

# Studying carrier frequency of spinal muscular atrophy in the State of Qatar and comparison to other ethnic groups: Pilot study

Faisal Ibrahim<sup>1</sup>  | Dinesh Velayutham<sup>2</sup> | Mohamed Alsharshani<sup>1</sup> | Usama AlAlami<sup>3</sup> | Manar AlDewik<sup>4</sup> | Tala Abuarja<sup>4</sup>  | Hilal Al Rifai<sup>5</sup> | Nader I. Al-Dewik<sup>4,5,6,7,8</sup> 

<sup>1</sup>Diagnostic Genetics Division (DGD), Department of Laboratory Medicine and Pathology (DLMP), Hamad Medical Corporation (HMC), Doha, Qatar

<sup>2</sup>Liberal Arts and Science, Hamad Bin Khalifa University (HBKU), Doha, Qatar

<sup>3</sup>School of Life Science, Manipal Academy of Higher Education (MAHE), Dubai, UAE

<sup>4</sup>Department of Research and Translational and Precision Medicine Research Lab, Women's Wellness and Research Center, Hamad Medical Corporation, Doha, Qatar

<sup>5</sup>Department of Pediatrics and Neonatology, Neonatal Intensive Care Unit, Newborn Screening Unit, Women's Wellness and Research Center, Hamad Medical Corporation, Doha, Qatar

<sup>6</sup>Translational Research Institute (TRI), Hamad Medical Corporation (HMC), Doha, Qatar

<sup>7</sup>Genomics and Precision Medicine (GPM), College of Health & Life Science (CHLS), Hamad Bin Khalifa University (HBKU), Doha, Qatar

<sup>8</sup>College of Health and Life Sciences, Hamad Bin Khalifa University, Education City, Doha, Qatar

## Correspondence

Nader I. Al-Dewik, Department of Research and Translational and Precision Medicine Research Lab, Women's Wellness and Research Center, Hamad Medical Corporation, P.O.BOX. 3050, Doha, Qatar.  
Email: [naldewik@hamad.qa](mailto:naldewik@hamad.qa); [nader.al-dewik@kingston.ac.uk](mailto:nader.al-dewik@kingston.ac.uk)

## Funding information

Qatar National Library

## Abstract

**Background:** Spinal muscular atrophy (SMA) is an autosomal recessive disease caused by mutations and deletions in *SMN1* at exon 7. The carrier frequency for *SMN1* mutations ranges from 2 to 4% in the general population.

**Methods:** We examined allelic, genotypic relatedness and copy number (CN) variations and frequencies of *SMN1* and *SMN2*, in 13,426 samples from Qatar biobank (QBB) to provide a precise estimation of SMA carrier frequency in Qatar in comparison to other populations.

**Results:** The SMA carrier frequency was found to be (2.8%) and the rs143838139 was found in 491/13426 (3.66%) of individuals. The SNP rs121909192, which is a pathogenic risk factor, was found in 321/13500 (2.38%). In Addition 242/11379 (2.13%) had two copies of *SMN1* and the rs143838139, which may explain the (2+0) silent carrier. Additionally, two participants were found to be SMA type 4 with 0 and 4 copy numbers in *SMN1* and *SMN2*, respectively.

**Conclusion:** The SMA carrier frequency in Qatar was found to be comparable to Saudi Arabia and Caucasians. The likely pathogenic variant, rs121909192, was found to be significantly higher when comparing with other in our study. The rs143838139 variant, which has a strong association with the silent carrier genotype, has been found. Consequently, testing for this SNP may enhance the

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Molecular Genetics & Genomic Medicine* published by Wiley Periodicals LLC.

precision of evaluating the likelihood of a patient having an affected child. We conclude that the frequency of SMA carriers varies within the Qatar population and other ethnic groups.

#### KEYWORDS

copy number (CN) variations, SMA, *SMN1*, *SMN2*, SNPs

## 1 | INTRODUCTION

Spinal muscular atrophy is an autosomal recessive neuromuscular disease characterized by degeneration of the anterior horn cells of the spinal cord, leading to progressive symmetrical limb and trunk paralysis associated with muscular atrophy (Prior, 2010). The incidence of SMA is 1 in 6000–10,000 live births, and the carrier frequency is 1:25–50 among different ethnic groups (Hendrickson et al., 2009; MacDonald et al., 2014; Ogino, Leonard, Rennert, Ewens, & Wilson, 2002; Sugarman et al., 2012). Its subtype (SMA I–SMA IV) is classified relative to the age of onset and severity of the disease (Lunn & Wang, 2008).

The survival motor neuron (SMN) gene is the main SMA determining gene and comprises *SMN1* and *SMN2* present on 5q13. *SMN1* and *SMN2* genes are typically distinguished by a single nucleotide difference (840 C>T) (Alías et al., 2009; Anhuf et al., 2003; Cusin et al., 2003; Feldkötter et al., 2002; Ogino, Leonard, Rennert, & Wilson, 2002). This single functional difference occurs at exon 7 (Al Jumah et al., 2022; Chen et al., 1999; Mailman et al., 2001, 2002). In most cases of SMA, individuals have a homozygous loss of function of the survival motor neuron gene *SMN1*, mainly due to a homozygous deletion of exon 7, with a much smaller quantity of residual SMN protein expression from each *SMN2* copy. The number of *SMN2* copies varies within the general population and is inversely associated with disease severity as more *SMN2* copies ensures that the absolute amount of the SMN protein produced is higher (D'Amico et al., 2011), such that affected individuals with three or more copies of *SMN2* typically have milder forms of the disease (Al Jumah et al., 2022). Furthermore, it is now established that SMA is caused by deletions or intragenic mutations of *SMN1*. Homozygous deletion of *SMN1* is found in more than 92% of SMA patients, and deleteriously mutated in the remaining patients (Lyahyai et al., 2012; Wirth, 2000). In those remaining patients, small mutations that abolish the production of the SMN protein are found, mostly in combination with an *SMN1* deletion (~4%) (Sugarman et al., 2012; Verhaart et al., 2017).

Testing for SMA is of great importance and through carrier screening the identification of asymptomatic carrier couples with no family history at risk of transmitting a genetic

disease to their future offspring is possible. The American College of Medical Genetics (ACMG) recommended routine carrier screening for SMA in the general population because of its high carrier frequency and the severity of the genetic disease (Kraszewski et al., 2018; Mailman et al., 2002; Prior et al., 2010). Carrier frequency for SMA ranges depending on ethnicity (Lyahyai et al., 2012; Ogino, Leonard, Rennert, Ewens, & Wilson, 2002; Prior et al., 2010; Sugarman et al., 2012). In Qatar, the premarital clinic issued a compulsory medical examination prior to marriage, by the Emiri decree in 2009, requiring all couples of Qatari citizens and residents to test for various diseases/disorders including genetic diseases such as homocystinuria, cystic fibrosis, and the option of SMA; such tests are performed at any genetic testing clinic in Qatar (Al-Dewik et al., 2018).

The utility of population-wide carrier screening has been demonstrated in pilot studies (Mailman et al., 2002). The key to screening for SMA is (1) determining the copy number of *SMN1* for SMA diagnosis and carrier testing and (2) determining the copy number of *SMN2* for clinical classification and prognosis. Traditionally, SMA testing, and carrier testing are done with polymerase chain reaction (PCR)-based assays, such as quantitative PCR (qPCR) (Kraszewski et al., 2018) multiplex ligation-dependent probe amplification (MLPA), and digital PCR (Anhuf et al., 2003; Arkblad et al., 2006; Ashley, 2015; Feldkötter et al., 2002; Huang et al., 2007; Kubo et al., 2015; Scarcioia et al., 2006; Stabley et al., 2015; Sutomo et al., 2002; Tomaszewicz et al., 2005; van der Steege et al., 1995; Zhong et al., 2011). Such methods primarily determine the copy number of *SMN1* based on the c.840C>T site that differs between *SMN1* and *SMN2*. With recent advances in next-generation sequencing (Chen et al., 2020), it is now possible to profile a large number of genes or even the entire genome at high throughput and in a clinically relevant time-frame. Driven by these advances, many countries are currently undertaking large-scale population sequencing efforts (Labrum et al., 2007; Turnbull et al., 2018). One of the most robust methods to detect SMA carriers is through whole genome sequencing where *SMN1* and *SMN2* copy numbers can be determined (Labrum et al., 2007; Turnbull et al., 2018). Sequencing across the whole genome and using different SNPs alleles and the

identification of genotypes can improve carrier detection and provide more accurate estimates of residual risk with respect to SMA carrier status (Labrum et al., 2007).

Population-specific studies on SMA provide insufficient information to support the calculation of carrier frequencies and risk assessments due to widespread carrier screening methods (Chen et al., 2005, 2020). Consequently, carrier screening must be performed by sensitive methods that can distinguish *SMN1* from *SMN2* (Turnbull et al., 2018). Therefore, we performed a wide SMA carrier genetic screening study using whole genome sequencing results from QBB to assess the prevalence of *SMN1* deletions, carrier frequency, and the frequency and combined impact of the *SMN1* and *SMN2* copy number (CN) in a study consisting of 11,383 healthy individuals from the state of Qatar and comparing our results to the carrier results of other population including Caucasians, Europeans, Africans, East Asians, South Asians, and admixed Americans consisting of Colombians, Mexican-Americans, Peruvians, and Puerto Ricans. The variants tested were rs143838139 (c.\*3+80T>G), a variant associated with a haplotype specific for *SMN1* duplication alleles, rs141760116 (c.835-2A>G), a splice acceptor site at *SMN1* and two missense *SMN1* variants rs1554066397 (c.5C>G) and rs121909192 (c.859G>C); all variants have previously been reported in association with cases of SMA on ClinVar.

## 2 | MATERIALS AND METHODS

### 2.1 | SMN copy number caller

The SMN copy number caller tool is capable of characterizing both whole-gene deletions/duplications and partial deletions of a region that includes exon 7 and 8 and can detect small variants linked to silent SMA carriers where two copies of *SMN1* are found on the same haplotype (c.\*3+80TG / g.27134T>G). It was employed in this study to determine the copy number of full-length *SMN1*, full-length *SMN2*, as well as *SMN2* delta7–8 (*SMN2* with a deletion of Exon7–8) from 13,426 out of 14,664 whole-genome sequencing (WGS) BAM files from the Qatar Biobank (QBB) participants (1238 results did not pass the quality control (QC) test and were omitted from our analysis).

This caller is designed to work with standard WGS sequencing depth ( $\geq 30\times$ ). It works by initially counting aligned reads to *SMN1* or *SMN2*. Any read counts in the 22.2 kb region (exons 1–6) are used to calculate total SMN CN (*SMN1*, *SMN2* and *SMN2* $\Delta 7-8$ ), while read counts in the 6.3 kb region (exons 7–8) are used to calculate the CN of intact SMN (*SMN1* and *SMN2*); the truncated SMN CN (*SMN2* $\Delta 7-8$ ) is calculated by subtracting the intact SMN CN from the total SMN CN.

After calculating the summed copy number, we differentiated *SMN1* from *SMN2* using supporting read counts at base differences between *SMN1* and *SMN2*. The individual CN of *SMN1* or *SMN2* at each site is calculated by considering the summed SMN CN and the fraction of *SMN1* or *SMN2* supporting reads out of all *SMN1*+*SMN2* reads. During the development of the caller, we called the CNs of *SMN1* and *SMN2* at the 16 different sites that were extracted from the reference genome, in the 1000 Genomes Project (1kGP) samples, and determined if the CN calls for each position were concordant with the CN calls at the c.840C>T or g.27134T>G or rs143838139 splice variant site.

### 2.2 | Visualizing the caller results

A visualization tool for producing dynamic representations of data and calls' QC was also created where the summed CNs of total (exons 1–6) and intact (exons 7–8) SMN (*SMN1*+*SMN2*) are displayed against the population distribution.

The CN of *SMN2* delta 7–8 is represented by the difference between the total and intact SMN CN. Individual CNs of *SMN1* and *SMN2* are determined using the sum of the intact CN and supporting read counts at eight base changes between *SMN1* and *SMN2*.

### 2.3 | SMA SNPs allele and genotype

The allelic and genotypic results of our individuals for rs121909192, rs1554066397, rs141760116, and rs143838139 were extracted from the QBB records. The rs121909192, rs1554066397, rs141760116, are reported as pathogenic/likely pathogenic, whereas the rs143838139 is a benign and known to be associated with silent carriers (2 + 0). The list of SNPs that was screened is shown in (Supplementary table 1)

## 3 | RESULTS

### 3.1 | SMN1 CN: SMN2 CN percentage

This is a population-based study designed to record data obtained from whole genome sequencing of the Qatar biobank (QBB) participants with the aim to gain insights into the genetic architecture of a clinically relevant disease and its carrier frequency in the Middle Eastern Qatar population. The present study is based on whole genome sequencing data obtained from 13,426 participants of QBB. The results of *SMN1* CN: *SMN2* CN amongst all carrier and non-carrier study individuals are shown in

**TABLE 1** Observed *SMN1* CN: *SMN2* CN in all study individuals ( $n = 13,426$ ).

<i>SMN1</i> CN: <i>SMN2</i> CN	No of subjects	Interpretation	Percentage
4:0	56	Noncarrier	0.42
4:1	31	Noncarrier	0.23
4:2	23	Noncarrier	0.17
4:3	1	Noncarrier	0.01
4:4	1	Noncarrier	0.01
3:0	241	Noncarrier	1.8
3:1	935	Noncarrier	6.96
3:2	313	Noncarrier	2.33
3:3	54	Noncarrier	0.402
3:4	3	Noncarrier	0.02
2:0	799	Noncarrier	5.6
2:1	3806	Noncarrier	28.35
2:2	6241	Noncarrier	46.48
2:3	504	Noncarrier	3.75
2:4	29	Noncarrier	0.22
1:0	8	SMA carrier	0.06
1:1	98	SMA carrier	0.73
1:2	195	SMA carrier	1.45
1:3	76	SMA carrier	0.57
1:4	4	SMA carrier	0.03

(Table 1). Carriers were defined as having only one copy of the *SMN1* gene. A total of 381/13426 individuals (2.8%) were identified as SMA carriers. The carrier rate was 2.8% indicating that the carrier prevalence was approximately 1:35 in the current study. The two copies of *SMN1* CN are seven times more prevalent (11379/13426) than the three copies of *SMN1* CN (1547/13426) with a frequency of 84.78% and 11.52 respectively. There were eight SMA carriers with one copy of *SMN1* and no copies of *SMN2* “1:00” (Table 1).

In our samples, the frequency of individuals with the SMA carrier “1:2” CN was the highest 195/13426 with a percentage of (1.45) compared with other individuals with the SMA carrier “1:1” 98/13426 and “1:3” 76/13426 with a percentage of 0.73% and 0.57% respectively (Table 1). The highest frequency of *SMN1* CN: *SMN2* CN was seen in individuals with the “2:2” CN 6241/13426 (46.48%) followed by individuals with the “2:1” 3806/13426 (28.35%) (Table 1).

Of note, four (0.03%) carrier individuals were found to have one *SMN1* CN and four *SMN2* CN “1:4” (Table 1). In addition, two participants were found to be SMA type 4 with 0 and 4 copy numbers in *SMN1* and *SMN2*, respectively (Figure 1).

Only one individual 1/13,426 had six copy number callers (CN) of the *SMN1* gene, (Table 2).

### 3.2 | Comparison of *SMN1*, *SMN2*, and *SMN2* delta7-8 CN frequency

Our study results of *SMN1*, *SMN2*, and *SMN2* delta 7–8 copy number and frequencies and the comparison regarding to other populations is reported in (Figure 2).

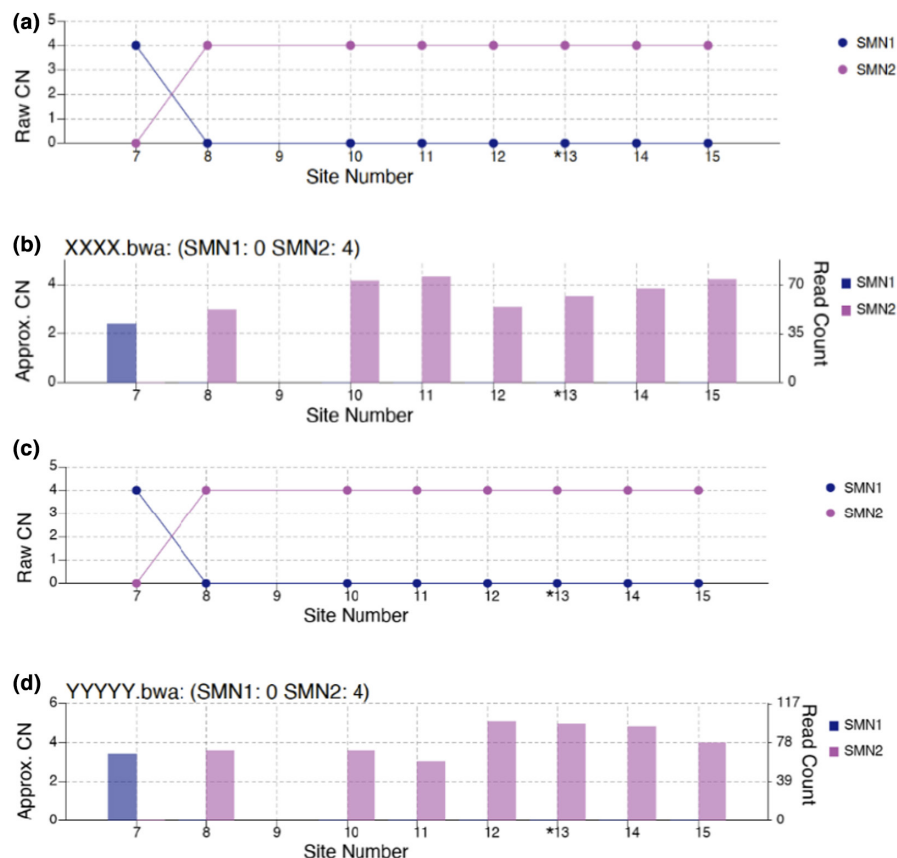
In *SMN1*, 2 CNs were the most prevalent across all populations including Qatar (84.78%) while 2 CNs were most commonly reported in *SMN2* except in the African cohort with 1 CN of *SMN2*. Calculations of CN from the exon 7–8 region for the 13,426 samples where we identified instances of *SMN2*  $\Delta$ 7–8 included 10845 samples with zero copies (80.78%), 2320 samples with one copy (17.30%) and 255 samples with two copies (1.91%); other populations also reported a distinctly greater prevalence of 0 CN in *SMN2*  $\Delta$ 7–8 relative to one and two copies.

### 3.3 | Analysis for the g.27134T > G polymorphism and detection of silent carriers

Our study detected the presence of the SNP (rs143838139), used to screen for potential silent carriers (2+0), which was found in 491/13,426 (3.66%) of all individuals and in 242/11379 (2.13%) of who had two copies of *SMN1* (Table 3).



**FIGURE 1** Visualization of SMN CN Caller result for two SMA type 4 cases. A + C. Raw CN values for *SMN1* and *SMN2* at 8 sites (site number 7–8 and 10–15). The raw CN of *SMN1* and *SMN2* at every site is calculated as the CN of intact SMN multiplied by the fraction of *SMN1*/*SMN2* supporting read counts by total supporting read counts. B + D. Raw read count is displayed on the right y axis and the left y-axis depicts estimated CN found by dividing the read count by the median haploid depth of the sample.



**TABLE 2** *SMN1* CN and *SMN2* CN amongst all study individuals ( $n = 13,426$ ).

	No of subjects	Percentage
<i>SMN1</i> CN		
0	2	0.015
1	381	2.84
2	11379	84.75
3	1547	11.52
4	112	0.83
5	3	0.022
6	1	0.0074
<i>SMN2</i> CN		
0	1105	8.23
1	4870	36.27
2	6772	50.44
3	638	4.76
4	39	0.29
5	1	0.0074

### 3.4 | SMA SNP genotyping results

SMA SNPs genotyping results of our tested individuals are listed in (Table 4).

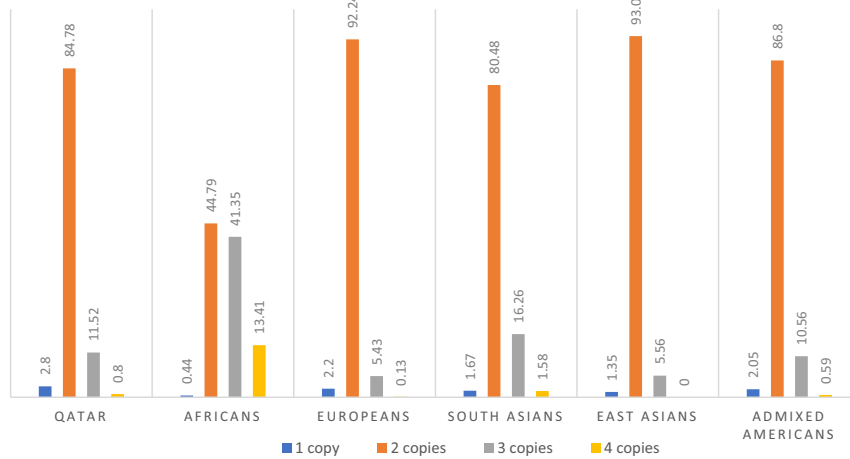
The SNP rs121909192 was present in 321/13500 (2.38%). This SNP showed the strongest association as a pathogenic/risk factor of SMA and was associated with increased risk of occurrence of SMA ( $p < 0.00001$ ) (Table 3). The SNP with the strongest association with SMA silent carriers' frequency in our study was found to be rs143838139 ( $p < 0.00001$ ) which is significant to other ethnicities. Of the 14,666 individuals tested, the SNP rs141760116 was present in five individuals (0.034%) and absent in the remaining 14,661 individuals (99.96%). Comparison of our results to other ethnic populations is recorded in Tables 3 and 4.

## 4 | DISCUSSION

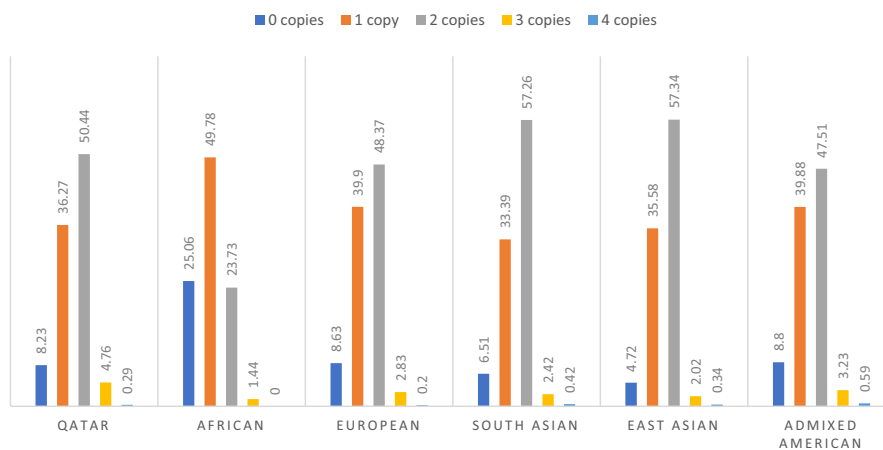
To date, there are a few studies on the prevalence/ incidence of SMA, fewer of which are recent with most being carried out in Europe (Chen et al., 1999).

Studies have shown that SMA is predominant in Middle Eastern countries, Iran, Egypt, Pakistan, and Saudi Arabia, where consanguineous marriages are common (Ibrahim et al., 2012; Thareja et al., 2021). Consanguineous marriages may play a role in the increased prevalence of SMA among the Qatari population as first-cousin marriages are present. Moreover, second and third-degree marriages are also present in the state of Qatar. Consanguineous

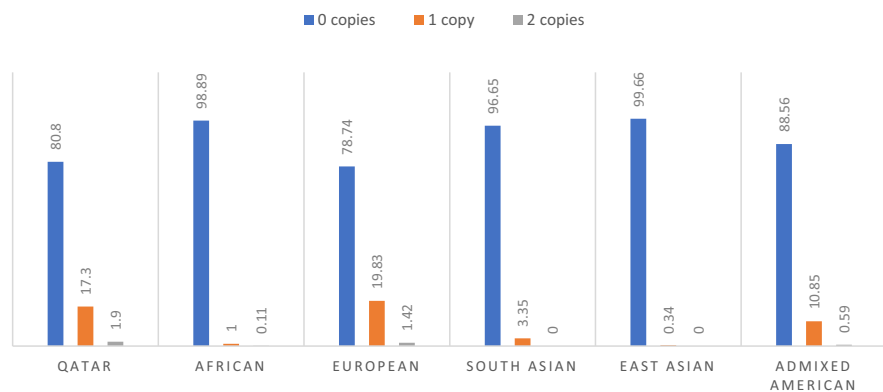
## SMN1 COPY NUMBER FREQUENCY



## SMN2 COPY NUMBER FREQUENCY



## SMN2Δ7-8 COPY NUMBER FREQUENCY



**FIGURE 2** *SMN1*, *SMN2*, and *SMN2Δ7-8* CN frequency across populations in comparison to our cohort in Qatar. The comparison of CN frequency of populations in Africa (902 individuals), Europe (9648 individuals), South Asia (1199 individuals), East Asia (593 individuals), and Admixed America (341 individuals) as a percentage relative to the cohort in Qatar (14,208 individuals/13,426 individuals for *SMN2Δ7-8*) (Adapted and reproduced data from Chen et al., 2020).

marriages were seen in 397/599 (66.2%) Qatari families, and first cousin groups account for 65% of the Qatari population (Ben-Omran et al., 2020). It is worth noting that our study individuals were recruited from QBB, which contained data from both Qatari nationals and long-term residents ( $\geq 15$  years) of Qatar. This explains the results and prevalence that is seen in this study.

We analyzed the CN distributions and recorded CN calls for *SMN1* and *SMN2* and *SMN1* CN: *SMN2* CN using

sequencing data from 13,426 samples from QBB and compared our results to other population (Europeans, Africans, East Asians, South Asians, and admixed Americans consisting of Colombians, Mexican Americans, Peruvians, and Puerto Ricans) (Table 3). Tables 1 and 2 show the number of individuals and percentage of *SMN1* CN and *SMN2* CN in addition to and *SMN1* CN: *SMN2* CN in our study population.

Moreover, Table 2 shows the number of carriers identified across our Qatar population. Our results calculated

T A B L E 3 SMA SNP genotype and allele results of the Qatar population.

Genetic variants	Pathogenicity	Genotype	QBB Qatar	All <sup>a</sup>	p value	Allele	QBB Qatar	All <sup>a</sup>	p value
rs141760116 N = 14,666	Likely pathogenic	AA	14,661 (99.97%)	2202 (99.96%)	0.793112	A	29,327 (99.98%)	248,209 (99.999597%)	0.1599
		AG	5 (0.034%)	1 (0.04%)		G	5 (0.0173%)	1 (0.00041%)	
		GG	0 (0.0%)	0 (0.0%)					
rs143838139* N = 13,841	Benign	TT	13,276 (96.0%)	2069 (82.6%)	<0.00001	T	27,042 (97.69%)	130,634 (92.1%)	0.6649
		TG	491 (3.55%)	415 (16.4%)		G	639 (2.31%)	11,270 (7.9%)	
		GG	74 (0.53%)	20 (0.8%)					
rs1554066397 N = 1480	Pathogenic	CC	1479 (99.932%)	2062 (99.4%)	0.868	C	2959 (99.97%)	1278 (99.922%)	0.6649
		CG	1 (0.068%)	0 (0.0%)					
		GG	0 (0.0%)	12 (0.578%)		G	1 (0.03%)	1 (0.0782%)	
rs121909192 N = 13,500	Likely pathogenic	GG	13,179 (97.62%)	2500 (99.84%)	<0.00001	G	26,627 (98.62%)	5003 (99.90016%)	<0.00001
		GC	269 (1.99%)	3 (0.12%)		C	373 (1.38%)	5 (0.09984%)	
		CC	52 (0.39%)	1 (0.04%)					

<sup>a</sup>Frequency results were obtained from Ensembl genome browser 108 <https://www.ensembl.org>.

T A B L E 4 SMA SNP genotype results for rs143838139 and rs121909192 in Qatar population compared with different ethnicities.

Genetic variants	Genotype	Qatari Total no 13841 and 13,500	AFR <sup>a</sup> Total no 661	AMR <sup>a</sup> Total no 347	EAS <sup>a</sup> Total no 504	EUR <sup>a</sup> Total no 503	SAS <sup>a</sup> Total no 489
rs143838139	TT	13,276 (96.0%)	250 (37.8%)	331 (95.4%)	504 (100.00%)	499 (99.2%)	485 (99.2%)
	TG	491 (3.55%)	391 (59.2%)	16 (4.6%)	<sup>b</sup>	4 (0.8%)	4 (0.8%)
	GG	74 (5.3%)	20 (3.02%)	<sup>b</sup>	<sup>b</sup>	<sup>b</sup>	<sup>b</sup>
rs121909192	GG	13,179 (97.62%)	611 (100.0%)	347 (100.0%)	<sup>b</sup>	501 (99.6%)	487 (99.6%)
	GC	269 (1.99%)	<sup>b</sup>	<sup>b</sup>	<sup>b</sup>	1 (0.2%)	2 (0.4%)
	CC	52 (0.39%)	<sup>b</sup>	<sup>b</sup>	<sup>b</sup>	1 (0.2%)	<sup>b</sup>

<sup>a</sup>Results were obtained from Ensembl genome browser 108 <https://www.ensembl.org>.

<sup>b</sup>Results were not detected.

the probability of being an SMA carrier when an individual without a family history of SMA carries one *SMN1* copy.

We identified a total of 381 samples (2.84%) with *SMN1* carrier (less than two copies) and 1663 (12.38%) with *SMN1* gains (more than two copies). 5975 (44.50%) samples consisted of *SMN2* losses and 678 (5.04%) with *SMN2* gains (Tables 1 and 2). The highest one-copy carrier rate was identified in specimens from the Caucasian group with a frequency of 1 in 37 samples (2.7%) (Anhuf et al., 2003); our results were comparatively similar to them (2.8%) with a frequency of almost 1 in 1:35 samples.

The results we recorded for SMA carriers' frequency was higher than those previously reported in other populations (Anhuf et al., 2003; Chen et al., 1999; Sugarman et al., 2012). (Figure 2), but similar to what was found in Saudi Arabia where their analysis showed the presence of one copy of the *SMN1* gene in 108 samples and two copies (normal) in 4090 samples resulted in a carrier frequency of 2.6% (Al Jumah et al., 2022). Analysis of 150 Moroccan newborns predicted a carrier frequency of approximately 1:25 (Lyahyai et al., 2012), which is higher than our results of a frequency of 1 in 37 samples. Results regarding other ethnicities found that Europeans have the highest carrier frequency at 2.2%, followed by admixed Americans (2.05%), South Asians (1.67%), and East Asians (1.35%). Africans have the lowest carrier frequency (0.44%), whereas the African group had significantly lower one-copy carrier genotypes when compared with our results, at 1 in 225 (0.44%,  $p < 0.05$ ) (Chen et al., 2020). Such results are currently the lowest reported *SMN1* carrier frequencies for any population or ethnic group (Table 4). It was also found that the 1.67% carrier frequency detected in the South-Asian sample is higher than that reported in East-Asian populations (1.35%) (Chen et al., 2020). Moreover, our results are close to those of Middle Eastern countries including Iran, Egypt, Pakistan, and Saudi Arabia (Al Jumah et al., 2022; Ibrahim et al., 2012).

It was found that for all reported ethnic groups, including our Qatar population study but excluding the African group, the two-copy genotype was over seven times more prevalent than the three-copy genotype group (Sugarman et al., 2012). This is consistent with all previously published data showing the two-copy genotype to be predominant. Surprisingly, the African population departed significantly from this genotype distribution, revealing similar frequencies for the two- and three-copy genotypes (44.79% and 41.35% respectively) (Sugarman et al., 2012), suggesting a much higher frequency of alleles with two or more *SMN1* copies relative to Qatar individuals and other ethnic groups (Table 4).

The most common combinations of *SMN1:SMN2* copy number seen in our results are 2:2 (46.48%) followed by 2:1 (28.35%). In general, individuals have

more copies of *SMN1* than *SMN2*. The most common combinations of *SMN1:SMN2* copy number in other ethnicities are 2:2 (44.9%) and 2:1 (33.4%). excluding the Africans that show higher variability in both *SMN1* and *SMN2* CN. This observation is consistent with what was found in our results. The variability of *SMN1* copy number is much lower than that of *SMN2* copy number. Conversely, 54.76% of Africans have three or more copies of *SMN1*, which is more than double of what is observed in any of the other four populations and in Qatar. Africans also have significantly lower *SMN2* CN than the other populations.

We reported the frequency of the exon 7–8 deletion (*SMN2* delta7-8) across our population compared with other populations in (Table 3). In our study, 17.3% had at least one copy of *SMN2* delta7-8. It was also found that 21.25% of Europeans and 11.44% of admixed Americans have at least one copy of *SMN2* delta7-8, while the frequency is lower in South Asians (3.35%), Africans (1.11%), and East Asians (0.34%) (Figure 2) (Chen et al., 2020).

The g.27134T>G SNP (rs143838139) is most strongly associated with two-copy *SMN1* alleles and is implemented for the detection of the g.27134T>G “2+0” silent carrier where one chromosome carries two copies of *SMN1* (either by *SMN1* duplication or gene conversion of *SMN2* to *SMN1*), and the other chromosome has no copies of *SMN1*. In our study, the rs143838139 SNP was found in 491 of the studied individuals (3.55%), and the silent carrier frequency was 242/11379 (2.13%) and absent in the remaining 13,350 individuals (96.45%) (Table 4). Based on the previous ethnic results, we found that the linkage of this SNP was the highest among African 59.2% (Table 4). Only 0.8% of individuals of Europeans and South Asian were carriers of the rs143838139 SNP (Table 4). This SNP is most strongly associated with two-copy *SMN1* alleles in Africans. The preponderance of the two-copy allele in the African group suggests a much higher frequency of individuals with the SMA silent carrier “2+0” genotype compared with other ancestries, where 33% of African individuals with two copies of *SMN1* also have the rs143838139 SNP (Sugarman et al., 2012).

The SNP rs121909192 was present in 321/13500 (2.38%) of our study population. This SNP showed the strongest association as pathogenic/risk factor with SMA and was associated with increased risk of occurrence of SMA. The SNP was most prevalent in the individuals of our cohort study. The SNP was present at relatively low levels in the European and South-East Asian populations, 0.2% and 0.4%, respectively, and was absent in the African, American, and East-Asian population (Table 4), which were limited in sample size compared with our Qatari study population. In further research, the use of a precision medicine strategy (Al-Dewik et al., 2022; Al-Dewik



et al., 2019) and population health analyses (Zhai et al., 2023) should be taken into consideration for SMA.

## 5 | CONCLUSION

In conclusion, the current research sheds light on the wide-spread of genetic diseases in our community and to the frequency of SMA carriers in Qatar. The SMA testing is highly recommended be done compulsory as part of premarital and/or genomic newborn screening. As a of premarital screening due to its high carrier rate in the community. DNA sequencing has become the preferred tool for testing disease-causing variants throughout the genome. Large, multigene targeted sequencing panels are the future of clinical carrier testing, and there is a wide critical need for a comprehensive NGS test that includes SMA carrier detection worldwide. It is noteworthy to mention that our data are leaning toward the role of consanguineous marriages in causing this disease, but the limitation is that our study included individuals who are long-term residents ( $\geq 15$  years living in Qatar). This is a pilot study and future work to strengthen validity will be implemented through independent validation of CN and variants associated with SMA and silent carrier using both q-PCR and microarray.

## AUTHOR CONTRIBUTIONS

**Nader I. Al-Dewik:** Conceptualization, data curation, formal analysis, investigation, methodology, resources, validation, writing-original draft, writing and editing review, supervision, funding acquisition, correspondence. **Faisal Ibrahim:** Writing and editing review, critical scientific review, formal analysis, and resources validation. **Dinesh Velayutham:** Formal analysis. **Mohamed Alsharshani:** Supervision. **Manar R. AlDweik:** Writing-original draft, writing and editing review, and formal analysis. **Usama AlAlami:** Critical scientific and manuscript editing. **Tala Abuarja:** Manuscript editing and referencing.

All authors contributed to manuscript writing and proofreading. All authors have read and agreed to the published version of the manuscript.

## ACKNOWLEDGMENTS

The authors want to thank their respective institutions for their continuous support and special thanks for the Qatar Biobank (QBB) and Qatar Genome Programme (QGP).

## FUNDING INFORMATION

The publication of this article is funded by the Qatar National Library, Doha, Qatar.

## CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

## DATA AVAILABILITY STATEMENT

This is a research article, and all data generated or analyzed during this study are included in this published article [and its supplementary information files]. All enquiries should be directed to the corresponding author's Email: [naldewik@hamad.qa](mailto:naldewik@hamad.qa); [nader.al-dewik@kingston.ac.uk](mailto:nader.al-dewik@kingston.ac.uk).

## INSTITUTIONAL REVIEW BOARD STATEMENT

This study was approved by the QBB.

## INFORMED CONSENT STATEMENT

Not applicable.

## ORCID

Faisal Ibrahim  <https://orcid.org/0009-0004-0431-5744>

Tala Abuarja  <https://orcid.org/0009-0003-4883-2543>

Nader I. Al-Dewik  <https://orcid.org/0000-0001-5739-1135>

## REFERENCES

- Al Jumah, M., Al Rajeh, S., Eyaid, W., Al-Jedai, A., Al Mudaiheem, H., Al Shehri, A., Hussein, M., & Al Abdulkareem, I. (2022). Spinal muscular atrophy carrier frequency in Saudi Arabia. *Molecular Genetics & Genomic Medicine*, 10(11), e2049. <https://doi.org/10.1002/mgg3.2049>
- Al-Dewik, N., Al-Mureikhi, M., Shahbeck, N., Ali, R., Al-Mesaifri, F., Mahmoud, L., Othman, A., AlMulla, M., Sulaiman, R. A., Musa, S., Abdoh, G., El-Akouri, K., Solomon, B. D., & Ben-Omran, T. (2018, Sep). Clinical genetics and genomic medicine in Qatar. *Molecular Genetics & Genomic Medicine*, 6(5), 702–712. <https://doi.org/10.1002/mgg3.474>
- Al-Dewik, N. I., & Qoronfleh, M. W. (2019). Genomics and precision medicine: molecular diagnostics innovations shaping the future of healthcare in Qatar. *Advances in Public Health*, 1–11. <https://doi.org/10.1155/2019/3807032>. [Crossref], [Google Scholar].
- Al-Dewik, N. I., Younes, S. N., Essa, M. M., Pathak, S., & Qoronfleh, M. W. (2022). Making biomarkers relevant to healthcare innovation and precision medicine. *Processes*, 10(6), 1107. <https://doi.org/10.3390/pr10061107>. [Crossref], [Google Scholar].
- 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>
- Anhuf, D., Eggermann, T., Rudnik-Schöneborn, S., & Zerres, K. (2003). Determination of SMN1 and SMN2 copy number using TaqMan technology. *Human Mutation*, 22(1), 74–78. <https://doi.org/10.1002/humu.10221>
- 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>
- Ashley, E. A. (2015). The precision medicine initiative: A new national effort. *JAMA*, 313(21), 2119–2120. <https://doi.org/10.1001/jama.2015.3595>
- Ben-Omran, T., Al Ghanim, K., Yavarna, T., El Akoum, M., Samara, M., Chandra, P., & Al-Dewik, N. (2020). Effects of consanguinity in a cohort of subjects with certain genetic disorders in Qatar. *Molecular Genetics & Genomic Medicine*, 8(1), e1051. <https://doi.org/10.1002/mgg3.1051>
- Chen, K. L., Wang, Y. L., Rennert, H., Joshi, I., Mills, J. K., Leonard, D. G., & Wilson, R. B. (1999). Duplications and de novo deletions of the SMN1 gene demonstrated by fluorescence-based carrier testing for spinal muscular atrophy. *American Journal of Medical Genetics*, 85(5), 463–469.
- Chen, W. J., Wu, Z. Y., Wang, N., Lin, M. T., & Mu-rong, S. X. (2005). Quantitative studies on SMN1 gene and carrier testing of spinal muscular atrophy. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi*, 22(6), 559–602.
- 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>
- Cusin, V., Clermont, O., Gérard, B., Chanterreau, D., & Elion, J. (2003). Prevalence of SMN1 deletion and duplication in carrier and normal populations: Implication for genetic counselling. *Journal of Medical Genetics*, 40(4), e39. <https://doi.org/10.1136/jmg.40.4.e39>
- D'Amico, A., Mercuri, E., Tiziano, F. D., & Bertini, E. (2011). Spinal muscular atrophy. *Orphanet Journal of Rare Diseases*, 6, 71. <https://doi.org/10.1186/1750-1172-6-71>
- Feldkötter, M., Schwarzer, V., Wirth, R., Wienker, T. F., & Wirth, B. (2002). 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. *American Journal of Human Genetics*, 70(2), 358–368. <https://doi.org/10.1086/338627>
- Hendrickson, B. C., Donohoe, C., Akmaev, V. R., Sugarman, E. A., Labrousse, P., Boguslavski, L., Flynn, K., Rohlf, E. M., Walker, A., Allitto, B., Sears, C., & Scholl, T. (2009). Differences in SMN1 allele frequencies among ethnic groups within North America. *Journal of Medical Genetics*, 46(9), 641–644. <https://doi.org/10.1136/jmg.2009.066969>
- Huang, C. H., Chang, Y. Y., Chen, C. H., Kuo, Y. S., Hwu, W. L., Gerdes, T., & Ko, T. M. (2007). Copy number analysis of survival motor neuron genes by multiplex ligation-dependent probe amplification. *Genetics in Medicine*, 9(4), 241–248. <https://doi.org/10.1097/gim.0b013e31803d35bc>
- Ibrahim, S., Moatter, T., & Saleem, A. F. (2012). Spinal muscular atrophy: Clinical spectrum and genetic mutations in Pakistani children. *Neurology India*, 60(3), 294–298. <https://doi.org/10.4103/0028-3886.98514>
- Kraszewski, J. N., Kay, D. M., Stevens, C. F., Koval, C., Haser, B., Ortiz, V., Albertorio, A., Cohen, L. L., Jain, R., Andrew, S. P., Young, S. D., LaMarca, N. M., De Vivo, D. C., Caggana, M., & Chung, W. K. (2018). Pilot study of population-based newborn screening for spinal muscular atrophy in New York state. *Genetics in Medicine*, 20(6), 608–613. <https://doi.org/10.1038/gim.2017.152>
- 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>
- Labrum, R., Rodda, J., & Krause, A. (2007). The molecular basis of spinal muscular atrophy (SMA) in south African black patients. *Neuromuscular Disorders*, 17(9–10), 684–692. <https://doi.org/10.1016/j.nmd.2007.05.005>
- Lunn, M. R., & Wang, C. H. (2008). Spinal muscular atrophy. *Lancet*, 371(9630), 2120–2133. [https://doi.org/10.1016/s0140-6736\(08\)60921-6](https://doi.org/10.1016/s0140-6736(08)60921-6)
- Lyahyai, J., Sbati, A., Barkat, A., Ratbi, I., & Sefiani, A. (2012). Spinal muscular atrophy carrier frequency and estimated prevalence of the disease in Moroccan newborns. *Genetic Testing and Molecular Biomarkers*, 16(3), 215–218. <https://doi.org/10.1089/gtmb.2011.0149>
- MacDonald, W. K., Hamilton, D., & Kuhle, S. (2014). SMA carrier testing: A meta-analysis of differences in test performance by ethnic group. *Prenatal Diagnosis*, 34(12), 1219–1226. <https://doi.org/10.1002/pd.4459>
- Mailman, M. D., Heinz, J. W., Papp, A. C., Snyder, P. J., Sedra, M. S., Wirth, B., Burghes, A. H., & Prior, T. W. (2002). Molecular analysis of spinal muscular atrophy and modification of the phenotype by SMN2. *Genetics in Medicine*, 4(1), 20–26. <https://doi.org/10.1097/00125817-200201000-00004>
- Mailman, M. D., Hemingway, T., Darsey, R. L., Glasure, C. E., Huang, Y., Chadwick, R. B., Heinz, J. W., Papp, A. C., Snyder, P. J., Sedra, M. S., Schafer, R. W., Abuelo, D. N., Reich, E. W., Theil, K. S., Burghes, A. H., de la Chapelle, A., & Prior, T. W. (2001). Hybrids monosomal for human chromosome 5 reveal the presence of a spinal muscular atrophy (SMA) carrier with two SMN1 copies on one chromosome. *Human Genetics*, 108(2), 109–115. <https://doi.org/10.1007/s004390000446>
- Ogino, S., Leonard, D. G., Rennert, H., Ewens, W. J., & Wilson, R. B. (2002). Genetic risk assessment in carrier testing for spinal muscular atrophy. *American Journal of Medical Genetics*, 110(4), 301–307. <https://doi.org/10.1002/ajmg.10425>
- Ogino, S., Leonard, D. G., Rennert, H., & Wilson, R. B. (2002). Spinal muscular atrophy genetic testing experience at an academic medical center. *The Journal of Molecular Diagnostics*, 4(1), 53–58. [https://doi.org/10.1016/s1525-1578\(10\)60680-0](https://doi.org/10.1016/s1525-1578(10)60680-0)
- Prior, T. W. (2010). Perspectives and diagnostic considerations in spinal muscular atrophy. *Genetics in Medicine*, 12(3), 145–152. <https://doi.org/10.1097/GIM.0b013e3181c5e713>
- Prior, T. W., Snyder, P. J., Rink, B. D., Pearl, D. K., Pyatt, R. E., Mihal, D. C., Conlan, T., Schmalz, B., Montgomery, L., Ziegler, K., Noonan, C., Hashimoto, S., & Garner, S. (2010). Newborn and carrier screening for spinal muscular atrophy. *American Journal of Medical Genetics. Part A*, 152a(7), 1608–1616. <https://doi.org/10.1002/ajmg.a.33474>
- Scarciolla, O., Stuppia, L., De Angelis, M. V., Murru, S., Palka, C., Giuliani, R., Pace, M., Di Muzio, A., Torrente, I., Morella, A., Grammatico, P., Giacanelli, M., Rosatelli, M. C., Uncini, A., & Dallapiccola, B. (2006). Spinal muscular atrophy genotyping by gene dosage using multiple ligation-dependent

- probe amplification. *Neurogenetics*, 7(4), 269–276. <https://doi.org/10.1007/s10048-006-0051-3>
- 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., Scavina, M., Sol-Church, K., & Butchbach, M. E. (2015). SMN1 and SMN2 copy numbers in cell lines derived from patients with spinal muscular atrophy as measured by array digital PCR. *Molecular Genetics & Genomic Medicine*, 3(4), 248–257. <https://doi.org/10.1002/mgg3.141>
- Sugarman, E. A., Nagan, N., Zhu, H., Akmaev, V. R., Zhou, Z., Rohlf, 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>
- Sutomo, R., Akutsu, T., Takeshima, Y., Nishio, H., Sadewa, A. H., Harada, Y., & Matsuo, M. (2002). Rapid SMN1 deletion test using DHPLC to screen patients with spinal muscular atrophy. *American Journal of Medical Genetics*, 113(2), 225–226. <https://doi.org/10.1002/ajmg.10744>
- Thareja, G., Al-Sarraj, Y., Belkadi, A., Almotawa, M., Suhre, K., & Albagha, O. M. E. (2021). Whole genome sequencing in the middle eastern Qatari population identifies genetic associations with 45 clinically relevant traits. *Nature Communications*, 12(1), 1250. <https://doi.org/10.1038/s41467-021-21381-3>
- Tomaszewicz, K., Kang, P., & Wu, B. L. (2005). Detection of homozygous and heterozygous SMN deletions of spinal muscular atrophy in a single assay with multiplex ligation-dependent probe amplification. *Beijing Da Xue Xue Bao Yi Xue Ban*, 37(1), 55–57.
- Turnbull, C., Scott, R. H., Thomas, E., Jones, L., Murugaesu, N., Pretty, F. B., Halai, D., Baple, E., Craig, C., Hamblin, A., Henderson, S., Patch, C., O'Neill, A., Devereau, A., Smith, K., Martin, A. R., Sosinsky, A., McDonagh, E. M., Sultana, R., ... Caulfield, M. J. (2018). The 100000 genomes project: Bringing whole genome sequencing to the NHS. *BMJ*, 361, k1687. <https://doi.org/10.1136/bmj.k1687>
- van der Steege, G., Grootsholten, P. M., van der Vlies, P., Draaijers, T. G., Osinga, J., Cobben, J. M., Scheffer, H., & Buys, C. H. (1995). PCR-based DNA test to confirm clinical diagnosis of autosomal recessive spinal muscular atrophy. *Lancet*, 345(8955), 985–986.
- Verhaart, I. E. C., Robertson, A., Wilson, I. J., Aartsma-Rus, A., Cameron, S., Jones, C. C., Cook, S. F., & Lochmüller, H. (2017). Prevalence, incidence and carrier frequency of 5q-linked spinal muscular atrophy – a literature review. *Orphanet Journal of Rare Diseases*, 12(1), 124. <https://doi.org/10.1186/s13023-017-0671-8>
- Wirth, B. (2000). An update of the mutation spectrum of the survival motor neuron gene (SMN1) in autosomal recessive spinal muscular atrophy (SMA). *Human Mutation*, 15(3), 228–237. [https://doi.org/10.1002/\(sici\)1098-1004\(200003\)15:3<228::Aid-humu3>3.0.Co;2-9](https://doi.org/10.1002/(sici)1098-1004(200003)15:3<228::Aid-humu3>3.0.Co;2-9)
- Zhong, Q. B. S., Kotsopoulos, S., Olson, J., Taly, V., Griffiths, A. D., Link, D. R., & Larson, J. W. (2011). Multiplex digital PCR: Breaking the one target per color barrier of quantitative PCR. *Europe PMC*, 11(13), 7.
- Zhai, K., Yousef, M. S., Mohammed, S., Al-Dewik, N. I., & Qoronfleh, M. W. (2023). Optimizing clinical workflow using precision medicine and advanced data analytics. *Processes*, 11(13), 939. [https://urldefense.com/v3/https://www.mdpi.com/2227-9717/11/3/939\\_\\_!N11eV2iwtfs!tzGT4-ss\\_bPd-p6FFySbo4MaJo3WQiUrYhRLtOIlw5Bf3V5dOd65\\_7gR1bqh-wH8c24f65aPVPC-EMUh7sGDK\\_\\$](https://urldefense.com/v3/https://www.mdpi.com/2227-9717/11/3/939__!N11eV2iwtfs!tzGT4-ss_bPd-p6FFySbo4MaJo3WQiUrYhRLtOIlw5Bf3V5dOd65_7gR1bqh-wH8c24f65aPVPC-EMUh7sGDK_$)

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Ibrahim, F., Velayutham, D., Alsharshani, M., AlAlami, U., AlDewik, M., Abuarja, T., Al Rifai, H., & Al-Dewik, N. I. (2023). Studying carrier frequency of spinal muscular atrophy in the State of Qatar and comparison to other ethnic groups: Pilot study. *Molecular Genetics & Genomic Medicine*, 11, e2184. <https://doi.org/10.1002/mgg3.2184>