recommendations. *Pediatr Neurol*. 2021;123:50-66. doi:10.1016/j.pediatrneurol.2021.07.011
2. Roach ES. Applying the Lessons of Tuberous Sclerosis: The 2015 Hower Award Lecture. *Pediatr Neurol*. 2016;63:6-22. doi:10.1016/J.PEDIATRNEUROL.2016.07.003
3. Rodriguez gomez M, Sampson JR, Holets Whittemore V. *Tuberous Sclerosis Complex*. 3th ed.; 1999.
4. European Chromosome 16 Tuberous Sclerosis Consortium. Identification and characterization of the tuberous sclerosis gene on chromosome 16. *Cell*. 1993;75(7):1305-1315. doi:10.1016/0092-8674(93)90618-Z
5. Van Slegtenhorst M, De Hoogt R, Hermans C, et al. Identification of the tuberous sclerosis gene TSC1 on chromosome 9q34. *Science*. 1997;277(5327):805-808. doi:10.1126/SCIENCE.277.5327.805
6. Osborne JP, Fryer A, Webb D. Epidemiology of tuberous sclerosis. *Ann N Y Acad Sci*. 1991;615(1):125-127. doi:10.1111/J.1749-6632.1991.TB37754.X
7. Ebrahimi-Fakhari D, Mann LL, Poryo M, et al. Incidence of tuberous sclerosis and age at first diagnosis: new data and emerging trends from a national, prospective surveillance study. *Orphanet J Rare Dis*. 2018;13(1). doi:10.1186/S13023-018-0870-Y
8. Ali SA, Mulita F. *Hamartoma*. StatPearls Publishing; 2022. Accessed February 8, 2023. https://www.ncbi.nlm.nih.gov/books/NBK562298/
9. Roach ES, Gomez MR, Northrup H. Tuberous sclerosis complex consensus conference: revised clinical diagnostic criteria. *J Child Neurol*. 1998;13(12):624-628. doi:10.1177/088307389801301206
10. Northrup H, Krueger DA. Tuberous Sclerosis Complex Diagnostic Criteria Update: Recommendations of the 2012 International Tuberous Sclerosis Complex Consensus Conference. Published online 2013. doi:10.1016/j.pediatrneurol.2013.08.001

11. Amin S, Lux A, Calder N, Laugharne M, Osborne J, O'callaghan F. Causes of mortality in individuals with tuberous sclerosis complex. *Dev Med Child Neurol*. 2017;59(6):612-617. doi:10.1111/dmcn.13352

12. Davis PE, Filip-Dhima R, Sideridis G, et al. Presentation and Diagnosis of Tuberous Sclerosis Complex in Infants. *Pediatrics*. 2017;140(6). doi:10.1542/PEDS.2016-4040

13. Randle SC. Tuberous sclerosis complex: A review. *Pediatr Ann*. 2017;46(4):e166-e171. doi:10.3928/19382359-20170320-01

14. Ewalt DH, Sheffield E, Sparagana SP, Delgado MR, Roach ES. RENAL LESION GROWTH IN CHILDREN WITH TUBEROUS SCLEROSIS COMPLEX. *J Urol*. 1998;160(1):141-145. doi:10.1016/S0022-5347(01)63072-6

15. Northrup H, Koenig MK, Pearson DA, Au KS. Tuberous Sclerosis Complex. *GeneReviews®*. Published online 1993. Accessed November 8, 2022. https://pubmed.ncbi.nlm.nih.gov/20301399/

16. Rose VM, Au KS, Pollom G, Roach ES, Prashner HR, Northrup H. Germline mosaicism in tuberous sclerosis: How common? *Am J Hum Genet*. 1999;64(4):986-992. doi:10.1086/302322

17. Van Slegtenhorst M, Nellist M, Nagelkerken B, et al. Interaction between hamartin and tuberin, the TSC1 and TSC2 gene products. *Hum Mol Genet*. 1998;7(6):1053-1057. doi:10.1093/HMG/7.6.1053

18. Rosengren T, Nanhoe S, de Almeida LGD, et al. Mutational analysis of TSC1 and TSC2 in Danish patients with tuberous sclerosis complex. *Sci Rep*. 2020;10(1):1-9. doi:10.1038/s41598-020-66588-4

19. Han JM, Sahin M. TSC1/TSC2 Signaling in the CNS. *FEBS Lett*. 2011;585(7):973. doi:10.1016/J.FEBSLET.2011.02.001

20. Huang J, Manning BD. The TSC1-TSC2 complex: a molecular switchboard controlling cell growth. *Biochem J*. 2008;412(2):179-190. doi:10.1042/BJ20080281

21. Laplante M, Sabatini DM. mTOR signaling in growth control and disease. *Cell*. 2012;149(2):274-293. doi:10.1016/J.CELL.2012.03.017

22. Fokkema IFAC, Taschner PEM, Schaafsma GCP, Celli J, Laros JFJ, den Dunnen JT. LOVD v.2.0: the next generation in gene variant databases. *Hum Mutat*. 2011;32(5):557-563. doi:10.1002/HUMU.21438

23. Northrup H, Wheless JW, Bertin TK, Lewis RA. Variability of expression in tuberous sclerosis. *J Med Genet*. 1993;30(1):41-43. doi:10.1136/JMG.30.1.41

24. Tyburczy ME, Dies KA, Glass J, et al. Mosaic and Intronic Mutations in TSC1/TSC2 Explain the Majority of TSC Patients with No Mutation Identified by Conventional Testing. *PLoS Genet*. 2015;11(11). doi:10.1371/JOURNAL.PGEN.1005637

25. Au KS, Williams AT, Roach ES, et al. Genotype/phenotype correlation in 325 individuals referred for a diagnosis of tuberous sclerosis complex in the United States. *Genet Med*. 2007;9(2):88-100. doi:10.1097/GIM.0B013E31803068C7

26. Landrum MJ, Lee JM, Benson M, et al. ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res*. 2018;46(D1):D1062-D1067. doi:10.1093/NAR/GKX1153

27. Kozlowski P, Roberts P, Dabora S, et al. Identification of 54 large deletions/duplications in TSC1 and TSC2 using MLPA, and genotype-phenotype correlations. *Hum Genet*. 2007;121(3-4):389-400. doi:10.1007/S00439-006-0308-9

28. Dabora SL, Kwiatkowski DJ, Franz DN, et al. Mutational Analysis in a Cohort of 224 Tuberous Sclerosis Patients Indicates Increased Severity of TSC2, Compared with TSC1, Disease in Multiple Organs. *Am J Hum Genet*. 2001;68(1):64. doi:10.1086/316951

29. Sancak O, Nellist M, Goedbloed M, et al. Mutational analysis of the TSC1 and TSC2 genes in a diagnostic setting: genotype--phenotype correlations and comparison of diagnostic DNA techniques in Tuberous Sclerosis Complex. *Eur J Hum Genet*. 2005;13(6):731-741. doi:10.1038/SJ.EJHG.5201402

30. Nellist M, Brouwer RWW, Kockx CEM, et al. Targeted Next Generation Sequencing reveals previously unidentified TSC1 and TSC2 and mutations. *BMCNellist, M, Brouwer, R W W, Kockx, C E M, van Veghel-Plandsoen, M, Withagen-Hermans, C, Prins-Bakker, L, Hoogeveen-Westerveld, M, Mrsic, A, van den Berg, M M P, Koopmans, A E, Wit, M C, Jansen, F E, Maat-Kievit, A J A, van*. 2015;16(1):1-11. doi:10.1186/s12881-015-0155-4

31. Kolb S, Kissel J. Spinal Muscular Atrophy. *Neurol Clin*. 2015;33(4):831-

846. doi:10.1016/j.ncl.2015.07.004

32. Dubowitz V. Ramblings in the history of spinal muscular atrophy. *Neuromuscul Disord*. 2009;19(1):69-73. doi:10.1016/j.nmd.2008.10.004

33. Wohlfart G, Fex J, Eliasson S. Hereditary proximal spinal muscular atrophy, a clinical entity simulating progressive muscular dystrophy . *Acta Psychiatr Neurol Scand*. 1955;30:395-406. Accessed March 17, 2023. https://pubmed.ncbi.nlm.nih.gov/13301901/

34. Kugelberg E, Welander L. Heredofamilial Juvenile muscular Atrophy Simulating muscular Dystrophy. *AMA Arch Neurol Psychiatry*. 1956;75(5):500-509. doi:10.1001/ARCHNEURPSYC.1956.02330230050005

35. Dubowitz V. Infantile muscular atrophy. A prospective study with particular reference to a slowly progressive variety. *Brain*. 1964;87(4):707-718. doi:https://doi.org/10.1093/brain/87.4.707

36. Lefebvre S, Bürglen L, Reboullet S, et al. Identification and characterization of a spinal muscular atrophy-determining gene. *Cell*. 1995;80(1):155-165. doi:10.1016/0092-8674(95)90460-3

37. Cusin V, Clermont O, Gérard B, Chantereau D, Elion J. Prevalence of SMN1 deletion and duplication in carrier and normal populations: implication for genetic counselling. *J Med Genet*. 2003;40(4):e39.

38. Sugarman EA, Nagan N, Zhu H, et al. Pan-ethnic carrier screening and prenatal diagnosis for spinal muscular atrophy: clinical laboratory analysis of >72 400 specimens. *Eur J Hum Genet 2012 201*. 2011;20(1):27-32. doi:10.1038/ejhg.2011.134

39. Abiusi E, Vaisfeld A, Fiori S, et al. Experience of a 2-year spinal muscular atrophy NBS pilot study in Italy: towards specific guidelines and standard operating procedures for the molecular diagnosis. *J Med Genet*. Published online 2022. doi:10.1136/JMG-2022-108873

40. Boemer F, Caberg JH, Beckers P, et al. Three years pilot of spinal muscular atrophy newborn screening turned into official program in Southern Belgium. *Sci Rep*. 2021;11(1). doi:10.1038/S41598-021-99496-2

41. Pane M, Donati MA, Cutrona C, et al. Neurological assessment of newborns with spinal muscular atrophy identified through neonatal

screening. *Eur J Pediatr*. 2022;181(7):2821. doi:10.1007/S00431-022-04470-3

42. Fischer U, Liu Q, Dreyfuss G. The SMN-SIP1 complex has an essential role in spliceosomal snRNP biogenesis. *Cell*. 1997;90(6):1023-1029. doi:10.1016/S0092-8674(00)80368-2

43. Liu Q, Fischer U, Wang F, Dreyfuss G. The Spinal Muscular Atrophy Disease Gene Product, SMN, and Its Associated Protein SIP1 Are in a Complex with Spliceosomal snRNP Proteins. *Cell*. 1997;90(6):1013-1021. doi:10.1016/S0092-8674(00)80367-0

44. Burghes AHM, Beattie CE. Spinal muscular atrophy: why do low levels of survival motor neuron protein make motor neurons sick? *Nat Rev Neurosci 2009 108*. 2009;10(8):597-609. doi:10.1038/nrn2670

45. Detering NT, Zambon A, Hensel N, et al. 264th ENMC International Workshop: Multi-system involvement in spinal muscular atrophy Hoofddorp, the Netherlands, November 19th – 21st 2021. *Neuromuscul Disord*. 2022;32(8):697-705. doi:10.1016/J.NMD.2022.06.005

46. Munsat TL, Davies KE. International SMA Consortium Meeting (26-28 June 1992, Bonn, Germany). *Neuromuscul Disord*. 1992;2(5-6):423-428. doi:10.1016/S0960-8966(06)80015-5

47. Talbot K, Tizzano EF. The clinical landscape for SMA in a new therapeutic era. *Gene Ther*. 2017;24(9):529-533. doi:10.1038/gt.2017.52

48. Tizzano EF. Treating neonatal spinal muscular atrophy: A 21st century success story? *Early Hum Dev*. 2019;138(xxxx):104851. doi:10.1016/j.earlhumdev.2019.104851

49. Serra-Juhe C, Tizzano EF. Perspectives in genetic counseling for spinal muscular atrophy in the new therapeutic era: early pre-symptomatic intervention and test in minors. *Eur J Hum Genet*. 2019;27:1774-1782. doi:10.1038/s41431-019-0415-4

50. Oskoui M, Darras BT, De Vivo DC. Spinal Muscular Atrophy: 125 Years Later and on the Verge of a Cure. *Spinal Muscular Atrophy Dis Mech Ther*. Published online January 1, 2017:3-19. doi:10.1016/B978-0-12-803685-3.00001-X

51. Zerres K, Rudnik-Schöneborn S, Forrest E, Lusakowska A, Borkowska J, Hausmanowa-Petrusewicz I. A collaborative study on the natural history

of childhood and juvenile onset proximal spinal muscular atrophy (type II and III SMA): 569 patients. *J Neurol Sci*. 1997;146(1):67-72. doi:10.1016/S0022-510X(96)00284-5

52. Cuscó I, López E, Soler-Botija C, Barceló MJ, Baiget M, Tizzano EF. A genetic and phenotypic analysis in Spanish spinal muscular atrophy patients with c.399_402del AGAG, the most frequently found subtle mutation in the SMN1 gene. *Hum Mutat*. 2003;22(2):136-143. doi:10.1002/HUMU.10245

53. Bürglen L, Lefebvre S, Clermont O, et al. Structure and organization of the human survival motor neurone (SMN) gene. *Genomics*. 1996;32(3):479-482. doi:10.1006/GENO.1996.0147

54. Rochette CF, Gilbert N, Simard LR. SMN gene duplication and the emergence of the SMN2 gene occurred in distinct hominids: SMN2 is unique to Homo sapiens. *Hum Genet*. 2001;108(3):255-266. doi:10.1007/S004390100473

55. Calucho M, Bernal S, Alías L, et al. Correlation between SMA type and SMN2 copy number revisited: An analysis of 625 unrelated Spanish patients and a compilation of 2834 reported cases. *Neuromuscul Disord*. 2018;28(3):208-215. doi:10.1016/j.nmd.2018.01.003

56. Hendrickson BC, Donohoe C, Akmaev VR, et al. Differences in SMN1 allele frequencies among ethnic groups within North America. *J Med Genet*. 2009;46(9):641-644. doi:10.1136/JMG.2009.066969

57. Wirth B. An update of the mutation spectrum of the survival motor neuron gene (SMN1) in autosomal recessive spinal muscular atrophy (SMA). *Hum Mutat*. 2000;15(3):228-237. doi:10.1002/(SICI)1098-1004(200003)15:3<228::AID-HUMU3>3.0.CO;2-9

58. Cuscó I, Barceló MJ, Del Rio E, et al. Characterisation of SMN hybrid genes in Spanish SMA patients: De novo, homozygous and compound heterozygous cases. *Hum Genet*. 2001;108(3):222-229. doi:10.1007/s004390000452

59. Alías L, Bernal S, Fuentes-Prior P, et al. Mutation update of spinal muscular atrophy in Spain: Molecular characterization of 745 unrelated patients and identification of four novel mutations in the SMN1 gene. *Hum Genet*. 2009;125(1):29-39. doi:10.1007/s00439-008-0598-1

60. Vijzelaar R, Snetselaar R, Clausen M, et al. The frequency of SMN gene variants lacking exon 7 and 8 is highly population dependent. *PLoS One*. 2019;14(7). doi:10.1371/JOURNAL.PONE.0220211

61. Chen X, Sanchis-Juan A, French CE, et al. Spinal muscular atrophy diagnosis and carrier screening from genome sequencing data. *Genet Med*. 2020;22(5):945-953. doi:10.1038/s41436-020-0754-0

62. Ruhno C, McGovern VL, Avenarius MR, et al. Complete sequencing of the SMN2 gene in SMA patients detects SMN gene deletion junctions and variants in SMN2 that modify the SMA phenotype. *Hum Genet*. 2019;0(0):0. doi:10.1007/s00439-019-01983-0

63. Mendonça R de H, Matsui C, Polido J, et al. Intragenic variants in the SMN1 gene determine the clinical phenotype in 5q spinal muscular atrophy. *Neurol Genet*. 2020;1(6). doi:10.1212/NXG.0000000000000505

64. Qu Y jin, Bai J li, Cao Y yan, et al. Mutation Spectrum of the Survival of Motor Neuron 1 and Functional Analysis of Variants in Chinese Spinal Muscular Atrophy. *J Mol Diagn*. 2016;18(5):741-752. doi:10.1016/J.JMOLDX.2016.05.004

65. Lorson CL, Strasswimmer J, Yao JM, et al. SMN oligomerization defect correlates with spinal muscular atrophy severity. *Nat Genet 1998 191*. 1998;19(1):63-66. doi:10.1038/ng0598-63

66. Bernal S, Alías L, Barceló M, et al. The c.859G>C variant in the SMN2 gene is associated with types II and III SMA and originates from a common ancestor. *J Med Genet*. 2010;47(9):640-642. doi:10.1136/jmg.2010.079004

67. Prior TW, Krainer AR, Hua Y, et al. A positive modifier of spinal muscular atrophy in the SMN2 gene. *Am J Hum Genet*. 2009;85(3):408-413.

68. Monani UR, Lorson CL, Parsons DW, et al. A single nucleotide difference that alters splicing patterns distinguishes the SMA gene SMN1 from the copy gene SMN2. *Hum Mol Genet*. 1999;8(7):1177-1183.

69. Vitte J, Fassier C, Tiziano F, et al. Refined characterization of the expression and stability of the SMN gene products. *Am J Pathol*. 2007;171(4):1269-1280.

70. Lorson CL, Hahnen E, Androphy EJ, Wirth B. A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular

atrophy. *Proc Natl Acad Sci U S A*. 1999;96(11):6307-6311.

71. Wirth B, Garbes L, Riessland M. How genetic modifiers influence the phenotype of spinal muscular atrophy and suggest future therapeutic approaches. *Curr Opin Genet Dev*. 2013;23(3):330-338.

72. Soler-Botija C, Cuscó I, Caselles L, López E, Baiget M, Tizzano EF. Implication of fetal SMN2 expression in type I SMA pathogenesis: protection or pathological gain of function? *J Neuropathol Exp Neurol*. 2005;64(3):215-223.

73. Boza-Morán MG, Martínez-Hernández R, Bernal S, et al. Decay in survival motor neuron and plastin 3 levels during differentiation of iPSC-derived human motor neurons. *Sci Rep*. 2015;5(Jun):11696.

74. Cartegni L, Krainer AR. Disruption of an SF2/ASF-dependent exonic splicing enhancer in SMN2 causes spinal muscular atrophy in the absence of SMN1. *Nat Genet 2002 304*. 2002;30(4):377-384. doi:10.1038/ng854

75. Vezain M, Saugier-Veber P, Goina E, et al. A rare SMN2 variant in a previously unrecognized composite splicing regulatory element induces exon 7 inclusion and reduces the clinical severity of spinal muscular atrophy. *Hum Mutat*. 2010;31(1):1110-1125. doi:10.1002/humu.21173

76. Wu X, Wang SH, Sun J, Krainer AR, Hua Y, Prior TW. A-44G transition in SMN2 intron 6 protects patients with spinal muscular atrophy. *Hum Mol Genet*. 2017;26(14):2768-2780. doi:10.1093/HMG/DDX166

77. Wadman RI, Jansen MD, Stam M, et al. Intragenic and structural variation in the SMN locus and clinical variability in spinal muscular atrophy. *Brain Commun*. 2020;2(2):1-13. doi:10.1093/braincomms/fcaa075

78. Hahnen E, Schönling J, Rudnik-Schöneborn S, Zerres K, Wirth B. Hybrid survival motor neuron genes in patients with autosomal recessive spinal muscular atrophy: New insights into molecular mechanisms responsible for the disease. *Am J Hum Genet*. 1996;59(5):1057-1065.

79. Niba ETE, Nishio H, Wijaya YOS, et al. Clinical phenotypes of spinal muscular atrophy patients with hybrid SMN gene. *Brain Dev*. 2021;43(2):294-302. doi:10.1016/J.BRAINDEV.2020.09.005

80. Arkblad EL, Darin N, Berg K, et al. Multiplex ligation-dependent probe amplification improves diagnostics in spinal muscular atrophy.

*Neuromuscul Disord*. 2006;16(12):830-838. doi:10.1016/j.nmd.2006.08.011

81. Anhuf D, Eggermann T, Rudnik-Schöneborn S, Zerres K. Determination of SMN1 and SMN2 copy number using TaqMan™ technology. *Hum Mutat*. 2003;22(1):74-78. doi:10.1002/HUMU.10221

82. Feldkötter M, Schwarzer V, Wirth R, Wienker TF, Wirth B. Quantitative analyses of SMN1 and SMN2 based on real-time lightcycler PCR: Fast and highly reliable carrier testing and prediction of severity of spinal muscular atrophy. *Am J Hum Genet*. 2002;70(2):358-368. doi:10.1086/338627

83. Stabley DL, Harris AW, Holbrook J, et al. SMN1 and SMN2 copy numbers in cell lines derived from patients with spinal muscular atrophy as measured by array digital PCR. *Mol Genet Genomic Med*. 2015;3(4):248. doi:10.1002/MGG3.141

84. Milligan JN, Larson JL, Filipovic-Sadic S, et al. Multisite Evaluation and Validation of a Sensitive Diagnostic and Screening System for Spinal Muscular Atrophy that Reports SMN1 and SMN2 Copy Number, along with Disease Modifier and Gene Duplication Variants. *J Mol Diagn*. 2021;23(6):753. doi:10.1016/J.JMOLDX.2021.03.004

85. Cuscó I, Bernal S, Blasco-Pérez L, et al. Practical guidelines to manage discordant situations of SMN2 copy number in patients with spinal muscular atrophy . *Neurol Genet*. 2020;6(6):e530. doi:10.1212/nxg.0000000000000530

86. Alías L, Bernal S, Barceló MJ, et al. Accuracy of marker analysis, quantitative real-time polymerase chain reaction, and multiple ligation-dependent probe amplification to determine SMN2 copy number in patients with spinal muscular atrophy. *Genet Test Mol Biomarkers*. 2011;15(9):587-594. doi:10.1089/gtmb.2010.0253

87. Wang CH, Finkel RS, Bertini ES, et al. Consensus statement for standard of care in spinal muscular atrophy. *J Child Neurol*. 2007;22(8):1027-1049.

88. Mercuri E, Finkel RS, Muntoni F, et al. Diagnosis and management of spinal muscular atrophy: Part 1: Recommendations for diagnosis, rehabilitation, orthopedic and nutritional care. *Neuromuscul Disord*. 2018;28(2):103-115. doi:10.1016/J.NMD.2017.11.005

89. Finkel RS, Mercuri E, Meyer OH, et al. Diagnosis and management of spinal muscular atrophy: Part 2: Pulmonary and acute care; medications, supplements and immunizations; other organ systems; and ethics. *Neuromuscul Disord*. 2018;28(3):197-207. doi:10.1016/J.NMD.2017.11.004

90. Tiziano FD, Tizzano EF. 25 years of the SMN genes: the Copernican revolution of spinal muscular atrophy. *Acta Myol*. 2020;39(4):336. doi:10.36185/2532-1900-037

91. Finkel RS, Mercuri E, Darras BT, et al. Nusinersen versus Sham Control in Infantile-Onset Spinal Muscular Atrophy. *N Engl J Med*. 2017;377(18):1723-1732. doi:10.1056/NEJMOA1702752

92. De Vivo DC, Bertini E, Swoboda KJ, et al. Nusinersen initiated in infants during the presymptomatic stage of spinal muscular atrophy: Interim efficacy and safety results from the Phase 2 NURTURE study. *Neuromuscul Disord*. 2019;29(11):842. doi:10.1016/J.NMD.2019.09.007

93. Mendell JR, Al-Zaidy S, Shell R, et al. Single-Dose Gene-Replacement Therapy for Spinal Muscular Atrophy. *N Engl J Med*. 2017;377(18):1713-1722. doi:10.1056/NEJMOA1706198/SUPPL_FILE/NEJMOA1706198_DISCLOSURES.PDF

94. Lowes LP, Alfano LN, Arnold WD, et al. Impact of Age and Motor Function in a Phase 1/2A Study of Infants With SMA Type 1 Receiving Single-Dose Gene Replacement Therapy. *Pediatr Neurol*. 2019;98:39-45. doi:10.1016/J.PEDIATRNEUROL.2019.05.005

95. Dhillon S. Risdiplam: First Approval. *Drugs*. 2020;80(17):1853-1858. doi:10.1007/S40265-020-01410-Z/METRICS

96. Baranello G, Darras BT, Day JW, et al. Risdiplam in Type 1 Spinal Muscular Atrophy. *N Engl J Med*. 2021;384(10):915-923. doi:10.1056/nejmoa2009965

97. Tizzano EF, Finkel RS. Spinal muscular atrophy: A changing phenotype beyond the clinical trials. *Neuromuscul Disord*. 2017;27(10):883-889. doi:10.1016/j.nmd.2017.05.011

98. Behjati S, Tarpey PS. What is next generation sequencing? *Arch Dis Child Educ Pract Ed*. 2013;98(6):238. doi:10.1136/ARCHDISCHILD-2013-

304340

99. Franklin RE, Gosling RG. Molecular configuration in sodium thymonucleate. *Nature*. 1953;171(4356):740-741. doi:10.1038/171740A0

100. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci*. 1977;74(12):5463-5467. doi:10.1073/PNAS.74.12.5463

101. Maxam AM, Gilbert W. A new method for sequencing DNA (DNA chenistry/dimethyl sulfate cleavage/hydrazine/piperidine). *Proc Natl Acad Sci U S A*. 1977;74(2):560-564. Accessed February 12, 2023. https://www.pnas.org

102. Watson JD, Crick FHC. Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. J.D. Watson and F.H.C. Crick. Published in Nature, number 4356 April 25, 1953. *Nature*. 1974;248(5451):765. doi:10.1038/248765A0

103. Shendure J, Balasubramanian S, Church GM, et al. DNA sequencing at 40: Past, present and future. *Nature*. 2017;550(7676):345-353. doi:10.1038/NATURE24286

104. Hu T, Chitnis N, Monos D, Dinh A. Next-generation sequencing technologies: An overview. *Hum Immunol*. 2021;82(11):801-811. doi:10.1016/J.HUMIMM.2021.02.012

105. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature*. 2003;431:931-945. doi:https://doi.org/10.1038/nature03001

106. Margulies M, Egholm M, Altman WE, et al. Genome sequencing in microfabricated high-density picolitre reactors. *Nature*. 2005;437(7057):376-380. doi:10.1038/NATURE03959

107. Rothberg JM, Hinz W, Rearick TM, et al. An integrated semiconductor device enabling non-optical genome sequencing. *Nature*. 2011;475(7356):348-352. doi:10.1038/NATURE10242

108. Bentley D, Balasubramanian S, Swerdlow H, et al. Accurate whole human genome sequencing using reversible terminator chemistry. *Nat* . 2008;456:53-59. doi:10.1038/nature07517

109. Goodwin S, McPherson JD, McCombie WR. Coming of age: ten years of next-generation sequencing technologies. *Nat Rev Genet*.

2016;17(6):333-351. doi:10.1038/nrg.2016.49

110. Wiley GB, Kelly JA, Gaffney PM. Use of next-generation DNA sequencing to analyze genetic variants in rheumatic disease. *Arthritis Res Ther*. 2014;16(1):1-8. doi:10.1186/S13075-014-0490-4/TABLES/1

111. Koboldt DC, Steinberg KM, Larson DE, Wilson RK, Mardis ER. The Next-Generation Sequencing Revolution and Its Impact on Genomics. *Cell*. 2013;155(1):27-38. doi:10.1016/J.CELL.2013.09.006

112. Shendure J, Ji H. Next-generation DNA sequencing. *Nat Biotechnol*. 2008;26(10):1135-1145. doi:10.1038/NBT1486

113. Sikkema-Raddatz B, Johansson LF, de Boer EN, et al. Targeted Next-Generation Sequencing can Replace Sanger Sequencing in Clinical Diagnostics. *Hum Mutat*. 2013;34(7):1035-1042. doi:10.1002/HUMU.22332

114. Hartman P, Beckman K, Silverstein K, et al. Next generation sequencing for clinical diagnostics: Five year experience of an academic laboratory. *Mol Genet Metab Reports*. 2019;19. doi:10.1016/J.YMGMR.2019.100464

115. Suwinski P, Ong C, Ling MHT, Poh YM, Khan AM, Ong HS. Advancing Personalized Medicine Through the Application of Whole Exome Sequencing and Big Data Analytics. *Front Genet*. 2019;10(February):1-16. doi:10.3389/fgene.2019.00049

116. Martin KR, Zhou W, Bowman MJ, et al. ARTICLE The genomic landscape of tuberous sclerosis complex. *Nat Commun*. 2017;8:15816. doi:10.1038/ncomms15816

117. Kothare S V., Singh K, Chalifoux JR, et al. Severity of manifestations in tuberous sclerosis complex in relation to genotype. *Epilepsia*. 2014;55(7):1025-1029. doi:10.1111/EPI.12680/SUPINFO

118. Ogórek B, Hamieh L, Hulshof HM, et al. TSC2 pathogenic variants are predictive of severe clinical manifestations in TSC infants: results of the EPISTOP study. *Genet Med*. 2020;22(9):1489-1497. doi:10.1038/S41436-020-0823-4

119. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature*. 2020;581(7809):434-443. doi:10.1038/s41586-020-2308-7

120. Mone F, O'Connor C, Hamilton S, et al. Evolution of a prenatal genetic

clinic—A 10-year cohort study. *Prenat Diagn*. 2020;40(5):618-625. doi:10.1002/PD.5661

121. Togi S, Ura H, Hatanaka H, Niida Y. Genotype and Phenotype Landscape of 283 Japanese Patients with Tuberous Sclerosis Complex. *Int J Mol Sci*. 2022;23(19). doi:10.3390/IJMS231911175/S1

122. Schorling DC, Becker J, Pechmann A, Langer T, Wirth B, Kirschner J. Discrepancy in redetermination of SMN2 copy numbers in children with SMA. *Neurology*. 2019;93(6):267-269. doi:10.1212/WNL.0000000000007836

123. Kubo Y, Nishio H, Saito K. A new method for SMN1 and hybrid SMN gene analysis in spinal muscular atrophy using long-range PCR followed by sequencing. *J Hum Genet*. 2015;60(5):233-239. doi:10.1038/jhg.2015.16

# ANNEX

ARTICLE OPEN ACCESS

# Practical guidelines to manage discordant situations of *SMN2* copy number in patients with spinal muscular atrophy

Ivon Cuscó, PhD,* Sara Bernal, PhD,* Laura Blasco-Pérez, MSc, Maite Calucho, MSc, Laura Alias, PhD, Pablo Fuentes-Prior, PhD, and Eduardo F. Tizzano, MD, PhD

**Correspondence**
Dr. Tizzano
etizzano@vhebron.net

*Neurol Genet* 2020;6:e530. doi:10.1212/NXG.0000000000000530

## Abstract

### Objective

Assessment of *SMN2* copy number in patients with spinal muscular atrophy (SMA) is essential to establish careful genotype-phenotype correlations and predict disease evolution. This issue is becoming crucial in the present scenario of therapeutic advances with the perspective of SMA neonatal screening and early diagnosis to initiate treatment, as this value is critical to stratify patients for clinical trials and to define those eligible to receive medication. Several technical pitfalls and interindividual variations may account for reported discrepancies in the estimation of *SMN2* copy number and establishment of phenotype-genotype correlations.

### Methods

We propose a management guide based on a sequence of specified actions once *SMN2* copy number is determined for a given patient. Regardless of the method used to estimate the number of *SMN2* copies, our approach focuses on the manifestations of the patient to recommend how to proceed in each case.

### Results

We defined situations according to *SMN2* copy number in a presymptomatic scenario of screening, in which we predict the possible evolution, and when a symptomatic patient is genetically confirmed. Unexpected discordant cases include patients having a single *SMN2* copy but noncongenital disease forms, 2 *SMN2* copies compatible with type II or III SMA, and 3 or 4 copies of the gene showing more severe disease than expected.

### Conclusions

Our proposed guideline would help to systematically identify discordant SMA cases that warrant further genetic investigation. The *SMN2* gene, as the main modifier of SMA phenotype, deserves a more in-depth study to provide more accurate genotype-phenotype correlations.

*These authors are equal contributors with a major role in drafting and revising the manuscript and analyzing the data.

From the Medicine Genetics Group (I.C., L.B.-P., M.C., E.F.T.), Vall dHebron Research Institute (VHIR), Barcelona; Department of Clinical and Molecular Genetics (I.C., L.B.-P., M.C., E.F.T.), Hospital Vall dHebron, Barcelona; Department of Genetics (S.B., L.A.), Hospital de la Santa Creu i Sant Pau, Barcelona; Biomedical Research Institute Sant Pau (IIB Sant Pau) (S.B., L.A.), Hospital de la Santa Creu i Sant Pau, Barcelona; Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER-ISCIII, U-705 Barcelona) (S.B., L.A.), Madrid; Molecular Bases of Disease (P.F.-P.), Biomedical Research Institute Sant Pau (IIB Sant Pau), Hospital de la Santa Creu i Sant Pau, Barcelona, Spain.

Go to Neurology.org/NG for full disclosures. Funding information is provided at the end of the article.

The Article Processing Charge was funded by Fundació Hospital Universitari Vall d'Hebron Institut de Recerca (VHIR) from funds of grant no. FIS PI18/000687.

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND), which permits downloading and sharing the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Copyright © 2020 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of the American Academy of Neurology.

## Glossary

**FL-SMN** = full-length SMN; **MLPA** = multiplex ligation-dependent probe amplification; **NGS** = next-generation sequencing; **SMA** = spinal muscular atrophy; **SMN** = survival motor neuron; **SMN-del7** = *SMN2* transcripts lacking exon 7; **SNV** = single nucleotide variant.

Spinal muscular atrophy (SMA) is a neuromuscular disorder with a global incidence of approximately 1:11,000 live births and a worldwide carrier frequency of 1:51.[1] According to age at onset and achieved motor abilities, patients with SMA are usually classified into type I (never sit), II (never walk unaided), or III (achieve independent walking abilities). Independent of the clinical severity, all forms of SMA are caused by loss or homozygous loss-of-function pathogenic variants of the *SMN1* gene, located at 5q13.[2,3] The number of copies of *SMN2*, the highly homologous paralog of *SMN1*, is currently the most important modifier of disease phenotype; in most patients with SMA, this number varies between 1 and 5.[4] In fact, both *SMN1* and *SMN2* encode, in principle, the same survival motor neuron (SMN) protein. However, a single C→T transition in exon 7 disrupts an exon splicing enhancer and/or creates a splicing silencer, and as a consequence, *SMN2* works as a hypomorphic allele that produces mainly transcripts lacking exon 7 (SMN-del7).[5] The SMN-del7 protein is functionally compromised and unstable and therefore rapidly degraded by the ubiquitin-proteasome system.[6] Thus, the SMA phenotype is ultimately due to insufficient levels of full-length SMN (FL-SMN) protein.

On confirmation of biallelic deletion or pathogenic variants of the *SMN1* gene in a given patient, the number of *SMN2* copies is usually determined and reported. In previous years, this figure was mainly informative and mostly used to elaborate genotype-phenotype correlations rather than to predict a particular phenotype. However, recent advances in SMA therapeutics have strengthened the importance of estimating as accurately as possible the number of *SMN2* copies for all patients with SMA. Indeed, whereas genetic confirmation of SMA is relatively straightforward (95% of the patients can be diagnosed with a simple qualitative test), the assessment of *SMN2* copy number requires a quantitative methodology that is not easily implemented in most laboratories. Issues of DNA sample quality, calibration controls, and expertise to resolve ambiguous cases have been previously discussed.[4] Along these lines, around 40% of samples recently studied by the same methodology in different laboratories yielded discordant results.[7] Furthermore, intrinsic biological factors are also a source of discrepancies and add complexity to understanding how a specific *SMN2* genotype influences the final phenotype in a given patient.[4]

Numerous studies have shown that the higher the number of copies of *SMN2*, the larger the amount of FL-SMN protein produced, and thus the milder the associated SMA phenotype. However, this correlation is not absolute, and some patients with 2 copies of *SMN2* have mild SMA phenotypes, whereas some with 4 or more copies of the gene have been described as type I or II (reviewed in Calucho et al., 2018).[4] Thus, accurate estimation of *SMN2* copy number is essential in the present scenario of therapeutic advances with 3 specific SMA therapies already approved—nusinersen, *onasemnogene abeparvovec,* and risdiplam—and with the perspective of SMA neonatal screening and early diagnosis to initiate treatment.[8,9] We propose a practical guide for the management of discordant SMA cases based on systematic specified actions once *SMN2* copy number has been determined for a given patient. Our approach is independent of the method used to estimate *SMN2* copy number and focuses on the manifestations of the patient to decide how to proceed in each case.

## Methods

This guideline can be applied to the vast majority of genetically confirmed SMA cases with biallelic deletion of *SMN1* and to patients who may need further analysis (e.g., those with hybrid *SMN2-SMN1* genes or pathogenic *SMN1* variants). We base the current guideline on our previously published meta-analysis of SMA genotype-phenotype correlations and in our continued multidisciplinary experience with patients referred to our consultation, both national (Spain) and international.[4] Briefly, our approach considers the initial report of *SMN2* copy number for a given patient, which is in turn based on a quantitative analysis by multiplex ligation-dependent probe amplification (MLPA) using a mixture of specific probes for the SMA locus (P021-B SMA MLPA kit, a new version of the MLPA kit that includes probes for all exons of the *SMN* genes, in addition to introns 6 and 7).[10,11] However, our proposed guide can be applied to any report regardless of the method used for *SMN2* analysis. Starting with the estimated *SMN2* copy number reported, we then focus on the manifestations of the patient and how to proceed in case of an unexpected discordance. An unambiguous assignment of the SMA type by motor milestones criteria (0 "congenital," I "never sit," II "never walk," or III "walker") was initially widely established for simplicity. However, when necessary, these categories were further refined into subtypes Ia, b, and c, IIa and b, IIIa and b, and the milder type IV SMA and even with minimal manifestations, as previously defined.[9,12] Altogether, we distinguish up to 10 different clinical diagnostic categories to which genetically confirmed cases may be ascribed (table 1) to establish genotype-phenotype correlations and define possible discrepancies.

### Data availability

All data and scripts used to generate the analyses of this article are available on request unless the type of request compromises ethical standards or legal requirements.

**Table 1** Spinal muscular atrophy (SMA) major clinical diagnostic categories in genetically confirmed cases

| Clinical categories/ SMA type | Main clinical description |
|---|---|
| **PS** | Presymptomatic cases (identified at birth by newborn screening or previous affected sibling) |
| **0/Ia** | Congenital cases/patients with early manifestations within the first weeks of life |
| **Ib** | Patients with manifestations within first 3 mo of life |
| **Ic** | Children capable of head control, nonsitters |
| **IIa** | Sitters who are not able to stand up |
| **IIb** | Sitters who are able to stand up, but not to walk independently |
| **IIIa** | Onset before age 3 y, short-term walkers |
| **IIIb** | Onset after age 3 y, long-term walkers |
| **IV** | Walkers with weakness initiated in adult life |
| **MM** | Patients with only MMs (include also asymptomatics) |

Abbreviations: MM = minimal manifestation; PS = presymptomatic.
Based on references 4, 9, and 12.

# Results

We defined several discordant situations according to *SMN2* copy number in patients with a specific phenotype in 2 different scenarios: (1) presymptomatic diagnosis of a case detected in a newborn screening program or because of a previous SMA family history and (2) when a symptomatic patient is genetically confirmed. The spectrum of possible situations includes from 1 to 4 or more *SMN2* copies. A genetically confirmed neonate is considered presymptomatic based mainly on the absence of hypotonia, weakness, hypo- or areflexia, or fasciculations. Other manifestations may be more subtle and therefore not clearly noticeable.[9,13] In the second scenario, according to the patient's phenotype, different discrepancies are discussed. We defined recommendations according to the reported literature and our own experience, as follows.

## Guideline in a neonatal screening: asymptomatic context

The different situations that could be encountered when facing a presymptomatic patient, the number of *SMN2* copies, the predicted phenotypes and suggested actions in each situation, and their rationale are given in table 2. Patients with 1 *SMN2* copy usually present a congenital SMA form, and the discordance refers to their presenting without symptoms in the neonatal period. On the other hand, an apparently normal neonate should be expected to have at least 2 *SMN2* copies, and different predictions and actions are endorsed.

## Guidelines in a symptomatic context

The different situations of symptomatic patients, the number of *SMN2* copies, the observed phenotypes and the rationale, and actions suggested in each case are summarized in table 3. Unexpected discordant cases include patients having (1) a single *SMN2* copy but noncongenital disease forms (types Ib, II, or even III), (2) 2 *SMN2* copies with type II or III SMA, (3) 3 copies of the gene with severe disease forms (type Ia and b), and (4) at least 4 *SMN2* copies but more severe SMA (types I or II).

# Discussion

We have developed a practical guide for management and advice to help in the interpretation and resolution of discordant SMA cases according to the number of *SMN2* copies and phenotype. Our approach applies to virtually all genetically confirmed cases and is independent of the method used to determine *SMN2* copy number (table 4), but focuses instead on the manifestations of the patient. We suggest several recommendations to rapidly define the course of actions for a given SMA patient. *SMN2* copy number estimation is essential to establish accurate genotype-phenotype correlations, to predict disease evolution, to stratify patients for clinical trials, and to define those eligible for a given treatment. However, in some patients, this information may be insufficient to correlate with the observed phenotype. So far, the number of copies of the *SMN2* gene and the presence of rare *SMN2* variants (e.g., NM_017411.3:c.859G>C and NM_017411.3:c.835-44A>G) remain the major modifiers of SMA disease phenotype.[14–17]

The main characteristics of methods currently used to quantitate *SMN2* copies (TaqMan, LightCycler, MLPA, PCR-CE, and digital PCR) are given in table 4.[10,11,18–29] In a meta-analysis of 33 studies published from 1999 to 2017, in which *SMN2* copy number was reported for a total of 3,393 patients with SMA, MLPA was used in 54% of patients (n = 1870) followed by LightCycler in 21.4% (n = 741) and TaqMan in 6.5% (n = 228) and fewer patients with the remaining methodologies.[4,22,27] All these different methodologies have advantages and disadvantages, and there are technical aspects beyond the method itself that have to be considered such as DNA sample quality and interpretation and control issues. Digital PCR approaches[28] and novel protocols using next-generation sequencing (NGS) may help with the resolution of particularly difficult cases. Noteworthy, NGS methodologies allow a thorough analysis of *SMN2* copies at the genomic level including also introns and allowing a better investigation of the equivalency and quality of the *SMN2* copies. In addition, NGS provides valuable information that may be validated to establish more comprehensive genotype-phenotype correlations.[19–21]

A virtually asymptomatic neonate with a single *SMN2* copy is an obviously unexpected situation. As indicated in table 2, congenital type 0 cases have only 1 *SMN2* copy, which is insufficient to rescue the phenotype of the disease at the prenatal stage. In these patients, SMA manifests usually at

**Table 2** Suggested course of actions in SMA cases identified during newborn screening

| *SMN2* copy number | Manifestations at birth (clinical category) | Expected correlation | Rationale for recommended actions | Recommended actions and expected phenotype |
|---|---|---|---|---|
| **1** | Not observed (PS) | No[a] | Presence of only 1 *SMN2* copy is usually associated with congenital SMA. If a child is asymptomatic at birth and remains so for the first weeks of life, this would suggest an error in the previous *SMN2* quantitation or the presence of a positive modifier single nucleotide variant in the single gene copy. | Retest for *SMN2* copy number with a new sample and/or consider another method/laboratory. If the presence of a single *SMN2* copy is confirmed, test for rare positive variants associated with better-than-expected phenotypes (e.g., c.859G>C[b] and c.835-44A>G), e.g., by Sanger sequencing, or perform next-generation sequencing (NGS) analysis.[19–21] |
| **2** | Not observed (PS) | Yes | Neonates with 2 *SMN2* copies usually have a normal appearance. There is a latency period in which SMA symptoms may not be detectable. However, subtle manifestations of the disease might appear shortly after birth. | Test for rare positive variants in *SMN2* associated with better-than-expected phenotypes (e.g., c.859G>C[b] and c.835-44A>G). If negative, the patient has >90% probability of developing severe, type I SMA.[4] If positive, the patient will be virtually a sitter or walker later in their life[4,14] |
| **3** | Not observed (PS) | Yes | Neonates with 3 *SMN2* copies have a normal appearance and usually without manifestations at least for the first 3 mo of life. | Test for rare positive variants in *SMN2* associated with better-than-expected phenotypes (e.g., c.859G>C[b] and c.835-44A>G). If negative, the patient has about 60% probability of developing type II disease, 35% type III, and 5% type Ic. Similar Bayesian estimations can also be calculated.[4,25] If positive, the patient will be virtually a walker later in their life[4,14] |
| **≥4** | Not observed (PS) | Yes | Neonates with 4 *SMN2* copies have a normal appearance. About 14% of SMA cases worldwide have 4 *SMN2* copies.[4] | Retest for *SMN2* copy number with a new sample and/or consider another method/laboratory. If copy number is confirmed, the patient has >90% probability of being a walker later in their life (SMA types III or IV).[4] Test for the rare positive variants is an option. |

Abbreviations: PS = presymptomatic; SMA = spinal muscular atrophy.
[a] In this case, the expected category would be congenital (see text for further details and discussion). Percentages are calculated according to reference 4.
[b] A commercial test is available for the c.859G>C variant (table 4).

birth with at least marked hypotonia and weakness, but more commonly with a complex clinical picture that includes in addition respiratory problems, contractures, cardiac malformation, vascular necrosis,[30] and diffuse and progressive brain abnormalities.[31] If the patient does not manifest any of these symptoms, the most likely explanation is an erroneous determination of *SMN2* copy number, which should be excluded. Retesting with a new DNA sample, eventually using a different method or performing the analysis in a different laboratory, might solve the issue. However, if the presence of only 1 *SMN2* copy is confirmed, it is possible that single nucleotide variants (SNVs) of this single gene copy or a potential *SMN2-SMN1* hybrid structure[32] make it functionally superactive, i.e., capable of generating more full-length mRNA transcripts and FL-SMN protein than wild-type *SMN2* and thus to at least partly rescue the phenotype. Thus, testing for known positive variants such as NM_017411.3:c.859G>C[16] and NM_017411.3:c.835-44A>G[15] is recommended. If negative, it would be interesting to conduct an *SMN2* NGS study of the patient to unravel changes that may act as positive modifiers of disease severity. Along these lines, at least 10 SMA cases with 1 *SMN2* copy and type II or even III disease have been reported or personally communicated to date.[4,33] Unfortunately, these apparently discrepant cases have not been further studied, and it remains to be seen whether these phenotype-genotype discrepancies are due to technical or biological reasons.

Genetically confirmed SMA cases of newborns with 2 *SMN2* copies have a high probability (>90%) of developing type I disease, but they usually have a normal appearance at birth. There is a latency period—from 1 to several weeks—in which clear symptoms of weakness and hypotonia may not be detectable. However, subtle or less evident manifestations may appear early after birth such as hypo- or areflexia, weak cry, diaphragmatic breathing, feeding problems, and dysautonomic manifestations (i.e., increase of sweating and irregular skin responses to temperature changes).[9] On the other hand, exceptional cases with 2 *SMN2* copies may manifest overt disease at birth as usually occurs in type 0 cases.[34] Thus, and considering the continuous spectrum of phenotypes in SMA, it would be difficult to differentiate between congenital type 0 and type Ia disease, and both categories could be merged into type 0/Ia disease.[12]

To better predict the evolution of patients with 2 *SMN2* copies, it would be advisable to test for the presence of rare positive variants mentioned above. Indeed, in our experience, around 40% of cases with 2 *SMN2* copies and a milder phenotype (types II or III) may harbor one of these SNVs.[4,14]

**Table 3** Suggested course of actions in symptomatic SMA cases, for whom phenotypes and genotypes are not correlated (see text for further details and discussion)

| *SMN2* copy number | Observed manifestations/ milestones clinical category | Expected clinical category | Rationale for recommended actions | Recommended actions |
|---|---|---|---|---|
| **1** | Patients with type I, II, or III SMA | 0 | Patients with 1 *SMN2* copy usually present congenital SMA. Patients with typical type I, II, or even III disease forms might point to an error in the initial *SMN2* copy number determination or to the presence of a positive modifier in their single *SMN2* copy. | Retest for *SMN2* copy number with a new sample and/or consider another method/laboratory. If the presence of a single *SMN2* copy is confirmed, test for SNVs in the gene that have been previously associated with better-than-expected phenotypes (e.g., c.859G>C[a] and c.835-44A>G), e.g., by Sanger sequencing, or perform next-generation sequencing (NGS) analysis.[19–21] |
| **2** | Typical type II or type III patients | Ia, Ib, Ic | The vast majority of patients with 2 *SMN2* copies have typical type I disease. Exceptions are usually due to the presence of positive *SMN2* modifiers. | Retest for *SMN2* copy number with a new sample and/or consider another method/laboratory. If confirmed, test for rare variants in *SMN2* that have been previously associated with better-than-expected phenotypes (e.g., c.859G>C[a] and c.835-44A>G). If negative, perform NGS analysis to detect novel SNVs or other changes that could be positive modifiers of disease severity.[19–21] |
| **3** | Type I cases with disease onset before the age of 3 mo (Ia; Ib) | Ic, IIa, IIb, IIIa, IIIb | Type Ic cases usually manifest disease between 3 and 6 mo of life, and have 3 *SMN2* copies. A typical type Ib patient has 2 copies of the gene. Three copies are also detected in type II and III patients. | Retest for *SMN2* copy number with a new sample and/or consider another method/laboratory. If *SMN2* copy number is confirmed, perform further studies to identify SNVs or partial intragenic deletions that could act as negative phenotype modifiers (e.g., complete MLPA,[11] NGS[19–21]). |
| **≥4** | Non-walkers, either type I or II | IIIa, IIIb, IV, MM | Most reported cases of type I or II SMA patients with 4 *SMN2* copies are due to pitfalls in the quantitation of *SMN2* copy number. | Retest for *SMN2* copy number with a new sample and/or consider another method/laboratory. If *SMN2* copy number is confirmed, perform further studies to identify SNVs or partial intragenic deletions that could act as negative phenotype modifiers (e.g., complete MLPA,[11] NGS[19–21]). |

Abbreviations: MLPA = multiplex ligation-dependent probe amplification; MM = minimal manifestation; NGS = next-generation sequencing; SMA = spinal muscular atrophy; SNV = single nucleotide variant.
[a] A commercial test is available for the c.859G>C variant (table 4).

Negative variants in *SMN2* have not been discovered, but warrant further investigation.

In patients with 3 *SMN2* copies, our previous meta-analysis revealed that about 60% of cases develop type II disease, 35% type III, but 5% still had the more severe type Ic SMA.[4] Therefore, all neonates with 3 gene copies would be expected to have a normal appearance and to remain essentially asymptomatic at least for the first 3 months of life. The NURTURE study of presymptomatic patients with 2 or 3 *SMN2* copies treated with nusinersen has shown that patients with 3 gene copies treated in the neonatal period have in general a better evolution.[13] Again, here it is advisable to check for rare positive variants to better predict the expected outcomes.

The treatment recommendations for presymptomatic cases with 4 *SMN2* copies are still an evolving issue.[8,35,36] Based on available evidence, and in the absence of a reliable biomarker of disease evolution, in the United States, it has been recently recommended to initiate treatment of all infants with 4 copies of *SMN2*.[35] In our meta-analysis of 3,393 cases, patients with 4 copies accounted for less than 14% of all reported SMA cases.[4] In the light of this finding, it is rather surprising that in a recent pilot newborn screening study, 15 of 37 detected cases (40%) had 4 *SMN2* copies.[36] Excluding technical issues with *SMN2* quantitation, if these results are reproduced in other newborn screening studies, it would be tempting to speculate that a certain number of individuals in the general population with 0,4 genotype (i.e., no *SMN1* gene but 4 *SMN2* copies) remain with minimal symptoms or asymptomatic throughout their lives and thus undetected. Preliminary results of the SMA newborn screening program in Australia reported 9 positive cases, but none had 4 *SMN2* copies.[37] It is important to highlight that copy number studies in positive patients detected by newborn screening should be performed in expertise centers and with a validated methodology. In the shared decision to immediately start treatment of neonates with 4 *SMN2* copies or delay the initiation of treatment, several alternatives—each with advantages and disadvantages—have to be considered (outlined in table 5). Whatever decision is taken, it is important to recall that disease onset in these patients before the first year of life is rather unlikely, giving the health care team and the parents more time to weigh advantages and disadvantages of each

**Table 4** Major features of the more commonly used methods to determine *SMN2* copy number

| Method | Main characteristics | Advantages | Disadvantages |
|---|---|---|---|
| **Real-time PCR (TaqMan platform)**[22,23] | Multiplex TaqMan real-time quantitative PCR assay. | Fast, robust, and sensitive technique. Low cost. Requires only small amounts of DNA. Easy interpretation of results by automated variant reporting software. DNA quality important, but not as limiting as in other techniques. Neither a standard curve nor control samples are necessary. | Assays are performed in triplicate and an internal control is necessary to normalize results, as the method is based on relative quantitation. Only the number of copies of *SMN2* exon 7 is determined. Real-time PCR technology is not always available in a routine laboratory. |
| **Real-time PCR (LightCycler platform)**[25,26] | Quantitative assay on the basis of real-time PCR, performed with a LightCycler instrument (Roche Diagnostics, Basel, Switzerland) by using the fluorescence resonance energy transfer technique (PCR products are based on the use of SYBR Green). | Low cost. Easy interpretation of results by automated variant reporting software. | As a relative quantitative assay, it requires standard curves to normalize results. DNA quality is limiting, large amounts of DNA are necessary, and the process is laborious. Only the number of copies of *SMN2* exon 7 is determined. |
| **Multiple ligation-dependent probe amplification, version B1**[10,11,27] | Variation of the multiplex PCR assay that permits amplification of multiple target genes with a single primer pair. Quantitates gene doses. | Low cost. Robust and sensitive. Requires only small amounts of DNA and its quality is not as limiting as in other techniques. Does not require sophisticated logistic (a therrmocycler and a vertical electrophoresis sequencer). Easy interpretation of results by automated variant reporting, free software (Coffalyser). *SMN1* and *SMN2* exons analyzed in the same experiment, detection of exonic hybrid genes, and partial intragenic deletions. False positive deletions may result from mutations located in regions of hybridization of the probes | As a relative quantitative assay, normalization of results is necessary, for which control samples are required. Long assay time (24 h minimum). |
| **AmplideX PCR-CE *SMN1/SMN2* (asuragen.com)** | Quantitative method based on a multiplex PCR and separation by capillary electrophoresis to calculate copy number of exon 7 in *SMN1* and *SMN2* genes, using specific fluorescently labeled primers. Quantitation is based on the peak area ratio of the target gene to an endogenous control, normalized to a calibration sample. | Robust and sensitive technique. Requires only small amounts of DNA. Simple and rapid workflow (<4 h). Does not require sophisticated technology or structural logistics (a thermocycler and a vertical electrophoresis sequencer). Easy interpretation of results by automated variant reporting software. Scalable design that allows the study of additional variants, such as hybrid genes, silent carriers, and 1 *SMN2* modifier variant (PLUS KIT). | As a relative quantitative assay, normalization is necessary, which requires calibration samples. Analysis limited to exon 7 of *SMN1* and *SMN2* genes, therefore unable to detect partial deletions. |
| **Droplet digital PCR (ddPCR)**[18,28,29] | DNA is partitioned into thousands of droplets that are subsequently amplified, and fluorescently labeled probe signals within each droplet are recorded as either positive or negative, depending on the presence or absence of a nucleotide target. | Eliminates the need for standard curves by using references or endogenous controls. As an absolute quantitation method of exons 7 and 8 in *SMN1* or *SMN2*, there is no need for normalization. Requires extremely low DNA concentrations (e.g., from dried blood spots). | The special technology necessary to perform ddPCR is not available in most laboratories. High costs limit determination of the copy number of all *SMN2* exons. Therefore, partial intragenic *SMN2* deletions are not detected. |
| **Next-generation sequencing (NGS)**[19–21] | Non–Sanger-based high-throughput DNA sequencing technologies (several platforms are currently available). | Allows analysis of complete genes (exons and introns) and to detect rearrangements and point mutations. | Not available in several diagnostic laboratories. Laborious process and longer assay times. Interpretation of results requires specialized bioinformatics tools and usually a bioinformatician. Quantitative studies using NGS are not very robust when *SMN1* and *SMN2* genes coexist due to their extremely high homology. |

therapeutic alternative. Some parents may want to move forward without further testing, but it is crucial that an expert team adequately communicates about the disease and manages their expectations.[8] The implementation of neonatal screening in different regions will help to better define protocols of follow-up and validate biomarkers of disease progression such as levels of plasma phosphorylated neurofilament heavy chain in these patients.[38]

Different approaches should be considered when dealing with symptomatic cases. Here, most of the discrepancies should be initially faced with a retesting of the patient with a new sample, a different method or even in a second laboratory. According to the results of this second test, it might be advisable to continue testing for known variants in *SMN2*. A recent SMA test that includes testing for the NM_017411.3:c.859G>C variant has been made commercially available (table 4)

**Table 5** Factors to be considered when deciding to treat neonates with genetically confirmed spinal muscular atrophy with 4 *SMN2* copies

| Factors to consider | Treat presymptomatically | Treat when symptoms appear |
|---|---|---|
| **Disease appearance and complications** | Avoid possible long-term disease complications | Risk of long-term disease complications |
| **Opportunity of treatment** | Some patients might be unnecessarily treated for a long time | Depends on rescue of disease manifestations |
| **Time to initiate therapy** | Predicted large therapeutic window | Therapeutic window might be too short or lost |
| **Adverse events when continuous therapy** | Risk of treatment complications | Reduced risk of treatment complications |
| **Economic aspects** | Higher cost of therapy | Higher cost of managing morbidity |
| **Quality of life (QoL) issues** | Effect of years of treatment on QoL | Effect of disease on QoL |
| **Parent and family expectations** | Unease of treating a healthy baby | Increased stress during follow-up waiting for the imminent onset of manifestations |

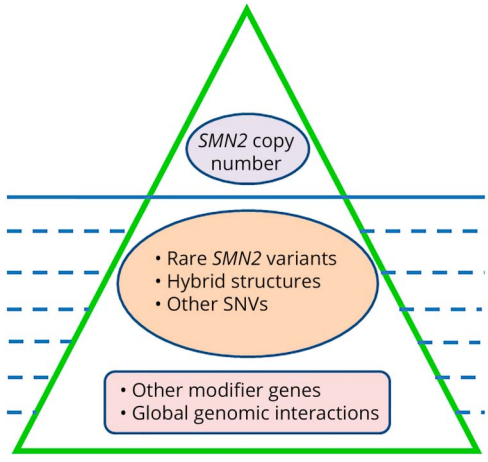
(asuragen.com). Furthermore, a new version of the SMA MLPA kit including all exons of the *SMN1* and *SMN2* genes has been reported. This new version of the kit would allow detection of some intragenic or 5′ terminal deletions that were previously extremely difficult to detect.[11] However, not all cases might be resolved with an accurate *SMN2* copy number assessment or checking for known variants by Sanger sequencing. If the results of all these studies are not categorical, *SMN2* NGS studies should be considered to determine whether the *SMN2* copies are functionally identical (table 4).[19–21]

Certainly, the *SMN2* gene, as the main modifier of SMA phenotype, deserves a more in-depth study beyond the current standard copy number determination. We believe that in terms of its impact on SMA phenotype, *SMN2* copy number might be considered as the tip of an iceberg of which other genetic and epigenetic features, most notably SNVs, represent the submerged part with relevant effects to phenotype of the patients with SMA (figure). A number of other genes have been proposed as candidate modifiers of the SMA phenotype including methylation status of *SMN2* (reviewed in Maretina et al., 2018),[39] although none of them are yet validated in clinical practice. Given that *SMN2* variants modify the disease phenotype and that transcripts derived from *SMN2* are targets for splicing modifiers in the therapeutic scenario, it is essential to gain a thorough insight into the complete *SMN2* sequences of discordant patients. Furthermore, we need to unveil possible linkages between specific *SMN2* variants and factors involved in *SMN2* splicing, on the one hand, and responses to treatment, on the other hand. In patients receiving expensive treatments, their efficacy should be periodically assessed to decide whether to continue treatment or to look for alternatives. Responses to treatment may vary in patients with SMA (from responders to slow responders to nonresponders),[40] but it is currently unknown whether specific features of their *SMN2* genes are directly correlated with these responses. Discovery and validation of positive and negative genetic markers remain thus an urgent matter in SMA research. New SMA classifications may need to be adopted in line with the current scenario of early genetic diagnosis, therapeutic intervention, and evolving phenotypes.[41] In this context, time to development of different manifestations and age at treatment initiation are becoming crucial as predictors of the trajectory of the disease.[9,42] In this envisaged perspective, a better and clearer definition of the *SMN2* genotype (copies and sequence) in each patient would be extremely relevant. Along these lines, our proposed guideline would help to systematically and rigorously identify discordant SMA cases that warrant further genetic investigation.

## Acknowledgment

The authors thank all colleagues for their confidence, information, and reference of patients for study.

**Figure** The iceberg representation of the genetic factors that influence SMA phenotype




*SMN2* copy number might be considered as the tip of an iceberg of which other *SMN2* genetic and epigenetic features, most notably *SMN2* SNVs, represent the submerged part. Other modifier genes and whole genomic data may complete possible influences. SMA = spinal muscular atrophy; SNV = single nucleotide variant.

## Study funding

This work was partially supported by grants from Fundación Daniel Bravo Andreu (to E.F.T. and P.F.-P. and supporting M.C.), SMA Europe (to E.F.T. and P.F.-P.), Biogen, and Spanish Instituto de Salud Carlos III, Fondo de Investigaciones Sanitarias and cofunded with ERDF funds (grant no. FIS PI18/000687) to E.F.T. and L.A.

## Disclosure

E.F. Tizzano discloses grant support to conduct CTs on SMA from Ionis/Biogen and serves as a consultant to AveXis, Novartis, Biogen, Biologix, Cytokinetics, and Roche. The other authors report no disclosures relevant to the manuscript. Go to Neurology.org/NG for full disclosures.

## Publication history

Received by *Neurology: Genetics* June 23, 2020. Accepted in final form September 29, 2020.

## Appendix Authors

| Name | Location | Contribution |
|---|---|---|
| **Ivon Cuscó, PhD** | Medicine Genetics Group, Vall d'Hebron Research Institute (VHIR); Department of Clinical and Molecular Genetics, Hospital Vall dHebron, Barcelona, Spain | Drafting/revision of the manuscript for content, including medical writing for content, and major role in the acquisition of data |
| **Sara Bernal, PhD** | Department of Genetics, Biomedical Research Institute Sant Pau (IIB Sant Pau), Hospital de la Santa Creu i Sant Pau, Barcelona; Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER-ISCIII, U-705 Barcelona), Madrid, Spain | Drafting/revision of the manuscript for content, including medical writing for content, and major role in the acquisition of data |
| **Laura Blasco-Pérez, MSc** | Medicine Genetics Group, Vall dHebron Research Institute (VHIR); Department of Clinical and Molecular Genetics, Hospital Vall dHebron, Barcelona, Spain | Drafting/revision of the manuscript for content, including medical writing for content |
| **Maite Calucho, MSc** | Medicine Genetics Group, Vall dHebron Research Institute (VHIR) 08035 Barcelona, Spain. Department of Clinical and Molecular Genetics, Hospital Vall dHebron, 08035 Barcelona, Spain | Drafting/revision of the manuscript for content, including medical writing for content |
| **Laura Alias, PhD** | Department of Genetics, Biomedical Research Institute Sant Pau (IIB Sant Pau), Hospital de la Santa Creu i Sant Pau, Barcelona; Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER-ISCIII, U-705 Barcelona), Madrid, Spain | Analysis or interpretation of data |
| **Pablo Fuentes-Prior, PhD** | Molecular Bases of Disease, Biomedical Research Institute Sant Pau (IIB Sant Pau), Hospital de la Santa Creu i Sant Pau, Barcelona, Spain | Drafting/revision of the manuscript for content, including medical writing for content, and major role in the acquisition of data |
| **Eduardo F. Tizzano, MD, PhD** | Medicine Genetics Group, Vall dHebron Research Institute (VHIR); Department of Clinical and Molecular Genetics, Hospital Vall dHebron, Barcelona, Spain | Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; and study concept or design |

## References

1. Sugarman EA, Nagan N, Zhu H, et al. Pan-ethnic carrier screening and prenatal diagnosis for spinal muscular atrophy: clinical laboratory analysis of >72,400 specimens. Eur J Hum Genet 2012;20:27–32.
2. Wirth B. An update of the mutation spectrum of the survival motor neuron gene (SMN1) in autosomal recessive spinal muscular atrophy (SMA). Hum Mutat 2000; 15:228–237.
3. Alías L, Bernal S, Fuentes-Prior P, et al. Mutation update of spinal muscular atrophy in Spain: molecular characterization of 745 unrelated patients and identification of four novel mutations in the SMN1 gene. Hum Genet 2009;125:29–39.
4. Calucho M, Bernal S, Alías L, et al. Correlation between SMA type and SMN2 copy number revisited: an analysis of 625 unrelated Spanish patients and a compilation of 2834 reported cases. Neuromuscul Disord 2018;28:208–215.
5. Monani UR, Lorson CL, Parsons DW, et al. A single nucleotide difference that alters splicing patterns distinguishes the SMA gene SMN1 from the copy gene SMN2. Hum Mol Genet 1999;8:1177–1183.
6. Vitte J, Fassier C, Tiziano FD, et al. Refined characterization of the expression and stability of the SMN gene products. Am J Pathol 2007;171:1269–1280.
7. Schorling DC, Becker J, Pechmann A, Langer T, Wirth B, Kirschner J. Discrepancy in redetermination of SMN2 copy numbers in children with SMA [online]. Neurology 2019;93:267–269.
8. Dangouloff T, Burghes A, Tizzano EF, et al. 244th ENMC international workshop: newborn screening in spinal muscular atrophy May 10–12, 2019, Hoofddorp, the Netherlands. Neuromuscul Disord 2020;30:93–103.
9. Tizzano EF. Treating neonatal spinal muscular atrophy: a 21st century success story? Early Hum Dev 2019;138:104851.
10. Alías L, Bernal S, Barceló MJ, et al. Accuracy of marker analysis, quantitative real-time polymerase chain reaction, and multiple ligation-dependent probe amplification to determine SMN2 copy number in patients with spinal muscular atrophy. Genet Test Mol Biomarkers 2011;15:587–594.
11. Vijzelaar R, Snetselaar R, Clausen M, et al. The frequency of SMN gene variants lacking exon 7 and 8 is highly population dependent. PLoS One 2019;14:e0220211.
12. Talbot K, Tizzano EF. The clinical landscape for SMA in a new therapeutic era. Gene Ther 2017;24:529–533.
13. De Vivo DC, Bertini E, Swoboda KJ, et al. Nusinersen initiated in infants during the presymptomatic stage of spinal muscular atrophy: interim efficacy and safety results from the Phase 2 NURTURE study. Neuromuscul Disord 2019;29:842–856.
14. Bernal S, Alías L, Barceló MJ, et al. The c.859G>C variant in the SMN2 gene is associated with types II and III SMA and originates from a common ancestor. J Med Genet 2010;47:640.
15. Wu X, Wang SH, Sun J, Krainer AR, Hua Y, Prior TW. A-44G transition in SMN2 intron 6 protects patients with spinal muscular atrophy. Hum Mol Genet 2017;26:2768–2780.
16. Prior TW, Krainer AR, Hua Y, et al. A positive modifier of spinal muscular atrophy in the SMN2 gene. Am J Hum Genet 2009;85:408–413.
17. Vezain M, Saugier-Veber P, Goina E, et al. A rare SMN2 variant in a previously unrecognized composite splicing regulatory element induces exon 7 inclusion and reduces the clinical severity of spinal muscular atrophy. Hum Mutat 2010;31: E1110–E1125.
18. Vidal-Folch N, Gavrilov D, Raymond K, et al. Multiplex droplet digital PCR method applicable to newborn screening, carrier status, and assessment of spinal muscular atrophy. Clin Chem 2018;64:1753–1761.
19. Chen X, Sanchis-Juan A, French CE, et al. Spinal muscular atrophy diagnosis and carrier screening from genome sequencing data. Genet Med 2020;22:945–953.
20. Ruhno C, McGovern VL, Avenarius MR, et al. Complete sequencing of the SMN2 gene in SMA patients detects SMN gene deletion junctions and variants in SMN2 that modify the SMA phenotype. Hum Genet 2019;138:241–256.
21. Blasco-Pérez L, Paramonov I, Iranzo L, et al. SMN2 genes study by NGS in spinal muscular atrophy:towards identification of new SMA modifiers. Eur J Hum Genet ESHG Congr 2020. Interactive e-poster session P10.53.C.

22. Anhuf D, Eggermann T, Rudnik-Schöneborn S, Zerres K. Determination of SMN1 and SMN2 copy number using TaqMan technology. Hum Mutat 2003;22:74–78.
23. Gómez-Curet I, Robinson KG, Funanage VL, Crawford TO, Scavina M, Wang W. Robust quantification of the SMN gene copy number by real-time TaqMan PCR. Neurogenetics 2007;8:271–278.
24. Kraszewski JN, Kay DM, Stevens CF, et al. Pilot study of population-based newborn screening for spinal muscular atrophy in New York state. Genet Med 2018;20: 608–613.
25. Feldkötter M, Schwarzer V, Wirth R, Wienker TF, Wirth B. Quantitative analyses of SMN1 and SMN2 based on real-time lightCycler PCR: fast and highly reliable carrier testing and prediction of severity of spinal muscular atrophy. Am J Hum Genet 2002; 70:358–368.
26. Cuscó I, Barceló MJ, Rojas-García R, et al. SMN2 copy number predicts acute or chronic spinal muscular atrophy but does not account for intrafamilial variability in siblings. J Neurol 2006;253:21–25.
27. Arkblad EL, Darin N, Berg K, et al. Multiplex ligation-dependent probe amplification improves diagnostics in spinal muscular atrophy. Neuromuscul Disord 2006;16: 830–838.
28. Stabley DL, Harris AW, Holbrook J, et al. SMN1 and SMN2 copy numbers in cell lines derived from patients with spinal muscular atrophy as measured by array digital PCR. Mol Genet Genomic Med 2015;3:248–257.
29. Zhong Q, Bhattacharya S, Kotsopoulos S, et al. Multiplex digital PCR: breaking the one target per color barrier of quantitative PCR. Lab Chip 2011;11: 2167–2174.
30. Grotto S, Cuisset J-M, Marret S, et al. Type 0 spinal muscular atrophy: further delineation of prenatal and postnatal features in 16 patients. J Neuromuscul Dis 2016; 3:487–495.
31. Mendonça RH, Rocha AJ, Lozano-Arango A, et al. Severe brain involvement in 5q spinal muscular atrophy type 0. Ann Neurol 2019;86:458–462.
32. Cuscó I, Barceló MJ, Del Rio E, et al. Characterisation of SMN hybrid genes in Spanish SMA patients: de novo, homozygous and compound heterozygous cases. Hum Genet 2001;108:222–229.
33. Sun Y, Grimmler M, Schwarzer V, Schoenen F, Fischer U, Wirth B. Molecular and functional analysis of intragenic SMN1 mutations in patients with spinal muscular atrophy. Hum Mutat 2005;25:64–71.
34. Kitaoka H, Shitara Y, Uchida Y, Kondo U, Omori I. Case of spinal muscular atrophy type 0 with mild prognosis. Pediatr Int 2020;62:106–107.
35. Glascock J, Sampson J, Connolly AM, et al. Revised recommendations for the treatment of infants diagnosed with spinal muscular atrophy via newborn screening who have 4 copies of SMN2. J Neuromuscul Dis 2020;7:97–100.
36. Müller-Felber W, Vill K, Schwartz O, et al. Infants diagnosed with spinal muscular atrophy and 4 SMN2 copies through newborn screening: opportunity or burden? J Neuromuscul Dis 2020;7:109–117.
37. Kariyawasam DST, Russell JS, Wiley V, Alexander IE, Farrar MA. The implementation of newborn screening for spinal muscular atrophy: the Australian experience. Genet Med 2020;22:557–565.
38. Darras BT, Crawford TO, Finkel RS, et al. Neurofilament as a potential biomarker for spinal muscular atrophy. Ann Clin Transl Neurol 2019;6:932–944.
39. Maretina MA, Zheleznyakova GY, Lanko KM, Egorova AA, Baranov VS, Kiselev AV. Molecular factors involved in spinal muscular atrophy pathways as possible disease-modifying candidates. Curr Genomics 2018;19:339–355.
40. Finkel RS, Mercuri E, Darras BT, et al. Nusinersen versus sham control in infantile-onset spinal muscular atrophy. N Engl J Med 2017;377:1723–1732.
41. Tizzano EF, Finkel RS. Spinal muscular atrophy: a changing phenotype beyond the clinical trials. Neuromuscul Disord 2017;27:883–889.
42. Dangouloff T, Servais L. Clinical evidence supporting early treatment of patients with spinal muscular atrophy: current perspectives. Ther Clin Risk Manag 2019;15: 1153–1161.

International Journal of
*Molecular Sciences*

MDPI

*Review*

# The Importance of Digging into the Genetics of *SMN* Genes in the Therapeutic Scenario of Spinal Muscular Atrophy

**Mar Costa-Roger [1,2], Laura Blasco-Pérez [1,2], Ivon Cuscó [1,2,3] and Eduardo F. Tizzano [1,2,*]**

[1] Medicine Genetics Group, Vall d'Hebron Research Institute (VHIR), 08035 Barcelona, Spain; mar.costa@vhir.org (M.C.-R.); laura.blasco@vhir.org (L.B.-P.); icusco@vhebron.net (I.C.)
[2] Department of Clinical and Molecular Genetics, Hospital Vall d'Hebron, 08035 Barcelona, Spain
[3] Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), 08035 Barcelona, Spain
* Correspondence: etizzano@vhebron.net; Tel.: +34-93274600 (ext. 6945)

**Abstract:** After 26 years of discovery of the determinant survival motor neuron 1 and the modifier survival motor neuron 2 genes (*SMN1* and *SMN2*, respectively), three SMN-dependent specific therapies are already approved by FDA and EMA and, as a consequence, worldwide SMA patients are currently under clinical investigation and treatment. Bi-allelic pathogenic variants (mostly deletions) in *SMN1* should be detected in SMA patients to confirm the disease. Determination of *SMN2* copy number has been historically employed to correlate with the phenotype, predict disease evolution, stratify patients for clinical trials and to define those eligible for treatment. In view that discordant genotype-phenotype correlations are present in SMA, besides technical issues with detection of *SMN2* copy number, we have hypothesized that copy number determination is only the tip of the iceberg and that more deepen studies of variants, sequencing and structures of the *SMN2* genes are necessary for a better understanding of the disease as well as to investigate possible influences in treatment responses. Here, we highlight the importance of a comprehensive approach of *SMN1* and *SMN2* genetics with the perspective to apply for better prediction of SMA in positive neonatal screening cases and early diagnosis to start treatments.

**Keywords:** spinal muscular atrophy; survival motor neuron 1; survival motor neuron 2; genotype-phenotype correlations; variants; hybrid structure

check for updates

**Citation:** Costa-Roger, M.; Blasco-Pérez, L.; Cuscó, I.; Tizzano, E.F. The Importance of Digging into the Genetics of *SMN* Genes in the Therapeutic Scenario of Spinal Muscular Atrophy. *Int. J. Mol. Sci.* **2021**, *22*, 9029. https://doi.org/10.3390/ijms22169029

Academic Editor: Merlin G. Butler

Received: 28 July 2021
Accepted: 19 August 2021
Published: 21 August 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

## 1. SMA Is a Disease of Two Genes, a Determinant *SMN1* and a Modifier *SMN2*

Spinal muscular atrophy (SMA) is a severe neuromuscular disease characterized by progressive proximal muscle weakness and atrophy as a result of alpha neuron degeneration and irreversible loss in the spinal cord anterior horn [1]. Overall estimated incidence is 1 in 11,000 live births with a carrier frequency around 1/54 [2]. Despite SMA clinically manifests as a continuum, based on age of onset, achieved motor milestones and clinical severity, SMA patients are divided into type 0-IV ranging from very severe congenital forms with short life expectancy due to respiratory failure at birth to adult-onset patients maintaining the ability to walk [3,4].

At the molecular level, SMA is an autosomal recessive disorder caused by loss of survival motor neuron 1 (*SMN1*, OMIM #600354) gene in the 5q13 locus [5]. Around 95% of cases are explained by homozygous deletion or gene conversion, while a minority of patients are compound heterozygous including intragenic pathogenic variants and deletion of one *SMN1* allele [6].

Besides *SMN1* gene there is *SMN2* (OMIM #601627), an almost equal centromeric paralog gene. Both *SMN* genes have an identical genomic organization consisting of nine exons interrupted by eight introns, which fits with the recent duplication of the *SMN1* gene, explaining why *SMN2* is only present in humans [5,7,8]. The exact structure of the SMA region remains unclear and initially *SMN1* and *SMN2* were described to be in opposite directions (head-to-head) [5], but more recently, evidence supports that the two genes are

oriented in the same direction [9]. More studies are needed to determine which structure is correct or if both orientations are present in the population.

Given that *SMN2* was originated from *SMN1*, their sequences only differ in 16 paralogous sequence variants (PSVs), which represent a total of 20 different nucleotides between both genes (15 SNVs and 1 indel) [10] as represented in Figure 1. The PSV c.840C>T, located in exon 7, causes exon skipping in the majority of *SMN2* pre-mRNA transcripts resulting in a truncated, nonfunctional and rapidly degraded protein that is not able to oligomerize (SMN-Δ7) [11] and can only produce the complete functional protein in 10–15% of cases [12–14]. Conversely, *SMN1* gene produces virtually full-length mRNA transcripts encoding the normal SMN protein. Bi-allelic alteration of *SMN1* is the rule to confirm SMA, however *SMN2* copies varying from 1 to 5 are present in all patients, as absence of both genes has never been reported in humans. An inverse correlation between *SMN2* copy number (*SMN2*_CN) and disease severity is currently accepted, being the number of *SMN2* copies the main modifier of the SMA phenotype (see Section 2) [5,15].

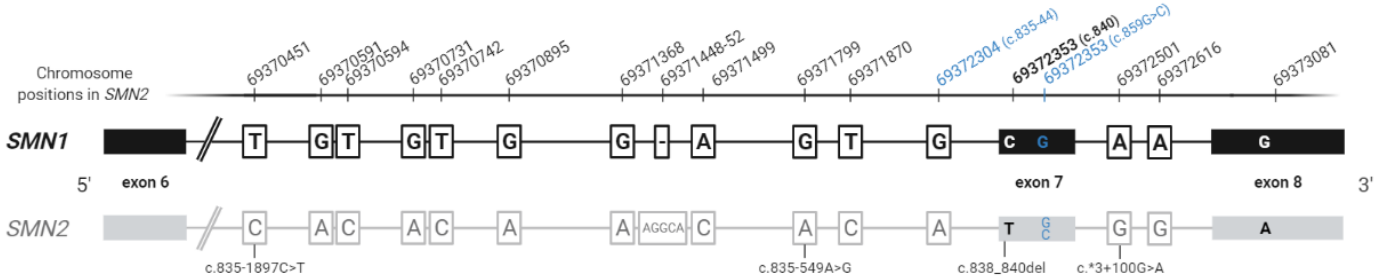



**Figure 1.** Schematic view of paralogous sequence variants (PSVs) between *SMN1* and *SMN2*. PSVs are represented in black in *SMN1* and in grey in *SMN2* and validated positive modifiers (c.835-44A>G and c.859G>C) in blue. The PSV that changes the splicing pattern of *SMN2* (c.840C>T) is highlighted in bold. Putative *SMN2* modifiers are indicated below the *SMN2* scheme. Note that c.859G>C is not a PSV, but a rare variant present only in *SMN2* gene. Chromosome positions in *SMN2* refer to hg19. Based on reference [10].

The SMA genomic region is highly polymorphic and dynamic, which is prone to unequal rearrangements leading to deletions, duplications or gene conversions [7]. In fact, the presence of *SMN1* or *SMN2* genes lacking exons 7 and 8 (*SMN1*/2Δ7-8) has been reported in the general population, with the breakpoint described in intron 6 [16,17]. The frequency of this variant varies greatly among populations including 15–21% in non-Finish Europeans, 7–11.5% in Americans and Finnish European individuals and 0.3–3% in Asian and African populations [17,18]. Several studies observed a strong inverse correlation between this partial deletion and *SMN2*_CN suggesting that the *SMN1*/2Δ7-8 variant is mainly derived from *SMN2* deletion events [17,19].

Apart from partial deletions of *SMN* genes, other structural variants have been characterized such as hybrid *SMN1-SMN2* genes. Around 5–10% of SMA patients show homozygous deletions of exon 7, but not of exon 8 of *SMN1* explained by the presence of hybrid genes [6,20,21]. This phenomenon could result from intrachromosomal deletions or more likely from gene conversion events in which part of the *SMN2* gene is fused to *SMN1* [20,22]. Although there is still some debate, many studies found that hybrid *SMN* genes appeared to be associated with a milder phenotype, mainly present in SMA type II and III patients [20,21,23].

Copy number of *SMN1* gene have also been described in the general population with differences across various ethnicities, with a higher average of *SMN1* copies in African American population [24]. In fact, 54.7% of Africans carry three or more *SMN1* copies according to a recent study [18]. In this context, the frequency of silent 2/0 carriers (individuals with two *SMN1* copies in cis) is also higher as it is directly related to the frequencies of *SMN1* deletions and duplications [25,26]. The detection of 2/0 carriers is challenging given the difficulty to differentiate them from 1/1 non-carriers. Interestingly, two *SMN1* variants have been associated with silent carriers in the Ashkenazi Jewish

population, including c.*3+80T>G and c.*211_*212del, which ultimately modify the SMA carrier risk being useful to detect around 20% of these special carriers [25,27]. However, it must be taken into account that there are several 2/0 cases without these variants [27].

Therefore, an accurately deep characterization of the SMA region is relevant not only for the detection of *SMN1* and *SMN2* copy number, but also for the different structural variants described. There are complex biological features of the SMN region that hinder the analysis of these genes, including the high homology between both genes, the multiple *SMN2* copies, the presence of partial deletions and hybrid structures and the effect of unknown intronic variants (Table 1). Technical limitations may include difficulties to ascribe a variant to *SMN1* or *SMN2*, quantitation of *SMN2* may not be always straightforward because of sample or methodological problems and structural changes are not usually detected with routine methods.

**Table 1.** Complex characteristics of *SMN* genomic region and associated technical limitations. Left column shows the biological issues and the right column the technical difficulties associated with the analysis of all these complex issues.

| Complex Biological Features of SMN Region | Technical Limitations |
| --- | --- |
| High homology between *SMN1* and *SMN2* | Difficulty in establishing if specific variants belong to *SMN1* or *SMN2* |
| Multiple copies of *SMN* genes | Inaccurate copy number determination |
| Partial deletions | |
| *SMN2/SMN1* hybrid structures | Undetectable by routine analysis (exonic sequencing and MLPA) |
| Unknown variants in deep intronic regions | |

## 2. The Known Validated Genotypes

Determination of *SMN2*_CN is a useful prognostic tool in order to establish accurate genotype-phenotype correlations, predict disease course and determine appropriate SMA patients for treatment [15]. Calucho et al. (2018) compiled a total of 3459 SMA patients and established quantitative *SMN2* correlations to predict disease evolution. Concretely, the higher number of *SMN2* copies, the milder the SMA phenotype, as most patients comply with the following rule: SMA type I patients had 2 *SMN2* genes, type II had 3 *SMN2*, type III had 3 or 4 *SMN2* copies and type IV patients had 4 *SMN2* genes [15]. Later on, Ruhno et al. (2019) proposed a model to classify patients based on their *SMN2* dosage, including concordant patients with an expected *SMN2*_CN for their disease severity, and discordant patients with either a milder or more severe phenotype. This model only differs from Calucho's correlations in type III patients, since it establishes the expected *SMN2*_CN to be 4 instead of 3 and 4 indistinctly [9]. In the recent literature, the correlation described by Calucho et al. (2018) is mainly maintained [28–32], although some exceptions can be found. Interestingly, the proportion of type I patients with 3 *SMN2* genes was increased in some cohorts, reaching 57% among SMA I patients, while Calucho's compilation reported only 23% [33,34]. In contrast, the cohort described by Sun et al. (2020) seems to have, in global numbers, a better-than-expected phenotype, as the majority of SMA type II patients presented 2 *SMN2* genes, type III patients had 2 or 3 *SMN2* copies and more than half of SMA type IV had only 3 *SMN2* genes [35]. Despite these differences in the literature, all studies agree that the higher the *SMN2*_CN, the less severe the SMA phenotype. This widely described correlation is coherent, since the higher number of *SMN2* copies, the higher amounts of SMN functional protein produced compensating the lack of *SMN1* gene and explaining the better prognosis of the patients.

Nevertheless, this correlation is not absolute and some discordant cases based on this rule are found, which can be further subdivided in "better-than-expected" or "worse-than-expected" patients [15]. In some of these individuals, apart from the *SMN2*_CN, different variants have been reported to modify the SMA phenotype, which can help inform prognostic outcomes. Two positive modifiers in *SMN2* gene have been described, both associated with a milder phenotype [36–38]. The first modifier described was the variant

c.859G>C (p.Gly287Arg) located in exon 7 of *SMN2*. It creates a novel exonic splicing enhancer site (SF2/ASF motif) predicted by ESEfinder 3.0 [36]. Through SMN splicing assays, it has been demonstrated to significantly increase *SMN2* exon 7 inclusion in vitro from 40–50% to 70% and subsequently the amount of full-length *SMN* transcript [36,37]. There is an additive effect of this variant, as the greater number of *SMN2* copies with c.859G>C, the better the phenotype of the patient. In addition, it has been postulated that this allele has originated from a common ancestor by haplotype analysis [39]. The second variant classified as a positive modifier is the variant c.835-44A>G, located in intron 6, and it is one of the 16 PSVs described between *SMN1* and *SMN2* [10,40]. Wu et al. (2017) demonstrated that this transition decreases the affinity of the RNA-binding protein HuR, which acts as a splicing repressor, increasing in ~20% the *SMN2* exon 7 inclusion [38]. Despite the rare frequency of at least c.859G>C variant (0.8%, 11/1345) [41], a recent SMA practical guideline recommends the evaluation of both variants in discordant SMA patients presenting a better-than-expected phenotype [42].

Other known modifier variants that could explain discordant cases based on *SMN2*_CN include *SMN1* intragenic variants. More than 80 pathogenic variants have been described in *SMN1* gene in compound heterozygous individuals, mainly located in the Tudor and C-terminal domains [43]. In general terms, in the Tudor domain, missense mutations appear to be associated with a more severe SMA phenotype, whereas in the context of frameshift and nonsense variants may be more dependent on the *SMN2* copies [44,45]. For instance, c.275G>C (p.Trp92Ser) variant has been reported in severe SMA type I patients with 3 *SMN2* copies and a reduced interaction with SMN target proteins has been shown using a protein binding assay [46–48]. On the other hand, mutations in the C-terminal domain appear mostly related to a worse-than-expected phenotype as is the case of the variant c.770_780dup (p.Leu261Alafs*5) [21,31,49]. In our experience, compound heterozygous patients with this pathogenic variant and the *SMN1* deletion carrying only one *SMN2* copy had congenital type 0 disease [50], whereas patients with two or three *SMN2* copies had type I or II disease, respectively [21].

In addition, some missense variants in exon 1 of *SMN1* are associated with a milder phenotype [44]. Two recurrent variants are c.5C>T (p.Ala2Val) and c.5C>G (p.Ala2Gly), which are considered hypomorphic alleles identified in SMA type III patients [51,52]. In fact, it has been shown in a SMA mouse model that the change p.Ala2Gly does not produce total loss of protein function [53] and no significant decrease of full-length *SMN1* transcripts [43]. All these cases highlight the relevance of (1) performing additional functional studies to further characterize *SMN1* pathogenic variants, both including novel and previously described variants, and (2) further characterize the *SMN2* copies in those patients to better explain their phenotypes.

## 3. The Unknown or Yet Non-Validated Genotypes

Aside from the previously described *SMN2* modifiers (c.859G>C and c.835-44A>G), other variants in this gene have been proposed to modify the SMA phenotype, although functional studies to demonstrate an effect in the SMN protein have not been performed or more cases have not yet been reported [9,38,54]. For instance, variants c.835-1897C>T and c.835-549A>G in intron 6 of *SMN2* and variant c.*3+100G>A in intron 7, later classified as a PSV [10], have been associated with a better-than-expected phenotype (Figure 1) [9,38]. Furthermore, a recent study suggested that variants c.81+45C>T in intron 1 of *SMN2* and c.838_840del in exon 7 were related to a more severe SMA phenotype. This work also identified a novel variant (c.-14C>T) in the promoter region of the *SMN2* gene in an SMA type I patient with apparently four *SMN2* copies indicating a possible association with a worse-than-expected phenotype [54]. In addition, variants in the *SMN1* promoter, which have been associated with non-functional *SMN1* alleles, have also been reported, even though they are infrequent findings [45]. It is important to bear in mind that variants located in deep intronic regions would not be detected through MLPA or exome sequencing, highlighting the importance of recently developed strategies for the entire sequencing of

*SMN1* or *SMN2* genes [10,18]. Another cause of discordance could be a clinical misclassification of the patient or an inaccurate *SMN2*_CN determination. On the one hand, the lack of clinical information or the modification of the patient's phenotype due to natural history or evolving trajectory because of current treatments could lead to reassign the SMA type, thus generating discordance with its *SMN2*_CN [8]. On the other hand, many factors can affect the determination of *SMN2* dosage, giving an inaccurate result. Although MLPA is considered the gold standard technique to detect the number of *SMN2* copies in SMA patients, real-time PCR (RT-PCR) and droplet digital PCR (ddPCR) are also commonly used. In MLPA and RT-PCR based on SYBR Green, DNA quality is crucial to achieve reliable results, and moreover, both methods can be affected by variants in the target region of primers or probes, leading to a misinterpretation of the *SMN2*_CN. In addition, control samples or references are needed in all the approaches since they are based on indirect quantification, and therefore the right choice of these controls is decisive to establish the correct *SMN2*_CN [42]. This fact is reflected in the work of Schorling et al. (2019), in which 20 SMA patients were retested for their *SMN2* dosage using new DNA samples and 45% of the results were discrepant in comparison with the initial ones [55]. Retesting of cohorts with discrepancies (i.e., [33–35]) would be interesting to confirm if some of the results are due to problems with *SMN2*_CN determination or clinical misclassification of patients. In fact, a guideline is proposed to manage the discordant situations that are present in SMA patients [42].

There are also some other genetic factors to consider that remain undetectable with routine techniques. First, different *SMN1-SMN2* hybrid structures have been described, but using the available MLPAs only those hybrids formed by exons 1 to 7 of *SMN2* and exon 8 of *SMN1* (more common) or vice versa are detectable (see Figure 2). Blasco-Pérez et al. (2021) described two SMA patients with hybrid structures consisting of the entire *SMN2* gene except for a region of intron 6, corresponding to *SMN1* [10]. This hybrid is undetectable using MLPA as this technique has only specific probes of *SMN1* and *SMN2* in exons 7 and 8 based on the exonic PSVs. Similarly, other uncommon hybrids (such as those reported by Cusco et al. (2001) [20], Qu et al. (2016) [43], Kubo et al. (2015) [51] or Blasco-Pérez et al. (2021) [10]) or partial deletions of *SMN2* copies, would be detectable by the complete sequencing of *SMN* genes or by the study of PSVs of the region [10,18]. Second, besides the common polymorphic *SMN1/2*Δ7-8, other partial deletions have been described, some of which are difficult to elucidate if they are located in *SMN1* or *SMN2* genes, further contributing to the complexity of the analysis [18,56]. Third, it has been described that *SMN2* can be hypermethylated, resulting in a partial inactivation of the gene expression, which translates into a worse-than-expected phenotype according to the patient's *SMN2*_CN [57]. Lastly, variants in regulatory regions outside the gene could modify its expression, either increasing or decreasing the amount of SMN protein produced by each copy of *SMN2* as occurs with the modifier variants described within the gene. In fact, several targets of transcriptional regulation in the *SMN2* locus are under study [58,59].

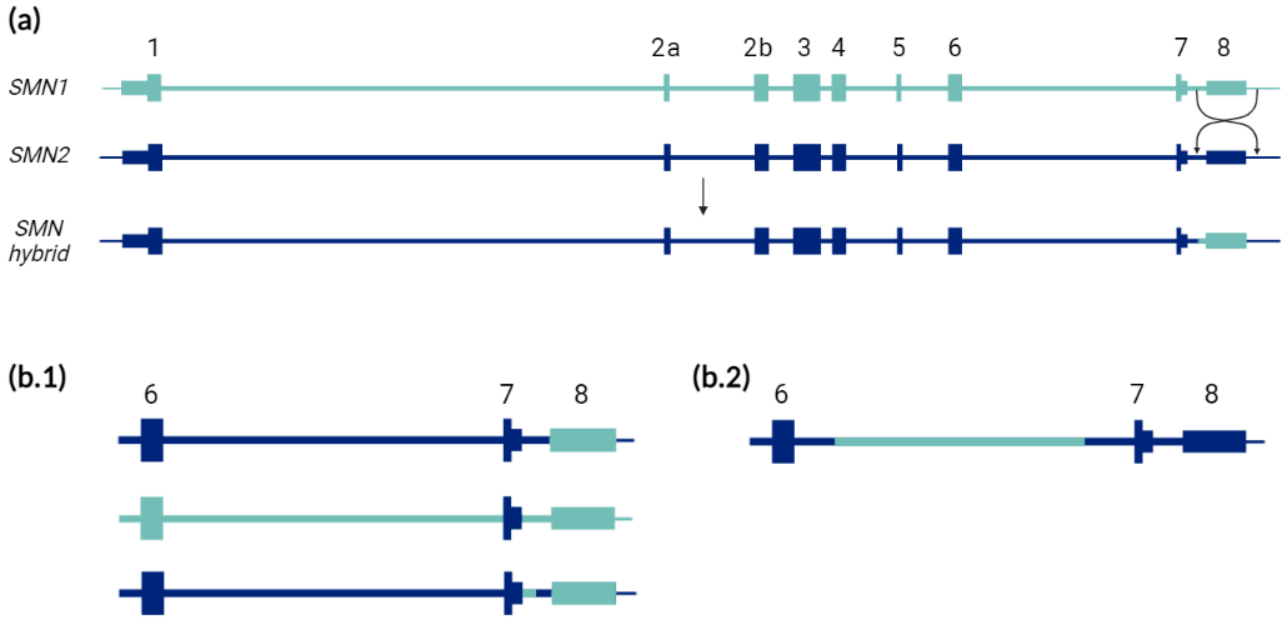



**Figure 2.** Schematic representation of different *SMN2/SMN1* hybrid structures reported in SMA patients. (**a**) Mechanism of formation of the classical *SMN2/SMN1* hybrid structure detectable through MLPA and other routine techniques and usually defined as "homozygous deletion of only exon 7 of the *SMN1* gene". (**b.1**) Previously reported *SMN2/SMN1* hybrid structures indistinguishable from the classical one using MLPA and other routine techniques [20,43,51]. (**b.2**) Previously reported *SMN2/SMN1* hybrid structure undetectable through MLPA and other routine techniques [10]. The hybrid structures from (**b**) are detectable by sequencing of the entire *SMN* genes. The different PSV sequences in each structure may determine the expression and functionality of each of the hybrid genes. Light blue represents *SMN1*, and dark blue *SMN2*.

Several discordant haploidentical SMA siblings have been described, presenting the same number of *SMN2* copies [60] and no differences by NGS analysis [9]. These cases are thought to be modulated by additional SMA genetic modifiers. Considering other possible SMA modifier genes, aside from *SMN2*, there are two candidates within the 5q13 region, *NAIP* (neuronal apoptosis inhibitory protein) and *SERF1A* (small EDRK-rich factor 1a). These genes have been found to be deleted in a proportion of patients, particularly with a more severe phenotype. However, results are inconclusive and it is more likely to be caused by a contiguity effect of the *SMN* deletion [61,62]. In addition, other factors outside the *SMN* gene locus may be involved in the *SMN2* response-activity and in the SMA phenotype's definition. These factors are usually divided into *SMN*-dependent factors, which directly alter the amount of SMN protein, and *SMN*-independent factors playing a role in many different functions such as actin polymerization, cytoskeleton dynamics or neurogenesis [63]. For instance, *PLS3* (plastin 3) gene has been proposed to possibly modulate SMA disease progression in discordant SMA siblings as higher expression levels were found in lymphoblasts in sisters with milder phenotypes [64]. However, up to date, no DNA markers or modifiers that lie outside the SMA region have been validated in clinical settings [9]. Nevertheless, a thorough discussion of these modifiers can be found elsewhere [63,65]. A representative list of possible modifiers of SMA phenotype and their references is summarized in Table 2.

**Table 2.** Possible modifiers of spinal muscular atrophy. This is a representative list to show the different lines of investigation to determine factors that may modify the SMN function and SMA phenotype. Even though we include several factors, the *SMN2* gene (copy number, sequence and structure) is the only validated as a DNA marker. Based on references [15,63–75]. Arrow indicates increase (up) or decrease (down).

| Modifier Type | Example | Effect | Reference |
|---|---|---|---|
| SMA locus | *SMN2* copies and variants | SMA types | Calucho et al. 2018 [15] |
| Splicing regulators | hnRNP-A1/Sam68 | Exon 7 inclusion | Pedrotti et al. 2010 [66] |
| SMN degradation | *UBA1* | ↑ Survival SMA mice | Powis et al. 2016 [67] |
| DNA methylation | *SLC23A2/NCOR2* | SMA types differences | Zheleznyakova et al. 2015 [68] |
| Actin polymerization | *PLS3* | Siblings differences | Oprea et al. 2008 [64] |
| Cytoskeleton dynamics | *ERK* | ↑ Survival SMA mice | Branchu et al. 2013 [69] |
| Endocytosis regulators | *NCALD* | Ameliorates SMA | Riessland et al. 2017 [70] |
| Neurogenesis regulators | *PTEN* | ↑ Survival SMA mice | Little et al. 2015 [71] |
| Axogenesis | *ZPR1* | ↓ in SMA patients | Helmken et al. 2003 [72] |
| Apoptosis | Bcl2 | ↓ SMA motor neurons | Soler-Botija et al. 2003 [73] |
| Hormones/growth factors | Prolactin | ↑ Survival SMA mice | Farooq et al. 2011 [74] |
| Environmental factors | Exercise | ↑ Survival SMA mice | Grondard et al. 2005 [75] |

## 4. Evolving Therapies and the Importance of *SMN2*

In the SMA therapeutic background, there are three approved therapies by the FDA (Food and Drug Administration) and EMA (European Medicines Agency): nusinersen, risdiplam and onasemnogene abeparvovec-xioi (OA). Nusinersen and risdiplam agents are designed to bind specifically to *SMN2* pre-mRNA in order to promote exon 7 inclusion increasing the amount of functional SMN protein [76,77]. The remainder approved SMA treatment OA, consists of a gene replacement therapy that restores the expression of normal *SMN1* using a viral vector (AAV9) expressing *SMN1* [78]. While the *SMN2* endogenous target regions comprise splicing regulators and intronic regions, the *SMN1* transgene in the AAV9 is an *SMN1* cDNA lacking intronic or other regulatory elements. At first sight, the identification of *SMN2* copies, variants and structures would be of particular interest for the approved SMA treatments targeting *SMN2.* However, this information could be also useful for patients treated with OA. Indeed, the c.859G>C variant was relevant in the AVXS-101 trial as it was defined as an exclusion criteria, albeit the approval was given without limitation on the genetic background [78]. For example, in the context of presymptomatic detection, a neonate with 2 copies of *SMN2* carrying the c.859G>C variant will have a better evolution than typical cases with 2 *SMN2* copies without this modifier variant [39]. In view of the current progress in the worldwide implementation of neonatal screening, when a neonate is genetically diagnosed with SMA, it is recommended to perform not only *SMN2*_CN determination, but also a more complete *SMN2* characterization, including variants and if possible, structural changes [42]. While the influence of *SMN* hybrid genes and partial deletions on the response of the different therapies has not yet been addressed, it can be speculated that therapeutic efficacy of *SMN2* modulators may be affected depending on particular hybrid structures [10]. The NGS approach of the complete SMN2 genes is also useful to determine regions of the *SMN2* that are apparently highly conserved within the patients. For example, the ISSN1 sequence of intron 7, an interesting region because it is the target of the 18bp oligonucleotide nusinersen, appears to be so far identical in SMA patients [10], but further studies with a larger number of cases should be performed to confirm this observation.

Upon approvals, the availability of different therapies is complicating the decision-making for treatment choices. Besides monotherapy, several SMA patients are receiving combinatorial therapies starting with nusinersen and after OA or vice versa [79–81], and combination could also include risdiplam (individual reports). In principle, the mechanisms of action of these therapies do not interfere specifically with each other but are rather complementary. For example, any patient receiving gene therapy, will produce SMN protein autonomously and theoretically indefinitely as an episome in the nucleus

of postmitotic cells and the endogenous *SMN2* could still be a possible target for *SMN2* modulators. However, the regulation and feedback that modulate SMN production by both mechanisms have not been elucidated yet. Recently, it has been reported that the overexpression of SMN protein by AAV9 has long-term neuronal toxic effects in a SMA mouse model [82]. Therefore, the putative overproduction of SMN with combinatorial therapies should also be cautiously considered in SMA patients, particularly in those who already received gene therapy. A recent consensus statement on gene replacement therapy for SMA does not recommend combinatorial therapy as part of routine care [83]. Further studies should be performed to determine when combinatorial therapies would be more effective than monotherapy. In the current scenario, combinatorial therapies may also include SMN-independent compounds, which are worth to be further investigated in SMA patients [84]. Interestingly, a recent study has showed promising results with the combinatorial use of an antisense oligonucleotide (ASO) mimicking nusinersen and an orally delivered histone deacetylase inhibitor (panobinostat) in SMA cell models. They have found that this compound increases the effects of the ASO on *SMN2* exon 7 inclusion enhancing the expression of *SMN2* [85]. As this is a constantly evolving field, regardless of the treatment received, all SMA cases may benefit from knowing their complete *SMN2* genotype to make better and tight correlations with each phenotype and more realistic outcomes after therapies.

## 5. Conclusions

The present therapeutic scenario highlights the importance to genetically confirm SMA patients in order to make them eligible for treatment options. Although the vast majority of patients can be straightforwardly diagnosed, it is important to be aware of rare particular *SMN1* deletions or variants that may be critical to genetically confirm a given patient [56]. Once bi-allelic *SMN1* alterations are confirmed, *SMN2* enters the scene for better prognostic and phenotype characterizations. Discordant situations of genotype-phenotype correlation in SMA exist including when a phenotype of a given patient is better-than-expected according to *SMN2* copies (fewer copies, better phenotype) or conversely, when the phenotype is worse-than-expected according to *SMN2* copies (more copies, but more severe phenotype). These discordances may be due to biological or technical issues and all discordant cases should be retested considering a new sample, a different methodology and/or even another laboratory (Figure 3). Testing for positive known variants that may influence the amount of complete SMN should be performed once the copy number results are confirmed. NGS studies of the entire gene allow further characterization of the quality of *SMN2* copies. Better-than-expected discordances are usually explained by known positive variants in *SMN2*, but some cases still remain unexplained and further genetic investigations may unravel potential causes that explain the phenotype. In our experience, worse-than-expected cases are usually due to technical pitfalls (*SMN2*_CN overestimation), although negative modifiers are under validation studies. Partial deletions of *SMN2* may be masked if *SMN2* sequencing or quantitation is not carefully evaluated. Copy number results may be complemented with NGS of the entire *SMN2* genes to define their sequence and structure and detect further modifiers. Confirmation of copy number can also be achieved with NGS studies [10]. Finally, the therapeutic context of SMA is becoming more complex and expanded with several SMN-dependent or SMN-independent therapeutic approaches. Thus, combinatorial therapies are expected to be protocolized in the future, when more evidence about their efficacy is available [84]. All these advances should consider *SMN2* copies, variants and structures as part of the integral characterization of patients receiving expensive and sometimes lifelong therapies. The *SMN2* gene, as the main modifier of SMA phenotype, warrants a deeper study beyond the copy number determination. In the near future, either in the presymptomatic neonatal screening scenario or in already symptomatic patients, routine analysis may be adapted to currently detect rare (modifier) variants, single-nucleotide polymorphisms and structural variants of the SMN locus.

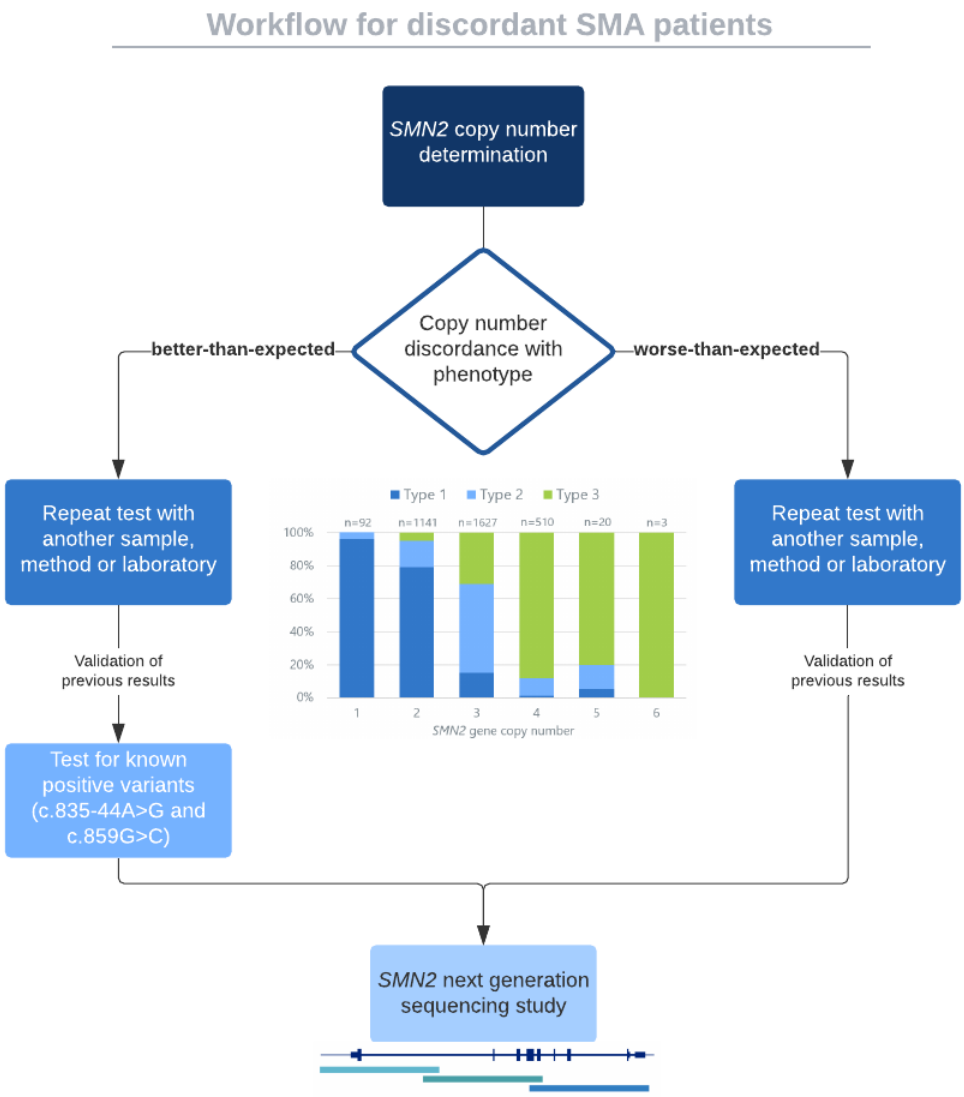



**Figure 3.** Workflow for discordant SMA symptomatic patients. Once *SMN2* is determined and discordance is found (according to the graphic in the center of the workflow based on the meta-analysis of 3459 cases from Calucho et al. (2018) [15]), a re-test should be performed with a new sample, another method or even another laboratory. If the discrepancy is confirmed, when the phenotype is better than expected, we may test for known and validated positive variants (c.859G>C and c.835-44A>G). If negative, we may continue with further NGS studies (as in [10]). In case that the phenotype is worse than expected, we should perform an NGS test in order to determine potential negative modifiers, hybrid structures or intragenic deletions that may explain the phenotype. Based on reference [42].

**Author Contributions:** M.C.-R.; L.B.-P. writing—original draft preparation and data curation. I.C., E.F.T. Conceptualization, writing review and editing. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was partially supported by Grants from Biogen and Roche (to E.F.T. supporting M.C.-R. and L.B.-P.), and from Spanish Instituto de Salud Carlos III, Fondo de Investigaciones Sanitarias and cofunded with ERDF funds (Grant No. FIS PI18/000687) (to E.F.T.).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable

**Conflicts of Interest:** E.F.T. discloses grant support to conduct CTs on SMA from Ionis/Biogen and serves as a consultant to AveXis, Novartis, Biogen, Biologix, Cytokinetics, and Roche. The other authors report no disclosures relevant to the manuscript.

## References


1. Kolb, S.J.; Kissel, J.T. Spinal Muscular Atrophy. *Neurol. Clin.* **2015**, *33*, 831–846. [CrossRef]
2. Sugarman, E.; Nagan, N.; Zhu, H.; Akmaev, V.; Zhou, Z.; Rohlfs, E.; Flynn, K.; Hendrickson, B.; Scholl, T.; Sirko-Osadsa, D.; et al. Pan-ethnic carrier screening and prenatal diagnosis for spinal muscular atrophy: Clinical laboratory analysis of >72,400 specimens. *Eur. J. Hum. Genet.* **2012**, *20*, 27–32. [CrossRef]
3. Prior, T.W.; Leach, M.E.; Finanger, E. Spinal Muscular Atrophy. In *GeneReviews®*; Adam, M.P., Ardinger, H.H., Pagon, R.A., Wallace, S.E., Bean, L.J.H., Mirzaa, G., Amemiya, A., Eds.; University of Washington: Seattle, WA, USA, 2000; updated 3 December 2020; Available online: Website (accessed on 28 July 2021).
4. Talbot, K.; Tizzano, E.F. The clinical landscape for SMA in a new therapeutic era. *Gene Ther.* **2017**, *24*, 529–533. [CrossRef]
5. Lefebvre, S.; Bürglen, L.; Reboullet, S.; Clermont, O.; Burlet, P.; Viollet, L.; Benichou, B.; Cruaud, C.; Millasseau, P.; Zeviani, M.; et al. Identification and characterization of a spinal muscular atrophy-determining gene. *Cell* **1995**, *80*, 155–165. [CrossRef]
6. Wirth, B. An update of the mutation spectrum of the survival motor neuron gene (SMN1) in autosomal recessive spinal muscular atrophy (SMA). *Hum. Mutat.* **2000**, *15*, 228–237. [CrossRef]
7. Bürglen, L.; Lefebvre, S.; Clermont, O.; Burlet, P.; Viollet, L.; Cruaud, C.; Munnich, A.; Melki, J. Structure and organization of the human survival motor neurone (SMN) gene. *Genomics* **1996**, *32*, 479–482. [CrossRef]
8. Rochette, C.; Gilbert, N.; Simard, L. SMN gene duplication and the emergence of the SMN2 gene occurred in distinct hominids: SMN2 is unique to Homo sapiens. *Hum. Genet.* **2001**, *108*, 255–266. [CrossRef] [PubMed]
9. Ruhno, C.; McGovern, V.L.; Avenarius, M.R.; Snyder, P.J.; Prior, T.W.; Nery, F.C.; Muhtaseb, A.; Roggenbuck, J.S.; Kissel, J.T.; Sansone, V.A.; et al. Complete sequencing of the SMN2 gene in SMA patients detects SMN gene deletion junctions and variants in SMN2 that modify the SMA phenotype. *Hum. Genet.* **2019**, *138*, 241–256. [CrossRef] [PubMed]
10. Blasco-Pérez, L.; Paramonov, I.; Leno, J.; Bernal, S.; Alias, L.; Fuentes-Prior, P.; Cuscó, I.; Tizzano, E.F. Beyond copy number: A new, rapid, and versatile method for sequencing the entire SMN2 gene in SMA patients. *Hum. Mutat.* **2021**, *42*, 787–795. [CrossRef]
11. Lorson, C.L.; Hahnen, E.; Androphy, E.J.; Wirth, B. A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 6307–6311. [CrossRef] [PubMed]
12. Soler-Botija, C.; Cuscó, I.; Caselles, L.; López, E.; Baiget, M.; Tizzano, E.F. Implication of fetal SMN2 expression in type I SMA pathogenesis: Protection or pathological gain of function? *J. Neuropathol. Exp. Neurol.* **2005**, *64*, 215–223. [CrossRef]
13. Wirth, B.; Garbes, L.; Riessland, M. How genetic modifiers influence the phenotype of spinal muscular atrophy and suggest future therapeutic approaches. *Curr. Opin. Genet. Dev.* **2013**, *23*, 330–338. [CrossRef]
14. Boza-Morán, M.G.; Martínez-Hernández, R.; Bernal, S.; Wanisch, K.; Also-Rallo, E.; Le Heron, A.; Alías, L.; Denis, C.; Girard, M.; Yee, J.; et al. Decay in survival motor neuron and plastin 3 levels during differentiation of iPSC-derived human motor neurons. *Sci. Rep.* **2015**, *5*, 11696. [CrossRef] [PubMed]
15. Calucho, M.; Bernal, S.; Alías, L.; March, F.; Venceslá, A.; Rodríguez-Álvarez, F.J.; Aller, E.; Fernández, R.M.; Borrego, S.; Millán, J.M.; et al. Correlation between SMA type and SMN2 copy number revisited: An analysis of 625 unrelated Spanish patients and a compilation of 2834 reported cases. *Neuromuscul. Disord.* **2018**, *28*, 208–215. [CrossRef] [PubMed]
16. Arkblad, E.L.; Darin, N.; Berg, K.; Kimber, E.; Brandberg, G.; Lindberg, C.; Holmberg, E.; Tulinius, M.; Nordling, M. Multiplex ligation-dependent probe amplification improves diagnostics in spinal muscular atrophy. *Neuromuscul. Disord.* **2006**, *16*, 830–838. [CrossRef] [PubMed]
17. Vijzelaar, R.; Snetselaar, R.; Clausen, M.; Mason, A.G.; Rinsma, M.; Zegers, M.; Molleman, N.; Boschloo, R.; Yilmaz, R.; Kuilboer, R.; et al. The frequency of SMN gene variants lacking exon 7 and 8 is highly population dependent. *PLoS ONE* **2019**, *14*, e0220211. [CrossRef]
18. Chen, X.; Sanchis-Juan, A.; French, C.E.; Connell, A.J.; Delon, I.; Kingsbury, Z.; Chawla, A.; Halpern, A.L.; Taft, R.J.; Bentley, D.R.; et al. Spinal muscular atrophy diagnosis and carrier screening from genome sequencing data. *Genet. Med.* **2020**, *22*, 945–953. [CrossRef]
19. Alías, L.; Bernal, S.; Barceló, M.J.; Also-Rallo, E.; Martínez-Hernández, R.; Rodríguez-Alvarez, F.J.; Hernández-Chico, C.; Baiget, M.; Tizzano, E.F. Accuracy of marker analysis, quantitative real-time polymerase chain reaction, and multiple ligation-dependent probe amplification to determine SMN2 copy number in patients with spinal muscular atrophy. *Genet. Test. Mol. Biomark.* **2011**, *15*, 587–594. [CrossRef] [PubMed]
20. Cuscó, I.; Barceló, M.J.; del Rio, E.; Martín, Y.; Hernández-Chico, C.; Bussaglia, E.; Baiget, M.; Tizzano, E.F. Characterisation of SMN hybrid genes in Spanish SMA patients: De novo, homozygous and compound heterozygous cases. *Hum. Genet.* **2001**, *108*, 222–229. [CrossRef]
21. Alías, L.; Bernal, S.; Fuentes-Prior, P.; Barceló, M.J.; Also, E.; Martínez-Hernández, R.; Rodríguez-Alvarez, F.J.; Martín, Y.; Aller, E.; Grau, E.; et al. Mutation update of spinal muscular atrophy in Spain: Molecular characterization of 745 unrelated patients and identification of four novel mutations in the SMN1 gene. *Hum. Genet.* **2009**, *125*, 29–39. [CrossRef]

22. Hahnen, E.; Schönling, J.; Rudnik-Schöneborn, S.; Zerres, K.; Wirth, B. Hybrid survival motor neuron genes in patients with autosomal recessive spinal muscular atrophy: New insights into molecular mechanisms responsible for the disease. *Am. J. Hum. Genet.* **1996**, *59*, 1057–1065. [PubMed]
23. Niba, E.T.E.; Nishio, H.; Wijaya, Y.O.S.; Lai, P.S.; Tozawa, T.; Chiyonobu, T.; Yamadera, M.; Okamoto, K.; Awano, H.; Takeshima, Y.; et al. Clinical phenotypes of spinal muscular atrophy patients with hybrid SMN gene. *Brain Dev.* **2021**, *43*, 294–302. [CrossRef] [PubMed]
24. Hendrickson, B.; Donohoe, C.; Akmaev, V.; Sugarman, E.; Labrousse, P.; Boguslavskiy, L.; Flynn, K.; Rohlfs, E.; Walker, A.; Allitto, B.; et al. Differences in SMN1 allele frequencies among ethnic groups within North America. *J. Med. Genet.* **2009**, *46*, 641–644. [CrossRef] [PubMed]
25. Luo, M.; Liu, L.; Peter, I.; Zhu, J.; Scott, S.A.; Zhao, G.; Eversley, C.; Kornreich, R.; Desnick, R.J.; Edelmann, L. An Ashkenazi Jewish SMN1 haplotype specific to duplication alleles improves pan-ethnic carrier screening for spinal muscular atrophy. *Genet. Med.* **2014**, *16*, 149–156. [CrossRef]
26. Alías, L.; Barceló, M.J.; Bernal, S.; Martínez-Hernández, R.; Also-Rallo, E.; Vázquez, C.; Santana, A.; Millán, J.M.; Baiget, M.; Tizzano, E.F. Improving detection and genetic counseling in carriers of spinal muscular atrophy with two copies of the SMN1 gene. *Clin. Genet.* **2014**, *85*, 470–475. [CrossRef]
27. Alías, L.; Bernal, S.; Calucho, M.; Martínez, E.; March, F.; Gallano, P.; Fuentes-Prior, P.; Abuli, A.; Serra-Juhe, C.; Tizzano, E.F. Utility of two SMN1 variants to improve spinal muscular atrophy carrier diagnosis and genetic counselling. *Eur. J. Hum. Genet.* **2018**, *26*, 1554–1557. [CrossRef]
28. Coratti, G.; Messina, S.; Lucibello, S.; Pera, M.C.; Montes, J.; Pasternak, A.; Bovis, F.; Escudero, J.E.; Mazzone, E.S.; Mayhew, A.; et al. Clinical Variability in Spinal Muscular Atrophy Type III. *Ann. Neurol.* **2020**, *88*, 1109–1117. [CrossRef]
29. Osredkar, D.; Jílková, M.; Butenko, T.; Loboda, T.; Golli, T.; Fuchsová, P.; Rohlenová, M.; Haberlova, J. Children and young adults with spinal muscular atrophy treated with nusinersen. *Eur. J. Paediatr. Neurol.* **2021**, *30*, 1–8. [CrossRef]
30. Ou, S.; Ho, C.; Lee, W.; Lin, K.; Jones, C.; Jong, Y. Natural history in spinal muscular atrophy Type I in Taiwanese population: A longitudinal study. *Brain Dev.* **2021**, *43*, 127–134. [CrossRef] [PubMed]
31. Souza, P.; Pinto, W.; Ricarte, A.; Badia, B.; Seneor, D.; Teixeira, D.; Caetano, L.; Gonçalves, E.; Chieia, M.; Farias, I.; et al. Clinical and radiological profile of patients with spinal muscular atrophy type 4. *Eur. J. Neurol.* **2021**, *28*, 609–619. [CrossRef]
32. Zhang, Y.; He, J.; Zhang, Y.; Li, L.; Tang, X.; Wang, L.; Guo, J.; Jin, C.; Tighe, S.; Zhang, Y.; et al. The analysis of the association between the copy numbers of survival motor neuron gene 2 and neuronal apoptosis inhibitory protein genes and the clinical phenotypes in 40 patients with spinal muscular atrophy: Observational study. *Medicine (Baltim.)* **2020**, *99*, e18809. [CrossRef] [PubMed]
33. Lusakowska, A.; Jedrzejowska, M.; Kaminska, A.; Janiszewska, K.; Grochowski, P.; Zimowski, J.; Sierdzinski, J.; Kostera-Pruszczyk, A. Observation of the natural course of type 3 spinal muscular atrophy: Data from the polish registry of spinal muscular atrophy. *Orphanet J. Rare Dis.* **2021**, *16*. [CrossRef] [PubMed]
34. Aragon-Gawinska, K.; Seferian, A.; Daron, A.; Gargaun, E.; Vuillerot, C.; Cances, C.; Ropars, J.; Chouchane, M.; Cuppen, I.; Hughes, I.; et al. Nusinersen in patients older than 7 months with spinal muscular atrophy type 1: A cohort study. *Neurology* **2018**, *91*, e1312–e1318. [CrossRef] [PubMed]
35. Sun, Y.; Kong, X.; Zhao, Z.; Zhao, X. Mutation analysis of 419 family and prenatal diagnosis of 339 cases of spinal muscular atrophy in China. *BMC Med. Genet.* **2020**, *21*, 133. [CrossRef]
36. Prior, T.W.; Krainer, A.R.; Hua, Y.; Swoboda, K.J.; Snyder, P.C.; Bridgeman, S.J.; Burghes, A.H.M.; Kissel, J.T. A Positive Modifier of Spinal Muscular Atrophy in the SMN2 Gene. *Am. J. Hum. Genet.* **2009**, *85*, 408–413. [CrossRef]
37. Vezain, M.; Saugier-Veber, P.; Goina, E.; Touraine, R.; Manel, V.; Toutain, A.; Fehrenbach, S.; Frébourg, T.; Pagani, F.; Tosi, M.; et al. A rare SMN2 variant in a previously unrecognized composite splicing regulatory element induces exon 7 inclusion and reduces the clinical severity of spinal muscular atrophy. *Hum. Mutat.* **2010**, *31*, E1110–E1125. [CrossRef]
38. Wu, X.; Wang, S.H.; Sun, J.; Krainer, A.R.; Hua, Y.; Prior, T.W. A-44G transition in SMN2 intron 6 protects patients with spinal muscular atrophy. *Hum. Mol. Genet.* **2017**, *26*, 2768–2780. [CrossRef]
39. Bernal, S.; Alías, L.; Barceló, M.J.; Also-Rallo, E.; Martínez-Hernández, R.; Gámez, J.; Guillén-Navarro, E.; Rosell, J.; Hernando, I.; Rodríguez-Alvarez, F.J.; et al. The c.859G>C variant in the SMN2 gene is associated with types II and III SMA and originates from a common ancestor. *J. Med. Genet.* **2010**, *47*, 640–642. [CrossRef]
40. Monani, U.R.; Lorson, C.L.; Parsons, D.W.; Prior, T.W.; Androphy, E.J.; Burghes, A.H.M.; McPherson, J.D. A single nucleotide difference that alters splicing patterns distinguishes the SMA gene SMN1 from the copy gene SMN2. *Hum. Mol. Genet.* **1999**, *8*, 1177–1183. [CrossRef]
41. Bowen, B.M.; Truty, R.; Aradhya, S.; Bristow, S.L.; Johnson, B.A.; Morales, A.; Tan, C.A.; Westbrook, M.J.; Winder, T.L.; Chavez, J.C. SMA Identified: Clinical and Molecular Findings From a Sponsored Testing Program for Spinal Muscular Atrophy in More Than 2,000 Individuals. *Front. Neurol.* **2021**, *12*, 676. [CrossRef]
42. Cuscó, I.; Bernal, S.; Blasco-Pérez, L.; Calucho, M.; Alias, L.; Fuentes-Prior, P.; Tizzano, E.F. Practical guidelines to manage discordant situations of SMN2 copy number in patients with spinal muscular atrophy. *Neurol. Genet.* **2020**, *6*, e530. [CrossRef] [PubMed]

43. Qu, Y.; Bai, J.; Cao, Y.; Wang, H.; Jin, Y.; Du, J.; Ge, X.; Zhang, W.; Li, Y.; He, S.; et al. Mutation Spectrum of the Survival of Motor Neuron 1 and Functional Analysis of Variants in Chinese Spinal Muscular Atrophy. *J. Mol. Diagn.* **2016**, *18*, 741–752. [CrossRef] [PubMed]
44. Wijaya, Y.; Ar Rohmah, M.; Niba, E.; Morisada, N.; Noguchi, Y.; Saito, K.; Nishio, H.; Shinohara, M. Phenotypes of SMA patients retaining SMN1 with intragenic mutation. *Brain Dev.* **2021**, *43*, 745–758. [CrossRef] [PubMed]
45. Cuscó, I.; Barceló, M.; del Río, E.; Baiget, M.; Tizzano, E. Detection of novel mutations in the SMN Tudor domain in type I SMA patients. *Neurology* **2004**, *63*, 146–149. [CrossRef] [PubMed]
46. Kotani, T.; Sutomo, R.; Sasongko, T.; Sadewa, A.; Gunadi; Minato, T.; Fujii, E.; Endo, S.; Lee, M.; Ayaki, E.; et al. A novel mutation at the N-terminal of SMN Tudor domain inhibits its interaction with target proteins. *J. Neurol.* **2007**, *254*, 624–630. [CrossRef]
47. Yamamoto, T.; Sato, H.; Lai, P.S.; Nurputra, D.K.; Harahap, N.I.F.; Morikawa, S.; Nishimura, N.; Kurashige, T.; Ohshita, T.; Nakajima, H.; et al. Intragenic mutations in SMN1 may contribute more significantly to clinical severity than SMN2 copy numbers in some spinal muscular atrophy (SMA) patients. *Brain Dev.* **2014**, *36*, 914–920. [CrossRef]
48. Takarada, T.; Rochmah, M.A.; Harahap, N.; Shinohara, M.; Saito, T.; Saito, K.; Lai, P.; Bouike, Y.; Takeshima, Y.; Awano, H.; et al. SMA mutations in SMN Tudor and C-terminal domains destabilize the protein. *Brain Dev.* **2017**, *39*, 606–612. [CrossRef]
49. Parsons, D.; McAndrew, P.; Monani, U.; Mendell, J.; Burghes, A.; Prior, T. An 11 base pair duplication in exon 6 of the SMN gene produces a type I spinal muscular atrophy (SMA) phenotype: Further evidence for SMN as the primary SMA-determining gene. *Hum. Mol. Genet.* **1996**, *5*, 1727–1732. [CrossRef]
50. Mendonça, R.; Rocha, A.; Lozano-Arango, A.; Diaz, A.; Castiglioni, C.; Silva, A.; Reed, U.; Kulikowski, L.; Paramonov, I.; Cuscó, I.; et al. Severe brain involvement in 5q spinal muscular atrophy type 0. *Ann. Neurol.* **2019**, *86*, 458–462. [CrossRef]
51. Kubo, Y.; Nishio, H.; Saito, K. A new method for SMN1 and hybrid SMN gene analysis in spinal muscular atrophy using long-range PCR followed by sequencing. *J. Hum. Genet.* **2015**, *60*, 233–239. [CrossRef]
52. Mendonça, R.; Matsui, C.; Polido, G.; Silva, A.; Kulikowski, L.; Dias, A.T.; Zanardo, E.; Solla, D.; Kok, F.; Reed, U.; et al. Intragenic variants in the SMN1 gene determine the clinical phenotype in 5q spinal muscular atrophy. *Neurol. Genet.* **2020**, *6*, e505. [CrossRef] [PubMed]
53. Monani, U.R.; Pastore, M.T.; Gavrilina, T.O.; Jablonka, S.; Le, T.T.; Andreassi, C.; DiCocco, J.M.; Lorson, C.; Androphy, E.J.; Sendtner, M.; et al. A transgene carrying an A2G missense mutation in the SMN gene modulates phenotypic severity in mice with severe (type I) spinal muscular atrophy. *J. Cell Biol.* **2003**, *160*, 41–52. [CrossRef]
54. Wadman, R.I.; Jansen, M.D.; Stam, M.; Wijngaarde, C.A.; Curial, C.A.D.; Medic, J.; Sodaar, P.; Schouten, J.; Vijzelaar, R.; Lemmink, H.H.; et al. Intragenic and structural variation in the SMN locus and clinical variability in spinal muscular atrophy. *Brain Commun.* **2020**, *2*, 1–13. [CrossRef] [PubMed]
55. Schorling, D.; Becker, J.; Pechmann, A.; Langer, T.; Wirth, B.; Kirschner, J. Discrepancy in redetermination of SMN2 copy numbers in children with SMA. *Neurology* **2019**, *93*, 267–269. [CrossRef] [PubMed]
56. Jedličková, I.; Přistoupilová, A.; Nosková, L.; Majer, F.; Stránecký, V.; Hartmannová, H.; Hodaňová, K.; Trešlová, H.; Hýblová, M.; Solár, P.; et al. Spinal muscular atrophy caused by a novel Alu-mediated deletion of exons 2a-5 in SMN1 undetectable with routine genetic testing. *Mol. Genet. Genom. Med.* **2020**, *8*, 1–10. [CrossRef] [PubMed]
57. Hauke, J.; Riessland, M.; Lunke, S.; Eyüpoglu, I.; Blümcke, I.; El-Osta, A.; Wirth, B.; Hahnen, E. Survival motor neuron gene 2 silencing by DNA methylation correlates with spinal muscular atrophy disease severity and can be bypassed by histone deacetylase inhibition. *Hum. Mol. Genet.* **2009**, *18*, 304–317. [CrossRef]
58. Woo, C.J.; Maier, V.K.; Davey, R.; Brennan, J.; Li, G.; Brothers, J.; Schwartz, B.; Gordo, S.; Kasper, A.; Okamoto, T.R.; et al. Gene activation of SMN by selective disruption of lncRNA-mediated recruitment of PRC2 for the treatment of spinal muscular atrophy. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E1509–E1518. [CrossRef] [PubMed]
59. Nizzardo, M.; Simone, C.; Dametti, S.; Salani, S.; Ulzi, G.; Pagliarani, S.; Rizzo, F.; Frattini, E.; Pagani, F.; Bresolin, N.; et al. Spinal muscular atrophy phenotype is ameliorated in human motor neurons by SMN increase via different novel RNA therapeutic approaches. *Sci. Rep.* **2015**, *5*, 1–13. [CrossRef]
60. Cuscó, I.; Barceló, M.; Rojas-García, R.; Illa, I.; Gámez, J.; Cervera, C.; Pou, A.; Izquierdo, G.; Baiget, M.; Tizzano, E. SMN2 copy number predicts acute or chronic spinal muscular atrophy but does not account for intrafamilial variability in siblings. *J. Neurol.* **2006**, *253*, 21–25. [CrossRef]
61. Medrano, S.; Monges, S.; Gravina, L.P.; Alías, L.; Mozzoni, J.; Aráoz, H.V.; Bernal, S.; Moresco, A.; Chertkoff, L.; Tizzano, E. Genotype–phenotype correlation of SMN locus genes in spinal muscular atrophy children from Argentina. *Eur. J. Paediatr. Neurol.* **2016**, *20*, 910–917. [CrossRef]
62. Brkušanin, M.; Kosać, A.; Jovanović, V.; Pešović, J.; Brajušković, G.; Dimitrijević, N.; Todorović, S.; Romac, S.; Rašić, V.M.; Savić-Pavićević, D. Joint effect of the SMN2 and SERF1A genes on childhood-onset types of spinal muscular atrophy in Serbian patients. *J. Hum. Genet.* **2015**, *60*, 723–728. [CrossRef]
63. Wirth, B. Spinal Muscular Atrophy: In the Challenge Lies a Solution. *Trends Neurosci.* **2021**, *44*, 306–322. [CrossRef]
64. Oprea, G.; Kröber, S.; McWhorter, M.; Rossoll, W.; Müller, S.; Krawczak, M.; Bassell, G.; Beattie, C.; Wirth, B. Plastin 3 is a protective modifier of autosomal recessive spinal muscular atrophy. *Science* **2008**, *320*, 524–527. [CrossRef] [PubMed]
65. Wirth, B.; Mendoza-Ferreira, N.; Torres-Benito, L. Chapter 12—Spinal muscular atrophy disease modifiers. In *Spinal Muscular Atrophy: Disease Mechanisms and Therapy*; Academic Press: Amsterdam, The Netherlands, 2017; pp. 191–210.

66. Pedrotti, S.; Bielli, P.; Paronetto, M.; Ciccosanti, F.; Fimia, G.; Stamm, S.; Manley, J.; Sette, C. The splicing regulator Sam68 binds to a novel exonic splicing silencer and functions in SMN2 alternative splicing in spinal muscular atrophy. *EMBO J.* **2010**, *29*, 1235–1247. [CrossRef] [PubMed]
67. Powis, R.A.; Karyka, E.; Boyd, P.; Côme, J.; Jones, R.A.; Zheng, Y.; Szunyogova, E.; Groen, E.J.; Hunter, G.; Thomson, D.; et al. Systemic restoration of UBA1 ameliorates disease in spinal muscular atrophy. *JCI Insight* **2016**, *1*, e87908. [CrossRef] [PubMed]
68. Zheleznyakova, G.; Nilsson, E.; Kiselev, A.; Maretina, M.; Tishchenko, L.; Fredriksson, R.; Baranov, V.; Schiöth, H. Methylation levels of SLC23A2 and NCOR2 genes correlate with spinal muscular atrophy severity. *PLoS ONE* **2015**, *10*, e0121964. [CrossRef]
69. Branchu, J.; Biondi, O.; Chali, F.; Collin, T.; Leroy, F.; Mamchaoui, K.; Makoukji, J.; Pariset, C.; Lopes, P.; Massaad, C.; et al. Shift from extracellular signal-regulated kinase to AKT/cAMP response element-binding protein pathway increases survival-motor-neuron expression in spinal-muscular-atrophy-like mice and patient cells. *J. Neurosci.* **2013**, *33*, 4280–4294. [CrossRef]
70. Riessland, M.; Kaczmarek, A.; Schneider, S.; Swoboda, K.J.; Löhr, H.; Bradler, C.; Grysko, V.; Dimitriadi, M.; Hosseinibarkooie, S.; Torres-Benito, L.; et al. Neurocalcin Delta Suppression Protects against Spinal Muscular Atrophy in Humans and across Species by Restoring Impaired Endocytosis. *Am. J. Hum. Genet.* **2017**, *100*, 297–315. [CrossRef]
71. Little, D.; Valori, C.F.; Mutsaers, C.A.; Bennett, E.J.; Wyles, M.; Sharrack, B.; Shaw, P.J.; Gillingwater, T.H.; Azzouz, M.; Ning, K. PTEN Depletion Decreases Disease Severity and Modestly Prolongs Survival in a Mouse Model of Spinal Muscular Atrophy. *Mol. Ther.* **2015**, *23*, 270–277. [CrossRef]
72. Helmken, C.; Hofmann, Y.; Schoenen, F.; Oprea, G.; Raschke, H.; Rudnik-Schöneborn, S.; Zerres, Z.; Wirth, B. Evidence for a modifying pathway in SMA discordant families: Reduced SMN level decreases the amount of its interacting partners and Htra2-beta1. *Hum. Genet.* **2003**, *114*, 11–21. [CrossRef]
73. Soler-Botija, C.; Ferrer, I.; Alvarez, J.; Baiget, M.; Tizzano, E. Downregulation of Bcl-2 proteins in type I spinal muscular atrophy motor neurons during fetal development. *J. Neuropathol. Exp. Neurol.* **2003**, *62*, 420–426. [CrossRef]
74. Farooq, F.; Molina, F.A.; Hadwen, J.; MacKenzie, D.; Witherspoon, L.; Osmond, M.; Holcik, M.; MacKenzie, A. Prolactin increases SMN expression and survival in a mouse model of severe spinal muscular atrophy via the STAT5 pathway. *J. Clin. Investig.* **2011**, *121*, 3050. [CrossRef] [PubMed]
75. Grondard, C.; Biondi, O.; Armand, A.-S.; Lécolle, S.; Gaspera, B.D.; Pariset, C.; Li, H.; Gallien, C.-L.; Vidal, P.-P.; Chanoine, C.; et al. Regular Exercise Prolongs Survival in a Type 2 Spinal Muscular Atrophy Model Mouse. *J. Neurosci.* **2005**, *25*, 7622. [CrossRef] [PubMed]
76. Finkel, R.S.; Mercuri, E.; Darras, B.T.; Connolly, A.M.; Kuntz, N.L.; Kirschner, J.; Chiriboga, C.A.; Saito, K.; Servais, L.; Tizzano, E.; et al. Nusinersen versus Sham Control in Infantile-Onset Spinal Muscular Atrophy. *N. Engl. J. Med.* **2017**, *377*, 1723–1732. [CrossRef] [PubMed]
77. Baranello, G.; Darras, B.T.; Day, J.W.; Deconinck, N.; Klein, A.; Masson, R.; Mercuri, E.; Rose, K.; El-Khairi, M.; Gerber, M.; et al. Risdiplam in Type 1 Spinal Muscular Atrophy. *N. Engl. J. Med.* **2021**, *384*, 915–923. [CrossRef] [PubMed]
78. Mendell, J.R.; Al-Zaidy, S.; Shell, R.; Arnold, W.D.; Rodino-Klapac, L.R.; Prior, T.W.; Lowes, L.; Alfano, L.; Berry, K.; Church, K.; et al. Single-Dose Gene-Replacement Therapy for Spinal Muscular Atrophy. *N. Engl. J. Med.* **2017**, *377*, 1713–1722. [CrossRef]
79. Lee, B.H.; Collins, E.; Lewis, L.; Guntrum, D.; Eichinger, K.; Voter, K.; Abdel-Hamid, H.Z.; Ciafaloni, E. Combination therapy with nusinersen and AVXS-101 in SMA type 1. *Neurology* **2019**, *93*, 640–641. [CrossRef]
80. Matesanz, S.E.; Curry, C.; Gross, B.; Rubin, A.I.; Linn, R.; Yum, S.W.; Kichula, E.A. Clinical Course in a Patient With Spinal Muscular Atrophy Type 0 Treated With Nusinersen and Onasemnogene Abeparvovec. *J. Child Neurol.* **2020**, *35*, 717–723. [CrossRef]
81. Harada, Y.; Rao, V.K.; Arya, K.; Kuntz, N.L.; DiDonato, C.J.; Napchan-Pomerantz, G.; Agarwal, A.; Stefans, V.; Katsuno, M.; Veerapandiyan, A. Combination molecular therapies for type 1 spinal muscular atrophy. *Muscle Nerve* **2020**, *62*, 550–554. [CrossRef]
82. Van Alstyne, M.; Tattoli, I.; Delestrée, N.; Recinos, Y.; Workman, E.; Shihabuddin, L.; Zhang, C.; Mentis, G.; Pellizzoni, L. Gain of toxic function by long-term AAV9-mediated SMN overexpression in the sensorimotor circuit. *Nat. Neurosci.* **2021**, *24*, 930–940. [CrossRef]
83. Kirschner, J.; Butoianu, N.; Goemans, N.; Haberlova, J.; Kostera-Pruszczyk, A.; Mercuri, E.; van der Pol, W.L.; Quijano-Roy, S.; Sejersen, T.; Tizzano, E.F.; et al. European ad-hoc consensus statement on gene replacement therapy for spinal muscular atrophy. *Eur. J. Paediatr. Neurol.* **2020**, *28*, 38–43. [CrossRef] [PubMed]
84. Tizzano, E.F.; Finkel, R.S. Spinal muscular atrophy: A changing phenotype beyond the clinical trials. *Neuromuscul. Disord.* **2017**, *27*, 883–889. [CrossRef] [PubMed]
85. Pagliarini, V.; Guerra, M.; Di Rosa, V.; Compagnucci, C.; Sette, C. Combined treatment with the histone deacetylase inhibitor LBH589 and a splice-switch antisense oligonucleotide enhances SMN2 splicing and SMN expression in Spinal Muscular Atrophy cells. *J. Neurochem.* **2020**, *153*, 264–275. [CrossRef] [PubMed]

scientific reports

Check for updates

OPEN

# Three years pilot of spinal muscular atrophy newborn screening turned into official program in Southern Belgium

François Boemer[1]✉, Jean-Hubert Caberg[2], Pablo Beckers[1], Vinciane Dideberg[2], Samantha di Fiore[1], Vincent Bours[3], Sandrine Marie[4], Joseph Dewulf[4], Lionel Marcelis[5], Nicolas Deconinck[6], Aurore Daron[7], Laura Blasco-Perez[8], Eduardo Tizzano[8], Mickaël Hiligsmann[9], Jacques Lombet[10], Tatiana Pereira[10], Lucia Lopez-Granados[10], Sarvnaz Shalchian-Tehran[11], Véronique van Assche[12], Arabelle Willems[12], Sofie Huybrechts[13], Bénédicte Mast[14], Rudolf van Olden[15], Tamara Dangouloff[7] & Laurent Servais[7,16]

**Three new therapies for spinal muscular atrophy (SMA) have been approved by the United States Food and Drug Administration and the European Medicines Agency since 2016. Although these new therapies improve the quality of life of patients who are symptomatic at first treatment, administration before the onset of symptoms is significantly more effective. As a consequence, newborn screening programs have been initiated in several countries. In 2018, we launched a 3-year pilot program to screen newborns for SMA in the Belgian region of Liège. This program was rapidly expanding to all of Southern Belgium, a region of approximately 55,000 births annually. During the pilot program, 136,339 neonates were tested for deletion of exon 7 of *SMN1*, the most common cause of SMA. Nine SMA cases with homozygous deletion were identified through this screen. Another patient was identified after presenting with symptoms and was shown to be heterozygous for the *SMN1* exon 7 deletion and a point mutation on the opposite allele. These ten patients were treated. The pilot program has now successfully transitioned into the official neonatal screening program in Southern Belgium. The lessons learned during implementation of this pilot program are reported.**

**Abbreviations**

| | |
|---|---|
| ABMM | Association Belge contre les Maladies neuro-Musculaires |
| CHMP | Committee for Medicinal Products for Human Use |
| CHOP-INTEND | Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders |
| DBS | Dried blood spot |
| EMA | European Medicines Agency |
| EMG | Electromyography |
| ERB | Ethical review board |

[1]Biochemical Genetics Laboratory, Human Genetics, CHU Sart-Tilman, University of Liège, B35, 4000 Liège, Belgium. [2]Molecular Genetics Lab, Department of Human Genetics, CHU of Liège, University of Liège, Liège, Belgium. [3]Department of Human Genetics, CHU of Liège, University of Liège, Liège, Belgium. [4]Department of Laboratory Medicine, Cliniques Universitaires Saint-Luc, UCLouvain, Brussels, Belgium. [5]Laboratory of Pediatric Research, Free University of Brussels, Brussels, Belgium. [6]Neuromuscular Reference Center, Hôpital des Enfants Reine Fabiola (HUDERF), Université Libre de Bruxelles, Brussels, Belgium. [7]Division of Child Neurology, Reference Center for Neuromuscular Diseases, Department of Pediatrics, CHR Citadelle, University of Liège, Liège, Belgium. [8]Department of Clinical and Molecular Genetics, Hospital Valle Hebron, Barcelona, Spain. [9]Department of Health Services Research, CAPHRI Care and Public Health Research Institute, Maastricht University, Maastricht, The Netherlands. [10]ONE, Office de la Naissance et de l'Enfance, Herstal, Belgium. [11]Department of Neurology and Ethic Committee, CHR Citadelle, Liège, Belgium. [12]ABMM, Association Belge contre les Maladies neuro-Musculaires ASBL, La Louvière, Belgium. [13]Biogen, Cambridge, MA, USA. [14]Roche, Basel, Switzerland. [15]Novartis Gene Therapies Switzerland GmbH, Zürich, Switzerland. [16]MDUK Neuromuscular Center, Department of Paediatrics, University of Oxford, Oxford, UK. ✉email: F.Boemer@chuliege.be

| | |
|---|---|
| FDA | Food and Drug Administration |
| HINE | Hammersmith Infant Neurologic Examination |
| MLPA | Multiplex ligation-dependent probe amplification |
| NBS | Newborn screening |
| NMRC | Neuro Muscular Reference Centers |
| ONE | Office de la Naissance et de l'Enfance |
| qPCR | Quantitative polymerase chain reaction |
| RUSP | Recommended Uniform Screening Panel |
| SMA | Spinal muscular atrophy |
| SMN | Survival of Motor Neuron |
| TAT | Turnaround time |

Spinal muscular atrophy (SMA) is a neuromuscular disorder characterized by muscle atrophy resulting from the degeneration of motor neurons in the spinal cord. SMA is caused by biallelic pathogenic variants in the *SMN1* gene, which encodes Survival of Motor Neuron (SMN), a protein essential for survival of motor neurons[1]. Approximately 95% of patients carry a homozygous deletion of exon 7 in the *SMN1* gene, the remaining 5% of cases are due to the deletion of exon 7 on one allele and a deleterious variant on the opposite allele. *SMN2* is a pseudogene that differs from *SMN1* by only a few nucleotides, including a C to T transition in exon 7. This variant results in the skipping of exon 7 in about 90% of *SMN2* transcripts, thereby encoding a truncated, unstable protein. The full-length, functional SMN protein results from approximately 10% of *SMN2* transcripts. The number of *SMN2* copies is inversely correlated with the severity of the phenotype. Patients with two copies usually present with the most severe and frequent form of spinal muscular atrophy, SMA1. In these patients, symptom onset usually occurs before the age of 6 months, and this type of SMA is associated with high mortality and morbidity[2]. Patients with a larger number of copies of *SMN2* may present with symptoms long after acquisition of ambulation; a limited few even develop symptoms in adulthood. Currently, SMA is classified into four types, SMA1, SMA2, SMA3, and SMA4, based on maximal motor ability achieved.

Over the last few years, several new treatments for SMA have dramatically improved the prognosis of affected patients[3]. Nusinersen[4] was the first drug to be approved by the United States Food and Drug Administration (FDA) and the European Medicines Agency (EMA) in December 2016 and June 2017, respectively. In Belgium, nusinersen has been reimbursed by the healthcare system since September 2018. More recently, onasemnogene abeparvovec-xioi gene therapy[5] also received FDA and EMA approval, in May 2019 and May 2020 respectively. The marketing authorization of a third drug, risdiplam[6], was granted by the FDA last year, and it also received a positive opinion from the EMA's Committee for Medicinal Products for Human Use (CHMP) in February 2021. Several other drugs are currently in development[7].

Based on these recent advances in SMA management and on evidence showing that patients treated presymptomatically have better outcomes[8,9], newborn screening (NBS) for SMA has begun in several countries[10–18]. Moreover, in 2018 SMA was included in the Recommended Uniform Screening Panel (RUSP), the list of disorders that the US Department of Health and Human Services recommends be screened for as part of NBS programs[19].

In early 2018, the authors of this paper and Neuromuscular Reference Centers (NMRCs) of Southern Belgium launched a 3-year NBS pilot program for SMA under the project title "Sun May Arise on SMA". The pilot project was done in close collaboration with our industry partners AveXis, Biogen, and Roche, who funded a significant part of the program, as well as with the governmental agency in charge of NBS in Southern Belgium, the Office of Birth and Childhood (Office de la Naissance et de l'Enfance, ONE)[20,21]. It should be noted that NBS is not a federal competency in Belgium, and therefore such initiatives are conducted by a separate government agency in Northern Belgium.

The initial pilot phase of the 'Sun May Arise on SMA' project transitioned into an official program in Southern Belgium on 1 March 2021. Northern Belgium has correspondingly made a political commitment to include SMA in their official program in 2022.

This manuscript reports the key insights gained during the pilot effort.

## Results

**Inclusion of SMA in the NBS program.** The process that led to implementation of the NBS program for SMA in Southern Belgium has been previously reported[20]. A key principle was involvement of all stakeholders from the beginning. Political, ethical, and clinical partners, including genetic and screening labs, were involved in the project's governance.

**Incidence.** Over the 3-year pilot study from March 2018 to February 2021, 136,339 neonates were tested for the *SMN1* exon 7 deletion using a previously described qRT-PCR test with fluorescence read-out[20]. The dispersion plot of the ratio of *SMN1* to the housekeeping gene *RPP30* allowed clear discrimination between positive (i.e., SMA patients with a homozygous deletion of exon 7) and negative results (Fig. 1).

Nine SMA cases were identified. To our knowledge, no newborn carrying a homozygous deletion was missed over this period. All patients with symptoms of neuromuscular disease in Belgium are referred to an NMRC, thus it is quite unlikely that such a case could happen without one of the centers being informed. Nevertheless, we cannot rule out the possibility that a patient with SMA3 or SMA4 born during the period of the pilot study may be diagnosed in the future.

One SMA1 patient was not be diagnosed through NBS. The neonate was heterozygous for the *SMN1* exon 7 deletion and had the c.815A>G (p.Tyr272Cys) point mutation on the opposite allele. This patient was referred to an NMRC at the age of 4 months, after the onset of symptoms compatible with SMA.

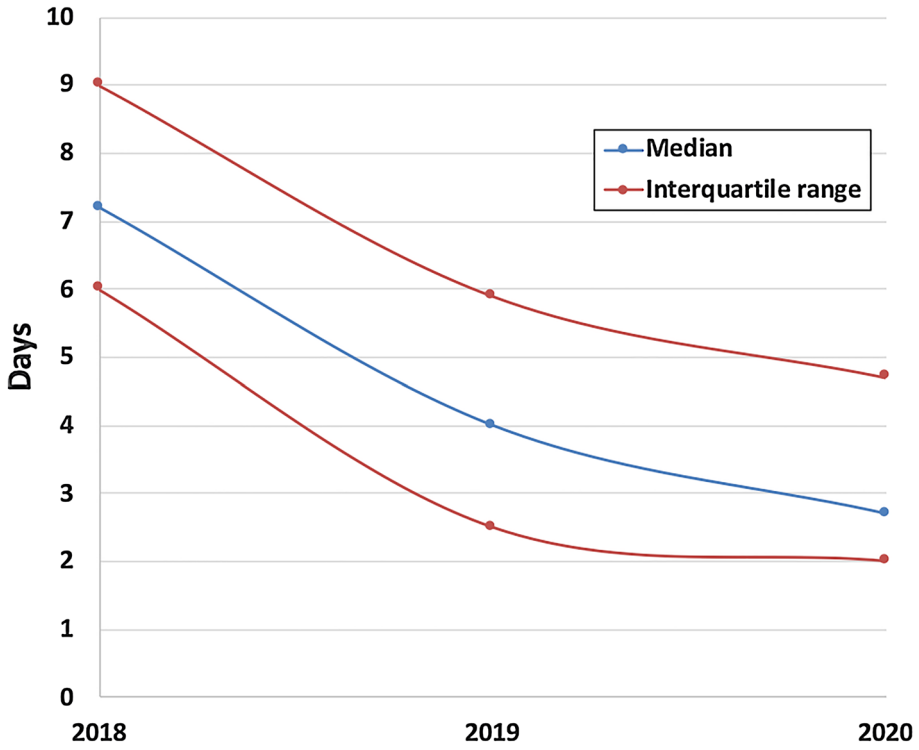



**Figure 1.** TAT improvement over the study period.

| ID | DBS sampling | DBS received by NBS center | DBS received by Liège lab | First-tier results | Second-tier results | Parents contacted | First visit | Treatment initiation | Delay between first visit and treatment initiation |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 3 | 4 | 4 | 11 | 18 | 20 | 21 | 32 | 11 |
| 2 | 3 | 8 | 8 | 27 | 30 | 30 | 31 | 38 | 7 |
| 3 | 4 | 5 | 9 | 13 | 13 | 13 | 14 | 41 | 27 |
| 4 | 4 | 13 | 19 | 27 | 27 | 31 | 32 | 54 | 22 |
| 5 | 4 | 9 | 29 | 31 | 35 | 35 | 37 | 49 | 12 |
| 6 | 3 | 4 | 11 | 18 | 22 | 20 | 21 | 39 | 18 |
| 7 | 3 | 7 | 15 | 17 | 21 | 18 | 20 | 29 | 9 |
| 8 | 3 | 5 | 15 | 18 | 19 | 22 | 23 | 32 | 9 |
| 9 | 3 | 6 | 6 | 9 | 10 | 9 | 10 | 30 | 20 |
| Median | 3 | 6 | 11 | 18 | 21 | 20 | 21 | 38 | 12 |

**Table 1.** Screening and diagnostic timeline (in post-natal days) for SMA patients identified by NBS.

This corresponds to an incidence for SMA in Southern Belgium of 1 in 13,634 newborns (95% confidence interval: 1/8417 to 1/35,858). The incidence of homozygous deletion is 1 in 15,149 individuals (95% confidence interval: 1/9163 to 1/43,696).

**Neonate referral.** Positive screening results were immediately communicated by the laboratory to both the neonate's pediatrician and to referent neurologists in NMRC. The parents were contacted on the same day by a referent neuro-pediatrician or by a pediatrician of the maternity ward and consultation was planned as soon as possible. Thanks to the second-tier MLPA testing performed on DBS-extracted DNA, the number of *SMN2* copies was available to the clinician at the patient's first visit, and therefore the clinician could immediately explain relevant therapeutic options to parents. The neonate's blood was then drawn to perform the MLPA confirmatory analysis. There were no false positives from the initial DBS testing.

The screening and diagnostic timelines for the ten SMA patients are detailed in Table 1. All nine patients identified through NBS began treatment before the age of 2 months. In order to ensure the most efficient management of patients, it is important to save time. Over the course of the project, the turnaround time (TAT) was considerably improved. For the first 9 months, the population coverage was limited to Liège NBS center, where about 300–350 samples were analyzed each week. The median TAT, calculated for the interval between DBS

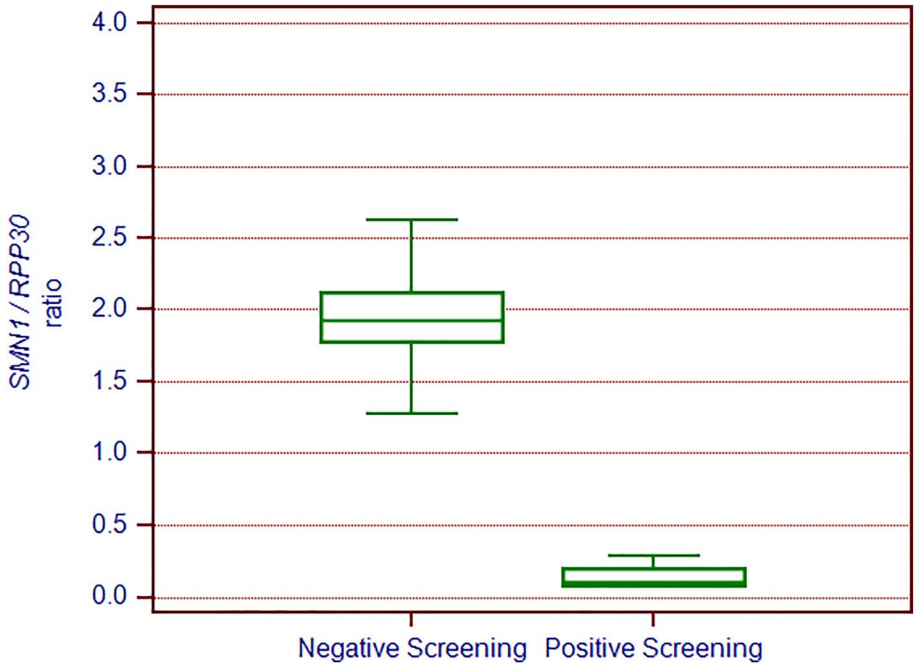



**Figure 2.** Box-and-whisker plot of the endpoint-fluorescence *SMN1* to *RPP30* ratio for negative (n = 136.330) and positive (n = 9) screening results.

reception in Liège's center and validation of the result, was 7.2 days (interquartile range: 6.0–9.0 days). At the beginning of 2019, the other two NBS centers in Southern Belgium joined the project, outsourcing their analytical process to Liège's center, and the number of samples analyzed increased to approximately 1200 samples per week. Early in 2019, acquisition of a dedicated qPCR instrument and hiring of a devoted lab technician permitted a considerable scale-up of our analytical throughput. Subsequently, TAT was reduced from 7.2 days in 2018 (interquartile range: 6.0–9.0 days) to 4.0 days later in 2019 (interquartile range: 2.5–5.9 days) and to 2.7 days in 2020 (interquartile range: 2.0–4.7 days) (Fig. 2).

**Patient treatment and outcomes.** Parents were informed about the different therapeutic options during first visit. Nusinersen was available in Belgium from the start of the study. Risdiplam and the gene therapy onasemnogene abeparvovec-xioi were not commercially available in the country during the pilot study but were accessible through several concurrent clinical trials in NMRC (Spr1nt: NCT03505099, STRIVE-EU: NCT03461289, Rainbowfish: NCT03779334). For the six patients who received nusinersen, treatment began an average of 10 days after the first consultation (7–20). Parents of Patient 9 initially refused the treatment, which explains the delay in initiation. The delay between the first consultation and the initiation of treatment was the longest for the three patients who participated in the therapeutic trials (18, 22, and 27 days) as participation in a trial required testing prior to inclusion. Patients who showed early clinical manifestations of the disease, even if weak (i.e., only areflexia), were those who had two copies of *SMN2*. These patients had developmental delays despite treatment. Patients with three or four copies of *SMN2* showed no symptoms at the time of treatment initiation and hit motor developmental milestones at the usual ages. *SMN2* copy number and modifier variants, treatment regimen, and evolution of symptoms in identified patients are summarized in Table 2.

**Lessons learned from individual cases.** *The case of treatment refusal.* The parents of one patient initially refused treatment. The child had three copies of *SMN2* and was asymptomatic at the time of diagnosis. The parents were not French speakers, and at the initial consultation were accompanied by a French-speaking cousin serving as a translator. This was not an optimal situation, as the translator was emotionally invested and only partially translated the physician's explanation to the parents. Following their refusal, they were offered a second consultation with two different child neurologists and a psychologist with a professional translator in attendance, and a further consultation was also proposed with a German-speaking neurologist. The parents stated several times that they would prefer to wait for their daughter to present with symptoms before discussing treatment. This prompted internal discussions among the clinical team to balance the right of parents to make decisions regarding the care of their child with the rights of the child given that clinical evidence clearly indicates that treatment before symptom onset is necessary to ensure the possibility of normal development[8,9].

After requesting several external medical and external opinions, we explained to the parents that the clinical team could not carry the responsibility of withholding care, and that the family court would have to be consulted. After receiving initial opinions from the prosecutor supportive of intervention, the parents accepted the necessity of treatment. Interestingly, the relationship between the clinical care team and the family remained positive,

| Id | Sex | *SMN2* copy number | *SMN2 polymorphism* | | Treatment | Treatment initiation in days | Phenotype at treatment start | Sitter (in months) | Walker (in months) | Age at last assessment (in months) | Max score on CHOP-INTEND scale[c] | Max score on HINE 2 scale[d] |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | c.859G > C | c.835-44A > G | | | | | | | | |
| 1 | M | 3 | Negative | Negative | Nusinersen | 32 | Asymptomatic | 7 | 13 | 33 | 64 | 26 |
| 2 | F | 2 | Negative | Negative | Nusinersen | 38 | Areflexia, discrete hypotonia, | 7 | 27 with help | 32 | 58 | 24 |
| 3 | M | 3 | Negative | Negative | OA[b] | 41 | Asymptomatic | 7 | 15 | 24 | 64 | 24 |
| 4 | M | 2 | / | / | OA[b] | 54 | Discrete hypotonia | 6,5 | Stand up alone | 22 | 51 | 20 |
| 5 | M | 4 | Negative | Negative | Nusinersen | 49 | Asymptomatic | 6 | 12 | 22 | 64 | – |
| 6 | F | 4 | Negative | Negative | Risdiplam | 39 | Asymptomatic | 5 | 12 | 20 | 64 | 26 |
| 7 | M | 2 | Negative | Negative | Nusinersen | 29 | Areflexia | 6 | No | 18 | 60 | 17 |
| 8 | M | 2 | Negative | Negative | Nusinersen | 32 | Areflexia | 6 | No | 14 | 54 | – |
| 9 | F | 3 | Negative | Negative | Nusinersen | 30 | Asymptomatic | 7 | 11 | 12 | 62 | 21 |
| 10[a] | M | 2 | / | / | Nusinersen | 150 | Proximal hypotonia, areflexia, tongue fasciculations | No | No | 17 | 34 | 2 |

**Table 2.** *SMN2* copy number and polymorphisms, treatment, and evolution of symptoms of SMA patients identified during the study period. [a]Compound heterozygous patient identified at the age of 4 months. [b]Onasemnogene abeparvovec-xioi. [c]CHOP-INTEND maximum score is 64. [d]HINE Sect. "Results" maximum score is 26. A dash indicates that the test was not given.

and 1 year after birth the mother stated that they had been in such an emotional state that they were 'unable to make the right decision' and now recognized that treatment was the best solution.

No other parents refused treatment. Some parents indicated their preference for a particular treatment. The choice to proceed with a treatment was always made in light of treatment availability, the child's clinical condition, and the scientific data available at the time, and with the mutual agreement of the treating physicians and the parents.

*Patients and siblings with four copies of SMN2.* As mentioned earlier, treatment of children is specifically discussed with the parents. In the two cases with four copies of *SMN2* identified during the pilot study, the parents promptly agreed to the proposal to initiate early treatment.

One of the patients identified with four copies of *SMN2* had two older siblings, aged 4 years and 6 years and 6 months, respectively. Interestingly, the mother presented with two copies of *SMN1* and the father with one copy. We then discovered that the maternal grandmother had three copies of *SMN1*, two on the same chromosome, and the paternal grandmother had only one copy. The mother was 2/0, which means that she would not have been identified as at-risk during carrier testing.

The initial clinical examination of the siblings of the patient indicated normal development, but the parents wished to have them tested. This was done, and we found that, like the infant, both children had the homozygous deletion of exon 7 of *SMN2* and four copies of *SMN2*. Their parents opted to delay treatment. Further evaluations of the siblings were performed after 3 months.

The physician had concerns regarding the potential muscle weakness of the older sibling, but the parents again opted to delay treatment. When the child was aged 7 years and 4 months, a video sent by the parents clearly confirmed a proximal weakness and fatigability. On examination, there was an absence of patellar reflex, and the need for the child to support himself with a hand on his leg when rising from the floor. The motor function measure and six-minute walk test were stable. The parents refused to treat at this stage.

At 7 years and 11 months, the electromyography (EMG) showed a 30% loss of motor amplitude. At 8 years, the same difficulties at the clinical examination were noticed with a complete absence of reflexes, and unchanged compound muscle action potential.

The second sibling, who was 4 years old at the time of diagnosis, showed no deficit in either the clinical examination, physiological tests or EMG. Follow-up is continuing with clinical and physiotherapy examinations every 6 months. To date, at the age of 5 years and 6 months, the second child is still wholly asymptomatic.

**Transition to health authorities: a strong partnership among stakeholders.** Retrospectively, the key element in the successful transition from the trial project to a government-sanctioned public health program was the involvement and unanimous support of all stakeholders from the beginning of the project and throughout its duration. Transitioning to an official program was an initial objective of the pilot program. The involvement of patient advocacy groups, neuromuscular reference centers, and newborn screening centers, as well as

| Screening period | 03/2018–02/2019 | 03/2019–02/2020 | 03/2020–02/2021 |
|---|---|---|---|
| Number of screened newborns | 22,930 | 57,607 | 55,802 |
| Expected number of SMA cases (λ) | 1.51 | 3.80 | 3.68 |
| Probability of 0 cases during period | 0.220 | 0.022 | **0.025** |
| Probability of 3 cases during period | **0.127** | 0.204 | 0.209 |
| Probability of 6 cases during period | 0.004 | **0.094** | 0.087 |

**Table 3.** Poisson probability of case occurrence in Southern Belgium based on annual periods. Bold values correspond to the number of SMA cases actually identified during the designated period.

public engagement through broadcast and social media (such as on the study's Facebook page, www.facebook.com/sunmayariseonsma) also significantly facilitated the rapid and smooth transition to an official program.

A clear governance structure helped to build a strong partnership between pilot study leaders, the regional agency in charge of NBS, and NBS centers. Public involvement gave rise to support from across the political spectrum in Belgium. The ordinance incorporating SMA into the NBS list for Southern Belgium was passed by the Parliament of Wallonia on 4 February 2021 for implementation on 1 March 2021, with immediate handover from the study team to the public health service after the completion of the 3-year pilot project. UCLouvain and ULBruxelles NBS centers are incorporating the SMA screening test into their own infrastructure.

## Discussion

The incidence of SMA of 1 in 15,149 determined during the NBS pilot study in Southern Belgium is broadly consistent with previous studies. The incidence reported in Taiwan was 1 in 17,181 neonates[12]. In Germany, 30 SMA cases were identified during screening of 213,279 DBS cards for a incidence of 1 in 7109 infants[17,22]. Australian NBS has identified nine SMA patients in 103,903 newborns screened for an incidence of 1 per 11,544[18]. New York State recently screened more than 225,000 neonates and reported a much lower incidence of 1 per 28,137[23]. The authors of that study argued that the low SMA incidence reported in their area is likely due to biased estimates, coupled with increased awareness and access to carrier screening, genetic counselling, cascade testing, prenatal diagnosis, and advanced reproductive technologies. A better understanding of this low incidence is of primary importance since it could have consequences on reimbursement for disease-modifying therapies and NBS funding decisions[24].

Surprisingly, we did not identify any SMA neonates during the third year of our pilot study. Based on the Poisson distribution of rare events, the probability of diagnosing no cases of SMA over 1 year is 2.5% (Table 3). Given the low probability that there should be no cases in a year, we hypothesized that carrier screening and prenatal testing had contributed to this outcome. We therefore contacted various molecular genetics centers in Southern Belgium to request the number of positive results for SMA based on pre-conceptional and prenatal diagnosis during the corresponding period. However, they reported no positive results that could explain this absence of cases over the previous year. Subsequently, three new cases were identified in the first 4 months following the end of the pilot, which further reinforces the hypothesis of a pure random distribution.

Our study is, to our knowledge, the first to report a SMA patient compound heterozygous for the *SMN1* exon 7 deletion and a point mutation on the opposite allele, in the context of NBS. Because the first-tier assays specifically target the homozygous *SMN1* deletion, this patient was not be identified during the screening process. Rather, the patient was identified at the age of 4 months, after referral for mild hypotonia. The clinical sensitivity of SMA NBS is estimated between 95 and 98%, as affected individuals who are compound heterozygotes (i.e., those with one *SMN1* allele lacking exon 7 and a point mutation on the second allele) are missed[11,25]. To date, no false negatives or false positives have been identified in our screening program.

The five neonates with either three or four copies of *SMN2* were all asymptomatic at treatment start (Table 2). Most presented the highest Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders (CHOP-INTEND) and Hammersmith Infant Neurologic Examination, Sect. "Results" (HINE-2) scores during their last motor assessment (age range: 12–33 months). The four newborns with two copies of *SMN2* showed a slight hypotonia and/or a discrete areflexia when the treatment was initiated. These patients did not get the highest scores on CHOP-INTEND and HINE-2 scales during their last motor assessment (age range: 14–32 months). Of these four patients, three were treated with the approved nusinersen therapy. Treatment initiation may thus be considered as relatively delayed (range: 29–54 days) when compared to first visit (range: 20–32 days). This lag may be a factor that has impaired the most favorable outcome for these patients. In the future, we hope that the recent transition of our pilot study into the official neonatal screening program will facilitate a more prompt care.

The overall evidence for the efficacy of early treatment of patients with SMA has been recently reviewed[26]. It is likely that the cost of the new SMA treatments initially hampered the implementation of NBS programs by the political authorities. Presently, the substantial cost burden of standard care for patients with SMA is estimated to be between US$ 75,047 and US$ 196,429 per year for SMA1 patients, and between US$ 27,157 and US$ 82,474 for other types of SMA[27]. Therefore, given the high cost-to-benefit ratio of drugs approved at current prices when administered to post-symptomatic patients[27], we know it is critical to identify patients prior to symptom onset. A medico-economic evaluation with assessment of patient quality of life is also currently ongoing to assess the cost-effectiveness of our NBS program[20]. Pre-treatment levels of phosphorylated neurofilaments are a validated marker of nerve cell damage in pre-symptomatic and in young SMA1 patients[28]. These levels decrease exponentially in pre-symptomatic SMA patients with two *SMN2* copies, indicating acute and severe neuronal loss[9].

These data indicate that it is critical to begin treatment of SMA1 patients with as little delay as possible. An NBS program is accordingly an ideal method for early identification of these infants.

There were several incidents encountered during this pilot program, the description of which may help other NBS programs more effectively communicate with the parents of recently diagnosed infants.

In one case, parents initially refused treatment. In hindsight, this might have been avoided if a professional translator had been present during the first consultation. In another case, three SMA-affected children of a mother with two copies of *SMN1* on the same allele were diagnosed as a result of NBS: the youngest through the NBS pilot program itself and his siblings following this initial positive identification. As the mother would not have been identified as at-risk during carrier testing, this clearly indicates that carrier screening should not be relied upon as the sole strategy against SMA.

Finally, we were faced with a case of a patient with symptoms that the parents refused to recognize. Political authorities must therefore put plans in place to deal with cases of refusal of treatment. Presently, some countries leave the decision of treatment to a multidisciplinary consultation meeting, whereas others leave all choice to the parents. The present authors believe that the interest of the child must take priority over parents' rights. A collegial discussion of these potential issues prior to implementation of an NBS program is necessary.

Our study suffers from the small size of the studied population. Southern Belgium has a total population of approximately 4.5 million people; therefore the number of cases identified in the neonate population remains low.

Today, nine countries around the world have started SMA NBS, with the number of newborns screened set to increase in the coming years as further countries embark on similar programs[29]. Our project confirms that a pilot program can be rapidly transitioned into the official NBS program. Given the effective treatments now available for SMA and the importance of treatment prior to the onset of symptoms, testing for SMA should be incorporated into screening of all newborns.

## Materials and methods

**Newborn samples.** NBS samples were collected on Whatman® 903 cards between 48 and 120 h of life either in maternity wards or at home, in accordance with legal requirements of the federal authority (Wallonie–Bruxelles Federation) in charge of NBS in Southern Belgium.

The dried blood spot (DBS) cards were sent to selected neonatal screening laboratories. No additional sampling was required to incorporate SMA testing in the standard NBS panel as the residual blood spots collected for conventional NBS were sufficient to test for SMA. After analysis, filter papers are stored at room temperature for 5 years.

As detailed in our previous manuscript[20], parental consent was not required for participation in this study. While strongly recommended, NBS is not mandatory in Southern Belgium and parents are informed that they have the right to refuse screening for their child. This opt-out option is not disease-specific; it applies to the neonatal screening panel as a whole. The project was approved by our ethical review board (reference number B412201734396), in accordance with the Declaration of Helsinki.

**NBS assay and confirmatory method.** The flow chart for screening for SMA is shown in Fig. 3. We designed a quantitative polymerase chain reaction (qPCR) assay to specifically detect homozygous deletions of *SMN1* exon 7 on DNA extracted from DBS[20]. DNA extraction was performed by alkaline denaturation at 98 °C. qPCR amplification was performed in 96-well plates, preloaded with primers, dye-labeled probes, and master mix provided by Eurogentec. This assay cannot identify heterozygous carriers of the deletion of exon 7 or *SMN1* point mutations, and the number of copies of *SMN2* were not determined in this first-tier assay. Given the importance of *SMN2* copy number in SMA management, qPCR-positive results were confirmed by the multiplex ligation-dependent probe amplification (MLPA) technique, which also provided information on *SMN2* status. For this purpose, we used the Salsa MLPA Probemix P021 SMA diagnostic kit (MRC Holland).

First-tier positive samples were re-analyzed twice from the same DBS. Simultaneously, a second-tier MLPA assay was performed from the same DNA extracted for the first-tier qPCR. Upon positive results from confirmatory testing, neonates were immediately referred to a neuro-pediatrician in one of the NMRCs involved in the trial. At the first visit, fresh blood was collected to confirm the positive screening result by MLPA on an independent sample. Additionally, we also sequenced the *SMN2* gene to look for the presence of both c.859G>C and c.835-44A>G intragenic modifier variants. A *SMN2*-specific PCR has been used to amplify exons 7 and 8 and study the presence or absence of the positive modifier variants. The primers (available on request) were designed based on the paralogous sequence variants described by Blasco-Pérez et al.[30], in order to achieve specificity towards *SMN2* (Blasco-Perez et al., in preparation).

**Population coverage.** There are approximately 55,000 annual births in Southern Belgium, and NBS for these infants is carried out by three independent academic centers. The current project was launched in March 2018 in Liège's NBS laboratory, which screens about 16,000 newborns per year. Due to strong support from the supervisory authorities and the efforts of the project management team to promote the project, the pilot study rapidly expanded to include the two other screening centers of Southern Belgium, UCLouvain and ULBruxelles. In order to rapidly implement the program in these two centers, DNA was extracted in the lab to which the DBS card was sent. Sealed microtiter plates containing samples for SMA screening were then transferred to the lab in Liège, which ran qPCR assays on all samples. SMA screening was offered to the entire neonate population of Southern Belgium beginning in early 2019.

**Clinical and therapeutic protocol.** All patients were examined by board certified neuro-paediatricians with expertise in SMA. The different therapeutic options were proposed to parents during the first visit. The

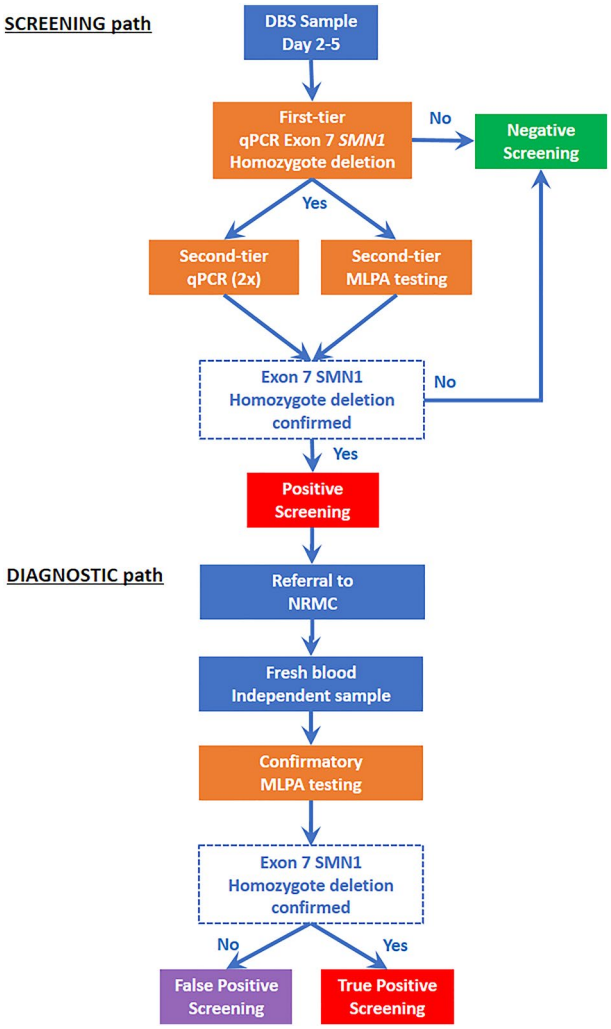



**Figure 3.** Screening and diagnostic flowchart.

phenotype at the start of treatment and the ages of sitting and walking acquisitions were recorded. Longitudinal motor milestone assessment was evaluated by trained physiotherapists, using CHOP-INTEND and HINE-2 scales.

**Statistical analyses.** Exact probability of rare event occurrence was estimated by a Poisson distribution in which the probability mass function is $p(x) = e^{-\lambda} \cdot \lambda^x / x!$, where $\lambda$ is the average number of events per year, and x is number of events in each interval.

**Ethics approval.** Ethical approval (reference B412201734396) was obtained from the Institutional Review Board (Ethical Committee of the Hospital CHR Citadelle, Liège, Belgium) in compliance with the Declaration of Helsinki.

## Data availability

The data that support the findings of this study are available from the corresponding author, FB, upon reasonable request.

Received: 13 April 2021; Accepted: 27 September 2021
Published online: 07 October 2021

## References

1. Lefebvre, S. *et al.* Identification and characterization of a spinal muscular atrophy-determining gene. *Cell* **80**, 155–165 (1995).
2. Kolb, S. J. *et al.* Natural history of infantile-onset spinal muscular atrophy. *Ann. Neurol.* **82**, 883–891 (2017).
3. Ramdas, S. & Servais, L. New treatments in spinal muscular atrophy: An overview of currently available data. *Expert Opin. Pharmacother.* **21**, 307–315 (2020).

4. Finkel, R. S. *et al.* Nusinersen versus sham control in infantile-onset spinal muscular atrophy. *N. Engl. J. Med.* **377**, 1723–1732 (2017).
5. Mendell, J. R. *et al.* Single-dose gene-replacement therapy for spinal muscular atrophy. *N. Engl. J. Med.* **377**, 1713–1722 (2017).
6. Baranello, G. *et al.* Risdiplam in type 1 spinal muscular atrophy. *N. Engl. J. Med.* **384**, 915–923 (2021).
7. Servais, L., Baranello, G., Scoto, M., Daron, A. & Oskoui, M. Therapeutic interventions for spinal muscular atrophy: Preclinical and early clinical development opportunities. *Expert Opin. Investig. Drugs* https://doi.org/10.1080/13543784.2021.1904889 (2021).
8. Govoni, A., Gagliardi, D., Comi, G. P. & Corti, S. Time is motor neuron: Therapeutic window and its correlation with pathogenetic mechanisms in spinal muscular atrophy. *Mol. Neurobiol.* **55**, 6307–6318 (2018).
9. De Vivo, D. C. *et al.* Nusinersen initiated in infants during the presymptomatic stage of spinal muscular atrophy: Interim efficacy and safety results from the Phase 2 NURTURE study. *Neuromuscul. Disord.* **29**, 842–856 (2019).
10. Pyatt, R. E. & Prior, T. W. A feasibility study for the newborn screening of spinal muscular atrophy. *Genet. Med.* **8**, 428–437 (2006).
11. Prior, T. W. *et al.* Newborn and carrier screening for spinal muscular atrophy. *Am. J. Med. Genet. Part A* **152**, 1608–1616 (2010).
12. Chien, Y.-H. *et al.* Presymptomatic diagnosis of spinal muscular atrophy through newborn screening. *J. Pediatr.* **190**, 124-129.e1 (2017).
13. Kraszewski, J. N. *et al.* Pilot study of population-based newborn screening for spinal muscular atrophy in New York state. *Genet. Med.* **20**, 608–613 (2018).
14. Dangouloff, T. *et al.* 244th ENMC international workshop: Newborn screening in spinal muscular atrophy May 10–12, 2019, Hoofdorp, the Netherlands. *Neuromuscul. Disord.* **30**, 93–103 (2020).
15. McMillan, H. J. *et al.* Newborn screening for spinal muscular atrophy: Ontario testing and follow-up recommendations. *Can. J. Neurol. Sci.* (2020). https://doi.org/10.1017/cjn.2020.229
16. Shinohara, M. *et al.* A novel system for spinal muscular atrophy screening in newborns: Japanese pilot study. *Int. J. Neonatal Screen.* **5**, 41 (2019).
17. Vill, K. *et al.* One year of newborn screening for SMA—results of a German pilot project. *J. Neuromuscul. Dis.* **6**, 503–515 (2019).
18. Kariyawasam, D. S. T. T., Russell, J. S., Wiley, V., Alexander, I. E. & Farrar, M. A. The implementation of newborn screening for spinal muscular atrophy: The Australian experience. *Genet. Med.* **22**, 557–565 (2020).
19. Kemper, A. R. *et al.* Evidence-based review of newborn screening for spinal muscular atrophy (SMA): Final report (v5.2). *USA Matern. Child Heal. Bur. United States Secr. Heal. Hum. Serv. Advis. Comm. Heritable Disord. Newborns Child.* **5**, 108 (2018).
20. Boemer, F. *et al.* Newborn screening for SMA in Southern Belgium. *Neuromuscul. Disord.* **29**, 343–349 (2019).
21. Boemer, F. *et al.* (S)un (M)ay (A)rise on SMA: The hope of a region without spinal muscular atrophy. *Rev. Med. Liège* **74**, 461 (2019).
22. Czibere, L. *et al.* High-throughput genetic newborn screening for spinal muscular atrophy by rapid nucleic acid extraction from dried blood spots and 384-well qPCR. *Eur. J. Hum. Genet.* **28**, 23–30 (2020).
23. Kay, D. M. *et al.* Implementation of population-based newborn screening reveals low incidence of spinal muscular atrophy. *Genet. Med.* **22**, 1296–1302 (2020).
24. Dangouloff, T., Boemer, F., Caberg, J.-H. & Servais, L. Correspondence on: "Discrepancy in Spinal Muscular Atrophy Incidence findings in newborn screening programs: The influence of carrier screening?" by Kay et al. *Genet. Med.* (2020). https://doi.org/10.1038/s41436-020-0887-1
25. Wirth, B. An update of the mutation spectrum of the survival motor neuron gene (SMN1) in autosomal recessive spinal muscular atrophy (SMA). *Hum. Mutat.* **15**, 228–237 (2000).
26. Dangouloff, T. & Servais, L. Clinical evidence supporting early treatment of patients with spinal muscular atrophy: Current perspectives. *Ther. Clin. Risk Manag.* **15**, 1153–1161 (2019).
27. Dangouloff, T., Botty, C., Beaudart, C., Servais, L. & Hiligsmann, M. Systematic literature review of the economic burden of spinal muscular atrophy and economic evaluations of treatments. *Orphanet J. Rare Dis.* **16**, 47 (2021).
28. Darras, B. T. *et al.* Neurofilament as a potential biomarker for spinal muscular atrophy. *Ann. Clin. Transl. Neurol.* **6**, 932–944 (2019).
29. Dangouloff, T., Vrščaj, E., Servais, L., Osredkar, D. & Group, S. N. W. S. Newborn screening programs for spinal muscular atrophy worldwide: Where we stand and where to go. *Neuromusc. Disord.* **31**(6), 574–582 (2021).
30. Blasco-Pérez, L. *et al.* Beyond copy number: A new, rapid, and versatile method for sequencing the entire SMN2 gene in SMA patients. *Hum. Mutat.* **42**, 787–795 (2021).


## Acknowledgements

Authors thank the paramedical teams of Neuromuscular Centers of CHU of Liège and the technical teams of newborn screening and molecular laboratories of CHU of Liège. We thank Dominic Tromans for the English edition of the manuscript.

## Author contributions

F.B. wrote the manuscript, and contributed to study design and to method development. J.H.C. contributed to study design and to method development; he reviewed the manuscript. P.B. and V.D. contributed to method development and reviewed the manuscript. S.D.F., L.B.P. and E.T. performed technical experiments and revised technical aspects of the manuscript. V.B. provided genetic advice on method development and reviewed the manuscript. S.M., J.D., and L.M. contributed to sample collection and reviewed the manuscript. N.D., A.D., S.S.T., V.V.A., and A.W. ensured SMA patient follow-up and reviewed the manuscript. M.H., S.H., B.M., and R.V.O. contributed to study design and reviewed the manuscript. T.D. is the project manager; she contributed to the study design, ensured SMA patient follow-up, and reviewed the manuscript. L.S. is the project leader; he contributed to study design, ensured SMA patient follow-up, and reviewed the manuscript. All authors reviewed and approved the final manuscript.

## Funding

This pilot study is supported by AveXis, Biogen, Roche, the ABMM (Association Belge contre les Maladies neuro-Musculaires), Minister's Office Alda GREOLI (Wallonia-Brussels Community) and donations from individuals.

## Competing interests

L.S. is member of Biogen, AveXis, Roche, and Cytokinetics scientific advisory boards and has provided consultancy to Roche, AveXis, and Biogen. E.T. has received grant support to conduct clinical trials on SMA from Ionis/Biogen and serves as a consultant to AveXis, Novartis, Biogen, Biologix, Cytokinetics, Roche. A.D. is investigator of SMA studies for Roche and Novartis Gene Therapies and received honoraria for consultancy for the scientific advisory board of AveXis Belgium. B.M. is an employee of Roche. S.H. is an employee of and has stock/stock

options in Biogen. R.V.O. is employee of Novartis Gene Therapies and owns Novartis stock or other equities. T.D. and F.B. have given lectures sponsored by Biogen, Novartis and Roche. The other authors have no financial disclosures relevant to this article.

## Additional information

**Correspondence** and requests for materials should be addressed to F.B.

**Reprints and permissions information** is available at www.nature.com/reprints.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2021

genes

*Review*

# Recommendations for Interpreting and Reporting Silent Carrier and Disease-Modifying Variants in SMA Testing Workflows

**John N. Milligan [1,*], Laura Blasco-Pérez [2], Mar Costa-Roger [2], Marta Codina-Solà [2] and Eduardo F. Tizzano [2]**

[1] Asuragen, A Bio-Techne Brand, Austin, TX 78721, USA
[2] Department of Clinical and Molecular Genetics, University Hospital Vall d'Hebron, 08035 Barcelona, Spain
* Correspondence: john.milligan@bio-techne.com

**Abstract:** Genetic testing for SMA diagnosis, newborn screening, and carrier screening has become a significant public health interest worldwide, driven largely by the development of novel and effective molecular therapies for the treatment of spinal muscular atrophy (SMA) and the corresponding updates to testing guidelines. Concurrently, understanding of the underlying genetics of SMA and their correlation with a broad range of phenotypes and risk factors has also advanced, particularly with respect to variants that modulate disease severity or impact residual carrier risks. While testing guidelines are beginning to emphasize the importance of these variants, there are no clear guidelines on how to utilize them in a real-world setting. Given the need for clarity in practice, this review summarizes several clinically relevant variants in the *SMN1* and *SMN2* genes, including how they inform outcomes for spinal muscular atrophy carrier risk and disease prognosis.

**Keywords:** spinal muscular atrophy; carrier screening; diagnosis; *SMN1*; *SMN2*

check for updates

**Citation:** Milligan, J.N.; Blasco-Pérez, L.; Costa-Roger, M.; Codina-Solà, M.; Tizzano, E.F. Recommendations for Interpreting and Reporting Silent Carrier and Disease-Modifying Variants in SMA Testing Workflows. *Genes* **2022**, *13*, 1657. https://doi.org/10.3390/genes13091657

Academic Editor: Christopher Grunseich

Received: 1 September 2022
Accepted: 13 September 2022
Published: 15 September 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

## 1. Spinal Muscular Atrophy Disease Etiology

Spinal muscular atrophy (SMA) is an autosomal recessive neuromuscular disease caused by loss of survival motor neuron 1 (*SMN1*) gene function and is a primary genetic cause of infant death [1]. SMA is a rare disease with a pan-ethnic incidence of ~1/11,000 live births and a high carrier rate of ~1/54 [2]. SMA is divided into clinical types based on the age of onset and maximum motor milestone achievement, with a gradient of phenotypes ranging from never sitting unassisted, with onset prior to six months of age, to adult-onset mild muscular weakness. Most SMA patients are classified into three main types in order of decreasing severity: type 1 (~60% of patients), type 2 (~30% of patients), and type 3 (~10% of patients). Rarer SMA types, such as type 0 and type 4, also exist [3–5].

Bi-allelic loss of the *SMN1* gene is the cause of disease in ~95% of patients with SMA. The remaining 5% of patients are compound heterozygotes, with an *SMN1* deletion on one chromosome and a loss-of-function point mutation in *SMN1* on the other chromosome. The vast majority (~98%) of SMA patients inherit the *SMN1* alterations from their parents [6,7]. SMA carriers lack a functional *SMN1* copy on a single chromosome and frequently have one functional copy on the other (1 + 0). However, a *cis* carrier genotype with two *SMN1* copies on a single chromosome (2 + 0), commonly referred to as a silent carrier, is also well -documented [8]. In one study examining a large North American population, the detection rate of SMA carriers using *SMN1* copy number alone varied from ~71% to 95% depending on ethnicity [9]. Most of the missed carriers were due to silent carriers (2 + 0) that cannot be resolved from wild-type (1 + 1) individuals solely based on copy number, since results would be 2 *SMN1* copies for both genotypes [9]. While gene conversion from *SMN2* to *SMN1* is known to occur and is one potential cause for the silent carrier (2 + 0) genotype [8], the clinical significance of gene conversions is not fully understood. Recent studies have shown that variants c.*3+80T>G and c.*211_*212del in *SMN1* (Figure 1A) are associated

with *SMN1* duplication in many ethnic groups and their presence informs the risk of silent carrier *SMN1* genotypes (2 + 0) to varying degrees depending on ethnicity [10,11].

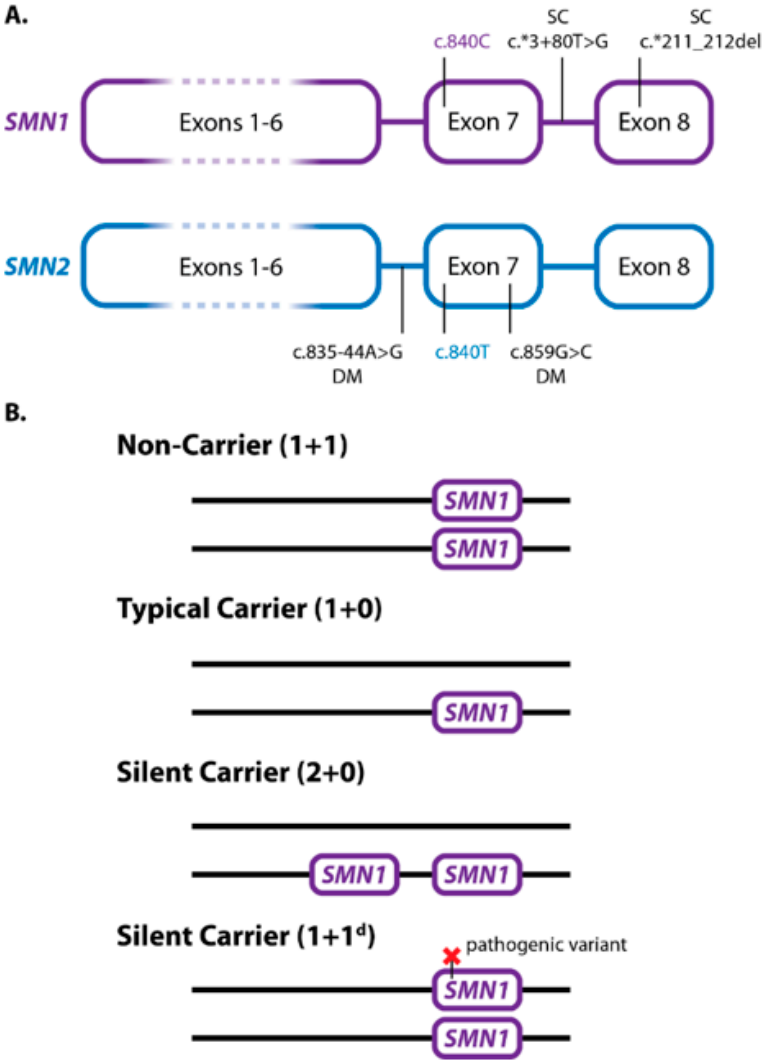



**Figure 1.** Genetics of *SMN1*, *SMN2*, and SMA Carriers. (**A**) Silent carriers and disease-modifying variants in *SMN1* and *SMN2*. Nucleotides at position c.840 in exon 7, typically used to distinguish *SMN1* and *SMN2*, are indicated by color (PSVs). Gene duplication variants in *SMN1* associated with 2 + 0 silent carriers are indicated by the letters SC. Common disease modifier variants in *SMN2* are indicated by the letters DM. (**B**) SMA carrier genetics. Non-carriers typically have one copy of *SMN1* on each chromosome. Typical carriers have only one *SMN1* copy, lacking *SMN1* on the other chromosome. Silent carriers (2 + 0) often have two copies of *SMN1* on a single chromosome, lacking *SMN1* on the other chromosome. Silent carriers can also have one copy of *SMN1* on both chromosomes but with a pathogenic variant in one copy.

SMA disease severity inversely correlates with *SMN2* copy number, meaning the more copies of *SMN2*, the less severe the phenotype [5]. *SMN1* and *SMN2* differ in 16 paralogue sequence variants (PSVs) [12]. One PSV, c.840C>T, disrupts a splice enhancer that decreases the number of exon 7 containing mRNAs to 10–20%, which results in a significantly reduced amount of functional SMN protein compared to that made from a functional *SMN1* gene. However, due to complete homology with the *SMN1*-associated SMN protein sequence, *SMN2*-generated SMN protein levels offer a compensatory effect, thus resulting in lessened disease severity with increased *SMN2* copies. Though the *SMN2* copy number is vital for assessing disease severity, there are also a few variants known to be SMA disease

modifiers. Specifically, c.859G>C in *SMN2* (Figure 1A) is linked to improved splicing efficiency of *SMN2* by 20%, which also leads to reduced disease severity [13,14]. Indeed, 44 SMA patients carrying the c.859G>C variant have been described, all of whom presented a milder phenotype than expected according to their *SMN2* copies. This variant has been described in various populations, showing a common haplotype that points towards a common ancestral origin [12]. Thus, *SMN1* is associated with molecular SMA diagnosis and carrier status, whereas *SMN2* is associated with the severity of the disease.

## 2. SMA Diagnostic and Carrier Screening Testing

Copy number analysis for *SMN1* and *SMN2* genes associated with SMA can be difficult, as the copy number of these varies much more than other regions within the genome. Furthermore, rapid turnaround time for SMA diagnostic testing is important for timely administration of therapies which halt neuron degeneration [15,16]. SMA genetic testing for *SMN1* and *SMN2* exon 7 copy numbers is accomplished using a variety of methods, including PCR followed by capillary electrophoresis (PCR/CE), quantitative PCR (qPCR), digital droplet PCR (ddPCR), multiplex ligation-dependent probe amplification (MLPA), and next-generation sequencing (NGS). These methods have recently been extensively described, including the strengths and weaknesses of each approach [17]. PCR-based systems are generally the fastest and simplest methods, though qPCR and ddPCR assays require separate reactions for each gene, and qPCR requires the generation of a standard curve, which can limit throughput. MLPA and PCR/CE both provide copy numbers for *SMN1* and *SMN2*, but MLPA has a longer and more complex workflow, requiring at least 24 h to complete as compared to PCR/CE, which can be completed in a few hours [17]. While PCR/CE is restricted to quantifying exon 7 and intron 7 from *SMN1* and *SMN2*, MLPA quantifies all exons in these genes, which can reveal partial gene deletions. NGS provides the most comprehensive analysis for variants, hybrid genes, and partial deletions, but the workflow can be laborious, time-intensive, and requires complex instrumentation. Furthermore, NGS analysis and interpretation requires significant hardware resources and bioinformatics expertise, especially for *SMN1* and *SMN2* analysis, given the high homology between the genes and high variability in potential copy numbers [18]. Recently, a more focused NGS method to analyze these genes provides full characterization of the *SMN* region in an affordable manner [19].

Traditionally, testing for *SMN1* exon 7 copy number alone is used for SMA diagnosis. However, a deletion of exon 8 alone has been reported in milder SMA types in two patients [20]. In addition, exon 8 information may have utility for the detection of hybrid genes, depending on the testing methodology [21,22]. Although typical *SMN1/2* hybrids involving exon 7 and exon 8 are the most common reported in the literature [6,23], hybrid genes may also be detected using other loci that differentiate *SMN1* and *SMN2*, for example, by comparing exon 7 and intron 7 [24] or involving intron 6 to exon 8 [19].

In addition to copy numbers, some methods are also able to detect variants in the *SMN1* and *SMN2* genes associated with silent carrier risk and disease severity, as detailed in the following sections. In short, the test methodology should balance the need for the right information to guide clinical care in the shortest possible timeframe with practical constraints such as the availability of instrumentation, personnel, and other resources.

## 3. SMA Carrier Genotypes, Testing, and Reporting

An SMA carrier is an asymptomatic individual lacking a functional copy of *SMN1* on one chromosome. Most SMA carriers have an *SMN1* deletion on one chromosome and one functional *SMN1* copy on the other (1 + 0), representing a heterozygous deletion (Figure 1B). Silent carriers, in contrast, have a (2 + 0) genotype, whereas others may have another type of pathogenic variant in *SMN1* on one chromosome and two *SMN1* copies ($1^d$ + 1), or rarer genotypes with higher *SMN1* copy numbers ($1^d$ + 2, 3 + 0) [8]. Due to these multiple genotypes, the detection rate of SMA carriers using the *SMN1* copy number alone to detect (1 + 0) genotypes varies from ~71% up to 95% depending on ethnicity [9].

Thus, there is a proportion of false-negative results for carrier status when reporting only the *SMN1* copy number. Residual carrier risk estimations based on the *SMN1* copy number alone have been calculated by compiling results across multiple studies and ethnicities (Table 1, first four columns) [25]. Since the total *SMN1* copy number is used to assess carrier risk, the limitations of such testing, specifically the inability to detect silent carriers using *SMN1* copy number alone, should be described when reporting results [8].

In addition to the *SMN1* copy number, data has shown that the presence of *SMN1* gene duplication variants c.*3+80T>G in intron 7 and c.*211_*212del in exon 8 (Figure 1A) can be indicative of the silent carrier (2 + 0) genotype in many ethnicities [10,11]. Several test methods can detect these variants, including MLPA (P-460), NGS, PCR/Sanger, and PCR/CE [12,18,19,24,26,27]. Typically, these variants co-occur [10]; however, individuals with only one of the two variants have been identified [11]. Detection of either c.*3+80T>G or c.*211_*212del alone is generally considered indicative of *SMN1* gene duplication, and thus associated with increased silent carrier risk [10]. However, c.*211_*212del in exon 8 has been detected in *SMN2* hybrid genes in SMA patients with no copies of *SMN1*, indicating that it is possible that an isolated occurrence of either can be associated with a hybrid gene [11].

In response to characterization of the *SMN1* gene duplication variants across multiple ethnicities, guidelines have been updated to reflect that these variants improve residual risk estimates [28]. Table 1 (last two columns) summarizes these results across several studies, which can be used to provide an estimate of residual risk based on ethnicity. The impact of these variants has not been evaluated in all ethnicities, and some studies show varying residual risk levels within an ethnicity [10,18,29]. This is likely due to both the broad range of ethnic backgrounds included in each category and the fact that ethnicities are often self-reported, which creates ambiguity in how these groups are classified and reported [30]. Consequently, the numbers shown here represent risk estimations from studies with the largest number of individuals analyzed for each ethnicity, recognizing that while these are the best estimations available, they are not exact figures. Continued research is needed to further refine both diagnostic interpretations and residual risk values for different genetic ancestries, so literature should be reviewed regularly [31].

**Table 1.** Residual SMA carrier risk estimates by ethnicity based on *SMN1* copy number and gene duplication variant status. Carrier frequency represents carrier risk without testing by ethnicity. Subsequent columns estimate residual risk based on *SMN1* copy number alone. The last two columns estimate the residual risk with two copies of *SMN1* with additional information on the presence of *SMN1* gene duplication variants (*SMN1* c.*3+80T>G and c.*211_*212del), where "positive" indicates presence of one or both variants, and "negative" indicates absence of both variants. Values are rounded to the nearest integer. Asian includes groups with South Asian and East Asian ancestry.

| Ethnicity | Carrier Frequency | 2 Copies *SMN1* Exon 7 | 3 Copies *SMN1* Exon 7 | 2 Copies *SMN1*, Variant Status "Negative" | 2 Copies *SMN1*, Variant Status "Positive" |
|---|---|---|---|---|---|
| Ashkenazi Jewish | 1:56 [a] | 1:514 [a] | 1:5899 [a] | 1:580 [b] | ~1 [b] |
| Asian | 1:50 [a] | 1:719 [a] | 1:5185 [a] | 1:779 [c] | 1:57 [c] |
| African American/Black | 1:71 [a] | 1:132 [a] | 1:6997 [a] | 1:375 [d] | 1:39 [d] |
| Caucasian/European | 1:45 [a] | 1:604 [a] | 1:4719 [a] | 1:814 [c] | 1:12 [c] |
| Hispanic | 1:83 [a] | 1:641 [a] | 1:7574 [a] | 1:906 [d] | 1:99 [d] |
| Spanish | 1:40 [e] | 1:781 [e] | Not Reported | 1:888 [e] | ~1 [e] |
| Israeli Jewish | 1:38 [a] | 1:450 [a] | 1:4004 [a] | Not Reported | Not Reported |
| Asian Indian | 1:50 [a] | 1:428 [a] | 1:5252 [a] | Not Reported | Not Reported |
| Iranian | 1:16 [a] | 1:96 [a] | 1:1604 [a] | Not Reported | Not Reported |

Data for risk estimates adapted from references as indicated with letters. a: [25]. b: [10]. c: [18]. d: [29]. e: [11].

The absence of these gene duplication variants does not rule out the possibility of a carrier (2 + 0) genotype, nor does their presence definitively diagnose silent carriers across different ancestries. In these cases, the analysis of copy number in the progenitors of the carrier under study would help to determine the cis or trans configuration of *SMN1* genes, though this implies extra testing that is not always possible [11]. Nevertheless, resolution of *SMN1* gene duplication variants modifies the residual risk of SMA carrier status in all ethnicities studied to date (Table 1). Therefore, co-occurrence of these variants with two copies of *SMN1* indicates increased carrier risk, while absence of the variants with two copies of *SMN1* indicates reduced carrier risk compared to using *SMN1* copy number alone, regardless of ethnicity [10,11,18,28,29].

For reporting purposes, *SMN1* gene duplication variant information is relevant only when two copies of *SMN1* are present; variant interpretation is not necessary when a one *SMN1* copy carrier genotype (1 + 0) is identified through *SMN1* copy number testing. Furthermore, when three or more copies of *SMN1* are present, interpretation of these variants is unnecessary given the extremely low likelihood of being a carrier [25]. In cases where ethnicity is unknown, uncertain, or unreported, a range of possible risk values may be provided and discussed in counseling patients, while noting that risk varies depending on ethnicity and, more specifically, ancestry [30]. To clarify potential reporting, examples of *SMN1* copy number and gene duplication variant status results in a carrier screening setting are provided in Table 2 based on available guidelines [8,28]. See also Prior et al. 2011 for an example report [8].

**Table 2.** Carrier Results Interpretation Examples. The examples provided here are interpretations based on relevant guidelines [8,11] and literature [10,11,18,25,29]. When interpreting and presenting results, all relevant local guidelines and regulations should be followed.

| Example Results | *SMN1* Copies | c.*3+80T>G | c.*211_*212del | Interpretation |
|---|---|---|---|---|
| Case 1 | 1 | Not indicated | Not indicated | Carrier<br>The *SMN1* copy number indicates a carrier of SMA. Genetic counseling is recommended and carrier testing should be made available to other at-risk family members. |
| Case 2 | 2 | Positive | Negative | Increased Carrier Risk<br>The *SMN1* copy number is two, ruling out a typical carrier genotype (1 + 0). However, the presence of one or more variants indicates an increased risk of being a silent carrier. The residual risk of SMA carrier status based on genotype alone is between 1:99 to ~1 depending on ethnicity. Ethnic-specific risk values based on these results are provided (see Table 1, last column). Parental testing should be considered to elucidate the presence of a silent carrier (2 + 0). Genetic counseling is recommended and carrier testing should be made available to other at-risk family members. |
| Case 3 | 2 | Positive | Positive | Increased Carrier Risk<br>Refer to Case 2 for example language. |
| Case 4 | 2 | Negative | Negative | Reduced Carrier Risk<br>The *SMN1* copy number and variant status indicate reduced, but not eliminated, carrier risk. The residual risk of SMA carrier status based on genotype alone is between 1:375 and 1:906 depending on ethnicity. Ethnic-specific risk values based on these results are provided (see Table 1, 2nd to last column). Genetic counseling is recommended. |

**Table 2.** *Cont.*

| Example Results | *SMN1* Copies | c.*3+80T>G | c.*211_*212del | Interpretation |
|---|---|---|---|---|
| Case 5 | 3 | At genetic counselor's discretion | At genetic counselor's discretion | Reduced Carrier Risk<br>The *SMN1* copy number indicates a significantly reduced, but not eliminated, carrier risk. The residual risk of SMA carrier status based on genotype is low. Ethnic-specific risk values based on these results are provided (see Table 1, Column 4). Genetic counseling is recommended. |

Since gene conversions are another mechanism that can lead to silent carriers [8], evidence of conversion from *SMN2* to *SMN1* (*SMN1/2* hybrids) could inform silent carrier risk. However, this possibility has not been sufficiently investigated clinically, and hybrid genes have a variable gene architecture [32]. As a result, there is insufficient evidence to determine carrier risk based on hybrid genes.

## 4. Disease Prognosis Genotypes, Testing, and Reporting

While the *SMN2* copy number is not relevant for the diagnosis of SMA, guidelines recommend that *SMN2* copy number results be reported to inform prognosis and treatment decisions [17,33–35]. The *SMN2* copy number is strongly correlated with SMA type, but the copy number alone is not sufficient to predict SMA type. This limitation should be clearly communicated when reporting *SMN2* copy number results.

Additionally, the c.859G>C variant is a positive disease modifier associated with reduced disease severity and improved prognosis. Several test methods can detect this variant, including NGS, specific PCR/Sanger, and PCR/CE [19,24]. Evidence indicates that c.859G>C improves *SMN2* splicing, exon 7 inclusion, and full-length SMN protein production, leading to improved phenotypic outcomes [13,14]. For instance, while 90% of individuals with SMA and two copies of *SMN2* exon 7 typically have SMA type 1, individuals with SMA that have two copies of *SMN2* exon 7 and the c.859G>C variant typically have SMA type 2 or type 3, with no known cases of SMA type 1 in individuals where this variant is present [13,14,33]. A similar effect has been observed in patients with three copies of *SMN2* exon 7 and the c.859G>C variant, typically resulting in SMA type 3 [12,33]. The number of *SMN2* copies with c.859G>C also correlates with phenotype, with multiple copies leading to milder phenotypes [12]. While the c.859G>C variant has not been reported in patients with one or four copies of *SMN2*, available evidence suggests that any individual with this variant would have a milder phenotype than expected based on *SMN2* copy number alone.

In addition to c.859G>C, another positive modifier known as c.835-44A>G has been described (Figure 1A), albeit with limited investigation in SMA patients to date. This variant is one of the PSV differentiating *SMN1* from *SMN2*, and its presence in intron 6 of *SMN2* increases the inclusion of exon 7 [36]. This modifier can be detected with specific PCR/Sanger or NGS methods [12,19]. Other putative positive and negative disease modifiers have been described [15,17,32]. However, these variants have been identified only in a small number of patients without a clear genotype-phenotype correlation [19].

Aside from SNP and INDEL variants that impact disease prognosis, several recent publications have mentioned *SMN1/2* hybrids as another positive disease modifier [15,37–39]. These hybrid genes arise when *SMN1* is partially converted to *SMN2* or vice versa. Since they retain elements of *SMN1*, some hybrids can increase exon 7 inclusion in SMN mRNAs compared to typical *SMN2*, producing greater quantities of full length SMN protein that lead to a milder phenotype [37,38]. However, *SMN1/2* hybrids are heterogeneous, and their impacts on full-length SMN transcript and protein quantity are likely dependent on which *SMN1* elements are retained [37]. More data are needed to inform the interpretation of hybrid genotypes beyond the general observation that *SMN1/2* hybrids can be associated with milder phenotypes.

For reporting purposes, likely prognosis can be interpreted using *SMN2* copy number alone when disease-modifying variants are not detected, noting that the correlation between genotype and phenotype is not absolute [8,34,35]. A positive result for c.859G>C may be reported as a marker associated with reduced severity and/or improved prognosis in comparison with the typical presentation based on the *SMN2* copy number genotype. To clarify probable SMA types based on *SMN2* copy number and c.859G>C, a summary of published treatment guidelines and peer-reviewed studies is provided in Table 3. This prognostic information is relevant only for individuals diagnosed with SMA. Examples for reporting *SMN2* copy number and c.859G>C status when providing test results are provided. Other disease modifier variants such as c.835-44A>G or the presence of *SMN2* hybrids can be reported when further research genetic studies are performed, mainly in discordant patients [15,17].

## 5. Newborn Screening for SMA

With multiple treatment options available and compelling data showing the value of early treatment to maximize patient benefit, SMA newborn screening (NBS) has become an increasing priority. In the US, this screening is included in the RUSP (Recommended Uniform Screening Panel) and other NBS recommendations [34]. In the same line, the SMA NBS Alliance promotes the implementation of NBS in all of Europe by 2025 (www.sma-screening-alliance.org/ (accessed on 12 September 2022)).

In SMA NBS, *SMN1* is the primary indicator of disease status. Given the throughput and cost restrictions necessary for NBS, testing is often limited to the presence or absence of *SMN1* exon 7 using DNA isolated from dried blood spots (DBS) and is frequently combined with testing for severe combined immunodeficiency (SCID) in a single assay [34,40]. When positive screening results are identified, follow-up testing is performed to confirm diagnosis and obtain *SMN2* copy number results to infer disease prognosis. However, recent studies have provided data supporting the reporting of *SMN2* copy numbers along with initial screening results, as it is beneficial for SMA patients with two copies of *SMN2* where treatment timing is most crucial [16]. Others have suggested that disease modifier variant testing is also important to further refine the likely prognosis for SMA patients identified through NBS with two or three copies of *SMN2* [17]. As NBS programs and our understanding of the intersection of screening and treatment continue to expand, it is likely that NBS testing will move toward providing as much genetic information as possible to maximize treatment benefits in newborns with SMA [41]. As the complexity of NBS is increasing, genetic programs in newborns should come along with adequate pre-test genetic counseling to provide more precise information to the families.

**Table 3.** Likely SMA prognosis based on *SMN2* copy number and variant status. *SMN1* copy numbers are presumed to be 0, consistent with diagnosis. Genotypes not referenced below (e.g., 3 copies *SMN2* with two or more c.859G>C alleles) have not yet been reported. The reporting examples provided here are interpretations based on consensus recommendations published by the American College of Medical Genetics (ACMG), Cure SMA, and the SMA Care group [8,34,35], as well as other relevant guidelines and literature [13,14,17,33]. For recommendations on follow-up testing and management of SMA cases as well as probability estimations of SMA type based on results, see [17]. When interpreting and presenting results, all relevant local guidelines and regulations should be followed.

| SMN2 Copy Number | c.859G>C Variant Status | Interpretation and Reporting Example |
|---|---|---|
| 1 | Negative | SMA (Type 0 probable) [a]<br>Most individuals with SMA and one *SMN2* copy present with Type 0 congenital disease. While the relationship between *SMN2* copy number and disease outcomes is strongly correlated, it is not absolute, and individual exceptions do occur. Genetic counseling is recommended. |

**Table 3.** *Cont.*

| SMN2 Copy Number | c.859G>C Variant Status | Interpretation and Reporting Example |
|---|---|---|
| 2 | Negative | SMA (Type 1 probable) [a]<br>Most individuals with SMA and two *SMN2* copies present with Type 1 SMA. Refer to other examples with Negative c.859G>C Variant Status for example language. |
| 2 | Detected in one copy | SMA (Type 2/3 probable) [b,c]<br>Whereas most individuals with SMA and two *SMN2* copies present with Type 1 SMA, the presence of the c.859G>C variant in one *SMN2* copy is associated with reduced severity consistent with SMA Type 2/3. Genetic counseling is recommended. |
| 2 | Detected in two copies | SMA (Type 3/4 probable) [c,d]<br>Whereas most individuals with SMA and two *SMN2* copies present with Type 1 SMA, the presence of the c.859G>C variant in two *SMN2* copies is associated with reduced severity consistent with SMA Type 3/4. Genetic counseling is recommended. |
| 3 | Negative | SMA (Type 2/3 probable) [a]<br>Refer to other examples with negative c.859G>C variant status for an example language. |
| 3 | Detected in one copy | SMA (Type 3 probable) [c,e]<br>Whereas most individuals with SMA and three *SMN2* copies present with Type 2/3 SMA, the presence of the c.859G>C variant in one *SMN2* copy is associated with reduced severity consistent with SMA Type 3. Genetic counseling is recommended. |
| ≥4 | Negative | SMA (Type 3/4 probable) [a]<br>Refer to other examples with negative c.859G>C variant status for example language. |

Interpretation of phenotype and source data adapted from references as indicated with letters. a: [17,34]. b: [13,14,33]. c: [12] d: [42] e: [33].

## 6. Conclusions

While understanding of the impact of *SMN1* and *SMN2* variants on SMA carrier status and disease prognosis continues to evolve, a solid foundation of clinical studies demonstrates the utility of identifying several variants in addition to copy numbers. More specifically, when variants predicting *SMN1* copies in cis are present, it is possible to adjust the risk of silent carrier status, which can help inform reproductive decisions for couples. Additionally, disease modifier testing can improve prognostic predictions in individuals diagnosed with SMA, explaining some of the discrepancies between observed *SMN2* copy numbers and expected SMA disease progression. The information provided by these variants can benefit laboratories and clinicians interested in providing more accurate SMA carrier screening and prognostic predictions.

**Funding:** This work was partially supported by Grants from Biogen ESP-SMG-17-11256 (to E.F.T. supporting L.B.-P.), Roche and Spanish Instituto de Salud Carlos III, Fondo de Investigaciones Sanitarias and co-funded with ERDF funds (Grant No. FIS PI18/000687) (to E.F.T.).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** John N. Milligan is an employee of Asuragen, a Bio-Techne Brand.

## References

1. Stabley, D.L.; Harris, A.W.; Holbrook, J.; Chubbs, N.J.; Lozo, K.W.; Crawford, T.O.; Swoboda, K.J.; Funanage, V.L.; Wang, W.; Mackenzie, W.; et al. SMN1 and SMN2 copy numbers in cell lines derived from patients with spinal muscular atrophy as measured by array digital PCR. *Mol. Genet. Genom. Med.* **2015**, *3*, 248–257. [CrossRef] [PubMed]
2. Sugarman, E.A.; Nagan, N.; Zhu, H.; Akmaev, V.R.; Zhou, Z.; Rohlfs, E.M.; Flynn, K.; Hendrickson, B.C.; Scholl, T.; Sirko-Osadsa, D.A.; et al. Pan-ethnic carrier screening and prenatal diagnosis for spinal muscular atrophy: Clinical laboratory analysis of >72,400 specimens. *Eur. J. Hum. Genet.* **2012**, *20*, 27–32. [CrossRef] [PubMed]
3. Talbot, K.; Tizzano, E.F. The clinical landscape for SMA in a new therapeutic era. *Gene Ther.* **2017**, *24*, 529–533. [CrossRef] [PubMed]

4. Sugarman, E.; Nagan, N.; Zhu, H.; Akmaev, V.R.; Zhou, Z.; Rohlfs, E.M.; Flynn, K.; Hendrickson, B.C.; Scholl, T.; Sirko-Osadsa, D.A.; et al. Molecular analysis of spinal muscular atrophy and modification of the phenotype by SMN2. *Genet. Med.* **2002**, *4*, 20–26.
5. Wadman, R.I.; Stam, M.; Gijzen, M.; Lemmink, H.H.; Snoeck, I.N.; Wijngaarde, C.A.; Braun, K.P.J.; Schoenmakers, M.A.C.G.; van den Berg, L.H.; Dooijes, D.; et al. Association of motor milestones, SMN2 copy and outcome in spinal muscular atrophy types 0–4. *J. Neurol. Neurosurg. Psychiatry* **2017**, *88*, 365–367. [CrossRef] [PubMed]
6. Alías, L.; Bernal, S.; Fuentes-Prior, P.; Barceló, M.J.; Also, E.; Hernandez, R.M.; Rodríguez-Alvarez, F.J.; Martín, Y.; Aller, E.; Grau, E.; et al. Mutation update of spinal muscular atrophy in Spain: Molecular characterization of 745 unrelated patients and identification of four novel mutations in the SMN1 gene. *Hum. Genet.* **2009**, *125*, 29–39. [CrossRef] [PubMed]
7. Wirth, B. An update of the mutation spectrum of the survival motor neuron gene (SMN1) in autosomal recessive spinal muscular atrophy (SMA). *Hum. Mutat.* **2000**, *15*, 228–237. [CrossRef]
8. Prior, T.W.; Nagan, N.; Sugarman, E.A.; Batish, S.D.; Braastad, C. Technical standards and guidelines for spinal muscular atrophy testing. *Genet. Med.* **2011**, *13*, 686–694. [CrossRef]
9. Hendrickson, B.C.; Donohoe, C.; Akmaev, V.R.; Sugarman, E.A.; Labrousse, P.; Boguslavskiy, L.; Flynn, K.; Rohlfs, E.M.; Walker, A.; Allitto, B.; et al. Differences in SMN1 allele frequencies among ethnic groups within North America. *J. Med. Genet.* **2009**, *46*, 641–644. [CrossRef]
10. Luo, M.; Liu, L.; Peter, I.; Zhu, J.; Scott, S.A.; Zhao, G.; Eversley, C.; Kornreich, R.; Desnick, R.J.; Edelmann, L. An Ashkenazi Jewish SMN1 haplotype specific to duplication alleles improves pan-ethnic carrier screening for spinal muscular atrophy. *Genet. Med.* **2014**, *16*, 149–156. [CrossRef]
11. Alías, L.; Bernal, S.; Calucho, M.; Martínez, E.; March, F.; Gallano, P.; Fuentes-Prior, P.; Abuli, A.; Serra-Juhe, C.; Tizzano, E.F. Utility of two SMN1 variants to improve spinal muscular atrophy carrier diagnosis and genetic counselling. *Eur. J. Hum. Genet.* **2018**, *26*, 1554–1557. [CrossRef] [PubMed]
12. Blasco-Pérez, L.; Costa-Roger, M.; Leno-Colorado, J.; Bernal, S.; Alias, L.; Codina-Solà, M.; Martínez-Cruz, D.; Castiglioni, C.; Bertini, E.; Travaglini, L.; et al. Deep Molecular Characterization of Milder Spinal Muscular Atrophy Patients Carrying the c.859G>C Variant in SMN2. *Int. J. Mol. Sci.* **2022**, *23*, 8289. [CrossRef] [PubMed]
13. Prior, T.W.; Krainer, A.R.; Hua, Y.; Swoboda, K.J.; Snyder, P.C.; Bridgeman, S.J.; Burghes, A.H.; Kissel, J.T. A positive modifier of spinal muscular atrophy in the SMN2 gene. *Am. J. Hum. Genet.* **2009**, *85*, 408–413. [CrossRef] [PubMed]
14. Vezain, M.; Saugier-Veber, P.; Goina, E.; Touraine, R.; Manel, V.; Toutain, A.; Fehrenbach, S.; Frébourg, T.; Pagani, F.; Tosi, M.; et al. A rare SMN2 variant in a previously unrecognized composite splicing regulatory element induces exon 7 inclusion and reduces the clinical severity of spinal muscular atrophy. *Hum. Mutat.* **2010**, *31*, E1110–E1125. [CrossRef] [PubMed]
15. Costa-Roger, M.; Blasco-Pérez, L.; Cuscó, I.; Tizzano, E.F. The Importance of Digging into the Genetics of SMN Genes in the Therapeutic Scenario of Spinal Muscular Atrophy. *Int. J. Mol. Sci.* **2021**, *22*, 9029. [CrossRef] [PubMed]
16. Lee, B.H.; Deng, S.; Chiriboga, C.A.; Kay, D.M.; Irumudomon, O.; Laureta, E.; Delfiner, L.; Treidler, S.O.; Anziska, Y.; Sakonju, A.; et al. Newborn Screening for Spinal Muscular Atrophy in New York State: Clinical Outcomes from the First 3 Years. *Neurology* **2022**. [CrossRef]
17. Cuscó, I.; Bernal, S.; Blasco-Pérez, L.; Calucho, M.; Alias, L.; Fuentes-Prior, P.; Tizzano, E.F. Practical guidelines to manage discordant situations of SMN2 copy number in patients with spinal muscular atrophy. *Neurol. Genet.* **2020**, *6*, e530. [CrossRef]
18. Chen, X.; Sanchis-Juan, A.; French, C.E.; Connell, A.J.; Delon, I.; Kingsbury, Z.; Chawla, A.; Halpern, A.L.; Taft, R.J.; Bentley, D.R.; et al. Spinal muscular atrophy diagnosis and carrier screening from genome sequencing data. *Genet. Med.* **2020**, *22*, 945–953. [CrossRef]
19. Blasco-Pérez, L.; Paramonov, I.; Leno, J.; Bernal, S.; Alías, L.; Fuentes-Prior, P.; Cuscó, I.; Tizziano, E.F. Beyond copy number: A new, rapid, and versatile method for sequencing the entire SMN2 gene in SMA patients. *Hum. Mutat.* **2021**, *42*, 787–795. [CrossRef]
20. Gambardella, A.; Mazzei, R.; Toscano, A.; Annesi, G.; Pasqua, A.; Annesi, F.; Quattrone, F.; Oliveri, R.I.; Valentino, P.; Bono, F.; et al. Spinal muscular atrophy due to an isolated deletion of exon 8 of the telomeric survival motor neuron gene. *Ann. Neurol.* **1998**, *44*, 836–839. [CrossRef]
21. Scheffer, H.; Cobben, J.M.; Matthijs, G.; Wirth, B. Best practice guidelines for molecular analysis in spinal muscular atrophy. *Eur. J. Hum. Genet.* **2001**, *9*, 484–491. [CrossRef] [PubMed]
22. Niba, E.T.E.; Nishio, H.; Wijaya, Y.O.S.; Lai, P.S.; Tozawa, T.; Chiyonobu, T.; Yamadera, M.; Okamoto, K.; Awano, H.; Takeshima, Y.; et al. Clinical phenotypes of spinal muscular atrophy patients with hybrid SMN gene. *Brain Dev.* **2021**, *43*, 294–302. [CrossRef] [PubMed]
23. Cusco, I.; Barceló, M.; Del Rio, E.; Martín, Y.; Hernández-Chico, C.; Bussaglia, E.; Baiget, M.; Tizzano, E. Characterisation of SMN hybrid genes in Spanish SMA patients: De novo, homozygous and compound heterozygous cases. *Hum. Genet.* **2001**, *108*, 222–229. [CrossRef] [PubMed]
24. Milligan, J.N.; Larson, J.L.; Filipovic-Sadic, S.; Laosinchai-Wolf, W.; Huang, Y.-W.; Ko, T.-M.; Abbott, K.M.; Lemmink, H.H.; Toivonen, M.; Schleutker, J.; et al. Multisite Evaluation and Validation of a Sensitive Diagnostic and Screening System for Spinal Muscular Atrophy that Reports SMN1 and SMN2 Copy Number, along with Disease Modifier and Gene Duplication Variants. *J. Mol. Diagn.* **2021**, *23*, 753–764. [CrossRef] [PubMed]
25. Macdonald, W.K.; Hamilton, D.; Kuhle, S. SMA carrier testing: A meta-analysis of differences in test performance by ethnic group. *Prenat. Diagn.* **2014**, *34*, 1219–1226. [CrossRef] [PubMed]

26. Park, J.E.; Yun, S.A.; Roh, E.Y.; Yoon, J.H.; Shin, S.; Ki, C.S. Carrier Frequency of Spinal Muscular Atrophy in a Large-scale Korean Population. *Ann. Lab. Med.* **2020**, *40*, 326–330. [CrossRef]
27. Lopez-Lopez, D.; Loucera, C.; Carmona, R.; Aquino, V.; Salgado, J.; Pasalodos, S.; Miranda, M.; Alonso, Á.; Dopazo, J. SMN1 copy-number and sequence variant analysis from next-generation sequencing data. *Hum. Mutat.* **2020**, *41*, 2073–2077. [CrossRef]
28. Prior, T.W.; Nagan, N.; Sugarman, E.A.; Batish, S.D.; Braastad, C. ADDENDUM: Technical standards and guidelines for spinal muscular atrophy testing. *Genet. Med.* **2016**, *18*, 752, Addendum to *Genet. Med.* **2011**, *13*, 686–694. [CrossRef]
29. Feng, Y.; Ge, X.; Meng, L.; Scull, J.; Li, J.; Tian, X.; Zhang, T.; Jin, W.; Cheng, H.; Wang, X.; et al. The next generation of population-based spinal muscular atrophy carrier screening: Comprehensive pan-ethnic SMN1 copy-number and sequence variant analysis by massively parallel sequencing. *Genet. Med.* **2017**, *19*, 936–944. [CrossRef]
30. Kaseniit, K.E.; Haque, I.S.; Goldberg, J.D.; Shulman, L.P.; Muzzey, D. Genetic ancestry analysis on >93,000 individuals undergoing expanded carrier screening reveals limitations of ethnicity-based medical guidelines. *Genet. Med.* **2020**, *22*, 1694–1702. [CrossRef]
31. Davidson, D.; Kaseniit, K.; Haque, I. Duplication Tag SNP g.27134T>G should not be considered diagnostic of SMA carrier status. In Proceedings of the ACMG Annual Meeting, Phoenix, AZ, USA, 22–24 March 2017.
32. Ruhno, C.; McGovern, V.L.; Avenarius, M.R.; Snyder, P.J.; Prior, T.W.; Nery, F.C.; Muhtaseb, A.; Roggenbuck, J.S.; Kissel, J.T.; Sansone, V.A.; et al. Complete sequencing of the SMN2 gene in SMA patients detects SMN gene deletion junctions and variants in SMN2 that modify the SMA phenotype. *Hum. Genet.* **2019**, *138*, 241–256. [CrossRef] [PubMed]
33. Calucho, M.; Bernal, S.; Alías, L.; March, F.; Venceslá, A.; Rodríguez-Álvarez, F.J.; Aller, E.; Fernández, R.M.; Borrego, S.; Millán, H.M.; et al. Correlation between SMA type and SMN2 copy number revisited: An analysis of 625 unrelated Spanish patients and a compilation of 2834 reported cases. *Neuromuscul. Disord.* **2018**, *28*, 208–215. [CrossRef] [PubMed]
34. Glascock, J.; Sampson, J.; Haidet-Phillips, A.; Connolly, A.; Darras, B.; Day, J.; Finkel, R.; Howell, R.R.; Klinger, K.; Kuntz, N.; et al. Treatment Algorithm for Infants Diagnosed with Spinal Muscular Atrophy through Newborn Screening. *J. Neuromuscul. Dis.* **2018**, *5*, 145–158. [CrossRef] [PubMed]
35. Mercuri, E.; Finkel, R.S.; Muntoni, F.; Wirth, B.; Montes, J.; Main, M.; Mazzone, E.S.; Vitale, M.; Snyder, B.; Quijano-Roy, S.; et al. Diagnosis and management of spinal muscular atrophy: Part 1, Recommendations for diagnosis, rehabilitation, orthopedic and nutritional care. *Neuromuscul. Disord.* **2018**, *28*, 103–115. [CrossRef]
36. Wu, X.; Wang, S.H.; Sun, J.; Krainer, A.R.; Hua, Y.; Prior, T.W. A-44G transition in SMN2 intron 6 protects patients with spinal muscular atrophy. *Hum. Mol. Genet.* **2017**, *26*, 2768–2780. [CrossRef] [PubMed]
37. Butchbach, M.E.R. Genomic Variability in the Survival Motor Neuron Genes (SMN1 and SMN2): Implications for Spinal Muscular Atrophy Phenotype and Therapeutics Development. *Int. J. Mol. Sci.* **2021**, *22*, 7896. [CrossRef]
38. Keinath, M.C.; Prior, D.E.; Prior, T.W. Spinal Muscular Atrophy: Mutations, Testing, and Clinical Relevance. *Appl. Clin. Genet.* **2021**, *14*, 11–25. [CrossRef]
39. Maretina, M.A.; Zheleznyakova, G.Y.; Lanko, K.M.; Egorova, A.A.; Baranov, V.S.; Kiselev, A.V. Molecular Factors Involved in Spinal Muscular Atrophy Pathways as Possible Disease-modifying Candidates. *Curr. Genom.* **2018**, *19*, 339–355. [CrossRef]
40. Taylor, J.L.; Lee, F.K.; Yazdanpanah, G.K.; Staropoli, J.F.; Liu, M.; Carulli, J.P.; Sun, C.; Dobrowolski, S.F.; Hannon, W.H.; Vogt, R.F. Newborn blood spot screening test using multiplexed real-time PCR to simultaneously screen for spinal muscular atrophy and severe combined immunodeficiency. *Clin. Chem.* **2015**, *61*, 412–419. [CrossRef]
41. Pane, M.; Donati, M.A.; Cutrona, C.; De Sanctis, R.; Pirinu, M.; Coratti, G.; Ricci, M.; Palermo, C.; Berti, B.; Leone, D.; et al. Neurological assessment of newborns with spinal muscular atrophy identified through neonatal screening. *Eur. J. Pediatr.* **2022**, *181*, 2821–2829. [CrossRef]
42. Souza, P.V.S.; Pinto, W.B.V.D.R.; Ricarte, A.; Badia, B.D.M.L.; Seneor, D.D.; Teixeira, D.T.; Caetano, L.; Gonçalves, E.A.; Chieia, M.A.T.; Farias, I.B.; et al. Clinical and radiological profile of patients with spinal muscular atrophy type 4. *Eur. J. Neurol.* **2021**, *28*, 609–619. [CrossRef] [PubMed]