BRAIN COMMUNICATIONS

Intragenic and structural variation in the *SMN* locus and clinical variability in spinal muscular atrophy

Renske I Wadman,[1] Marc D Jansen,[1] Marloes Stam,[1] Camiel A Wijngaarde,[1] Chantall A D Curial,[1] Jelena Medic,[1] Peter Sodaar,[1] Jan Schouten,[2] Raymon Vijzelaar,[2] Henny H Lemmink,[3] Leonard H van den Berg,[1] Ewout J N Groen[1,*] and W Ludo van der Pol[1,*]

*These authors contributed equally to this work.

Clinical severity and treatment response vary significantly between patients with spinal muscular atrophy. The approval of therapies and the emergence of neonatal screening programmes urgently require a more detailed understanding of the genetic variants that underlie this clinical heterogeneity. We systematically investigated genetic variation other than *SMN2* copy number in the *SMN* locus. Data were collected through our single-centre, population-based study on spinal muscular atrophy in the Netherlands, including 286 children and adults with spinal muscular atrophy Types 1–4, including 56 patients from 25 families with multiple siblings with spinal muscular atrophy. We combined multiplex ligation-dependent probe amplification, Sanger sequencing, multiplexed targeted resequencing and digital droplet polymerase chain reaction to determine sequence and expression variation in the *SMN* locus. *SMN1*, *SMN2* and *NAIP* gene copy number were determined by multiplex ligation-dependent probe amplification. *SMN2* gene variant analysis was performed using Sanger sequencing and RNA expression analysis of *SMN* by droplet digital polymerase chain reaction. We identified *SMN1–SMN2* hybrid genes in 10% of spinal muscular atrophy patients, including partial gene deletions, duplications or conversions within *SMN1* and *SMN2* genes. This indicates that *SMN2* copies can vary structurally between patients, implicating an important novel level of genetic variability in spinal muscular atrophy. Sequence analysis revealed six exonic and four intronic *SMN2* variants, which were associated with disease severity in individual cases. There are no indications that *NAIP1* gene copy number or sequence variants add value in addition to *SMN2* copies in predicting the clinical phenotype in individual patients with spinal muscular atrophy. Importantly, 95% of spinal muscular atrophy siblings in our study had equal *SMN2* copy numbers and structural changes (e.g. hybrid genes), but 60% presented with a different spinal muscular atrophy type, indicating the likely presence of further inter- and intragenic variabilities inside as well as outside the *SMN* locus. *SMN2* gene copies can be structurally different, resulting in inter- and intra-individual differences in the composition of *SMN1* and *SMN2* gene copies. This adds another layer of complexity to the genetics that underlie spinal muscular atrophy and should be considered in current genetic diagnosis and counselling practices.

1 UMC Utrecht Brain Center, Department of Neurology and Neurosurgery, University Medical Center Utrecht, 3584 CX Utrecht, the Netherlands
2 MRC Holland BV, 1057 DL Amsterdam, the Netherlands
3 Department of Genetics, University Medical Center Groningen, 9713 GZ Groningen, the Netherlands

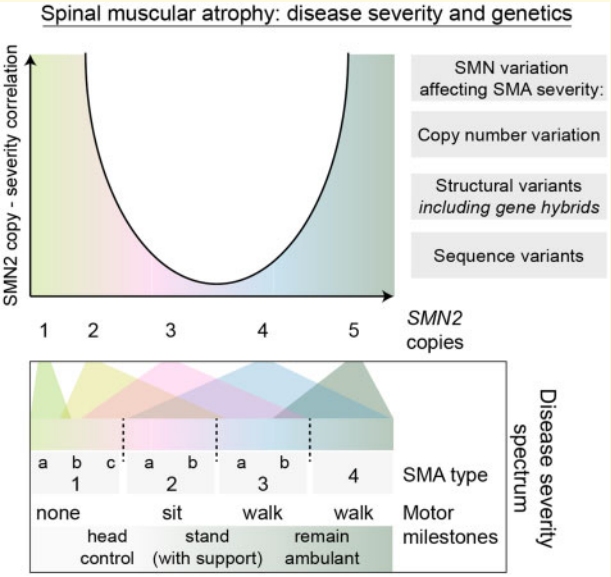
Correspondence to: Renske I. Wadman, MD, PhD University Medical Center Utrecht, F02.230, Heidelberglaan 100, 3584 CX Utrecht, the Netherlands
E-mail: r.i.wadman@umcutrecht.nl

Received February 19, 2020. Revised April 17, 2020. Accepted April 22, 2020. Advance Access publication June 8, 2020

Keywords: spinal muscular atrophy; SMA; SMN2; SMN1

Abbreviations: MLPA = multiplex ligation-dependent probe amplification; SMA = spinal muscular atrophy; SMN = survival motor neuron

Graphical Abstract



Introduction

Proximal hereditary spinal muscular atrophy (SMA) is an important genetic cause of mortality in infants and progressive motor impairment in children and adults (Mercuri *et al.*, 2012; Wadman *et al.*, 2017). It is caused by deficiency of the survival motor neuron (SMN) protein due to the homozygous loss of function of the *SMN1* gene (HGNC:11117; OMIM600354). The second *SMN* gene, *SMN2* (HGNC:11118; OMIM601627), differs only at five nucleotide positions from *SMN1*. One nucleotide substitution in Exon 7 critically influences mRNA splicing, leading to the absence of Exon 7 in the large majority of *SMN2* mRNA transcripts (delta7 *SMN2*) and the production of limited quantities of full-length SMN protein (Lefebvre *et al.*, 1995).

SMA has a striking range of severity with onset from infancy to adulthood. This is reflected in the clinical classification system that distinguishes Types 1–4 (Mercuri *et al.*, 2012). More *SMN2* copies are associated with relatively higher SMN protein levels in tissues from patients with SMA and with milder phenotypes. However, variation is only partially explained by copy number variation in *SMN2* (Lefebvre *et al.*, 1995). For example, severity in patients with three *SMN2* copies ranges from infantile onset with limited motor development (Type 1) to childhood onset with the ability to walk (Type 3). SMA severity-modifying genes outside the *SMN* locus, including *plastin 3* (*PLS3*) and *neurocalcin delta* (*NCALD*), which, when overexpressed, may substitute specific cellular SMN functions, have been identified in specific families but are unlikely to explain clinical variation at the population level (Oprea *et al.*, 2008; Hosseinibarkooie *et al.*, 2016; Riessland *et al.*, 2017; Wadman *et al.*, 2020).

The architecture of the human *SMN* locus on chromosome 5q is highly complex due to multiple duplications and inversions and has, therefore, not yet been completely elucidated (Burghes, 1997; Wirth, 2000; Rochette *et al.*, 2001; Arkblad *et al.*, 2006; Lunn and Wang, 2008; Thauvin-Robinet *et al.*, 2012). Rare intragenic variants in *SMN2* have been described (Prior *et al.*, 2009; Bernal *et al.*, 2010; Wu *et al.*, 2017; Calucho *et al.*, 2018; Ruhno *et al.*, 2019) that modify disease severity, and it has been suggested that variation within the *SMN2* locus, such as deletions of the adjacent *NAIP1*, modifies severity (Burlet *et al.*, 1996; Watihayati *et al.*, 2009; Amara *et al.*, 2012; Ruhno *et al.*, 2019; Vorster *et al.*, 2020). Variation in the sequence of *SMN2* and the *SMN* locus requires further study in large and well-defined patient cohorts.

The relevance of elucidating genetic variability in the *SMN* locus has further increased with the approval of the first *SMN2* splicing modulating therapy and the expectation that more such therapies will become available soon. First experiences with the *SMN2*-specific antisense oligonucleotide therapy, nusinersen, suggest that not all patients respond equally well to treatment. This could

partially be explained by currently unidentified genetic variation (Harahap *et al.*, 2015; Wu *et al.*, 2017).

To further improve our understanding of the correlation between genetic and clinical variation, we performed a detailed analysis of the structure, sequence and expression of the *SMN* locus in 286 SMA patients (Wadman *et al.*, 2017; Wadman *et al.*, 2018). We identified an additional level of genetic heterogeneity of the *SMN* locus and its association with the clinical phenotype.

Materials and methods

We enrolled patients with SMA Types 1–4 between September 2010 and August 2018 from our single-centre prevalence cohort study in the Netherlands.

The Medical Ethical Committee of the University Medical Center Utrecht approved the study protocol (09-307/NL29692.041.09). This study was registered at the Dutch registry for clinical studies and trials (http://www.ccmo-online.nl). All patients gave written informed consent. Informed consent was obtained from all participants and/or each subject and additionally from their parents if children were younger than 18 years.

The reporting of this study conforms to the Strengthening the Reporting of Observational Studies in Epidemiology statement (von Elm *et al.*, 2007).

Patients

Details of the population-based prevalence cohort study on SMA Types 1–4 in the Netherlands have been described previously (Wadman *et al.*, 2017; Wadman *et al.*, 2018). Inclusion criteria were a clinical diagnosis of SMA Types 1–4 and genetic confirmation of a homozygous deletion of *SMN1* or heterozygous deletion with a point mutation on the other allele of *SMN1* (HGNC:11117; OMIM600354). There was no age restriction for inclusion. All included patients visited the outpatient clinic for (paediatric) neurology at our centre and were evaluated by one of the medical doctors (R.I.W., C.A.W., M.S.). We interviewed all patients and/or their parents and examined muscle strength using the Medical Research Council scale and motor function using the Hammersmith functional motor scale expanded (Wadman *et al.*, 2017). We used the SMA classification system based on age at onset and the best of two achieved milestones (independent sitting and walking) (Table 1) (Munsat and Davies, 1992; Zerres and Rudnik-Schoneborn, 1995; Zerres *et al.*, 1997; Dubowitz, 1999; Rudnik-Schoneborn *et al.*, 2009; Mercuri *et al.*, 2012; Wadman *et al.*, 2017).

Concordant and discordant patients were defined to analyse the predictive value of *SMN2* copy numbers for the clinical phenotype. We used the same model as described previously to define the expected copy number (Ruhno *et al.*, 2019): SMA Type 1 has two copies of *SMN2*, Type 2 has three *SMN2* gene copies, and Type 3 has four *SMN2* gene copies. With this model, we selected discordant patients with a milder or more severe phenotype in relation to their *SMN2* copy number.

Genetic analysis

Copy number analysis

SMN1, *SMN2* and *NAIP* copy number status was performed at Medical Research Council Holland using SALSA multiplex ligation-dependent probe amplification (MLPA) kit P021 (version B1). All MLPA reactions were carried out according to the manufacturer's protocol (www.mlpa.com; www.mrcholland.com). A reference sample with two copies of *SMN1* and two copies of *SMN2* was used in every reaction. The MLPA products were analysed using an ABI Prisma 310 genetic analyser (Applied Biosystems), with LIZ 500 as the internal size standard. Data analysis and interpretation were performed using Coffyalyser.Net software (www.mrcholland.com). Repeated experiments showed good reproducibility of data. Seventy samples were analysed four times, 60 samples were analysed three times and 23 samples were analysed twice in different certified laboratories (i.e. Medical Research Council Holland, Department of Medical Genetics UMC Utrecht, Netherlands, and Department of Medical Genetics UMC Groningen, Netherlands) with various sets of MLPA probe mixes (P021 versions A1 and A2; P060 versions B1 and B2). With regard to inter-experimental differences, a different *SMN2* copy number was found in only eight samples out of 286 (3%), all with a borderline of three or four *SMN2* copies, a third analysis always confirming one of the previous results.

We used MLPA data to determine *SMN2* copy number and other structural variants.

SMN2 copy number was determined using dosage analysis of Exon 7. The recently developed P021 MLPA probe set allows for a detailed interrogation of the structural composition of *SMN1* and *SMN2* genes (Fig. 1) (Vijzelaar *et al.*, 2019).

A hybrid *SMN1–SMN2* gene was suspected in case of a discrepant copy number of Exons 7 and 8. A single hybrid gene consists of one persistent *SMN1* Exon 8 copy and a corresponding, inverse downgrade of the copy number of *SMN2* Exon 8. A double hybrid consists of two *SMN1* Exon 8 copies and a two copies downgrade of *SMN2* Exon 8 compared to *SMN2* Exon 7. The presence of these hybrid *SMN1–SMN2* genes was confirmed with Sanger sequencing. An extra Exon 8 was defined as an increased number of copies of *SMN2* Exon 8 compared to *SMN2* Exon 7 copy number.

Dosage analysis of *SMN2* Exons 1–6 was also performed. A partial *SMN2* deletion or duplication was suspected in case of a higher or lower copy number (dosage $\geq$1 increase compared to the other copies) compared to

Table 1 SMA classification

SMA type and subtypes	Age at onset	Highest achieved motor milestones
1	0–6 months	Never acquires ability to sit unsupported
1a	Prenatal/neonatal	Symptoms in prenatal and/or neonatal (first month) period, no head control
1b	1–3 months	No head control and no ability to roll over
1c	3–6 months	Will usually acquire additional motor skills, such as head control or rolling from supine to prone, or at least to one side at any stage in life. Patients with SMA Type 1c are reported to survive into adulthood with or without respiratory support
2	6–18 months	Able to sit unsupported, not able to walk unsupported
2a		Unsupported sitting but not able to stand or walk with help
2b		In addition to unsupported sitting also able to stand or walk with help, but not unassisted
3	>18 months	Able to walk unsupported
3a	18–36 months	
3b	>36 months	
4	During adulthood, i.e. ≥18 years	Able to walk unsupported

the number of Exon 7 copies. Distinction between Exons 1–6 *SMN1* or *SMN2* was not possible based on homologous region of the two genes.

If no DNA was available for the MLPA experiment, confirmation of *SMN1* deletion and *SMN2* copy number was retrieved from a previously performed MLPA for diagnosis ($n = 13$).

NAIP1 copy number was detected using the *NAIP* Exon 5 sequence, as this exon is absent in *NAIP2*. The copy number was analysed by comparing the signal with the *SMN* dosage.

Mutational analysis

SMN2 was analysed by Sanger sequencing of all eight exons and flanking intronic regions as described previously (Koppers *et al.*, 2013). Primers for polymerase chain reaction amplification were designed using ENST00000380743 (*SMN2*) and ENST00000517649; ENST00000523981 (*NAIP*) (Ensemble GRCh37) (Supplementary Table 1 and 2), and optimal annealing temperature for each primer set was determined by a temperature gradient polymerase chain reaction. Each identified mutation was confirmed by an independent polymerase chain reaction and sequencing reaction on genomic DNA.

NAIP mutations were determined using multiplexed targeted resequencing, carried out on a MiSeq high-throughput next-generation sequencing platform (Illumina). We used DesignStudio (Illumina) to create a Truseq Custom Amplicon project applying the Standard Truseq Custom Amplicon Library preparation protocol (amplicon library available on request). The amplicons targeting coding, non-coding, and 5′- and 3′-untranslated regions covered 96% of the regions of interest with good quality (quality score >30). Bar-coded paired-end sequencing libraries with 2 × 250 base pair read length per amplicon were created using prepared Truseq Custom Amplicon Kit (Illumina). Sequencing reads were mapped to the human genome reference build GRCh37 using Burrows Wheeler Aligner (BWA 6.1). Base calling accuracy, measured by the Phred quality score (Q score), was presumed to be 'good' from a score of 30. Subsequent depth of coverage, quality filters, variant calling and variant annotation were performed using SAMtools v0.1.19, GATKv3.2 and the 1000 Genomes project. All variants thus identified were confirmed using Sanger sequencing.

The impact of the mutation on the structure and function of the protein was predicted by *in silico* analysis using PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/

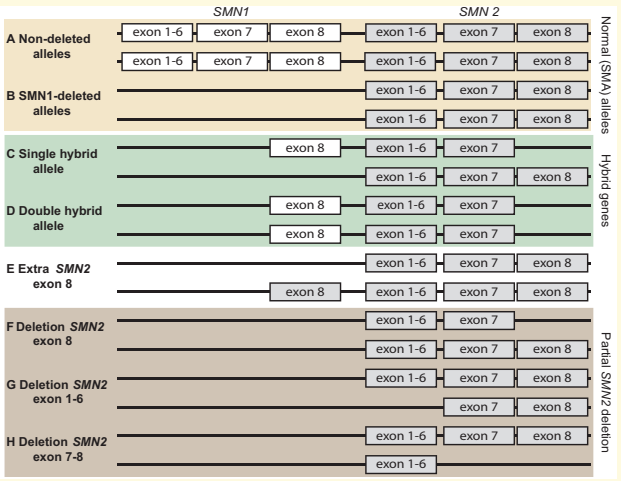


Figure 1 Representation *SMN* alleles including hybrid *SMN1–SMN2* genes. (**A**) Alleles with both *SMN1* and *SMN2* copies, representing non-carrier and non-disease status. (**B**) Alleles deleted of *SMN1* resulting in SMA. *SMN2* copy number can vary between 1 and 6 copies. (**C**) Single hybrid gene with a deletion of *SMN1* Exons 1–7 and a persistence of the non-coding *SMN1* Exon 8. *SMN2* shows an unequal number of copies of Exons 7 and 8. (**D**) Double hybrid gene with a deletion of *SMN1* Exons 1–7 and persistence of two non-coding *SMN1* Exon 8. (**E**) Extra *SMN2* Exon 8 compared to copy number of *SMN2* Exons 1–7. (**F**) Deletion of *SMN2* Exon 8 compared to the copy number of *SMN2* Exons 1–7. (**G**) Deletion of *SMN2* Exons 1–6 in one or more of the *SMN2* copies. (**H**) Deletion of *SMN2* Exons 7–8 in one or more of the *SMN2* copies.

bgi.shtml) and HumanSpliceFinder (http://www.umd.be/HSF3/technicaltips.html). The possible effect of intronic variants was also analysed with HumanSpliceFinder. The impact of synonymous mutations was predicted using relative synonymous codon usage (Sharp *et al.*, 1986; Bonekamp and Jensen, 1988; Folley and Yarus, 1989; Komar *et al.*, 1999; Sauna and Kimchi-Sarfaty, 2011).

Droplet digital PCR analysis

We used PAXgene blood RNA tubes (BD Biosciences, San Jose, CA, USA) for the storage and stabilization of RNA from peripheral blood. RNeasy Mini Kit (Qiagen, Dusseldorf, Germany) was used to extract RNA from blood. The RNA was DNase-digested with TURBO DNA-free kit (Ambion). RNA concentration was determined by spectrophotometer absorbance determination, and quality was assessed using nanodrop (NanoDrop 2000; Thermo Scientific) analysis. Quality and integrity control of PAXgene samples were performed using an Agilent 2200 TapeStation. We used a RNA Integrity Number (Rin^e) cut-off value of >5.6. We used the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems, No. 4368814) for the reverse transcription of 75 ng RNA to cDNA.

The *SMN1*, *SMN2*, *SMN2 delta 7* [as published in Wadman *et al.* (2016)], *TPB* assay and *HPRT1* assay (TBP = dHsaCPE505863; HPRT1 = dHsaCPE5192872; Bio-Rad, Hercules, CA, USA) were validated by a temperature gradient on control cDNA (Wadman *et al.*, 2016). The assays were carried out using QX200™ Droplet Digital PCR System (Bio-Rad). In brief, 22-µl reactions contained 1 µl of cDNA, 1 µl of 20× assay mix (TBP/HPRT1, HEX labelled), 227 nM *SMN* probe (FAM), 818 nM of forward and reverse *SMN* primer, 11 µl of 2× droplet digital PCR Supermix for probes (no dUTP) and 6.95 µl of RNase/DNase free water. We mixed the reaction mix with droplet generation oil (# 186-4110; Bio-Rad) and partitioned its droplets in an automated droplet generator (Bio-Rad). Polymerase chain reaction amplification for *SMN1*, *SMN2* and *SMN2Δ7* in combination with *TBP* and *HPRT1* reference genes was performed using a Bio-Rad T100 thermal cycler. After amplification, we analysed the droplets in a QX200 droplet reader as per the manufacturer's protocol. mRNA concentrations were calculated as copies per nanogram of cDNA. Reference probes were used to check the stability of probe measures in each plate. We decided to use both measures to calculate the mean levels of *SMN*, although *TBP* showed interpolate variability and *HRPT1* showed variation correlated to age. Levels of each *SMN* product (*SMN1*, *SMN2*, *SMN2 delta7*) were analysed using the geometric mean of the *SMN* levels of two separate experiments (Vandesompele *et al.*, 2002).

Statistical analysis

Normality was tested with the Kolmorogov–Smirnov and Shapiro–Wilk tests. Multivariate analyses were checked and corrected for co-linearity. Univariate and multivariate tests, including dichotomous data, were performed using (multivariate) logistic regression. Comparison of data between SMA types, (dis)cordant patients and/or *SMN2* copy number and variants was performed using Kruskal–Wallis (KW), Jonckheere Terpstra (JT) or Mann–Whitney (MW) *U*-test (continuous data) or Chi-square/Fisher's exact analysis (dichotomous data). We used IBM SPSS v23 for all statistical analyses.

Data availability statement

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

Results

We enrolled 286 patients with a genetically confirmed diagnosis of SMA and 53 parents (24 trios: both parents and child and 5 pairs: single parent and child). A total of 56 patients in our cohort were analysed as part of 25 families, which either included one or more siblings or one or more second-degree relatives. *SMN1* and *SMN2* copy numbers were determined in all patients and in all parents whose DNA was available (Table 2).

SMN1 copy number status and gene variations

Two hundred eighty-four patients (99%) had a homozygous deletion of *SMN1* Exon 7. Two patients had a heterozygous deletion of *SMN1* with a small mutation of *SMN1* on the other allele (Fig. 2). One of these patients, with SMA Type 1c, had a heterozygous deletion of *SMN1* on one allele and an 11-nucleotide duplication in Exon 6 (c.770-780dup p. Gly261Leufs*8) leading to a frame shift mutation on the other allele (Parsons *et al.*, 1996; Parsons *et al.*, 1998; Martin *et al.*, 2002; Clermont *et al.*, 2004; Alias *et al.*, 2009). The other patient, with SMA Type 3a, had a heterozygous deletion of *SMN1* and a point mutation in Exon 4 (c.542A>G; pAsp181Gly) in the other allele. Using *in silico* mRNA analysis, the c.542A>G mutation was predicted to create a new splice-donor site within Exon 4 of *SMN1* leading to an truncated transcript, introducing a preliminary stop codon (Wadman *et al.*, 2017). Three of 53 (6%) parents were carriers of two *SMN1* copies. After confirmation of parental status, this suggests the presence of two *SMN1* copies on one allele and a deletion of *SMN1* on the other allele or a *de novo SMN1* deletion (Wirth, 2000). One parent of a patient with severe SMA Type 1a had zero copies of *SMN2*.

SMN2 copy number status

SMN2 copy numbers varied from one to five gene copies. Copy number prevalence in the patient cohort was 1%,

Table 2 Clinical characteristics and *SMN* copy number

	Total SMA (*n* = 286)	SMA Type 1 (*n* = 59)	SMA Type 2 (*n* = 120)	SMA Type 3 (*n* = 98)	SMA Type 4 (*n* = 9)	Parents (*n* = 53)
Gender (F:M)	151:135	28:32	73:47	46:51	4:5	29:24
Median age in years at inclusion (range)	14.9 (0.2–78)	1.3 (0.2–62)	13.3 (0.4–67)	32.7 (2–77.5)	47.4 (36–70)	NA
Median age in years at onset (range)	1 (0–43)	0.3 (0–1.5)	0.8 (0.3–8.8)	2.2 (1–17.5)	31 (21–43)	NA
SMN1 copy number, *n* (%)						
0	284 (99.5)	59 (98)	119 (100)	96 (99)	9 (100)	0 (0)
1	2 (0.5)	1[a] (2)	0 (0)	1[b] (1)	0 (0)	50 (94)
2	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	3 (6)
SMN2 copy number, *n* (%)						
0	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (2)
1	3 (1)	3 (6)	0 (0)	0 (0)	0 (0)	8 (15)
2	30 (11)	26 (41)	2 (2)	3[b] (3)	0 (0)	24 (45)
3	165 (57)	29[a] (49)	109 (91)	27 (27)	0 (0)	20 (38)
4	84 (30)	1 (2)	9 (7)	64 (66)	9 (100)	0 (0)
5	4 (1)	0 (0)	0 (0)	4 (4)	0 (0)	0 (0)
Hybrid *SMN1–SMN2*, *n* (%)						
None	258 (90)	55 (93)	110 (92)	87 (89)	6 (67)	44 (83)
Single	25 (9)	4 (7)	10 (8)	8 (8)	3 (33)	9 (17)
Double	3 (1)	0 (0)	0 (0)	3 (3)	0 (0)	0 (0)
Partial duplication or deletion *SMN2*, *n* (%)						
None	276 (96)	56 (95)	116 (97)	95 (97)	9 (100)	45 (85)
Deletion Exons 1–6	1 (0.5)	1 (1)	0 (0)	0 (0)	0 (0)	0 (0)
Deletion Exons 7–8	5 (2)	1 (1)	1 (1)	3 (3	0 (0)	9 (17)
Extra SMN2 Exon 8	1 (0.5)	0 (0)	1 (1)	0 (0)	0 (0)	0 (0)
Deletion SMN2 Exon 8	3 (1)	2 (3)	1 (1)	0 (0)	0 (0)	0 (0)
NAIP1 copy number, *n* (%)						
0	25 (9)	13 (28)	10 (9)	2 (2)	0 (0)	0 (0)
1	161 (59)	30 (64)	88 (74)	41 (42)	2 (22)	26 (50)
2	75 (27.5)	2 (4)	19 (16)	47 (49)	7 (78)	18 (35)
3	10 (4)	2 (4)	2 (2)	6 (6)	0 (0)	7 (13)
4	1 (0.5)	0 (0)	0 (0)	1 (1)	0 (0)	1 (2)

F = female; M = male; NA = not applicable; NAIP = NLR family apoptosis inhibitor protein.
[a]Including one patient with the deletion of *SMN1* on one allele and frame shift mutation in Exon 6 (c.770-780dup; G261Lfs*8) in the other allele.
[b]Including one patient with the deletion of *SMN1* on one allele and a point mutation in Exon 4 (c.542A>G; D181G).

11%, 58%, 29% and 1% for 1–5 copies, respectively (Table 2). In 201 patients (70%), the *SMN2* copy number corresponded with the expected clinical phenotype, i.e. 1–4 copies with SMA Types 1a, 1b, 2 and 3 or 4, respectively. *SMN2* copy number correlated with SMA type (X^2 $P < 0.01$), age at onset (Spearman's rho 0.7, $P < 0.01$) and *NAIP1* copy number (X^2 $P < 0.01$).

Sequence variation in SMN2

We used Sanger sequencing to determine variation in *SMN2* in 252 patients. Sequencing revealed six exonic and four intronic *SMN2* variants (Fig. 2). *In silico* analysis of these variants suggested effects ranging from benign to likely damaging (Supplementary Table 3). Variants in Intron 1 (c.1 − 14C>T; c.81 + 45C>T) resulted in an altered *SMN2* copy composition with a lower copy number of exons (1–6 compared to Exons 7–8), correlating with a more severe phenotype. The two variants in Exon 7 were associated with a more severe (c.838_840del) or a benign (c.859G>C) clinical phenotype, in comparison to what was expected based upon *SMN2* copy number. There was no clear association between the other *SMN2* variants and SMA phenotype.

NAIP1 gene copy number and mutation analysis

NAIP1 copy number varied between zero and four copies (Table 2). *NAIP1* copy number correlated with SMA type and *SMN2* copy number (X^2 $P < 0.05$). In addition, compared to *SMN2* copy number, the *NAIP* copy number had no additional value in predicting the SMA phenotype. *NAIP* sequencing revealed two mutations [c.134A>G(H45Y); c.3503C>T(R1168K/R1330K)] in two unrelated patients presenting with different degrees of severity.

Family analysis of SMA type and SMN2 copy number

Next, we investigated the relationship between type of SMA and *SMN2* copy number in related patients. We included 25 different families, including 56 siblings and first-degree relatives (Fig. 3). Fifty-three patients (95%) shared the same number of *SMN2* gene copies, but clinical phenotypes were discordant in 34 patients (60%) from 14 families (e.g. siblings with SMA Types 2a and 2b or siblings with SMA Types 2b or 3a).

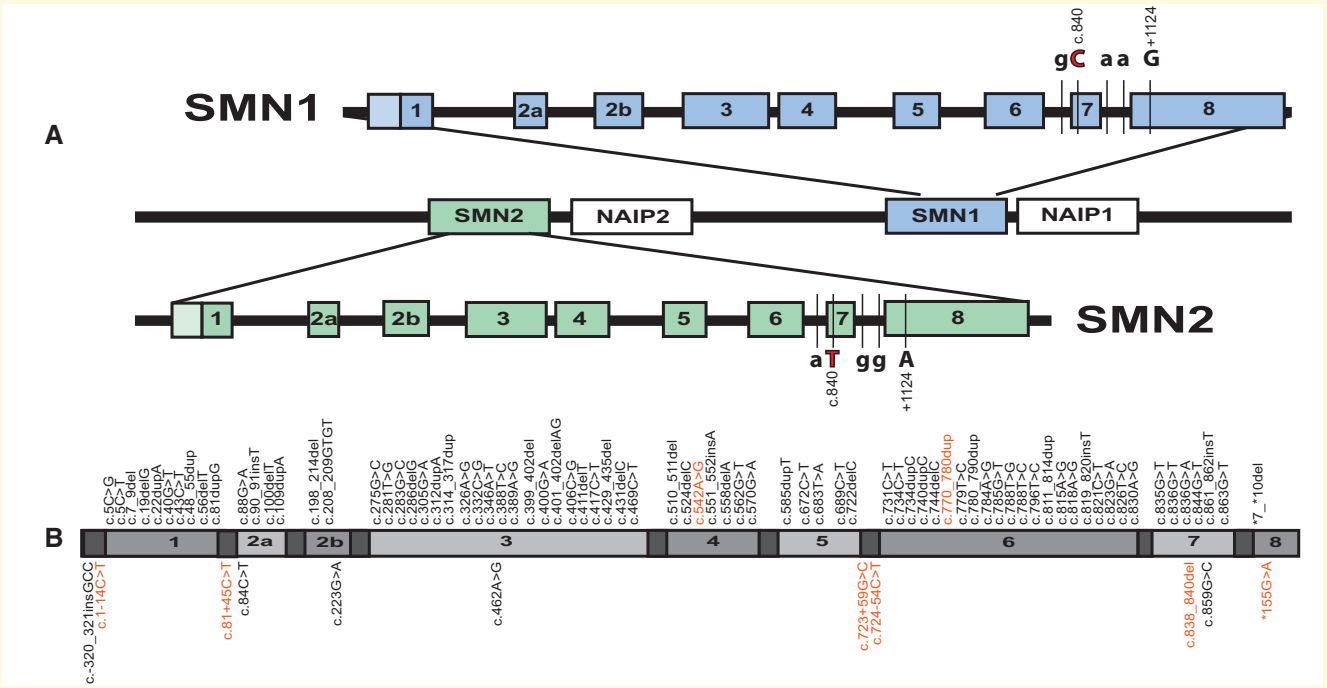


Figure 2 Genetic variation in *SMN1* and *SMN2*. (**A**) *SMN* locus and base pair differences between *SMN1* and *SMN2*. The exact location of the *SMN* and *NAIP* genes in relation to each other is still unclear. (**B**) Representation of *SMN1* and *SMN2*. Mutations are shown for *SMN1* (upper notations) and *SMN2* (lower notations). Mutations shown in red are novel variants reported in this article. Numbering refers to standard. Exon and intron sizes are not to scale (for a full list of references to previously published variants, see Supplementary Table 4).

Partial deletions and conversions of SMN2

We used MLPA data and Sanger sequencing ($n = 3$) to investigate the presence of partial gene deletions and conversions (see Materials and Methods section and Fig. 1). We found a single hybrid gene copy of *SMN1–SMN2* in 25 patients and a double hybrid gene in 3 patients, as confirmed by Sanger sequencing (Fig. 3 and Table 2). We could confirm paternal or maternal inheritance of hybrid gene copies in 10 cases, but this could not be determined in the other patients because insufficient DNA was available from the parents. Two patients carried double hybrid gene copies with one hybrid gene copy inherited from each parent. Moreover, we identified structural abnormalities other than hybrid genes. One patient with SMA Type 1c had a deletion of Exons 1–6 in two *SMN2* gene copies and an additional two *SMN2* copies with Exons 1–8, probably because of a mutation in the promoter region (c.1 – 14C>T) and Intron 1 (c.81 + 45C>T) in two copies. Five patients (carrying 2–4 *SMN2* copies) had a deletion of Exons 7–8 in one of their *SMN2* copies. We also detected this partial deletion of *SMN* (i.e. Exons 7–8) in nine parents (17%) and two controls (5%). Both parents of a patient with SMA Type 1a (harbouring one copy of *SMN2*) carried only one *SMN1* copy with one functional *SMN2* copy (Exons 1–8). Their other *SMN2* copies contained only *SMN* Exons 1–6.

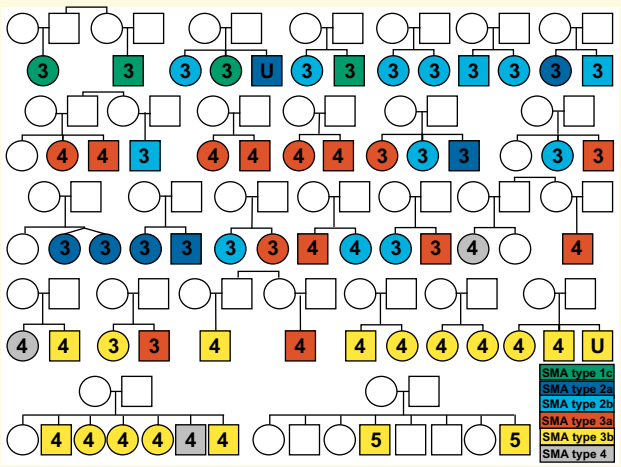


Figure 3 Family trees of dis- and concordant families. Pedigree chart of 25 families with 2 or more children affected with SMA who were included in this study. Colours of the pedigrees reflect SMA Types 1c–4. The numbers inside the pedigree reflect *SMN2* copy numbers. 'U' indicates that the *SMN2* copy number status is unknown. *SMN2* copy number was the same in 53 patients (95%), but clinical phenotypes were discordant in 34 patients (60%).

SMN expression analysis

We next investigated the effect of genetic variation (including the presence of partial deletions and hybrid

genes) on *SMN* mRNA expression in a cohort of 109 patients (Fig. 4). *SMN1* expression was completely absent in all patients with a homozygous deletion of *SMN1* copy number status. Mean levels of *SMN1* expression differed between carriers and controls (MW $P < 0.01$) (Fig. 4A). Full-length *SMN2* and delta7 *SMN2* expression levels were higher in patients compared to controls or carriers (KW $P < 0.01$). Age correlated with *SMN2* full-length levels (Spearman's rho −0.26, $P < 0.01$), but not delta7 *SMN2* (Spearman's rho −0.18, $P = 0.06$). There was no correlation between *SMN* expression levels and SMA type (KW *SMN2* FL KW $P = 0.9$; delta7 *SMN2* KW $P = 0.7$) but levels of full-length *SMN2* and delta7 *SMN2* differed between patients with varying *SMN2* copy numbers (JT *SMN2* FL $P < 0.05$; delta7 *SMN2* $P < 0.01$) (Fig. 4B). Full-length *SMN2* expression was higher in a double hybrid gene background than in a single hybrid gene background in 3 versus 33 patients with SMA, respectively (MW $P = 0.03$) (Fig. 4C). *SMN2* expression levels in two patients with the c.859G>C mutation were not significantly different from those in patients without this variant (MW $P = 0.2$) (Fig. 4D).

SMN2 variation in relation to clinical phenotype and disease course

Two hundred one patients (70%) had an *SMN2* copy number that corresponded with the expected clinical phenotype. One copy of *SMN2* was associated with neonatal onset SMA Type 1a ($n = 3$) and two *SMN2* copies with SMA Type 1b if c.859G>C was absent (95%). On a two or three *SMN2* copy background, neither the presence of a hybrid *SMN1–SMN2* gene nor the *NAIP* copy number was predictive or correlated with a milder (1c) or more severe (1b) than expected phenotype. At the milder end of the SMA clinical spectrum, four or five copies of *SMN2* were almost always associated with SMA Types 3 or 4 (87%). Deleted, converted or duplicated *NAIP* copies (e.g. 0, 1/2 or 3/4 copies) were identified across all SMA types and were not associated with a specific phenotype.

Eighty-two patients (29%) had a more severe (51%) or milder (49%) phenotype than expected based on *SMN2* copy number (see Materials and Methods section). All patients with SMA Type 4 in our cohort ($n = 9$) carried four *SMN2* copies, which are usually associated with SMA Type 3 (Piepers *et al.*, 2008). Three out of four patients with two *SMN2* copies who did not have SMA Type 1 but SMA Types 2a, 2b or 3b all had a c.859G>C mutation. The fourth patient had SMA Type 3a and an extra copy of *SMN2* with only Exons 1–6. Twenty-eight patients with SMA Type 1c carried three copies of *SMN2*, and one had even four copies with a double mutation in the promoter region.

Patients with a hybrid *SMN1–SMN2* gene ($n = 28$) showed a milder disease course compared to patients with the same *SMN2* copy number, but no statistical analysis was possible using these individual clinical parameters. None of the patients with SMA Types 2 or 3 on a three or four copy *SMN2* background with a hybrid *SMN1–SMN2* gene needed respiratory support (mean age 29 years; median 9 years; range 2–69), in contrast to 20% ($n = 24$) of patients without a hybrid gene with SMA Types 2 or 3 and 3 or 4 *SMN2* copies (start of ventilation: mean age of 21 years, median 14 years; range 2–62 years).

Discussion

The approval of therapies and the emergence of neonatal screening programmes urgently require a better understanding of genetic variants that underlie clinical heterogeneity in SMA. Our study aimed to explore the variability in the *SMN* locus in more detail than before, including an analysis of *SMN2* and *NAIP1* sequences, copy number variation, (partial) deletions or duplications and their relation to SMA severity. We show that *SMN2* copies are structurally different between patients and identified *SMN2* variants that explain clinical variability in individual cases. More importantly, we identified *SMN1–SMN2* hybrid genes as a relatively frequent and important structural variation in *SMN2* copies, between and even within patients.

Our study confirms that *SMN2* copy number is the most important severity modifier in SMA. We observed the expected association of *SMN2* copy numbers with specific SMA types (i.e. SMA Type 1: two copies; SMA Type 2: three copies; SMA Types 3 and 4: four copies) in 70% of cases (Lefebvre *et al.*, 1995; Feldkotter *et al.*, 2002; Wirth *et al.*, 2006; Rudnik-Schoneborn *et al.*, 2009; Calucho *et al.*, 2018). The strongest correlation of *SMN2* copy number and SMA type is present at both ends of the severity spectrum (Calucho *et al.*, 2018). For example, neonatal onset (SMA Type 0/1a) is virtually always associated with one *SMN2* copy and the majority of children with SMA Type 1b carry two *SMN2* copies (Mercuri *et al.*, 2012; Calucho *et al.*, 2018). Patients with late-onset and milder SMA (Types 3b and 4) mostly have four or more *SMN2* copies. In patients with three *SMN2* copies, the most prevalent copy number in this cohort, clinical variation is much more pronounced, ranging from patients with no ability to sit independently (SMA Type 1c) to ambulant patients with early onset (SMA Type 3a). The fact that *SMN2* copy number variation is insufficient to explain all relevant clinical variation is further illustrated by the 60% of siblings with discordant phenotypes but similar *SMN2* copy numbers in 95% of our families. It suggests the presence of other genetic variants that influence SMA severity, either within or outside the *SMN* locus (Jones *et al.*, 2019).

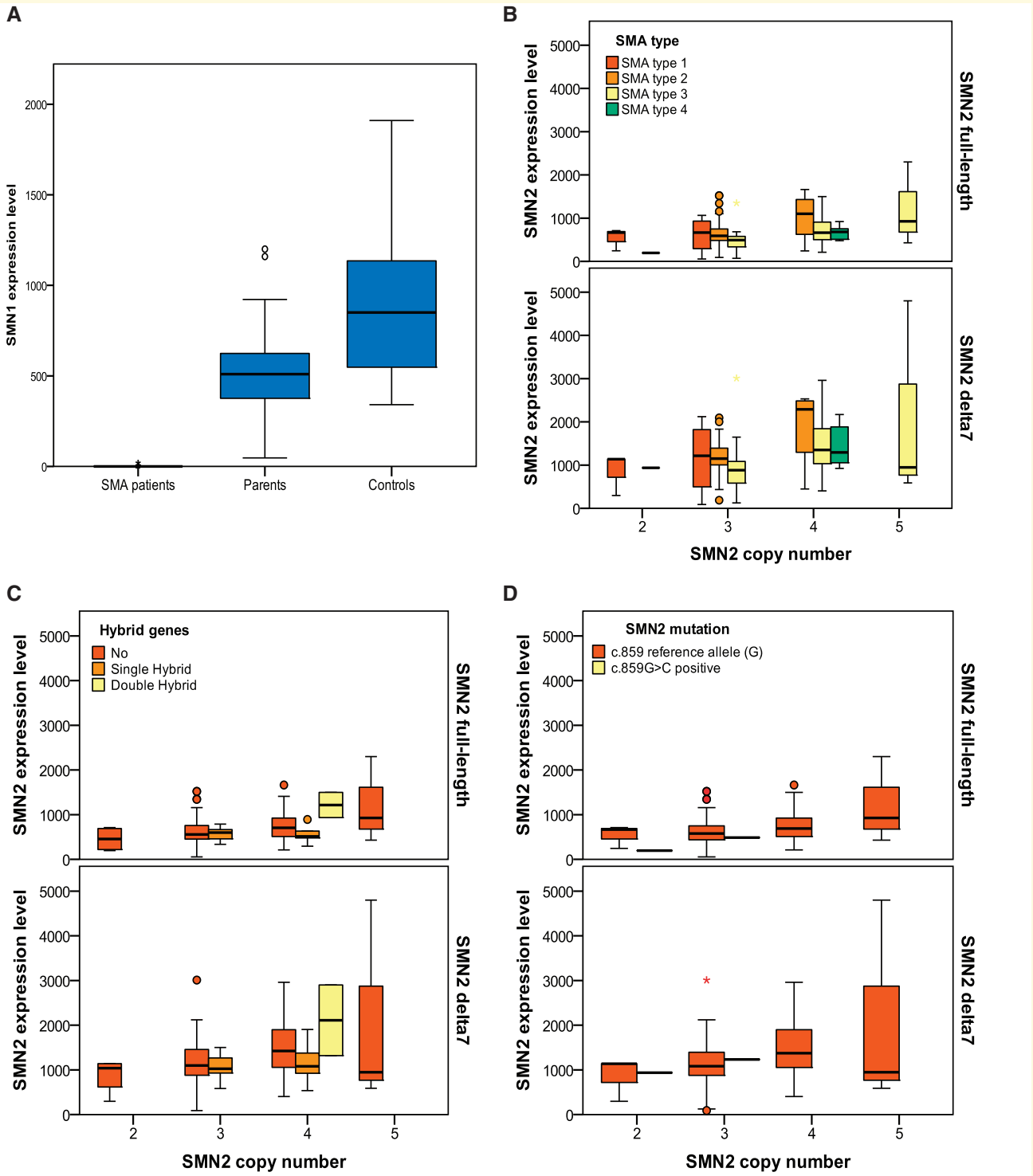


Figure 4 Expression levels of *SMN1* and *SMN2*. (**A**) *SMN1* expression levels differ between patients, carriers (= parents) and controls ($P < 0.01$). *SMN1* levels were non-detectable in patients with a homozygous deletion of *SMN1*. (**B**) *SMN2* expression levels (*SMN2* full-length upper panel, *SMN2* delta7 lower panel) show a correlation with the *SMN2* copy numbers (e.g. higher *SMN2* copy number correlates with higher *SMN2* expression levels) (*SMN2* FL KW $P = 0.02$; *SMN2* delta7 KW $P = 0.09$), also when analysed within the SMA types. (**C**) Hybrid genes resulted in higher levels of *SMN2* full length if analysed within the same *SMN2* copy number (KW $P = 0.06$). *SMN2* full-length levels (upper panel) were higher in a double hybrid gene background compared to levels on a single hybrid background in patients with SMA (MW $P = 0.03$). (**D**) No difference was found in expression levels of patients with ($n = 110$) or without ($n = 2$) a c.859G>C mutation (MW $P = 0.2$). *SMN* expression levels were presented as number of copies per 75 ng RNA. Panels **B**–**D** present data of SMA patients only.

Because specific mutations in *SMN2* that modify severity have been reported, we first assessed intragenic variation in a relatively large cohort of well-defined patients (Prior *et al.*, 2009; Bernal *et al.*, 2010; Harahap *et al.*, 2015; Wu *et al.*, 2017; Ruhno *et al.*, 2019). We identified 10 single-nucleotide variants in *SMN2*, including five novel ones that are *SMN2* specific (i.e. they have not been reported in the *SMN1* sequence) (Hahnen and Wirth, 1996; Wirth *et al.*, 1997; Alias *et al.*, 2009; Jędrzejowska *et al.*, 2014). Four of these variants had severity-modifying effects. We found previously described polymorphisms in *SMN2* Exons 2a and 3 (c.84C>T and c.462A>G, respectively) in 30% ($n = 88$) of patients without a clear correlation with the phenotype (Ruhno *et al.*, 2019). We documented a strategic mutation in the promoter region of the *SMN2* gene (c.1 − 14C>T) that explained the clinical phenotype (SMA Type 1) in the presence of four *SMN2* copies. With extended MLPA analyses, we confirmed that this mutation abrogated the function of at least two *SMN2* copies. Mutations in Exon 7 of *SMN2* showed more clear associations with the clinical phenotype. We detected a deletion of three nucleotides (c.838_840del) in a child with SMA Type 1b and two *SMN2* copies. The early disease onset (≤1 month) may implicate a slightly more severe disease course than expected. The c.859G>C in Exon 7 mutation has previously been found to be associated with milder phenotypes in patients with low *SMN2* copy numbers. We found this mutation in four patients, increasing the number of reported cases to a total of 21 (Prior *et al.*, 2009; Bernal *et al.*, 2010; Calucho *et al.*, 2018; Ruhno *et al.*, 2019). This mutation alters an exonic splicing enhancer, thereby probably interfering with splicing and transcription of the *SMN2* gene (Prior *et al.*, 2009). In contrast to a previous report, however, we were unable to confirm a positive effect on *SMN2* expression levels in our patients with a c.859G>C mutation (Vezain *et al.*, 2010). The presence of heterozygous mutations at c.859, as shown in our current and previous studies, implies that not all patients' *SMN2* copies contain this SNP and copies of *SMN2* are, therefore, different (Prior *et al.*, 2009; Bernal *et al.*, 2010; Calucho *et al.*, 2018). The modifying effect of the c.859G>C mutation occurred in patients with fewer *SMN2* copies than expected (e.g. two copies and SMA Type 3b). The lack of correlation between the *SMN2* variants and SMN protein expression levels may suggest the presence of other isoforms of SMN, which we are currently unable to detect (Singh *et al.*, 2012; Harahap *et al.*, 2018).

Our MLPA results show structural heterogeneity of the *SMN* locus beyond copy number variation. We identified hybrid *SMN* genes and partial deletions of *SMN2* in 12% of our patients. Trio analysis showed that this variation was often inherited. Hybrid *SMN1–SMN2* genes were found in patients with a relatively mild disease course compared to patients with the same *SMN2* copy number. The correlation with a better clinical phenotype was supported by the observation that patients with a double hybrid gene showed higher expression levels of full-length *SMN2*, suggesting a more efficient transcription of the hybrid gene. The mechanism behind the upregulation of SMN protein expression is currently not well understood. The molecular architecture of hybrid genes may also vary (Wu *et al.*, 2017), e.g. conversion of an exon with or without intronic sequences, which may have additional effects on the transcription and clinical phenotype. Other patients carried partial deletions of *SMN2* Exons 1–6 or 7–8, strongly suggesting further structural heterogeneity between *SMN2* copies. Since these deletions are not rare, we think that similar *SMN2* copy numbers encompass a much larger genetic and functional heterogeneity that provides a likely explanation for clinical variation. Deletion junctions resulting in the partial deletion of Exons 7–8 have recently been described in 10% of patients in a cohort of 217 SMA patients (Ruhno *et al.*, 2019). We detected deletions of Exons 7–8 in a much lower percentage, i.e. 2% of patients in our cohort. *SMN* locus rearrangements vary considerably between populations, including the loss of Exons 7–8, which provides a likely explanation for this discrepancy (Vijzelaar *et al.*, 2019; Vorster *et al.*, 2020).

Our findings are of particular relevance in relation to present genetic therapies. Although inter-sample variation of the MLPA in our repeated analysis was very low (3%), for eight out of 286 patients (3%), it was not possible to determine whether they had three or four copies. This may raise difficulties in some countries, where there is no reimbursement of antisense-oligonucleotide (ASO) therapy for patients with more than three *SMN2* copies, or in prenatal screening programmes where a similar cut-off might be used (Baker *et al.*, 2019; Muller-Felber *et al.*, 2020). Intronic sequence variation is a possible explanation for differences in treatment response. We analysed flanking intronic regions of up to 100 reads but did not detect intronic variation, including the previously described positive modifier in Intron 6 (−44A>G) (Wu *et al.*, 2017; Ruhno *et al.*, 2019). Moreover, none of the 10 detected *SMN2* mutations in our cohort were located in the flanking regions of Introns 6 or 7, which represent the target sites of *SMN2* splicing modulating ASOs or small molecules currently in development (Singh *et al.*, 2006; Calder *et al.*, 2016; Fletcher *et al.*, 2017). Although we cannot exclude the presence of other deep intronic variations (in)directly influencing these targeted therapies, our current data suggest that genetic variation at the target sites of therapies is rare. Structural variability, as illustrated by the presence of hybrid genes, may, however, reflect the presence of DNA sequences in patients who are more or less susceptible to gene-targeted therapies. The exact DNA sequences and mechanisms associated with this variability remain to be determined. Recent technological advances allow for the increasingly detailed analysis of highly complex genetic regions such as the *SMN* locus, including approaches based on

improved analysis of current short-read sequencing methods, optical mapping and long-read sequencing (van Dijk *et al.*, 2018; Ho *et al.*, 2020). Indeed, these approaches have already been shown to be applicable to the *SMN* locus and SMA in proof-of-concept studies (Ebbert *et al.*, 2019; Chen *et al.*, 2020). Combining these novel methods with a large, well-phenotyped cohort of patients in future studies will be required to obtain a complete picture of the genetic variability that underlies clinical variation in SMA.

This study shows that gene copies of *SMN2* are structurally different between and also within patients. This may have implications for current counselling and treatment practices. With currently available sequencing and genotyping methods, obtaining genotype–phenotype correlations and predictions for individual patients remains a challenge.

Supplementary material

Supplementary material is available at *Brain Communications* online.

Acknowledgements

The authors wish to thank Yana van der Weegen and Elske Mak-Nienhuis for their contributions to the sample preparation and sequencing of *SMN2*.

Funding

This work was supported by a grant from the Prinses Beatrix Spierfonds (WAR08-24).

Competing interests

R.I.W., M.D.J., M.S., C.A.W., C.A.D.C., J.M., P.S., J.S., R.V., H.H.L. and E.J.N.G. report no disclosures or competing interests. L.H.v.d.B. reports grants from ALS Foundation Netherlands, The Netherlands Organization for Health Research and Development [Vici scheme; and funded through the EU Joint Programme—Neurodegenerative Disease Research, JPND (SOPHIA, STRENGTH, ALS-CarE projects)], personal fees from Shire, Biogen, Cytokinetics and Treeway, outside the submitted work. W.L.v.d.P. received grants from non-profit entities Prinses Beatrix Spierfonds, Stichting Spieren voor Spieren, Vriendenloterij. His employer received fees for his membership of the Novartis Data Monitoring Committee (Branaplam) and *ad hoc* consultancy for Biogen and Avexis.

References

Alias L, Bernal S, Fuentes-Prior P, Barcelo MJ, Also E, Martinez-Hernandez R, 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.

Amara A, Adala L, Ben Charfeddine I, Mamai O, Mili A, Lazreg TB, et al. Correlation of SMN2, NAIP, p44, H4F5 and Occludin genes copy number with spinal muscular atrophy phenotype in Tunisian patients. Eur J Paediatr Neurol 2012; 16: 167–74.

Arkblad EL, Darin N, Berg K, Kimber E, Brandberg G, Lindberg C, et al. Multiplex ligation-dependent probe amplification improves diagnostics in spinal muscular atrophy. Neuromuscul Disord 2006; 16: 830–8.

Baker M, Griggs R, Byrne B, Connolly AM, Finkel R, Grajkowska L, et al. Maximizing the benefit of life-saving treatments for Pompe disease, spinal muscular atrophy, and Duchenne muscular dystrophy through newborn screening: essential steps. JAMA Neurol 2019; 76: 978.

Bernal S, Alias L, Barcelo MJ, Also-Rallo E, Martinez-Hernandez R, Gamez 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: 640–2.

Bonekamp F, Jensen KF. The AGG codon is translated slowly in E. coli even at very low expression levels. Nucleic Acids Res 1988; 16: 3013–24.

Burghes AH. When is a deletion not a deletion? When it is converted. Am J Hum Genet 1997; 61: 9–15.

Burlet P, Burglen L, Clermont O, Lefebvre S, Viollet L, Munnich A, et al. Large scale deletions of the 5q13 region are specific to Werdnig-Hoffmann disease. J Med Genet 1996; 33: 281–3.

Calder AN, Androphy EJ, Hodgetts KJ. Small molecules in development for the treatment of spinal muscular atrophy. J Med Chem 2016; 59: 10067–83.

Calucho M, Bernal S, Alias L, March F, Vencesla A, Rodriguez-Alvarez FJ, 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: 208–15.

Chen X, , Sanchis-Juan A, French CE, Connell AJ, Delon I, Kingsbury Z, et al. Spinal muscular atrophy diagnosis and carrier screening from genome sequencing data. Genet Med 2020; 22: 945–53.

Clermont O, Burlet P, Benit P, Chanterau D, Saugier-Veber P, Munnich A, et al. Molecular analysis of SMA patients without homozygous SMN1 deletions using a new strategy for identification of SMN1 subtle mutations. Hum Mutat 2004; 24: 417–27.

Dubowitz V. Very severe spinal muscular atrophy (SMA type 0): an expanding clinical phenotype. Eur J Paediatr Neurol 1999; 3: 49–51.

Ebbert MTW, Jensen TD, Jansen-West K, Sens JP, Reddy JS, Ridge PG, et al. Systematic analysis of dark and camouflaged genes reveals disease-relevant genes hiding in plain sight. Genome Biol 2019; 20: 97.

Feldkotter 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: 358–68.

Fletcher S, Bellgard MI, Price L, Akkari AP, Wilton SD. Translational development of splice-modifying antisense oligomers. Expert Opin Biol Ther 2017; 17: 15–30.

Folley LS, Yarus M. Codon contexts from weakly expressed genes reduce expression in vivo. J Mol Biol 1989; 209: 359–78.

Hahnen ET, Wirth B. Frequent DNA variant in exon 2a of the survival motor neuron gene (SMN): a further possibility for distinguishing the two copies of the gene. Hum Genet 1996; 98: 122–3.

Harahap NI, Takeuchi A, Yusoff S, Tominaga K, Okinaga T, Kitai Y, et al. Trinucleotide insertion in the SMN2 promoter may not be related to the clinical phenotype of SMA. Brain Dev 2015; 37: 669–76.

Harahap NIF, Niba ETE, Ar Rochmah M, Wijaya YOS, Saito T, Saito K, et al. Intron-retained transcripts of the spinal muscular atrophy genes, SMN1 and SMN2. Brain Dev 2018; 40: 670–7.

Ho SS, Urban AE, Mills RE. Structural variation in the sequencing era. Nat Rev Genet 2020; 21: 171–89.

Hosseinibarkooie S, Peters M, Torres-Benito L, Rastetter RH, Hupperich K, Hoffmann A, et al. The power of human protective modifiers: PLS3 and CORO1C unravel impaired endocytosis in spinal muscular atrophy and rescue SMA phenotype. Am J Hum Genet 2016; 99: 647–65.

Jędrzejowska M, Gos M, Zimowski JG, Kostera-Pruszczyk A, Ryniewicz B, Hausmanowa-Petrusewicz I. Novel point mutations in survival motor neuron 1 gene expand the spectrum of phenotypes observed in spinal muscular atrophy patients. Neuromuscul Disord 2014; 24: 617–23.

Jones CC, Cook SF, Jarecki J, Belter L, Reyna SP, Staropoli J. Spinal muscular atrophy (SMA) subtype concordance in siblings: findings from the cure SMA cohort. J Neuromuscul Dis 2019; 7: 33–40.

Komar AA, Lesnik T, Reiss C. Synonymous codon substitutions affect ribosome traffic and protein folding during in vitro translation. FEBS Lett 1999; 462: 387–91.

Koppers M, Groen EJ, van Vught PW, van Rheenen W, Witteveen E, van Es MA, et al. Screening for rare variants in the coding region of ALS-associated genes at 9p21.2 and 19p13.3. Neurobiol Aging 2013; 34: 1518.e5–7.

Lefebvre S, Burglen L, Reboullet S, Clermont O, Burlet P, Viollet L, et al. Identification and characterization of a spinal muscular atrophy-determining gene. Cell 1995; 80: 155–65.

Lunn MR, Wang CH. Spinal muscular atrophy. Lancet 2008; 371: 2120–33.

Martin Y, Valero A, del Castillo E, Pascual SI, Hernandez-Chico C. Genetic study of SMA patients without homozygous SMN1 deletions: identification of compound heterozygotes and characterisation of novel intragenic SMN1 mutations. Hum Genet 2002; 110: 257–63.

Mercuri E, Bertini E, Iannaccone ST. Childhood spinal muscular atrophy: controversies and challenges. Lancet Neurol 2012; 11: 443–52.

Muller-Felber W, Vill K, Schwartz O, Glaser D, Nennstiel U, Wirth B, et al. Infants diagnosed with spinal muscular atrophy and 4 SMN2 copies through newborn screening—opportunity or burden? J Neuromuscul Dis 2020; 7: 109–17.

Munsat TL, Davies KE. International SMA consortium meeting (26-28 June 1992, Bonn, Germany). Neuromuscul Disord 1992; 2: 423–8.

Oprea GE, Krober S, McWhorter ML, Rossoll W, Muller S, Krawczak M, et al. Plastin 3 is a protective modifier of autosomal recessive spinal muscular atrophy. Science 2008; 320: 524–7.

Parsons DW, McAndrew PE, Allinson PS, Parker WD, Burghes AH, Prior TW. Diagnosis of spinal muscular atrophy in an SMN non-deletion patient using a quantitative PCR screen and mutation analysis. J Med Genet 1998; 35: 674–6.

Parsons DW, McAndrew PE, Monani UR, Mendell JR, Burghes AH, Prior TW. 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: 1727–32.

Piepers S, Berg LH, Brugman F, Scheffer H, Ruiterkamp-Versteeg M, Engelen BG, et al. A natural history study of late onset spinal muscular atrophy types 3b and 4. J Neurol 2008; 255: 1400–4.

Prior TW, Krainer AR, Hua Y, Swoboda KJ, Snyder PC, Bridgeman SJ, et al. A positive modifier of spinal muscular atrophy in the SMN2 gene. Am J Hum Genet 2009; 85: 408–13.

Riessland M, Kaczmarek A, Schneider S, Swoboda KJ, Lohr H, Bradler C, 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.

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: 255–66.

Rudnik-Schoneborn S, Berg C, Zerres K, Betzler C, Grimm T, Eggermann T, et al. Genotype-phenotype studies in infantile spinal muscular atrophy (SMA) type I in Germany: implications for clinical trials and genetic counselling. Clin Genet 2009; 76: 168–78.

Ruhno C, McGovern VL, Avenarius MR, Snyder PJ, Prior TW, Nery FC, 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: 241–56.

Sauna ZE, Kimchi-Sarfaty C. Understanding the contribution of synonymous mutations to human disease. Nat Rev Genet 2011; 12: 683–91.

Sharp PM, Tuohy TM, Mosurski KR. Codon usage in yeast: cluster analysis clearly differentiates highly and lowly expressed genes. Nucleic Acids Res 1986; 14: 5125–43.

Singh NN, Seo J, Rahn SJ, Singh RN. A multi-exon-skipping detection assay reveals surprising diversity of splice isoforms of spinal muscular atrophy genes. PLoS One 2012; 7: e49595.

Singh NK, Singh NN, Androphy EJ, Singh RN. Splicing of a critical exon of human Survival Motor Neuron is regulated by a unique silencer element located in the last intron. Mol Cell Biol 2006; 26: 1333–46.

Thauvin-Robinet C, Drunat S, Saugier Veber P, Chantereau D, Cossee M, Cassini C, et al. Homozygous SMN1 exons 1-6 deletion: pitfalls in genetic counseling and general recommendations for spinal muscular atrophy molecular diagnosis. Am J Med Genet A 2012; 158A: 1735–41.

van Dijk EL, Jaszczyszyn Y, Naquin D, Thermes C. The third revolution in sequencing technology. Trends Genet 2018; 34: 666–81.

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol 2002; 3: RESEARCH0034.

Vezain M, Saugier-Veber P, Goina E, Touraine R, Manel V, Toutain A, 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: E1110–25.

Vijzelaar R, Snetselaar R, Clausen M, Mason AG, Rinsma M, Zegers M, et al. The frequency of SMN gene variants lacking exon 7 and 8 is highly population dependent. PLoS One 2019; 14: e0220211.

von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. PLoS Med 2007; 4: e296.

Vorster E, Essop FB, Rodda JL, Krause A. Spinal muscular atrophy in the Black South African population: a matter of rearrangement? Front Genet 2020; 11: 54.

Wadman RI, Jansen MD, Curial CAD, Groen EJN, Stam M, Wijngaarde CA, et al. Analysis of FUS, PFN2, TDP-43, and PLS3 as potential disease severity modifiers in spinal muscular atrophy. Neurol Genet 2020; 6: e386.

Wadman RI, Stam M, Gijzen M, Lemmink HH, Snoeck IN, Wijngaarde CA, et al. Association of motor milestones, SMN2 copy and outcome in spinal muscular atrophy types 0-4. J Neurol Neurosurg Psychiatry 2017; 88: 365–7.

Wadman RI, Stam M, Jansen MD, van der Weegen Y, Wijngaarde CA, Harschnitz O, et al. A comparative study of SMN protein and mRNA in blood and fibroblasts in patients with spinal muscular atrophy and healthy controls. PLoS One 2016; 11: e0167087.

Wadman RI, Wijngaarde CA, Stam M, Bartels B, Otto LAM, Lemmink HH, et al. Muscle strength and motor function throughout life in a cross-sectional cohort of 180 patients with spinal muscular atrophy types 1c-4. Eur J Neurol 2018; 25: 512–8.

Watihayati MS, Fatemeh H, Marini M, Atif AB, Zahiruddin WM, Sasongko TH, et al. Combination of SMN2 copy number and NAIP deletion predicts disease severity in spinal muscular atrophy. Brain Dev 2009; 31: 42–5.

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: 228–37.

Wirth B, Brichta L, Schrank B, Lochmuller H, Blick S, Baasner A, et al. Mildly affected patients with spinal muscular atrophy are partially protected by an increased SMN2 copy number. Hum Genet 2006; 119: 422–8.

Wirth B, Schmidt T, Hahnen E, Rudnik-Schoneborn S, Krawczak MM, Myhsok, B, et al. De novo rearrangements found in 2% of index patients with spinal muscular atrophy: mutational mechanisms, parental origin, mutation rate, and implications for genetic counseling. Am J Hum Genet 1997; 61: 1102–11.

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: 2768–80.

Zerres K, Rudnik-Schoneborn 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: 518–23.

Zerres K, Rudnik-Schoneborn 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: 67–72.