INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI®
MOLECULAR ANALYSIS OF PATIENTS WITH RECESSIVE 5q-SPINAL MUSCULAR ATROPHY: INCREASED DIAGNOSTIC AND PROGNOSTIC SENSITIVITY

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Matthew David Mailman, M.S.

The Ohio State University

2001

*****Dissertation Committee:

Professor Thomas W. Prior, Adviser
Professor Arthur H.M. Burghes
Professor Sissy M. Jhiang
Professor Ralf Krahe

Approved by

Advisor
Department of Pathology
ABSTRACT

With an incidence of about 1 in 10,000 live births, autosomal recessive proximal spinal muscular atrophy (SMA) is the leading inherited cause of infant death. SMA is a neuromuscular disease characterized by hypotonia, muscle weakness, and loss of ambulation due to a defect of α-motor neurons in the anterior horn of the spinal cord. The age of onset and the presentation vary greatly among individuals afflicted with this disease. Between 92-98% of patients with SMA are homozygously deleted for the SMN1 gene on chromosome 5q13. A further 1-4% of patients are compound heterozygotes, and the remaining patients have an unknown etiology. A duplication in chromosome 5q has created a gene that is 99% homologous with SMN1 called SMN2. Although the amino acid sequence of the two genes is identical, a nucleotide change in exon 7 of SMN2 results in an alternative splice variant that yields a truncated protein in the large majority of product. The small percentage of full-length protein produced from SMN2 modifies the phenotype, such that more copies of SMN2 are found in patients with milder presentation. This dissertation presents the results of studies designed to increase our understanding of the prognostic value of SMN2 gene copy number in the modification of the SMA phenotype, as well as to increase the diagnostic sensitivity of molecular analysis of spinal muscular atrophy patients.
Although the literature states that 92-98% of patients with SMA lack the SMN1 gene, it is the experience of our molecular diagnostic laboratory that less than half of the patients with symptoms resembling spinal muscular atrophy are homozygously deleted for the gene. There are many contributing factors for this low detection rate, including the fact that the SMA test is often one of exclusion when a patient has non-specific symptoms such as generalized muscle weakness. This finding has important implications for carrier testing, because although a patient may have SMA-like symptoms, unless a mutation has been identified in a proband, testing for SMN1 carrier status will provide meaningless negative results in greater than 50% of individuals.

For autosomal recessive disorders, parents of affected patients are usually considered obligate carriers with only a single functional copy of the gene. We have found that 4% of parents of SMA patients have two copies of the SMN1 gene based on a quantitative dosage assay. This phenomenon is explained in about 2% of cases by a new mutation occurring between generations. By separating individual chromosomes in mouse/human hybrid clones, we were able to show that some of the remaining parents have two copies of SMN1 on a single chromosome. This presents the disturbing possibility that a parent who seems not to be a carrier can have a 50% chance of transmitting a null chromosome to their offspring. Incorporating this novel test into the screening regimen of SMA patients would distinguish de novo mutations from carrier parents with two copies of SMN1 in cis and would allow for more accurate genetic counseling of such families.
Intragenic \textit{SMN1} mutations occur in 1-4\% of patients with spinal muscular atrophy. Because of the presence of multiple copies of the highly homologous \textit{SMN2} gene, the process of identifying these mutations in \textit{SMN1} is technically challenging. Additionally, there are no extremely common mutations, and these mutations can be found throughout the gene. For these reasons, intragenic mutations are not screened for in the diagnostic laboratory. The consequence of this is that 1-4\% of patients with SMA are not diagnosed. We have developed a new test that uses allele-specific fluorescent primers to amplify seven of the most common mutations. This increases the sensitivity of testing for SMA patients. We were able to identify one new mutation in a patient using this assay.

Increased \textit{SMN2} copy number is known to correlate with a milder SMA phenotype. A large-scale study had not previously been reported, and we were interested in the prognostic value of \textit{SMN2} gene copy number. We compared copy number between 52 type I and 90 type III patients and found that there is a striking difference. No type III patients had less than three copies of \textit{SMN2}, while only 4\% of type I patients had more than two copies. As for prognosticating, we can state that it is unlikely that patients with one or two \textit{SMN2} genes will walk. We can also predict that patients with three or four copies of \textit{SMN2} will probably have onset of symptoms after six months of age, will likely live for many years, and should be able to sit up on their own.

Approximately 5\% of patients with symptoms of SMA have an unknown etiology. It is possible that mutations in another gene cause a phenotype similar to SMA.
We tested two genes, *SIP1* and *Sm N*, whose products interact with SMN, for mutations. These candidate genes were not mutated in 203 SMA-like patients with no *SMN1* mutations. It is possible that mutations in other genes or that mutations in non-coding regions of *SMN1* could cause symptoms of SMA.
Dedicated to Stefanie
ACKNOWLEDGMENTS

First, I would like to thank Dr. Thomas Prior for advising me through this program. He has helped me become a better scientist, a better writer, and has provided me a comfortable environment in which to work for the past several years.

Thank you to Dr. Arthur Burghes, Dr. Sissy Jhiang, and Dr. Ralf Krahe for their time and effort in serving as members of my dissertation committee.

I will take many warm memories with me of my time spent in the Molecular Pathology Laboratory. Special thanks to Rob Pyatt and Pam Snyder who made me laugh and helped me stay sane. Thanks to Dr. Will Parsons, Dr. Patty McAndrew, Dr. Claire Bartolo, Audrey Papp, Mary Sedra, Dr. John Heinz, Michelle Fuchik, Scott Bridgeman, Bob Schafer, and Ashley Hejmanowski who all made coming to work educational and fun. I am grateful for the help that Kiswar Alam provided in examining the Sm N gene and for Tamara Hemingway's aid in the production of mouse/human hybrids.

Many thanks to Dr. Jim Waldman who commits so much of his time and energy to the Pathology graduate students. Thanks to Dr. Sedmak, the Chairperson of the Department of Pathology. This research has been funded by grants from the National Institute of Health, the Muscular Dystrophy Association, and the Families of SMA.
I need to thank my family: Philip, Barbara, and Brian Mailman, who have given me unlimited amounts of unconditional love and support throughout my life and through my "extended" education. Thanks to my cats: Leo, Jackson, and Fern who all spent many hours keeping me company during the writing of this dissertation and made many useful suggestions. Finally, and most of all, thank you to my wife Stefanie whose boundless love and support keeps me strong and motivates me every day.
VITA

January 18, 1971..............Born – Brooklyn, New York, USA

1994..................................B.S. Biology, State University of New York at Binghamton

1996.................................M.S. Oncology, University of Wisconsin at Madison

1997-present.....................Graduate Research Associate, The Ohio State University

PUBLICATIONS

Research Publication


FIELDS OF STUDY

Major Field: Pathology
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Dedication</td>
<td>vi</td>
</tr>
<tr>
<td>Acknowledgments</td>
<td>vii</td>
</tr>
<tr>
<td>Vita</td>
<td>ix</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xiii</td>
</tr>
<tr>
<td>List of Tables</td>
<td>xv</td>
</tr>
<tr>
<td>List of Abbreviations</td>
<td>xvi</td>
</tr>
<tr>
<td><strong>Chapters:</strong></td>
<td></td>
</tr>
<tr>
<td>1. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Background</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Hypothesis and research objectives</td>
<td>5</td>
</tr>
<tr>
<td>1.3 Organization</td>
<td>6</td>
</tr>
<tr>
<td>2. Literature Review</td>
<td>8</td>
</tr>
<tr>
<td>2.1 Clinical aspects of SMA</td>
<td>8</td>
</tr>
<tr>
<td>2.1.1 Clinical diagnosis</td>
<td>8</td>
</tr>
<tr>
<td>2.1.1.1 Type I (severe) SMA</td>
<td>9</td>
</tr>
<tr>
<td>2.1.1.2 Type II (intermediate) SMA</td>
<td>10</td>
</tr>
<tr>
<td>2.1.1.3 Type III (mild) SMA</td>
<td>10</td>
</tr>
<tr>
<td>2.1.1.4 Atypical SMA</td>
<td>11</td>
</tr>
<tr>
<td>2.1.2 Management of the disease</td>
<td>13</td>
</tr>
<tr>
<td>2.1.3 Inheritance</td>
<td>14</td>
</tr>
<tr>
<td>2.1.4 Familial heterogeneity and modifying genes</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>x</td>
</tr>
</tbody>
</table>
2.2 SMN1 gene ................................................................................................... 18
  2.2.1 Linkage to chromosome 5q13 and identification of SMN1 as the
      SMA-determining gene ........................................................................... 18
  2.2.2 Mutations in the SMN1 gene ................................................................ 19
  2.2.3 SMN protein function ....................................................................... 22
      2.2.3.1 Identified domains of the SMN protein ........................................ 22
      2.2.3.2 Gems and the spliceosome ........................................................... 25
  2.3 Role of SMN2 .......................................................................................... 29
      2.3.1 Background .................................................................................. 29
      2.3.2 mRNA alternative splicing and protein expression ......................... 30
      2.3.3 Mouse model ................................................................................ 33
  2.4 Chimeric SMN genes .............................................................................. 35
  2.5 Figures and tables .................................................................................. 39

3 Analysis of SMN1 Clinical Testing Data .................................................. 45
  3.1 Introduction .............................................................................................. 45
  3.2 Results and discussion ......................................................................... 46
      3.2.1 SMN1 deletions ............................................................................ 46
      3.2.2 Carrier testing ............................................................................... 47
  3.3 Conclusions ............................................................................................ 48
  3.4 Figures and tables .................................................................................. 50
  3.5 Materials and methods .......................................................................... 52

4 Hybrids Monosomal for Chromosome 5 Demonstrate the Presence of Two-Copy
   SMN1 Carriers ............................................................................................. 53
  4.1 Introduction .............................................................................................. 53
  4.2 Results and discussion ......................................................................... 54
      4.2.1 Parent 1: two-copy SMN1 carrier ................................................ 54
      4.2.2 Parent 2: de novo mutation .......................................................... 57
  4.3 Conclusions ............................................................................................ 57
  4.4 Figures and tables .................................................................................. 60
  4.5 Materials and methods .......................................................................... 67
      4.5.1 Genomic DNA isolation and identification of SMN1 copy
         number .................................................................................................. 67
      4.5.2 Creation of mouse/human hybrid clones ....................................... 67
      4.5.3 PCR amplification of SMN exons and APC .................................. 68
      4.5.4 Modified SMN1 dosage assay used on hybrid clones ................. 69
      4.5.5 Haplotype analysis ....................................................................... 69
      4.5.6 Probe labeling and fluorescent in situ hybridization (FISH) ......... 70

5 SMN1 Allele-Specific Intragenic Mutation Panel Increases the Sensitivity of
   Clinical Testing for SMA ........................................................................ 71
LIST OF FIGURES

Figure | Page
-------|------
2.1 Neurogenic atrophy of skeletal muscle in spinal muscular atrophy | 39
2.2 Map of duplicated genes in the SMA region | 40
2.3 Structure of the SMN protein | 41
2.4 SMN protein complex | 42
2.5 Schematic of the role of SMN | 43
3.1 Profile of the population tested for SMN1 deletions | 50
4.1 Hypothetical representation of two possible chromosomal orientations | 60
4.2 Representation of the two possible genotypes | 61
4.3 Autoradiogram of the dosage assay for parent 1 | 62
4.4 Dual color fluorescent in situ hybridization | 63
5.1 Distribution of intragenic mutations in the SMN1 gene | 75
5.2 Fluorescent allele-specific multiplex PCR panel | 76
6.1 Variation of SMN2 copy number in normal individuals | 84
7.1 Vectorette PCR procedure | 95
7.2 Vectorette amplification product | 96
7.3 Exon/intron borders of SmN gene | 97
7.4 Strategy for screening patients.................................................................98
7.5 Sequence of $SIP1$ cDNA........................................................................99
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Intragenic mutations identified in the <em>SMNI</em> gene</td>
</tr>
<tr>
<td>3.1</td>
<td>Comparison of carrier data with recessive model</td>
</tr>
<tr>
<td>4.1</td>
<td>PCR amplification of parent 1 clones</td>
</tr>
<tr>
<td>4.2</td>
<td>Haplotype analysis of parent 1 clones</td>
</tr>
<tr>
<td>5.1</td>
<td>Allele-specific primer sequence</td>
</tr>
<tr>
<td>6.1</td>
<td><em>SMN2</em> copy number in type I compared with type III patients</td>
</tr>
</tbody>
</table>
LIST OF ABBREVIATIONS

A adenine or adenosine
ALS amyotrophic lateral sclerosis
APC adenomatous polyposis coli
bp base pair
°C degrees celsius
C cytosine or cytidine
CFTR cystic fibrosis transmembrane regulator
CO₂ carbon dioxide
DNA deoxyribonucleic acid
dNTP deoxyribonucleotide triphosphate
EDTA ethylenediamine tetraacetic acid
EMG electromyography
ESE exonic splice enhancer
F forward
FISH fluorescent in situ hybridization
FVC forced vital capacity
G guanosine or guanidine
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>H₂O</td>
<td>water</td>
</tr>
<tr>
<td>HAT</td>
<td>hypoxanthine, aminopterin, thymidine</td>
</tr>
<tr>
<td>kb</td>
<td>kilobase pair</td>
</tr>
<tr>
<td>kDal</td>
<td>kiloDalton</td>
</tr>
<tr>
<td>kV</td>
<td>kilovolts</td>
</tr>
<tr>
<td>µg</td>
<td>microgram</td>
</tr>
<tr>
<td>µL</td>
<td>microliter</td>
</tr>
<tr>
<td>MDE</td>
<td>mutation detection enhancement</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>magnesium chloride</td>
</tr>
<tr>
<td>mL</td>
<td>milliliter</td>
</tr>
<tr>
<td>mmol</td>
<td>millimole</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>NAIP</td>
<td>neuronal apoptosis inhibitory protein</td>
</tr>
<tr>
<td>NAIP⁵</td>
<td>exon 5 of the NAIP gene</td>
</tr>
<tr>
<td>ng</td>
<td>nanogram</td>
</tr>
<tr>
<td>NSE</td>
<td>neuron-specific enolase</td>
</tr>
<tr>
<td>NTP</td>
<td>nucleoside triphosphate</td>
</tr>
<tr>
<td>ψ</td>
<td>pseudogene</td>
</tr>
<tr>
<td>p</td>
<td>short arm of chromosome</td>
</tr>
<tr>
<td>P</td>
<td>probability</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>pmol</td>
<td>picomole</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>q</td>
<td>long arm of chromosome</td>
</tr>
<tr>
<td>R</td>
<td>reverse</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>rRNA</td>
<td>ribosomal RNA</td>
</tr>
<tr>
<td>rt-PCR</td>
<td>reverse-transcription PCR</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulfate</td>
</tr>
<tr>
<td>SIP1</td>
<td>SMN-interacting protein</td>
</tr>
<tr>
<td>SMA</td>
<td>spinal muscular atrophy</td>
</tr>
<tr>
<td>SMN</td>
<td>survival motor neuron protein</td>
</tr>
<tr>
<td>snoRNA</td>
<td>small nucleolar RNA</td>
</tr>
<tr>
<td>snRNA</td>
<td>small nuclear RNA</td>
</tr>
<tr>
<td>snRNP</td>
<td>small nuclear ribonucleoprotein</td>
</tr>
<tr>
<td>SSC</td>
<td>sodium chloride, sodium citrate buffer</td>
</tr>
<tr>
<td>SSCA</td>
<td>single-stranded conformation analysis</td>
</tr>
<tr>
<td>SSCP</td>
<td>single-stranded conformation polymorphism</td>
</tr>
<tr>
<td>T</td>
<td>thymine or thymidine</td>
</tr>
<tr>
<td>TAE</td>
<td>tris acetic acid EDTA buffer</td>
</tr>
<tr>
<td>TBE</td>
<td>tris boric acid EDTA buffer</td>
</tr>
<tr>
<td>U</td>
<td>units</td>
</tr>
<tr>
<td>V</td>
<td>volts</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

1.1 Background

Autosomal recessive proximal spinal muscular atrophy (SMA) is caused by homozygous loss or mutation of the telomeric copy of the survival motor neuron (SMN1) gene (Lefebvre et al. 1995). SMN1 (previously referred to as SMN1, SMNT, and telSMN) is located on chromosome 5q13 (Brzustowicz et al. 1990; Gilliam et al. 1990; Melki et al. 1990; Daniels et al. 1992; Simard et al. 1992; Francis et al. 1993; MacKenzie et al. 1993; Soares et al. 1993; Brahe et al. 1994; Burghes et al. 1994; Melki et al. 1994; Wirth et al. 1994, 1995). The incidence of 5q-SMA is approximately 1/10,000 live births, with a carrier frequency of about 1/50 (Roberts et al. 1970; Peam 1973; Peam 1978a; Melki et al. 1996, McAndrew et al. 1997). Clinically, SMA patients are classified as type I, II, or III based on the severity of the disease and the age of onset (Munsat 1991; International SMA Consortium 1992; Munsat and Davies 1992; Dubowitz 1995; Zerres and Schoneborn 1995). SMA type I, also known as Werdnig-Hoffmann disease (Werdnig 1891; Hoffmann 1892), is the most severe presentation. Type I infants are hypotonic and have severe proximal muscle wasting due to a lack of α-motor neurons...
in the anterior horn of the spinal cord. These patients are diagnosed prenatally or within six months after birth. They almost never sit unaided and usually die within the first two years of life, most often due to respiratory muscle weakness. Patients with intermediate type II SMA develop the disease between 6 and 18 months of age. They are able to sit unaided, but usually do not stand or walk. SMA type III, also called Kugelberg-Welander disease, is less severe (Kugelberg and Welander 1954, 1956). Type III patients are diagnosed after 18 months, can have children, and usually live at least into the second or third decade (Russman et al. 1996).

Deletion and gene conversion events result in 92-98% rate of homozygous loss of the \textit{SMN1} gene in patients with classic SMA (Bussaglia et al. 1995; Cobben et al. 1995; Hahnen et al. 1995; Lefebvre et al. 1995; Rodrigues et al. 1995; Devriendt et al. 1996; Hahnen et al. 1996; Matthijs et al. 1996; van der Steege et al. 1996; Velasco et al. 1996; Burghes 1997; Campbell et al. 1997; DiDonato et al. 1997; Talbot et al. 1997a; Wirth et al. 1999; Wirth 2000). This high rate of loss of \textit{SMN1} in well-characterized SMA patients has enabled the development of an effective PCR-based assay for the molecular diagnosis of SMA (Lefebvre et al. 1995; van der Steege et al. 1995), which has made muscle biopsy unnecessary. Despite the fact that more than 92-98% of very well phenotyped SMA patients lack any intact \textit{SMN1}, it has been our experience that less than 43% of patients referred for diagnostic testing have the common \textit{SMN1} deletion (described in chapter 3). Although the remaining patients share some of the typical features of SMA patients, including hypotonia, progressive muscle weakness, and loss of
ambulation, they do not have chromosome 5q-linked proximal spinal muscular atrophy. For this reason, the molecular diagnosis is absolutely necessary for the clinical diagnosis of SMA and for accurate genetic counseling.

Studies of SMA families have revealed two unusual occurrences. First, approximately 2% of SMA cases arise as the result of de novo mutation events (Wirth et al. 1997a). This is unusual for a recessive disorder. Second, the copy number of SMN1 can vary on a chromosome; we have observed that about 4% of the normal population possesses three copies of SMN1. Additional indirect evidence for the occurrence of two SMN1 genes on one chromosome was provided by the observation of McAndrew et al. (1997) of a patient who had normal two-copy SMN1 dosage and multiple affected children. It is, therefore, possible for an individual to possess one chromosome with two copies of SMN1 and a second chromosome with zero copies. Such a parent would appear normal by a quantitative SMN1 dosage assay, but in reality would be a carrier with a 25% recessive risk of having another affected child with a carrier spouse. Distinguishing between the normal recessive risk of a two-copy SMN1 carrier and the very low risk of recurrence after a de novo mutation in apparent non-carrier parents of affected SMA children is extremely important for genetic counseling. In chapter 4, I discuss experiments that provided definitive proof of the existence of SMA carriers with two SMN1 copies on a single chromosome (Mailman et al. 2001a).

Between 1 and 4% of SMA patients, including types I, II, and III patients, are compound heterozygotes, meaning that they are deleted for one allele of the SMN1 gene and have an intragenic mutation in the other allele (reviewed in Parsons et al. 1998a;
Wirth et al. 1999; Wirth 2000). Four “common” mutations (occurred in more than one family) have been described by our lab (Parsons et al. 1998a). Three mutations not identified in this population are common in a European population (Wirth et al. 1999). Clinically, intragenic mutations are not screened for because there is no overwhelmingly prevalent mutation, mutations are spread throughout the gene, and the presence of the highly homologous $SMN2$ gene makes the identification of intragenic mutations technically challenging. Chapter 5 describes the development of a panel of allele-specific fluorescent primers that preferentially amplify the mutant allele and allow for the simple, rapid, and cost-effective detection of any of these seven common mutations. This assay increases the sensitivity of mutation testing for SMA patients.

Because of a 500 kb inverted duplication, which occurred sometime in our genomic history, many of the genes in the chromosome 5q13 region have copy genes. $SMN1$ and the centromeric copy $SMN2$ differ by a critical nucleotide in exon 7 that alters the splicing of exon 7 in the mRNA (Lorson et al. 1999; Monani et al. 1999a). While the majority of $SMN1$ transcript is translated to full-length SMN protein, only a very small percentage of $SMN2$ results in full-length protein. While $SMN1$ is absent in 92-98% of SMA patients as a result of deletion or gene conversion events (Bussaglia et al. 1995; Cobben et al. 1995; Hahnen et al. 1995; Lefebvre et al. 1995; Rodrigues et al. 1995; Devriendt et al. 1996; Hahnen et al. 1996; Matthijs et al. 1996; van der Steege et al. 1996; Velasco et al. 1996; Burghes 1997; Campbell et al. 1997; DiDonato et al. 1997a; Talbot et al. 1997a), loss of $SMN2$ does not cause SMA. Although approximately 5% of normal controls homozygously lack the $SMN2$ gene (Lefebvre et al. 1995), SMA patients always
have a copy of the SMN2 gene, as loss of all SMN genes is believed to result in embryonic lethality. Mouse models of SMA have been produced by expressing the SMN2 gene in a mouse lacking murine Smn (Schrank et al. 1997, Hsieh-Li et al. 2000, Monani et al. 2000a). These studies clearly demonstrate that the copy number of SMN2 is instrumental in determining phenotypic severity. This study compares SMN2 copy number between 79 type I and 90 type III patients and demonstrates that there is a significant difference in the number of SMN2 genes present (chapter 6).

1.2 Hypotheses and research objectives

The large majority of SMA cases are explained by homozygous loss of the SMN1 gene. The working hypothesis throughout this dissertation is that the sensitivity of diagnostic SMA testing could be increased by addressing the minority of SMA cases caused by compound heterozygosity and mutations at a novel locus. We also hypothesized that the prognostic value of SMA deletion and carrier testing could be improved by the determination of SMN2 copy number and by the identification of carriers with two copies of SMN1 on a single chromosome, respectively. Specifically, we set out to determine:

1. the percentage of patients with symptoms of SMA that are deleted for SMN1.
2. if any two-copy SMN1 parents with affected children are carriers with two copies of SMN1 on a single chromosome.
3. if a rapid and inexpensive allele-specific PCR panel could be used to identify common intragenic mutations in SMN1.
4. if $SMN2$ copy number correlated with SMA type.

5. if the $SMN$-associated genes $SIP1$ or $SmN$ were mutated in SMA-like patients

1.3 Organization

This dissertation is divided into eight chapters:

1. Introduction

2. Literature Review

3. Analysis of $SMN1$ Clinical Testing Data

4. Hybrids Monosomal for Chromosome 5 Demonstrate the Presence of Two-Copy $SMN1$ Carriers

5. $SMN1$ Allele-Specific Intragenic Mutation Panel Increases the Sensitivity of Clinical Testing for SMA

6. Analysis of $SMN2$ Copy Number in Patients and Normal Controls Demonstrates a Modifying Effect on Phenotype

7. Genetics of Proximal SMA Unassociated with $SMN1$

8. Conclusions

Chapter 2 briefly describes spinal muscular atrophy including incidence, clinical description, and genetic etiology. An extensive literature review of background information important for the experiments performed in subsequent chapters is provided. Chapter 3 is a summary of deletion and carrier testing results collected in the diagnostic laboratory. Chapter 4 describes the first direct proof of the existence of carriers with two copies of $SMN1$ on the same chromosome. A novel PCR-based assay that detects common intragenic mutations in compound heterozygotes is discussed in chapter 5. $SMN2$ copy number is compared between types I and III SMA in chapter 6. Chapter 7
includes mutation screens of the *SMN*-associated genes *SIP1* and *SmN* in patients with symptoms resembling SMA. Finally, chapter 8 summarizes our findings from these studies.
CHAPTER 2

LITERATURE REVIEW

2.1 Clinical aspects of SMA

2.1.1 Clinical diagnosis

Spinal muscular atrophy is a disease with a large spectrum of severity ranging from prenatal onset with death in the first months of life to adult onset with a normal lifespan. The common clinical feature is progressive proximal muscle weakness and there is a relationship between age of onset and prognosis (Dubowitz 1964). Based on age of onset, patients are categorized as having type I, II, or III SMA (Munsat 1991; International SMA Consortium 1992; Dubowitz 1995). The definitive test, now that the \textit{SMN1} gene has been determined to be mutated in 92 to 98\% of SMA cases, is a molecular test involving PCR amplification of exon 7 of the \textit{SMN} genes followed by \textit{Dra} 1 and \textit{Dde} 1 digestion, respectively (van der Steege et al. 1995). \textit{SMN2} is cut by both enzymes, thereby distinguishing exons 7 and 8 of \textit{SMN1} and \textit{SMN2}. Electrophoresis through a polyacrylamide gel demonstrates the presence or homozygous absence of the
critical exon 7 of the \textit{SMN1} gene. Single-stranded conformation analysis (SSCA) is also used by some laboratories to detect \textit{SMN1} exon 7 (Lefebvre et al. 1995). If the molecular test does not show the homozygous absence of \textit{SMN1} exon 7, direct sequencing can be used to detect intragenic mutations in a compound heterozygotic state, or physiological measurements can be taken. A requirement for the diagnosis of SMA is the presence of normal nerve conduction velocity. Acute denervation (presence of fibrillation potentials) should be present by EMG. Muscle biopsy should indicate neurogenic atrophy with evidence of reinnervation (Figure 2.1). In infants, large round fibers in denervated fascicles are common. After mid-childhood, angular fibers (evidence of acute denervation) and fiber type grouping (evidence of reinnervation) should be seen (Iannaccone 1998)

\textbf{2.1.1.1 Type I (severe) SMA}

Type I SMA, also called Werdnig-Hoffmann disease (Werdnig 1891; Hoffmann 1892), has the earliest onset; prior to six months of age. These patients almost never sit unaided (Iannaccone et al. 1993), have severe muscle weakness and hypotonia, and usually die within the first two years of life due to respiratory insufficiency resulting from intercostal muscle wasting. The degree of respiratory muscle weakness is the most important prognostic indicator. Because of the intercostal muscle weakness, type I SMA patients often have a deformity that causes the ribcage to appear bell-shaped. Although proximal muscles are affected most, muscles of the limbs also atrophy. Hands and feet are sometimes the only muscles that can be moved voluntarily. Deep tendon reflexes are
absent. Tongue fasciculations are common and some patients have a tremor of the fingers called polymyoclonus. Sphincter muscles are unaffected. These patients fatigue quickly; and are prone to malnutrition and failure to thrive because they are too tired to feed themselves sufficiently. The age of onset within the cohort of type I patients is important for prognosis. The mortality rate for type I patients with onset of symptoms prior to three months of age is 90%, while type I babies with onset after three months can live into adulthood with severe motor handicaps (Russman et al. 1992).

2.1.1.2 Type II (intermediate) SMA

Classification as a type II patient requires onset between 6 and 18 months of age. Until six months, these patients achieve relatively normal motor milestones, however they are often hypotonic. The lower extremities are affected more than the upper, and although the majority of these patients are able to sit unaided, they are almost never able to walk. Some have a clubfoot deformity. The deep tendon reflexes vary, however they are normal in muscles that retain their strength. Some type II patients live until the third or fourth decade of life and the oldest known patient lived to be 72 years old (Russman et al. 1996).

2.1.1.3 Type III (mild) SMA

Type III SMA, also called Kugelberg-Welander or Wohlfhart-Kugelberg-Welander syndrome (Kugelberg and Welander 1954, 1956), is the mildest form of chromosome 5q-linked spinal muscular atrophy. Type III patients present with
symptoms after 18 months of age, are usually ambulant at some point, and often have a normal life expectancy (Russman et al. 1992). Seventy-five percent of type III patients stand and 25% walk by the age of two years (Byers and Banker 1961). The age at onset of symptoms is a prognostic indicator for whether or not these patients maintain the ability to walk (Zerres et al. 1997). If onset is before two years of age, the patient is likely to stop walking by 15 years, whereas onset after two years of age predicts that the patient will continue walking after the fifth decade of life (Russman et al. 1996). Type III patients who are ambulatory often have a waddling gait, lumbar lordosis, protuberant abdomen and are thin. Their deep tendon reflexes may or may not be normal (Iannaccone 1998). Serum creatine kinase levels are often elevated in type III patients (this does not occur in types I or II) (Kugelberg 1975). This disease is often confused with Becker or limb-girdle muscular dystrophy (Furukawa et al. 1968; Lunt et al. 1989; Topaloglu et al. 1989) and is excluded as amyotrophic lateral sclerosis (ALS) because of the benign course of the disease and the lack of involvement of the corticospinal tract (Iannaccone 1998). The most serious orthopedic issue that type III patients face is scoliosis (Shapiro and Specht 1993).

2.1.1.4 Atypical SMA

Types I, II, and III spinal muscular atrophy, as described above, are the most classical forms of the 5q-linked disease. Arthrogryposis is an uncommon finding in SMA, however patients have been described with this symptom and with homozygous deletion of the SMN1 gene (Bürghlen et al. 1996a; Bingham et al. 1997). Additionally
there is a rare X chromosome-linked severe form of spinal muscular atrophy with arthrogryposis (Greenberg et al. 1988; Kobayashi et al. 1995). Adult-onset type IV SMA has been described, however this is a heterogeneous disease as some cases are linked to chromosome 5q and others are not (Brahe et al. 1995; Clermont et al. 1995; Zerres et al. 1995). Multiple forms of autosomal dominant SMA have been described. Spinal muscular atrophy - type Finkel is a late onset (mean age = 49 years) form of the disease which has many features in common with 5q-recessive SMA including progressive muscle weakness and proximal muscle atrophy starting in the lower body. Tendon reflexes are impaired, fasciculations are common, and sensory and cranial nerves are unaffected. Unlike 5q-SMA, Finkel patients have cramps, fits of suffocation, and mytonia (Richieri-Costa et al. 1981). This disease has not been linked to chromosome 5.

There exist forms of dominantly inherited SMA presenting at the time of birth with arthrogryposis (Fleury and Hageman 1985) and without arthrogryposis with onset of symptoms between 0-8 years (Pearn 1978b). This last form, which represents 25% of childhood cases of SMA, has been shown to be unlinked to chromosome 5q (Kaufsch et al. 1991). An adult-onset dominant form presenting at a mean age of 37 years with rapid progression and very high penetrance has also been presented (Pearn 1978b).

Furthermore, a disease with the clinical and histological findings of recessive type II SMA was found in two siblings who were not deleted for SMN1 and inherited different markers flanking the SMN1 gene (Nevo et al. 1998).
2.1.2 Management of the disease

At this time, there is no cure for spinal muscular atrophy. Consequently, the symptoms are treated in an attempt to extend the life and the comfort of the patient. Weakness of the intercostal muscles causes respiratory insufficiency and restrictive lung disease. Assisted coughs, chest percussion therapy (Bach 1994), and intermittent positive breathing (Bach and Alba 1991) aid in promoting healthy lung capacity and in preventing pneumonia. Monitoring that the forced vital capacity (FVC) is greater than 50% and that the concentration of serum CO₂ is within appropriate limits allows physicians to apply the respiratory therapies described above at appropriate and necessary times. Severe thoracolumbar scoliosis in SMA patients is treated by the use of spinal orthotics, which do not prevent the progression of scoliosis, but instead increase the comfort of sitting and are only used if they do not inhibit the respiratory capacity of the patient. Surgery to fuse vertebrae corrects the scoliosis, helps the patient to breathe, and greatly increases their comfort level (Aprin et al. 1982; Daher et al. 1985; Piasecki et al. 1986). Clubfoot deformity is not common and usually does not require surgical intervention. Because of the limited mobility of many SMA patients, flexion contractures of the hips, knees, and ankles become a problem. Regular practice of range-of-motion exercises prevent and treat these contractures (Jannaccone 1998). SMA patients are known to have high cognitive function (Whelan 1987). It is important to ensure that despite their physical limitations, these patients have access to intellectually stimulating surroundings.
2.1.3 Inheritance

Chromosome 5q-linked spinal muscular atrophy is inherited in an autosomal recessive manner with very high penetrance. SMA is the most common genetic cause of infant mortality and is the second most common lethal autosomal recessive disorder after cystic fibrosis. The incidence of all three types of SMA combined is 1:10,000 and the carrier frequency is approximately 1:50 (Roberts et al. 1970, Pearn 1973; Pearn 1978a; Melki et al. 1996, McAndrew et al. 1997). Although the incidence of type I patients is much higher than for types II or III, the prevalence of type I patients is the lowest because of the high mortality rate in these patients. Unlike most recessive disorders, SMA has a significant de novo mutation rate of about 2% (Wirth et al. 1997a). Because of this fact, not all parents of affected patients are carriers of a SMN1 mutation. In fact, we have shown that 5% of the parents of probands have two copies of SMN1 (Mailman et al. 2001b) and have demonstrated that up to 3% of these parents have two copies of SMN1 on a single chromosome, and that a null chromosome can be transmitted to their child with a recurrence risk of 25% (Mailman et al. 2001a). Further complicating the inheritance pattern of spinal muscular atrophy was a report describing an example of maternal germ-line mosaicism, in which an apparent non-carrier gave birth to three affected siblings with the same mutation (Campbell et al. 1998).

2.1.4 Familial heterogeneity and modifying genes

SMN1 mutations are not 100% penetrant. Families exist in which siblings of affected patients have homozyous SMN1 mutations, yet are unaffected (Cobben et al.
1995; Wang et al. 1996). This may be explained by discordance in age of onset of the
disease, however a difference in age of onset of more than 15 years (as is the case here) is
extremely rare (Hausmanowa-Petrusewicz et al. 1968; Bouwsma and Leschot 1986:
Müller et al. 1992; Burghes et al. 1993; Rudnik-Schöneborn et al. 1994, 1996). In one of
these families an individual presented with symptoms at one year of age, while a
homozygously deleted brother is a 42 year old athletic professional diver (Cobben et al.
1995). In another study, a family with four affected siblings was described. Three
homozygously lacked SMN1, while the fourth was deleted for only a single copy of the
gene (Wirth et al. 1997b). Additionally, there are examples of sibs presenting with
different types of SMA (Müller et al. 1992; Burghes et al. 1994; Cobben et al. 1995;
Hahnen et al. 1995; Wang et al. 1996; DiDonato et al. 1997a) which were hypothesized
to be due to different alleles at the SMA locus on chromosome 5q (Becker 1964; Zerres
and Grimm 1983; Bouwsma and Leschot 1986; Zerres et al. 1987). The fact that in many
of those cases the sibs had identical haplotypes in the SMA region of chromosome 5
provided evidence that a modifying gene was the cause of the variation. Müller et al.
(1992) did a statistical analysis of four families that had at least two sibs with types II and
III SMA. They showed that the differences were not due to allelic variation (P<0.001).
and concluded that the differences in phenotype were likely due to environmental factors
or modifying genes.

All of the above examples of familial heterogeneity support the possibility of
modifying genes. Because they might segregate independently of the SMN1 gene,
modifying loci could account for the phenotypic differences seen within families. None
of those reports described \textit{SMN2} copy number, which is known to modify the SMA phenotype (Hahnen et al. 1996; Velasco et al. 1996; McAndrew et al. 1997; Burghes 1997; Campbell et al. 1997; Parsons et al. 1998a; Wirth et al. 1999; Schrank et al. 1997; Hsieh Li et al. 2000; Monani et al. 2000a; Mailman et al. 2001b). It is possible that within a family, differences in phenotype could be accounted for by varying numbers of \textit{SMN2} gene copies.

Other potential modifying genes would be those whose products interact with the \textit{SMN} protein, those that affect the expression of the \textit{SMN} gene, those that increase the percentage of full-length \textit{SMN} mRNA, or those that compensate for a function of \textit{SMN} protein. Tissue-specific expression of modifying genes also has the possibility of providing an explanation as to why the phenotype of SMA is limited to neurons. Profilin II is a protein that has been found to bind to polyproline motifs in \textit{SMN} and to colocalize in gems in neuron-specific nuclei (Giesemann et al. 1999). The \textit{Sm} core protein \textit{N} is a pre-mRNA splicing factor that is expressed only in neurons. \textit{Sm} \textit{N} interacts with \textit{SMN}, and is therefore a good candidate for modifying the SMA phenotype. I show in chapter 7 that mutations in \textit{Sm} \textit{N} do not cause spinal muscular atrophy. Still, it may modify the phenotype. The same is true for the \textit{SIP1} gene, which interacts with \textit{SMN} in subnuclear compartments called gems (see section 2.2.3.2). We found (chapter 7), as did Helmken et al. (2000), that mutations in \textit{SIP1} are not involved in a phenotype resembling SMA. Htra2-B1 is a protein that was found to bind an exonic splice enhancer site in exon 7 of \textit{SMN} resulting in increased fidelity of splicing and a concurrent increase in the amount of full-length \textit{SMN} produced (Hofmann et al. 2000). This was an interesting candidate to
modify the phenotype as it might be utilized clinically by inducing a greater amount of full-length transcript production from the \textit{SMN2} gene. \textit{Htra2-B1} was tested for mutations in families with a heterogeneous SMA phenotype, under the hypothesis that mutations in this splice enhancer protein might make the SMA phenotype worse and therefore, modifies the severity of SMA between sibs. No differences were found in the sequence of the gene between 36 sibs in 15 families, and so \textit{Htra2-B1} was discounted as a modifier of SMA (Helmken and Wirth 2000). The anti-apoptotic protein Bcl-2 was reported to bind to SMN (Iwahashi et al. 1997). This association with a protective function against Bax-mediated or Fas-mediated apoptosis was an intriguing possible explanation for the loss of motor neurons in SMA. However, this interaction was not confirmed in a subsequent study (Coovert et al. 2000) and so the connection between SMN and apoptosis awaits further evidence. Two transcription factors, FUSE binding protein and the papilomavirus E2 protein, have been found to interact with SMN (Williams et al. 2000; Strasswimmer et al. 1999). All of these proteins are potential modifiers of SMA.

Thus far, the only gene known to modify the phenotype of SMA is the centromeric \textit{SMN2} gene. It is likely that other genes or chemical compounds will be found in the near future which either upregulate expression or compensate for the splicing defect in \textit{SMN2}. Increasing the overall amount of full-length protein produced by \textit{SMN2} seems the most promising therapeutic avenue for spinal muscular atrophy.
2.2 SMN gene

The SMN gene contains 9 exons (1-2a, 2b-8) and encompasses a genomic region of 27 kb (Bürglen 1996a; Chen et al. 1998; Monani et al. 1999a). It produces a 1.7 kb transcript coding for 294 amino acids (Lefebvre et al. 1995). Approximately 90% of SMN transcript is full-length. The remaining 10% is deleted for exon 5 (Lefebvre et al. 1995; Gennarelli et al. 1995; Parsons et al. 1996). Full-length SMN transcript yields a 38 kDa protein which is ubiquitously expressed and is expressed at particularly high levels in the spinal cord (Coovert et al. 1997; Lefebvre et al. 1997). The SMN protein is found in both the cytoplasm and in the nucleus (Liu and Dreyfuss 1996). Recently, two distinct nuclear compartments have been found to contain SMN (Meister et al. 2000) implying the possibility of more than one nuclear function.

2.2.1 Linkage to chromosome 5q13 and identification of SMN as the SMA-determining gene

The initial linkage of recessive proximal progressive spinal muscular atrophy to chromosome 5 was in 1990 when Gilliam et al. localized the gene to chromosome 5q11.2-13.3. Further refinement of the region to 5q13 was subsequently carried out by several groups (Brzustowicz et al. 1990; Melki et al. 1990; Daniels et al. 1992; Simard et al. 1992; Francis et al. 1993; MacKenzie et al. 1993; Soares et al. 1993; Brahe et al. 1994; Burghes et al. 1994; Melki et al. 1994; Wirth et al. 1994, 1995a). The gene for SMA was determined to lie in a complicated and unstable region of chromosome 5 in which a 500 kb duplication has resulted in the presence of two copies of each of the genes present in
the region (Figure 2.2). For a time, investigators were not clear as to whether mutations in the SMN1 gene or the NAIP gene were responsible for spinal muscular atrophy.

Deletions of NAIP were found in 67% of SMA patients and in only 2% of controls (Roy et al. 1995). Deletions of the SMN1 gene were found in more than 98% of SMA patients (Lefebvre et al. 1995; Gilliam 1995) indicating that NAIP mutations likely marked the extent of the SMN1 deletion. The finding of intragenic mutations in the SMN1 gene in SMA patients, in the absence of NAIP mutations, made it clear that SMN1 was the SMA gene (Lefebvre et al. 1995; Bussaglia et al. 1995; Parsons et al. 1996; Bürglen et al. 1996c; Brahe et al. 1996; McAndrew et al. 1997; Clermont et al. 1997; Rochette et al. 1997; Hahnen et al. 1997; Talbot et al. 1997b; Gambardella et al. 1998; Parsons et al. 1998a, 1998b; Wang et al. 1998; Wirth et al. 1999). No intragenic mutations have yet been reported in NAIP in a patient with SMA.

2.2.2 Mutations in SMN1 gene

A summary of the data from several studies indicates that about 94% of all SMA patients are homozygously deleted for the SMN1 gene. The deletion frequency can be subdivided by SMA type as follows: 96% type I, 94% type II, and 86% type III (reviewed in Wirth 2000). A large study of 525 well-defined SMA patients indicated that 92% were homozygously deleted, 3.6% had intragenic mutations, and 4.6% had no mutation (Wirth et al. 1999). The exact extent of the deletions of the SMN1 gene are unknown, however approximately 67% of these mutations extend through the NAIP gene (Roy et al. 1995).
The penetrance of SMN1 mutations has not been determined, however there are reported cases of asymptomatic individuals with homozygous SMN1 deletions (Cobben et al. 1995; Wang et al. 1996), indicating that the penetrance is less than 100%.

Studies of different ethnic populations indicate that the SMN1 mutation rates vary across groups. Some of the variation could be due to differences in diagnostic criteria between groups, however it is likely that the incidence differs between populations. For example, the black population of South Africa has a homozygous deletion rate of only 58.7-65.5% (Stevens et al. 1999; Labrum et al. 2000) of SMA cases. Stevens et al. report that these patients have more facial muscle weakness than patients of European ancestry. In Turkish patients, 85-93% of patients are deleted (Erdem et al. 1999; Savas et al. 2000), as are 100% (48/48) of Chinese patients (Chang et al. 1995), 85% of Japanese patients (Nishio et al. 1999), 92.3% of Spanish patients (Velasco et al. 1996), 87% of Canadian (mostly French Canadian) patients (Simard et al. 1997), 98.6% of Korean patients (Shin et al. 2000), 88.4% of Slovakian patients (Zat'ková et al. 2000), 92.9% of Croatian patients (Barišić et al. 1998), 82% of Saudi patients (Al Rajeh et al. 1998), and 85% of Bulgarian patients (Jordanova et al. 1998). Deletion of the NAIP gene is often used to mark the extent of the deletion involving SMN1. Interestingly, whereas Chinese patients have a significantly lower rate of NAIP5 deletion (Lin et al. 1995; Chang et al. 1995) than the rate of <50% in patients of European and North American origin (Aubry et al. 1995; Lefebvre et al. 1995; Wirth et al. 1995c), Arabic patients have a much higher rate of NAIP5 deletion (Samichuk et al. 1996). Only 26% of Bulgarian patients are deleted for NAIP5 (Jordanova et al. 1998). These ethnic-specific NAIP5 and SMN1
deletion rates provide interesting backgrounds for founder effect studies of specific deletions as well as for the possible involvement of the *NAIP* gene in the modification of the SMA phenotype.

Compound heterozygotic patients involving a deleted allele and an allele with an intragenic mutation make up approximately 1-4% of SMA patients (Parsons et al. 1998; Wirth et al. 1999). Twenty-three different intragenic mutations have been reported and are described in Table 2.1. Worldwide, the most common mutations are Y272C (17%) and 800ins11 (13%). In the German population, Y272C makes up 37% of all intragenic mutations (Wirth 2000). Although there is not a predominantly common mutation within the *SMN1* gene, mutations cluster in exons 6 and 7 (Talbot et al. 1997b; Hahnen et al. 1997; Parsons et al. 1998a; Wirth et al. 1999) and results in a loss in the ability of the protein to self-oligomerize (Lorson et al. 1998); a process critical to the function of SMN.
2.2.3 SMN protein function

2.2.3.1 Identified domains of the SMN protein

SMN protein functions as an oligomer and has several domains that bind other proteins (Figure 2.3). An important domain of the SMN protein is the oligomerization domain in exon 6, which allows for dimerization of the protein (Lorson et al. 1998; Pellizzoni et al. 1999). A yeast two-hybrid screen for SMN pulled down itself, and a direct correlation was identified between the ability of the protein to oligomerize and the clinical severity of the disease. Furthermore, it was found that although exons 5 and 7 are not individually required for oligomerization, they greatly enhance dimerization in in vitro studies. The oligomerization domain is novel and is highly conserved among human, mouse, rat, and the nematode (Lorson et al. 1998).

Full-length SMN protein is found diffusely throughout the cytoplasm and in gems in the nucleus (Liu and Dreyfuss 1996; Francis et al. 1998). Constructs containing various SMN mutations were made by Le et al. (2000) and were transfected into cells of type I SMA patients to determine which domains of the protein determined cellular localization. The amino-terminal end of the protein was found to be important for proper intracellular distribution of SMN protein. As discussed in section 2.2.3.2, SMN is believed to have a role in recycling pre-mRNA splicing factors across the nuclear membrane. Mutations in the amino-terminus may prevent the release of SMN from its nuclear complex with spliceosomal proteins, increasing the proportional concentration of SMN in abnormal nuclear aggregates (Pellizzoni et al. 1998; Le et al. 2000). Constructs
lacking exon 7 affected SMN distribution by increasing the proportion of the protein in the nucleus. Exon 7 was found not to be required for gem formation if full-length SMN was present with which to heterodimerize.

SMN exons 2A and 2B contain an RNA-binding domain (Lorson and Androphy 1998). It is thought that SMN may function in the cytoplasm by transporting mRNA molecules important for motor neuron development from the perikarya to distal dendrites as SMN is associated with cytoskeletal elements, binds to RNA, and shows strong immunostaining patterns in the cytoplasm suggesting axoplasmic flow (Pagliardini et al. 2000). Immunoreactivity was seen in dendrites in association with synaptic microtubules (Bechade et al. 1999). Further correlation with an axoplasmic transport function for cytoplasmic SMN comes from the finding that exons 4, 5, and 6 contain proline-rich motifs that bind to small proteins called profilins which regulate actin dynamics (Giesemann et al. 1999). Together, this evidence suggests that SMN may play a role in transporting mRNA from the nucleus to postsynaptic dendrites during neural development.

The central part of the SMN protein (amino acids 90-160) contains a conserved motif called a tudor domain (Ponting 1997; Talbot et al. 1998). The tudor domain is required for the direct binding of Sm proteins, which in turn promotes assembly of the U snRNP complex (Buhler et al. 1999). Point mutations in the tudor domain of SMN were found by this group to prevent Sm protein binding, thus providing clinical support for the importance of this domain of SMN.
The C-terminal 15 amino acids appear to be important in targeting SMN to the nucleus. Such a mutation targeted to mouse Smn in neurons yielded an animal model with defects in the morphology of the nucleus of motor neurons and of severe skeletal muscle atrophy (Frugier et al. 2000).

Several proteins bind to SMN in yeast two-hybrid screens. One of these is SIP1 (SMN-interacting protein) which binds a highly conserved domain in exon 2A of SMN (Liu et al. 1997) and to a region of exon 2B (Young et al. 2000b). Sm core proteins bind to SMN amino acids 129-135 (Buhler et al. 1999). The functions of these two proteins are discussed in section 2.2.3.2.

There is preliminary evidence that SMN may act as an activator of transcription. The papillomavirus E2 transcriptional activator and SMN were shown to interact in vitro and in vivo, where SMN was found to facilitate E2 transcriptional activity. Naturally occurring missense mutations in SMN inhibited this activation (Strasswimmer et al. 1999). The transcriptional activator FUSE also binds to SMN (Williams et al. 2000). dp103 is a murine homolog of an RNA helicase (Grundhoff et al. 1999) which interacts with SMN and may play a role in neuronal gene regulation during development (Campbell et al. 2000). Together with the above-described cytoplasmic transport of mRNA to post-synaptic dendrites, this evidence points to a possibly important function of SMN in transcriptional control of development in neurons.
2.2.3.2 Gems and the spliceosome

Antibodies to SMN show that in HeLa cells SMN localizes diffusely in the cytoplasm and in intensely staining nuclear structures that they termed gems (Liu and Dreyfuss 1996). The name “gems” is short for “Gemini of coiled bodies”, because in the nucleus of HeLa cells, gems appear next to coiled bodies, which are thought to be involved in the pre-mRNA splicing processes of small nuclear ribonucleoprotein assembly, transport, and recycling (Lamond and Carmo-Fonseca 1993; Bohmann et al. 1995). Coiled bodies also function in maturation of histone mRNA and in the processing of rRNA molecules (Gall et al. 1995). Although transcription does not take place in coiled bodies, components of the transcriptional machinery and proteins involved in cell cycle control are stored there (Grande et al. 1997; Jordan et al. 1997; Schul et al. 1998). Gems were found in the nuclei of more than 90% of HeLa cells, with the other 10% likely accounted for by the fact that gems dissociate and reassociate during the mitotic cell cycle (Liu and Dreyfuss, 1996). Later work has suggested that while HeLa cells can have coiled bodies without gems, or vice versa, this is likely to be a cell culture artifact. In tissue sections, gems and coiled bodies are always found together (Matera and Frey 1998; Young et al. 2000a). Gems have been identified in motor neurons of fetal tissue at 14 and 18 weeks gestation, which is an important time for neural development (Lefebvre et al. 1997; Francis et al. 1998; Bechade et al. 1999). Between two and eight gems on average were present per nucleus depending on the cell type analyzed (Fischer et al. 1997). Gems are associated with, but distinct from coiled bodies in HeLa cells, implying a possible shared function between the two subnuclear organelles. In mouse and human
fibroblasts, coiled bodies were less common. The presence, organization, and size of
gems varies in response to the cell cycle and with exposure to transcriptional inhibitors,
indicating a potential role of gems in RNA metabolism (Liu and Dreyfuss 1996). Gems
are virtually undetectable in type I SMA patients, while types II and III patients have less
than the normal complement of gems (Coovert et al. 1997; Lefebvre et al. 1997).

Further evidence for a role of SMN and gems in RNA metabolism came from
yeast two-hybrid screens, in which SMN was found to interact with a part of hnRNP U
(Liu and Dreyfuss 1996); a member of a family of proteins which bind to, and are
involved in the processing and transport of pre-mRNA in the nucleus (Dreyfuss et al.
1993). SMN was also shown to bind fibrillarin (Liu and Dreyfuss 1996), which is found
in conjunction with small nucleolar (sno) RNA and is necessary for pre-rRNA processing
(Jansen et al. 1991; Tollervey et al. 1991). SMN binds to itself as well as several novel
proteins including the SIP (SMN-interacting protein) family (Liu and Dreyfuss 1996),
Gemin3 and Gemin4 (Charroux et al. 1999 and 2000), and to Sm core proteins (Liu et al
1997) (Figure 2.4), which along with snRNA are part of spliceosomal small nuclear
ribonucleoprotein (snRNP) molecules (DeRobertis 1983; Fisher et al. 1985; Mattaj 1988,

SMN binds to SIP1 \textit{in vivo} and \textit{in vitro} both in the cytoplasm and in gems.
The \textit{SIP1} gene encodes a 279 amino acid / 32 kDa protein (Liu et al. 1997). \textit{SMN} and
\textit{SIP1} expression is co-regulated in a mouse model indicating a possible functional
correlation. While immunohistochemical analysis of SMN and SIP1 in motor neurons
demonstrates complete co-localization in gems, the two proteins only partially overlap in
the cytoplasm, which suggests that S1P1 has functions that are independent of SMN (Jablonka et al. 2001). S1P1 shares limited homology with the yeast protein Brr1, which functions in the production of snRNP molecules of the spliceosome (Noble and Guthrie 1996a, 1996b). SMN and S1P1 bind to snRNP, and this complex is essential for proper splicing (Figure 2.5) (Fischer et al. 1997; Liu et al. 1997; Pellizzoni et al. 1998). SMN and S1P1 immunoprecipitate as a stable 300 kDa complex, suggesting that they work in conjunction with several other proteins. In Xenopus oocytes, antibodies blocking the complex containing SMN and S1P1 prevented snRNP assembly and blocked their nuclear transport. In this cell, S1P1 was found to function by assembling the Sm core domain on U snRNP molecules. It seems likely that SMN has a direct role in the recruitment of Sm proteins to snRNA, whereas S1P1 actually assembles the Sm proteins to the snRNA molecules (Fischer et al. 1997). These experiments performed in the egg cell of Xenopus laevis should be taken in light of the caveat that unlike in human somatic cells, SMN is exclusively cytoplasmic. It was thought that the interaction between Sm proteins and SMN only occurred in the cytoplasm (Liu et al. 1997; Fischer et al. 1997; Buhler et al. 1999), however this was found not to be the case by Meister et al. (2000) who noted that Sm associates with nuclear SMN. This led to several interesting comments concerning possible additional functions of SMN. The first was that SMN can remove Sm from U snRNP in the nucleus, making the role of SMN in regulation of splicing much more important. Secondly, it is conceivable that SMN acts to recycle inactivated U snRNP from the nucleus; the opposite function previously ascribed. And thirdly, SMN might transport free Sm in the nucleus to the cytoplasm. As previous experiments analyzed this
function in the *Xenopus* system, in which SMN is strictly cytoplasmic (Fischer et al. 1997), it is now thought that Sm loading onto snRNA to form snRNP may take place both in the cytoplasm and in the nucleus (Meister et al. 2000).

An N-terminal deletion was engineered to the SMN protein and was found to severely affect nuclear snRNP organization as well as to inhibit pre-mRNA splicing in a dominant manner. Whereas wild type SMN protein was found to promote pre-mRNA splicing (Pellizoni et al. 1998), *SMN1* mutations from SMA patients prevented this increase in splicing by a loss of function or recessive mechanism (Pellizoni et al. 1998, 1999). This group concluded that SMN functions in the recycling of snRNP molecules after they are inactivated by their role in pre-mRNA splicing. They also found that the importance of SMN oligomerization is to increase affinity for Sm core protein binding. Despite this evidence, heterozygous *Smn +/−* mice that show motor neuron loss do not demonstrate defective splicing machinery (Jablonka et al. 2000), which may indicate that a different function of SMN is important in the pathology of SMA.

Compared with other tissues, motor neurons possess very high concentrations of SMN and SIP1. Additionally, neurons contain the tissue-specific proteins called Sm N (Latchman 1990) and profilin II (Geisemann et al. 1999). High levels of SMN and SIP1 along with the presence of tissue-specific proteins that interact with SMN imply a possible explanation as to why neurons might be the specific targets of degradation in spinal muscular atrophy. *SIPL1* and *Sm N* appealed to me as possible candidate genes that might be mutated in patients with SMA that did not have an identified *SMN1* mutation.
After screening patients with symptoms of SMA, we determined that mutations in neither gene cause such a phenotype (chapter 7). Helmken et al. (2000) also found that SIP1 is not involved in an SMA-like phenotype.

2.3 Role of SMN2

2.3.1 Background

Sometime in genomic history prior to non-human primates, a 500 kb inverted duplication occurred in the chromosomal region around 5q13. As a result, several genes in this area have copies. Most of the copy genes are inactive pseudogenes, however the centromeric copy of SMN1, called SMN2 (formerly referred to as cBCD541, cenSMN, SMNc, and SMN\textsuperscript{CEN}), is a functional gene whose two exonic base variants do not affect the coding sequence compared with SMNL. 35 variants in the genomic sequence differentiate SMN2 from SMN1 (Monani et al. 1999a; Monani 1999b; Lefebvre et al. 1995; Bürglen et al. 1996a). With greater than 99% identity, these are the most highly homologous functional copy genes ever reported. SMN2 copy number and phenotypic severity are inversely correlated (Lefebvre et al. 1995; Velasco et al. 1996; McAndrew et al. 1997), however it is possible for SMA patients of different grades to have the same number of SMN2 copies. A critical nucleotide in exon 7 (+6) affects an exonic splice enhancer (ESE) site and causes an important difference in the profile of alternative splicing of the mRNA (Lorson et al. 1999; Monani et al. 1999a). Whereas SMNL produces 90% full-length transcript, SMN2 produces approximately 20-30% full-length
transcript and protein (Lefebvre et al. 1995; Gennarelli et al. 1995), and thus SMN2 copy number acts as an important modifier of the SMA phenotype by increasing the net amount of full-length SMN protein produced.

2.3.2 mRNA alternative splicing and protein expression

The SMN2 gene transcript was identified in Northern blots from RNA samples of SMA patients homozygously lacking the SMNI gene (Lefebvre et al. 1995). This indicated that the SMN2 gene was indeed transcribed. While only producing 20-30% full-length transcript, the remainder of mRNA from SMN2 lacks exon 5 and/or exon 7 (Lefebvre et al. 1995; Gennarelli et al. 1995). Promoter sequence variation has been ruled out as a cause of the difference in splicing (Monani et al. 1999b). By the transfection of SMN minigenes containing each of the five nucleotide differences observed by Lefebvre et al. (1995), it was observed that only the C to T base change at exon 7 (+6) was necessary and sufficient to cause the alternative splicing of exon 7 of SMN (Lorson et al. 1999). This particular base change does not alter an amino acid, and was previously believed to be a simple polymorphism. However, it has recently been shown that the single base change alters an exonic splice enhancer, which is a DNA element that regulates splicing of exon 7 (Lorson and Androphy 2000).

Certain elements define intron/exon borders. These include splice acceptor sites, splice donor sites, and polypyrimidine tracts. When all of these elements exactly match consensus sequence, introns are properly spliced in the pre-mRNA. When the sequences are less than perfect consensus, several problems can occur, including the skipping of an
exon (exon is spliced out with the surrounding introns) or the usage of cryptic splice sites, which may include intronic sequence into the mRNA. Regulatory mechanisms exist in order to assist the spliceosome in recognizing imperfect syntax at splice sites. Exonic splice enhancers are part of such a mechanism. ESE’s are conserved sequence motifs that serve as binding sites for elements of the spliceosome (Tian et al. 1992, 1993; Watakabe et al. 1993; Xu et al. 1993; Wang et al. 1995; Buvoli et al. 1997) and allow for the inclusion of exons with sub-optimal consensus splice sites (Berget 1995; Gersappe and Pintel 1999). Lorson and Androphy (2000) found that the critical nucleotide at exon 7 (+6) borders an AG-rich exonic splice enhancer site that they have shown to be required for inclusion of SMN exon 7. Mutagenized bases within this particular ESE, called SE2, resulted in the non-inclusion of exon 7 into the mRNA. Exon 7 (+6) is one base upstream of SE2 and it is still unclear exactly how or if mutation of this base affects the ESE. It is interesting to note that ESE’s are very sensitive to T (U in the RNA) base substitutions. The base change in exon 7 (+6) is C to T, providing further evidence that this particular base change may in fact affect the ESE.

SR proteins (contain a serine/arginine-rich domain) are the specific spliceosomal proteins recruited to ESE’s. Hofmann et al. (2000) described a SR-like splicing protein called Htra2-B1, which increases the ratio of full-length:Δ7 SMN transcript produced from the SMN2 gene in a dose-dependent manner and whose binding requires an intact SE2 splice enhancer in exon 7 for proper function. The observation that Htra2-B1 stimulates the activity of the exonic splice enhancer in exon 7 and allows for the inclusion of exon 7 made Htra2-B1 an interesting candidate for mutational analysis in
spinal muscular atrophy cases in which no SMN1 mutation was identified. This study determined that mutations in this gene were not present in such SMA patients (Helmken and Wirth 2000). Another interesting study that has yet to be reported would be whether or not Htra2-B1 overexpression in the SMA mouse model is able to rescue the phenotype by increasing the amount of full-length SMN produced from the SMN2 gene. It is conceivable that after birth the phenotype of SMA is already established and that treatment may be too late. Such up-regulation strategies might have to be attempted in utero for this reason. The results of this study would have extremely important implications for clinical trials and potential future therapy for SMA patients.

The full-length protein that is translated from the SMN gene is 294 amino acids (Lefebvre et al. 1995) and 38 kDa (Liu and Dreyfuss 1996). Western blot analysis determined that SMN protein was expressed at high levels in spinal cord, brain, kidney, and liver. Medium level of expression was found in skeletal and cardiac muscle, while low levels were seen in fibroblasts and lymphocytes. In normal fetal control tissue, SMN protein was detected by 8 weeks of gestation in developing neural tube, and from the second trimester through postnatal development, high levels are observed in motor neurons (Tizzano et al. 1998). SMN levels were more than 100 times lower in spinal cord of type I SMA patients than in disease control patients with a similar neuromuscular disorder that retained two copies of SMN1 (Co overt et al. 1997). Mirabella et al. (1999) showed by immunostudies that while type I SMA patients have very low levels of SMN protein in motor neurons, they have transcript at levels quantitatively comparable to normal controls.
A panel of antibodies was used to show that SMN protein was produced (albeit at lower levels than controls) in fibroblasts of types I, II, and III SMA patients (Cooveret et al. 1997). Because these patients lacked the SMN1 gene, it was concluded that the SMN2 gene does produce full-length protein. The decrease in SMN protein levels in fibroblasts from SMA patients was moderate, despite the fact that fibroblasts are not the target tissue of SMA. Cooveret et al. also demonstrated that SMN1 copy number significantly affects the amount of functional SMN protein by comparing the number of gems (discussed in section 2.2.3.2) in the nuclei of a type I carrier parent (one copy of SMN1) and of a normal control (two copies of SMN1). They also showed that SMN2 copies are not necessarily functionally equivalent. Examples of types I, II, and III patients that all had two copies of SMN2 (and zero copies of SMN1) were apparent, suggesting either that SMN2 copy number does not always correlate with phenotype or that a modifying gene affected the phenotype. Gem number was found to correlate very well with grade of SMA, which could indicate that certain alleles of SMN2 produce a protein product that is more functional than others. (Table 2.2).

2.3.3 Mouse model

Although the SMN protein is conserved through evolution and highly homologous genes can be found in S. pombe, C. elegans, and D. melanogaster (Talbot et al. 1997b; Owen et al. 2000; Hannus et al. 2000; Miguel-Aliaga et al. 2000), the murine animal model, possessing a mammalian central nervous system, is likely to be the most useful in answering questions as to why the phenotype of spinal muscular atrophy is so specific to spinal α-motor neurons. The duplication in the SMA region of chromosome 5

33
in humans yielded two copies of the SMN gene (SMN1 and SMN2). Humans also have copies of other genes in the region, however most copies are pseudogenes, such as PNAIP. Mice have only a single copy of Smn that is comparable to human SMN1 (DiDonato et al. 1997b; Viiolet et al. 1997), and homozygous deletion of Smn causes embryonic lethality (Schrank et al. 1997). Up to seven functional copies of the Naip gene exist in the mouse (Yaraghi et al. 1998; Growney et al. 2000), while only one copy exists in humans. Because the Smn-/- mouse is not viable, attempts were made to make a less severe mutation in the mouse. A conditional knockout of Smn was made using the Cre/LoxP system so that exon 7 was spliced out only in motor neurons (Frugier et al. 2000). This approach produced offspring that had symptoms of SMA. However, this model may not be appropriate, because full-length SMN is completely absent in most motor neurons of such mice, and is expressed at normal levels in neurons that escape the NSE promoter (Monani et al. 2000b). This is inconsistent with human SMA, in which every motor neuron produces a low level of full-length SMN.

Two groups reported a different mouse model that crossed a human SMN2 transgene onto the background of a Smn gene knockout (Hsieh-Li et al. 2000; Monani et al. 2000a). The presence of the SMN2 gene rescued the embryonic lethality of the Smn knockout, and provided a mouse model of spinal muscular atrophy in which the phenotype was less severe when more copies of SMN2 were present. Although Hsieh-Li et al. (2000) reported all three types of SMA being born in the same litter, Monani et al. (2000) found a more consistent phenotype within litters that correlated with SMN2 copy number. Mice with one copy of SMN2 were stillborn or lived for up to 12 hours after
birth. Two copies of SMN2 yielded mice that lived up to six days and eight copies completely rescued the SMA phenotype. Severe mice from both studies had low levels of full-length SMN protein and gems were not detected in motor neurons.

2.4 Chimeric SMN genes

Chimeric SMN genes, as defined in this dissertation, are hybrid genes with SMN2 sequence upstream of and including exon 7 with SMNI sequence downstream of exon 7. This situation can occur by several possible mechanisms including a deletion of all of the sequence between SMNI and SMN2 resulting in the fusion of the genes by unequal crossover between the SMNI and SMN2 genes, or by a gene conversion event where sequence from SMN2 is copied into SMNI. This process is often the result of the resolution of a heteroduplex formed by a double-stranded break repair mechanism during replication and is distinct from unequal recombination. Gene conversion was first demonstrated in a metazoan by Smith et al. (1970) in Drosophila, and has since been shown to play an important role in the molecular nature of spinal muscular atrophy.

Much indirect evidence pointed toward gene conversion as a mechanism that played a role in the loss of the SMNI gene in SMA patients with a concomitant gain in SMN2 gene copy number. Greater than 99% identity exists between the SMNI and SMN2 genes (Monani et al. 1999), which makes conversion a distinct possibility. Whereas 50% of type I patients have two chromosomes each with a single copy of SMN2 (DiDonato et al. 1994; Melki et al. 1994; Lefebvre et al. 1995; Wirth et al. 1995b), type II and type III patients usually have at least three copies of SMN2 (DiDonato et al. 1994: 35
Wirth et al. 1995b). The loss of NAIP\(^{5}\) in 50\% of type I patients (Roy et al. 1995), which is not observed in a high frequency in types II and III patients (Roy et al. 1995; Hahnen et al. 1995; Cobben et al. 1995; Thompson et al. 1995; Rodrigues et al. 1996; Velasco et al. 1996), indicated that large deletions are the likely mode of SMN1 loss in such type I patients. Velasco et al. (1996) performed a quantitative analysis of SMN1 versus SMN2 gene copy number in what they believed to be obligate carriers of types I, II, and III SMA. They found that parents of mild SMA patients were much more likely to have three or more copies of SMN2, while parents of severe patients most often had one or two copies. The quantitative dosage assay of McAndrew et al. (1997), which does not rely on the assumption of obligate carrier status, found that there were significantly more copies of SMN2 in type II and type III patients. The presence of the NAIP gene as well as the 5' markers of SMN1, despite the lack of the SMN1 polymorphism in exon 7, made gene conversion a viable mechanism to explain the loss of SMN1 with a concurrent increase in SMN2 copy number. Clearly, there was a strong correlation between SMN2 copy number and severity of the disease, and this nicely fit the model in which gene conversion changed SMN1 genes into extra copies of SMN2.

More definitive evidence of the occurrence of gene conversion came with the identification of chimeric SMN genes. Certain patient samples amplified exon 8 but not exon 7 of SMN1, while SMN2 exon 7 was present. Single strands of DNA containing a SMN2 polymorphism in exon 7 with SMN1 polymorphisms downstream were demonstrated by the following methods: PCR followed by sequencing (Talbot et al. 1997a; Nishio et al. 1999), PCR followed by digestion (Hahnen et al. 1996; van der
Steege et al. 1996; DiDonato et al. 1997a), and pulse field gel electrophoresis (Campbell et al. 1997). Two possible explanations of this observation were: 1) unequal recombination or a deletion of the sequence between $SMN2$ exon 7 and $SMNI$ exon 8. that would remove the sequence between $SMN2$ exon 7 and $SMNI$ exon 8, and 2) gene conversion which would change the variation in exon 7 from $SMNI$ to $SMN2$ (C to T).

The study by Campbell et al. (1996) using pulse-field gel electrophoresis and a NAIP probe provided direct evidence that the telomeric $SMNI$ gene was present in type II and III cases, while $SMN2$ sequence was found in exon 7. It is sometimes the case that the physical orientation of the genes in the SMA region of chromosome 5q is altered such that the $SMNI$ and $SMN2$ genes are juxtaposed (Burghes 1997). In such a case, a deletion of $SMNI$ would bring NAIP adjacent to $SMN2$, thus mimicking the situation observed during gene conversion.

Although Lorson et al. (1999) showed that a polymorphic change of C to T at the exon 7 (+ 6) position of the $SMN$ gene is necessary and sufficient to convert the splicing pattern of $SMNI$ to that of $SMN2$, chimeric $SMN2$ alleles are not always equivalent to normal $SMN2$ alleles. $SMN2$ protein in type I patients does not incorporate into gems (Coovert et al. 1997), which are subnuclear organelles that play a role in recycling members of the pre-mRNA splicing machinery. Types II and III patients can produce functional gems containing SMN. Thus far, the basis of a functional difference between normal and converted $SMN2$ genes has yet to be sufficiently explained, however it is conceivable that a position effect with respect to the NAIP gene may cause a difference in the expression of converted and normal $SMN2$ alleles (Burghes 1997). Our data,
presented in chapter 6, supports the notion that gems in type II and III patients may simply result from a greater SMN2 copy number and a consequently higher level of full-length SMN protein compared to type I patients.
2.5 Figures and tables

Figure 2.1: Neurogenic atrophy of skeletal muscle in spinal muscular atrophy. Large groups of atrophic fibers are present, often encompassing entire fascicles. Atrophic fibers are usually round, but are angular if between normal or hypertrophic fibers which are rare type 1 muscle cells compensating for the atrophy of other cells. They maintain their nuclei and display partial fiber-type grouping resulting from denervation/reinnervation by collateral sprouting of nearby axons. Large groups of fibers of the same type (>25) are most common in SMA type III.
Figure 2.2: Map of duplicated genes in the SMA region of chromosome 5q13
Figure 2.3: Structure of the SMN protein with binding sites for associated proteins.
Figure 2.4: SMN protein complex represents some of the known proteins that interact with SMN.
Figure 2.5: Schematic of the role of SMN in pre-mRNA splicing.
<table>
<thead>
<tr>
<th>Exon (E) / Intron (I)</th>
<th>Mutation</th>
<th>SMA Type</th>
<th>Number of Patients</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1</td>
<td>A2G</td>
<td>II, III</td>
<td>3</td>
<td>Parsons et al. 1998b</td>
</tr>
<tr>
<td>E1</td>
<td>Q15X</td>
<td>I, III</td>
<td>2</td>
<td>Wirth et al. 1999</td>
</tr>
<tr>
<td>E2a</td>
<td>124insT</td>
<td>I</td>
<td>1</td>
<td>Wirth et al. 1999</td>
</tr>
<tr>
<td>E2b</td>
<td>241-242ins4</td>
<td>III</td>
<td>1</td>
<td>Wirth et al. 1999</td>
</tr>
<tr>
<td>E3</td>
<td>430del4</td>
<td>I-III</td>
<td>4</td>
<td>Bussaglia et al. 1995</td>
</tr>
<tr>
<td>E3</td>
<td>E134K</td>
<td>I</td>
<td>1</td>
<td>Clermont et al. 1997</td>
</tr>
<tr>
<td>E3</td>
<td>472del5</td>
<td>I</td>
<td>1</td>
<td>Brahe et al. 1996</td>
</tr>
<tr>
<td>E4</td>
<td>542delGT</td>
<td>I-III</td>
<td>3</td>
<td>Parsons et al. 1998a, 1998b</td>
</tr>
<tr>
<td>E4</td>
<td>591delA</td>
<td>II</td>
<td>1</td>
<td>Wirth et al. 1999</td>
</tr>
<tr>
<td>E4</td>
<td>618insT</td>
<td>I</td>
<td>1</td>
<td>Clermont et al. 1997</td>
</tr>
<tr>
<td>E6</td>
<td>P245L</td>
<td>III</td>
<td>1</td>
<td>Rochette et al. 1997</td>
</tr>
<tr>
<td>E6</td>
<td>S262I</td>
<td>III</td>
<td>2</td>
<td>Hahnen et al. 1997; McAndrew et al. 1997</td>
</tr>
<tr>
<td>E6</td>
<td>Y272C</td>
<td>I-III</td>
<td>9</td>
<td>Lefebvre et al. 1995; Rochette et al. 1997; Wirth et al. 1999; Wirth 2000</td>
</tr>
<tr>
<td>E6</td>
<td>T274I</td>
<td>II, III</td>
<td>4</td>
<td>Hahnen et al. 1997; Parsons et al. 1998a; Wirth et al. 1999</td>
</tr>
<tr>
<td>E6</td>
<td>G275S</td>
<td>III</td>
<td>1</td>
<td>Bürglen et al. 1996c</td>
</tr>
<tr>
<td>E6</td>
<td>868-11del7</td>
<td>I</td>
<td>1</td>
<td>Lefebvre et al. 1995</td>
</tr>
<tr>
<td>E7</td>
<td>G279C</td>
<td>II, III</td>
<td>2</td>
<td>Wang et al. 1998</td>
</tr>
<tr>
<td>E7</td>
<td>G279V</td>
<td>I</td>
<td>2</td>
<td>Talbot et al. 1997b</td>
</tr>
<tr>
<td>E7</td>
<td>922+3del4</td>
<td>II</td>
<td>1</td>
<td>Lefebvre et al. 1995</td>
</tr>
<tr>
<td>I7</td>
<td>922+6 T→G</td>
<td>III</td>
<td>1</td>
<td>Wirth et al. 1999</td>
</tr>
</tbody>
</table>

Table 2.1: Intragenic mutations identified in the SMN1 gene
CHAPTER 3

ANALYSIS OF SMNI CLINICAL TESTING DATA

3.1 Introduction

This study presents the clinical experience of six years of SMA testing in the Molecular Pathology Laboratory at The Ohio State University. Despite the fact that more than 95% of 5q-linked SMA patients lack any intact SMN1, it has been our experience that less than 43% of patients referred for diagnostic testing have the common SMN1 deletion. Although the fact that the remaining patients share some of the common features of SMA patients, including hypotonia, progressive muscle weakness, and loss of ambulation they do not have chromosome 5q-linked proximal spinal muscular atrophy. For this reason, molecular testing is absolutely necessary for the clinical diagnosis of SMA and for accurate genetic counseling.

For autosomal recessive disorders, it is expected that parents will be obligate carriers of a mutant allele. However, there are two factors which complicate SMA carrier testing. One is the presence of de novo mutations, which occur in about 2% of probands (Wirth et al. 1997a). The other factor is the presence of single chromosomes with two copies of SMN1 confounding all dosage-based assays for SMN1 copy number.
We present the results of the carrier testing of 399 familial relations of SMA probands and describe the population most interested in being tested. Approximately 5% of parents of homozygously deleted patients had two non-deleted copies of the \( SMN1 \) gene.

3.2 Results and discussion

3.2.1 \( SMN1 \) deletions

Since 1995 our laboratory has offered diagnostic \( SMN1 \) deletion analysis to detect the homozygous absence of the gene (van der Steege et al. 1995). This deletion test was most often performed on DNA extracted from whole blood. Muscle biopsies, newborn blood spots, paraffin-embedded autopsy material, and fixed tissue on glass slides have been utilized for postmortem diagnosis. Postmortem molecular testing is often important for the confirmation of clinical diagnosis, thus allowing for accurate genetic counseling.

It has been the experience of this laboratory that the clinical symptoms of SMA are not always sufficiently specific to make a reliable diagnosis. Patient samples sent for the molecular test generally have similar symptoms including hypotonia, proximal muscle weakness, and loss of ambulation. These symptoms are not obviously specific to 5q-spinal muscular atrophy. Of the 610 patients tested, 232 (38.0%) were deleted for exons 7 and 8, and 29 (4.8%) were deleted only for exon 7. The remaining 349 (57.2%) were not homozygously deleted. Our data indicate that the clinical diagnosis of SMA
may not be straightforward and oftentimes the test is being ordered to exclude the disease. Muscle biopsies are sometimes performed on hypotonic infants prior to the molecular test.

### 3.2.2 Carrier testing

A carrier test for SMA was developed in our laboratory in 1997 (McAndrew et al.), which compares the ratios of the co-amplified products of the *SMN1* gene and the *CFTR* gene. This competitive radiolabeled PCR reaction is optimized such that the copy numbers of the *SMN1* and *SMN2* genes can be determined in comparison to the two copies of the *CFTR* gene. Currently, 399 individuals have been tested for carrier status. Parents were most often tested as shown in Figure 3.1. Unlike most autosomal recessive disorders, one cannot assume that parents of affected children are obligate carriers. *De novo* mutations (Wirth et al. 1997a) and the presence of chromosomes with two copies of *SMN1* (Mailman et al. 2001a) allow for the possibility of affected individuals born to parents with normal *SMN1* copy number. Ninety-five percent of the parents tested were shown to be carriers (Table 3.1). The remaining 4.7% present a diagnostic challenge. Because we cannot distinguish the *de novo* events (which confer a very low recurrence risk) from two-copy *SMN1* chromosomes (which confer the normal 25% recurrence risk) based solely on the *SMN1* dosage test, linkage studies, dosage testing of extended family members, or monosomal hybrid studies would be helpful in providing accurate genetic counseling to such parents.
The second most common population requesting the carrier test are the unrelated spouses of carriers. Three percent (3/100) of unrelated spouses were positive for carrier status indicating a carrier frequency of 3%, which is in agreement with the previously described frequency of 1/50 (Pearn 1978a). Eighty-eight aunts and uncles of probands were tested of which 41 (47%) were carriers. Additionally, we found that 73% (27/37) of sibs, 21% (3/14) of first cousins, and 40% (8/20) of grandparents were carriers. None of these values are different from what would be expected from a recessive model with a 2% de novo mutation rate (Table 3.1).

Because more than 57% of the patients tested are not deleted for the critical exon 7, we only performed carrier tests on families known to have an SMN1 deletion. There are two important reasons for our policy. It is possible that one would perform dosage testing on a point mutation, which may result in false negatives in approximately 1-4% of cases. Another reason is that the clinical diagnosis is not straightforward. One may, therefore, perform dosage testing on the carrier of a phenotypically similar disorder, thereby falsely reducing the recurrence risk.

3.3 Conclusions

There is a high frequency of patients with some SMA-like features that are negative for SMA testing. This fact emphasizes the importance of the molecular test in the diagnosis of spinal muscular atrophy. There is great interest in SMA carrier testing, particularly by parents of affected patients. This has uncovered the interesting finding that 5% of parents possess two intact copies of SMN1 and require linkage analysis or
monosomal hybrid studies for accurate risk assessment. With such a high rate of SMA-like patients testing negative for the common deletion, we recommend that carrier testing be reserved for at-risk relatives of patients that are positive for the SMN1 deletion. Our data indicates that if a molecular diagnosis of SMA is not made in a family, then carrier testing will be performed either on the wrong gene or on a non-hereditary disease in more than half of the cases.
Figure 3.1: Profile of the population tested for SMNI carrier status.
<table>
<thead>
<tr>
<th>Relation to proband</th>
<th>N</th>
<th>Mean</th>
<th>95% C.I.</th>
<th>Expected for relationship</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parents</td>
<td>107</td>
<td>0.953</td>
<td>(0.913, 0.994)</td>
<td>0.980</td>
</tr>
<tr>
<td>Unrelated spouse</td>
<td>100</td>
<td>0.030</td>
<td>(0, 0.064)</td>
<td>0.020</td>
</tr>
<tr>
<td>Aunts/uncles</td>
<td>88</td>
<td>0.466</td>
<td>(0.360, 0.572)</td>
<td>0.490</td>
</tr>
<tr>
<td>Sibs</td>
<td>37</td>
<td>0.730</td>
<td>(0.580, 0.880)</td>
<td>0.654</td>
</tr>
<tr>
<td>Grandparents</td>
<td>20</td>
<td>0.400</td>
<td>(0.165, 0.635)</td>
<td>0.490</td>
</tr>
<tr>
<td>Cousins</td>
<td>14</td>
<td>0.214</td>
<td>(0, 0.353)</td>
<td>0.245</td>
</tr>
</tbody>
</table>

Table 3.1: Comparison of carrier data with recessive model of inheritance
3.5 Materials and methods

DNA was extracted from peripheral venous blood from patients by a simple salting out technique (Miller et al. 1988). The presence or absence of exon 7 of SMN1 was tested using the protocol of van der Steege et al. (1995). DNA was analyzed for SMA carrier status by a SMN1 dosage assay (McAndrew et al. 1997).
CHAPTER 4

HYBRIDS MONOSOMAL FOR CHROMOSOME 5 DEMONSTRATE THE
PRESENCE OF TWO-COPY SMNI CARRIERS

4.1 Introduction

Studies of SMA families have revealed two unusual occurrences. First, approximately 2% of SMA cases arise as the result of de novo mutation events (Wirth et al. 1997a). This is unusual for a recessive disorder. Second, the copy number of SMNI can vary on a chromosome; we have observed that about 4% of the normal population possesses three copies of SMNI. Scheffer et al. (2000) has also observed an individual with three copies of SMNI. Additional indirect evidence for the occurrence of two SMNI genes on one chromosome was provided by the observation of McAndrew et al. (1997) of a patient who had normal two-copy SMNI dosage and multiple affected children. It is therefore, possible for an individual to possess one chromosome with two copies of SMNI and a second chromosome with zero copies. Such a parent would appear normal by a SMNI dosage assay, but in reality would be a carrier with a 25% recessive risk of having another affected child with a carrier spouse. Distinguishing between the normal
recessive risk of a two-copy \textit{SMN}\textit{I} carrier and the very low risk of recurrence after a \textit{de novo} mutation in apparent non-carrier parents of affected SMA children is extremely important for genetic counseling. This study provides definitive proof of the existence of SMA carriers with two \textit{SMN}\textit{I} copies on a single chromosome.

4.2 Results and discussion

In our study of 100 parents of children with homozygous \textit{SMN}\textit{I} deletions, 4 had normal two-copy \textit{SMN}\textit{I} dosage. Two of these parents were available for further study: they will be designated parent 1 and parent 2. To examine the events occurring in these parents (both of whom were fathers), mouse/human hybrids were prepared using the technique of Yan et al. (2000), which allows for the separation of the individual parental chromosomes. A dosage assay was then performed to determine the copy number of \textit{SMN}\textit{I} on each of the separated chromosomes. If either of these parents had a null chromosome, we would expect the dosage test to indicate that approximately half of the clones containing one copy of chromosome 5 would contain two copies of \textit{SMN}\textit{I}, while the other half would have zero copies (Figure 4.1 a). Conversely, if the affected child were the result of a \textit{de novo} mutation, we would expect that each clone positive for one copy of chromosome 5 would have one copy of \textit{SMN}\textit{I} (Figure 4.1 b).

4.2.1 Parent 1: two copy \textit{SMN}\textit{I} carrier

Previous \textit{SMN}\textit{I} dosage analysis indicated that the genomic DNA of parent 1 contained two copies of \textit{SMN}\textit{I} and one copy of \textit{SMN}\textit{2}. To determine which hybrid...
clones from parent 1 contained human chromosome 5, part of the APC gene (chromosome 5q21) was amplified. Of the 18 clones that contained chromosome 5, nine amplified both exons 7 and 8 of the SMN gene, and nine did not amplify exon 7 or 8 (Table 4.1). These results were consistent with a deletion of SMNI and SMN2 in half of the chromosome 5-positive clones. Interestingly, all 18 of the clones amplified SMN exon 6, indicating that all hybrids contained the 5'-end of at least one of the SMN genes (Figure 4.2). To test whether SMNI and/or SMN2 were present in the clones, exon 7 was amplified and digested with DraI (van der Steege et al. 1995). Although this test is not quantitative, it indicated that 8/9 clones contained at least one copy of both SMNI and SMN2. Clone 28 did not contain intact SMN2, presumably due to somatic mutation.

The SMN dosage assay described by McAndrew et al. (1997) is a competitive quantitative PCR reaction that compares the amplification ratio between the assumed two copies of the CFTR gene on chromosome 7 and the SMN1 gene on chromosome 5. Because chromosome 7 would not be present in every hybrid clone that contained chromosome 5 (there is a random assortment of human chromosomes), the SMN dosage assay was modified such that the APC gene was substituted for CFTR as the external standard. Using the APC gene as the endogenous standard enabled us to control for chromosome 5 copy number, thus allowing for accurate determination of SMN1 copy number.

The results of the dosage assay for parent 1 are shown in Figure 4.3. Control samples of known dosage were run in lanes 1, 2, and 3. The parent's non-hybrid genomic DNA is in lane 4, and two clones representing the two separated chromosomes
A and B (determined by haplotype analysis) are in lanes 5 and 6, respectively. Densitometric data (not shown) for the bands in Figure 4.3 indicated that chromosome A had 2 copies of $SMNI$ compared to 1 copy of $APC$. Chromosome B amplified $APC$, demonstrating the presence of chromosome 5, but did not amplify SMN. This shows that chromosome A contained two copies of $SMNI$, while chromosome B had no intact copies of $SMNI$, proving that parent 1 is a carrier of a null chromosome and a two-copy $SMNI$ chromosome. This dosage data was reproducible in different clones that contained the same chromosome, A or B. The data exactly correlates with the dosage data from the genomic DNA (lane 4), where in the presence of both chromosomes, we amplify two copies of $APC$, two copies of $SMNI$, and one copy of $SMN2$.

Haplotyping analysis was performed on each of the clones to differentiate those with either of the two copies of chromosome 5. Four markers determined whether a clone had either or both chromosomes designated A and B (Table 4.2). This analysis does not eliminate the possibility that some clones may have had more than one copy of either chromosome A or B, however having a dosage control on chromosome 5 allowed us to be sure that the amplification of $APC$ would remain proportional to that of SMN. This data was consistent with the PCR and dosage data.

Dual color FISH analysis probing chromosome 5 for the the $SMN$ region and for $APC$ was performed on clones from parent 1. These results indicated that a single copy of chromosome 5 was present in clones 1 and 14 (Figures 4.4 a and 4.4 b), which were found by haplotype analysis to contain chromosomes B and A, respectively. Clone 15, for which haplotype analysis indicated the presence of both chromosomes A and B, was
found to have two copies of chromosome 5 (Figure 4.4 c). This confirmation of the
haploid state of clones 1 and 14 supports the conclusion from the dosage data in Figure
4.3 that clone 14 contains a single copy of chromosome 5 with two copies of the \textit{SMNI}
gene, and that clone 1 has a single copy of chromosome 5 with zero copies of the \textit{SMNI}
gene. This data was reproducible in other clones shown to contain either chromosome A
or B. Hybrid clone 1 is deleted for exons 7 and 8 of both \textit{SMNI} and \textit{SMN2}.

\textbf{4.2.2 Parent 2: \textit{de novo} mutation}

The genomic DNA from parent 2 had previously been shown to have two copies
of \textit{SMNI} and two copies of \textit{SMN2}. PCR analysis showed that \textit{SMN} exons 6, 7, and 8 as
well as the \textit{APC} gene were all present in 27/38 clones (71.1\%). Both \textit{SMNI} and \textit{SMN2}
were present, based on results from D\textsubscript{r}a\textsubscript{1} digests of exon 7. Dosage testing of clones
containing each copy of chromosome 5, as determined by haplotyping, demonstrated that
one copy of \textit{SMNI} and one copy of \textit{SMN2} were present in each clone (data not shown).
This data indicates that parent 2 likely had an affected child as a result of a \textit{de novo}
deletion.

\textbf{4.3 Conclusions}

\textit{De novo} mutations have been identified by the loss of chromosome 5 markers
(Melki et al. 1994; Wirth et al. 1997\textsubscript{a}; Chen et al. 1999) when sufficient family members
could be recruited and informative markers were present. However, not all new
mutations can be identified by haplotype analysis since the rearrangement may not extend
into the region containing the marker, and because families may not be of sufficient size.
When a putative carrier is found to have normal dosage, it would be valuable to attempt to distinguish the two-copy SMNI chromosome from a de novo mutation. Nevertheless, what if this was not informative? How should a parent with an affected child, two copies of SMNI, and inconclusive marker results be counseled? This is a major diagnostic dilemma in the field of SMA diagnostics.

By analyzing individual chromosomes, the SMNI carrier status can be accurately determined and the parents can be appropriately counseled. As we have shown here, SMNI deletions inherited from parents with two-copy SMNI dosage should not be assumed to be de novo mutational events. In this study, parent 1 has been conclusively shown to carry a null chromosome and a two-copy SMNI chromosome. Mechanistically, there should not be any reason why two-copy SMNI chromosomes would not exist.

Conversion events in which SMNI sequence is replaced by SMN2 have been clearly demonstrated in SMA patients (Campbell et al. 1997; McAndrew et al. 1997; Taylor et al. 1998; Burghes 1997). Given the instability of this genomic region, it would also seem likely that SMN2 sequence could convert to SMNI. When copies of SMNI are lost as the result of unequal crossover, the reciprocal event would be expected to be a duplication on the other chromosome. If de novo mutations represent 2% of these cases (Wirth et al. 1997), then SMA carriers with two-copy/zero-copy chromosomes could explain the remaining 2% of parents with normal dosage who have affected children. The SMA recurrence risk for parent 1 is the same as for all carriers of recessive mutations (25%).
Parent 2 was shown to have \textit{SMN1} present in all hybrids. These results are most consistent with a \textit{de novo} deletion event, and therefore an extremely low recurrence risk. It should be noted that this parent could be a germline or somatic mosaic carrier.

The utilization of somatic cell hybrids allowed for the differentiation between the two transmission mechanisms of \textit{de novo} mutation and two-copy \textit{SMN1} carrier status, and therefore enabled us to provide accurate counseling information to these families. Presently, this technique is too labor-intensive for routine diagnostic testing. However, in the future this type of analysis may be important not only for parents, but for other family members as well. Lastly, to our knowledge, this is the most definitive proof of the existence of the two copy-\textit{SMN1} carrier.
Figure 4.1: Hypothetical representation of two possible chromosomal orientations of a parent with normal two-copy $SMN1$ dosage. \textit{a} A carrier with both copies of $SMN1$ on the same chromosome. \textit{b} A non-carrier with one copy of $SMN1$ on each chromosome. It should be noted that $SMN2$ copy number can vary from 0 to 3 in non-SMA controls.
Figure 4.2: Representation of the two possible genotypes for parent 1. One chromosome contains two copies of *SMN1* (*Chromosome A*). The other chromosome is entirely deleted for either *SMN1* or *SMN2* and is deleted for exons 7 and 8 of the other copy (*Chromosome B*). The orientation of *SMN1* and *SMN2* can vary on individual chromosomes (Burghes 1997).
Figure 4.3: Autoradiogram of the dosage assay for parent I. Lane 1: genomic 2-SMN1, 2-SMN2 control; lane 2: genomic 2-SMN1, 1-SMN2 control; lane 3: genomic 1-SMN1, 1-SMN2 control, lane 4: genomic DNA from parent I; lane 5: hybrid clone 14; lane 6: hybrid clone 1.
Figure 4.4: Dual color fluorescence in situ hybridization (FISH) with a PAC probe for SMN at 5q13 (125D9, red) and APC at 5q21 (554DG, green) allowed unequivocal identification of chromosome 5 copy number in mouse-human somatic cell hybrids from parent 1. a Hybrid clone 1 contained only one copy of human chromosome 5 (arrow) bearing one red and one green signal on its long arm. A small red signal noted on the short arm of chromosome 5 was visible in some metaphases and corresponds to pseudogenes and low copy repeats known to occur in the SMA region (Carpten et al. 1994; Thompsom et al. 1995; Chen et al. 1998). b Hybrid clone 14 also contained only one copy of human chromosome 5 (arrow) with the same signal pattern. c Clone 15 contained two copies of human chromosome 5 (arrows).
<table>
<thead>
<tr>
<th>Clone</th>
<th>SMN-exon 6</th>
<th>SMN-exon 7</th>
<th>SMN-exon 8</th>
<th>APC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>3</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>5</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>6</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>7</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>8</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>10</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>11</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>12</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>13</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>14</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>15</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>16</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>18</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>19</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>20</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>23</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>24</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>25</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>28</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>30</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 4.1: PCR amplification of parent 1 clones
<table>
<thead>
<tr>
<th>Clone</th>
<th>Chromosome</th>
<th>D5S630</th>
<th>D5S407</th>
<th>D5S644</th>
<th>D5S422</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>B</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>A</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>B</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>B</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>B</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>B</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>B</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>A</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>ND</td>
</tr>
<tr>
<td>15</td>
<td>A,B</td>
<td>1,2</td>
<td>1,2</td>
<td>1</td>
<td>1,2</td>
</tr>
<tr>
<td>18</td>
<td>B</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>20</td>
<td>A</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>23</td>
<td>A</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>24</td>
<td>A</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>28</td>
<td>B</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

NOTE - Alleles were assigned numerical values, 1 being the highest molecular weight allele, and 2 the lowest.

ND = not done

Table 4.2: Haplotype analysis of parent 1 clones
4.5 Materials and methods

4.5.1 Genomic DNA isolation and identification of SMNI copy number

DNA was extracted from peripheral venous blood from 100 parents of SMA patients by the salting out technique of Miller et al. (1988). DNA was analyzed for SMA carrier status by an SMN dosage assay (McAndrew et al. 1997), which compares the amplification ratio between SMNI and an exon of the cystic fibrosis transmembrane regulator (CFTR) gene in a competitive quantitative PCR reaction to determine the copy number of the SMNI gene.

4.5.2 Creation of mouse/human hybrid clones

Hybrid clones were created using the technique of Yan et al. (2000). Recipient murine cells derived from MSH2-deficient embryonic fibroblasts were immortalized with the EIA and ras oncogenes. These cells were maintained in Dulbecco’s modified Eagle’s Medium (DMEM) supplemented with 10% FCS and 10 μM 2-amino-6-mercaptopurine. An engineered mutation in HPRT allowed for HAT selection. 3 x 10^6 of a cell line called E2, selected for its growth and fusion abilities and its stable chromosomal content, were mixed with 12 x 10^6 fresh human lymphocytes. These cells were washed and spun two times in 0.25 M D-sorbitol, 0.1 mM calcium acetate, 0.5 mM magnesium acetate, and 0.1% bovine serum albumin (pH 7) and were then resususpended in 640 μL of the same solution. The cells were then electrically fused in a BTX ElectroCell Manipulator model ECM200 (BTX, San Diego) at 30V for 22 seconds, followed by three 300V pulses of 15 μsec each. Cells were then plated in 48-well plates in DMEM and 10% FCS.
Twenty-four hours later, the medium was exchanged for DMEM, 10% FCS, 0.5 mg/ml geneticin, and 1X HAT (Life Technologies, Gaithersberg, MD). HAT destroyed unfused murine embryonic fibroblasts and geneticin eliminated unfused human lymphocytes. Colonies were harvested for three to four weeks before genotyping.

4.5.3 PCR amplification of SMN exons and APC

All of the primer pairs were amplified using 30 ng each of primers in a 50 µL reaction containing 1 U of Taq polymerase (Applied Biosystems, Foster City, CA). 0.5 mM dNTP’s, 3 mM MgCl₂, and 5 µL of 10X PCR buffer. Samples were denatured for 5 minutes at 94°C, followed by 35 cycles of: [30 seconds at 94°C, 30 seconds at 55°C, and 1 minute at 72°C], terminating with a 7 minutes extension at 72°C in a GeneAmp PCR System 9600 (Applied Biosystems). SMN exon 6 was amplified using primers 541C618 (5'-CTCCCCATATGTCCAGATTCTTG-3') and EX63 (5'AAGAGTAATTTAAGCCTCAGACAG-3'). SMN exon 7 was amplified with primers R111 (Lefebvre et al. 1995) and X7-Dral (van der Steege et al. 1995). While exon 8 was amplified by primers 541C960 and 541C1120 (Lefebvre et al. 1995). Part of the APC gene was amplified using APCMDE-F and APCMDE-R described by Laken et al. (1997). PCR product was electrophoresed for 70 volt hours through a 2.5% agarose gel and then stained with ethidium bromide and visualized by ultraviolet light.
4.5.4 Modified SMNI dosage assay used on hybrid clones

The previously described dosage assay (McAndrew et al. 1997) relies on the comparison of SMNI, on chromosome 5, and CFTR which is on chromosome 7. A random assortment of human chromosomes become stably incorporated into the hybrid clones. We could not, therefore, depend on the presence of one copy each of chromosomes 5 and 7 for proper analysis of ratios between these two genes. To circumvent this problem, we used the APC gene, which is on chromosome 5q21, as the external control. 12.5 or 25 ng of DNA from each clone were analyzed using the same conditions as in the McAndrew et al. protocol with the following modifications: 3 ng/μL of SMN7 primer and 22.5 ng/μL of APCMDE (Laken et al. 1997) primer were amplified for 19 cycles.

4.5.5 Haplotype analysis

Determination of the genotypes of hybrid cell lines was done using fluorescently labeled microsatellite markers. Primers were obtained from Applied Biosystems and genetic locations from GeneMap99 (http://www.ncbi.nlm.nih.gov/genemap99/). Amplification of each microsatellite was done in 15 μL volumes using 10 ng of each respective genomic DNA, 8 pmol of each primer (5' primer fluorescently labeled), 100 mM of each dNTP, 0.6 U AmpliTaq Gold DNA Polymerase (Applied Biosystems), 10 mM pH 8.3 Tris-HCl, 50 mM KCl, and 2 mM MgCl₂. PCR products were loaded onto a 377XL sequencer (Applied Biosystems). Allele size and fluorescent intensity were
determined by Genescan and Genotyper software (Applied Biosystems). Cytogenetic locations for the markers are as follows: D5S630 (5p15.2), D5S407 (5q13), D5S644 (5q22), and D5S422 (5q34).

4.5.6 Probe labeling and fluorescence in situ hybridization (FISH)

The SMN probe (125D9) was labelled with Spectrum Orange™ and the APC probe (554DG) was labelled with SpectrumGreen™ using a nick translation kit from Vysis, Inc. (Downer’s Grove, IL) according to the manufacturer’s directions. The probes were hybridized to metaphase chromosomes overnight according to standard procedures as specified for locus-specific probes (Vysis, Inc.). Chromosomes were counterstained with DAPI H (Vysis, Inc.). Slides were examined on a Zeiss Axioskop epifluorescent microscope equipped with triple band pass filters and a filter wheel (Ludl). and images were obtained using MacProbe software and the PowerGene FISH System (Perceptive Scientific Systems, League City, TX).
CHAPTER 5

SMNI ALLELE-SPECIFIC INTRAGENIC MUTATION PANEL INCREASES THE SENSITIVITY OF CLINICAL TESTING FOR SMA

5.1 Introduction

About 1-4% of spinal muscular atrophy patients are compound heterozygotes (reviewed in Parsons et al. 1998a; Wirth et al. 1999; Wirth 2000). Although the intragenic mutations found in these patients occur throughout the gene, there is a cluster of mutations in exon 6 (Figure 5.1), which is a domain important for oligomerization (Lorson et al. 1998; Pellizzoni et al. 1999). Because testing for intragenic mutations is technically challenging, and because there are so many mutations to test for, routine testing is not offered in the diagnostic laboratory. The result is that 1-4% of SMA patients are not identified. In an attempt to increase diagnostic sensitivity, we describe an efficient and inexpensive fluorescent allele-specific PCR panel that allows for the rapid identification of the seven most common intragenic mutations found in compound heterozygotic SMA patients. This panel can easily be expanded or altered to meet the mutational profile of a specific population.
5.2 Results and discussion

Although many researchers have identified intragenic mutations in the SMN1 gene in compound heterozygotic patients (Wirth 2000), testing for these mutations is not performed on a routine diagnostic basis. The difficulties in screening for intragenic mutations include the fact that there is no highly prevalent mutation found in these patients, and that mutations are found throughout the gene. Adding to the complexity of finding intragenic mutations is the presence of SMN2, which makes for technical difficulties in identification of mutant sequence in SMN1 through the background of normal sequence from the copy gene. Also important is that newly identified mutations must be shown to be present on the SMN1 gene and not on SMN2. Several mutations have been reported in the literature more often than others and these mutations have not been found in the SMN2 gene (Lefebvre et al. 1995; Bussagliä et al. 1995; Parsons et al. 1996; Brahe et al. 1996; Bürglen et al. 1996c; McAndrew et al. 1997; Rochette et al. 1997; Clermont et al. 1997; Hahnen et al. 1997; Talbot et al. 1997b; Parsons et al. 1998a, 1998b; Wang et al. 1998; Wirth et al. 1999; Wirth 2000). This led us to develop a rapid and inexpensive assay that detects the more common intragenic mutations. The test entails the amplification of patient DNA using a panel of allele-specific primers that preferentially amplify the mutant allele. Seven mutations that have been observed more than once are detected (Figure 5.2) based on mismatches in the 3' end of the primer that are homologous to the mutant sequence (Table 5.1). Amplification of the APC gene is used as an internal control to confirm that sufficient template is present and amplification conditions are appropriate. The 800ins11 primer pair also acts as an internal control.
This primer set is not allele-specific. Rather it detects the mutation based on an 11 base pair size differential; a 153 bp band if a normal allele is present and a 164 bp band indicates a mutation. The 153 bp product acts as an amplification control, because this primer pair should amplify at least one normal SMN2 allele. If any of the seven mutations are identified, SMNI copy number is determined to confirm the compound heterozygosity of the sample.

This panel was used to screen 366 patients shown not to homozygously lack the SMNI gene based on results of the diagnostic deletion test (van der Steege et al. 1995). Because we estimate that only 44% of the patients tested by our laboratory actually have SMA, we expected to find between one and five intragenic mutations. One mutation was identified (T274I). This panel detects only the seven most common intragenic mutations, so it is possible that less common mutations were present in other patients. This multiplex panel was able to detect every patient previously shown to have one of these common mutations. For this reason, we believe that it is an effective and appropriate diagnostic assay to increase the sensitivity of SMA testing. Because of the low intragenic mutation rate, this application might be reserved for patients with a very high probability of having proximal spinal muscular atrophy.

5.3 Conclusions

To further increase the sensitivity of molecular testing, a cost-effective and rapid PCR-based intragenic mutation panel has been described. Our strategy was to screen every patient not homozygously deleted for SMNI, and perform an SMNI dosage test on
any patients with intragenic mutations to determine compound heterozygosity. Because of the relatively low rate of actual 5q-SMA patients screened in our diagnostic laboratory, it is recommended that intragenic mutation screening be reserved for patients for whom the diagnosis of 5q-linked spinal muscular atrophy is very confident. This panel is flexible such that additional mutations can be added if they are particularly common in a certain population.
5.4 Figures and tables

<table>
<thead>
<tr>
<th>1</th>
<th>2a</th>
<th>2b</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q15X</td>
<td>542delGT</td>
<td>618insT</td>
<td>591delA</td>
<td>800ins11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>124insT</td>
<td>472del5</td>
<td>430del14</td>
<td>542delGT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 5.1: Distribution of intragenic mutations in the SMNI gene. Nonsense and frameshift mutations are noted above the gene. Missense and splice site mutations are indicated below the gene.
Figure 5.2: Fluorescent allele-specific multiplex PCR panel detects seven intragenic mutations in the *SMNI* gene in patients but not in controls. The presence of a band indicates a mutation. *APC* and 800ins11 controls amplify in every sample and confirm the presence of sufficient template in the reaction.
<table>
<thead>
<tr>
<th>Mutation</th>
<th>Internal control</th>
<th>Forward primer</th>
<th>Reverse primer</th>
<th>Concentration (ng/µL)</th>
<th>Size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>APCMDE</td>
<td>APCMDE</td>
<td>5'-AAC AAG TTA CTA CAA A-3'</td>
<td>32</td>
<td>230</td>
<td></td>
</tr>
<tr>
<td>800ins11</td>
<td>5'-GAA AAG CAC CAA GTT TCA ACA CAT GAA AGA-3'</td>
<td>32</td>
<td>153/164 f</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S42AGT</td>
<td>S41C618</td>
<td>5'-AAC AAG TTA CTA CAA A-3'</td>
<td>45</td>
<td>199</td>
<td></td>
</tr>
<tr>
<td>A29G</td>
<td>S41C618</td>
<td>5'-AAC AAG TTA CTA CAA A-3'</td>
<td>45</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>S262I</td>
<td>S41C618</td>
<td>5'-AAC AAG TTA CTA CAA A-3'</td>
<td>16</td>
<td>194</td>
<td></td>
</tr>
<tr>
<td>Q15X</td>
<td>5'-GAA CAG GCA ATA CTC CTG-3'</td>
<td>16</td>
<td>123</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T274I</td>
<td>5'-GAA CAG GCA ATA CTC CTG-3'</td>
<td>16</td>
<td>107</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y272C</td>
<td>S41C618</td>
<td>5'-GAA CAG GCA ATA CTC CTG-3'</td>
<td>16</td>
<td>94</td>
<td></td>
</tr>
</tbody>
</table>

a 6-FAM labeled  
b TET-labeled  
c 153 bp for normal allele / 164 bp for 11 bp insertion

Table 5.1 Allele-specific primer sequence
5.5 Materials and methods

A multiplex panel of allele-specific primers (Table 5T) was used to preferentially amplify the mutant over the normal allele. For each primer pair, the non-allele-specific primer was fluorescently labeled. Part of the APC gene was used as an internal control, and was amplified in every sample regardless of the presence of mutations to confirm that sufficient template was present in each reaction. Every gel had the following controls: blank, normal, and each of the seven mutations being tested. Amplification was performed on a Perkin Elmer 9600 thermocycler (Applied Biosystems, Foster City, CA) starting with a 5 min denaturation at 94°C, followed by 30 cycles of: (1 min at 95°C, 2 min at 55°C, 3 min at 72°C), and concluding with a 7 min extension at 72°C. Product was electrophoresed through a 5% LongRanger (FMC Bioproducts, Rockland, ME) polyacrylamide gel on an ABI Prism 377 DNA Sequencer (Applied Biosystems) at 3kV/h for 1.5 hours. GeneScan TAMRA 500 (Applied Biosystems) was used as a size standard. Intensity of bands were analyzed using GeneScan software (Applied Biosystems). Although the A2G primer pair produces a 194 bp band, a smaller more visible product is used for diagnosis.
CHAPTER 6

ANALYSIS OF SMN2 COPY NUMBER IN PATIENTS AND NORMAL CONTROLS DEMONSTRATES A MODIFYING EFFECT ON PHENOTYPE

6.1 Introduction

SMN2 is a copy of the SMNI gene present in a duplicated region of chromosome 5q. Two exonic base changes differentiate SMN2 from SMNI (Lefebvre et al. 1995; Burgien et al. 1996a) and the variation in exon 7 of SMN2 is sufficient to greatly diminish the amount of full-length transcript and consequently the amount of full-length protein produced by the gene (Lorson et al. 1999; Monani et al. 1999a). This base change affects an exonic splice enhancer site (Lorson and Androphy 2000) resulting in the diminished ability to correctly splice out intron 7. Absence of SMN2 does not cause spinal muscular atrophy, however SMN2 copy number has been correlated with modification of the SMA phenotype in a transgenic mouse model (Schrank et al. 1997; Hsieh-Li et al. 2000; Monani et al. 2000a) and in humans, where increased copy number of SMN2 is associated with a more mild presentation of the disease (reviewed in Burghes 1997).
Several studies have shown that SMN2 gene copy number can modify the SMA phenotype (McAndrew et al. 1997; Hahnen et al. 1996; Velasco et al. 1996; Campbell et al. 1997; Wirth et al. 1999; Hsieh-Li et al. 2000; Monani et al. 2000a; Mailman et al. 2001b). We have performed a large-scale study comparing the SMN2 copy number between 52 type I and 90 type III patients. Our results conclusively demonstrate a greater number of copies of the SMN2 gene in type III versus type I patients.

Chimeric SMN alleles (exon 7$^{SMN2}$, exon 8$^{SMNI}$) were found by DiDonato et al. (1997a) to be more common in patients with mild SMA, suggesting that these alleles might be more mild than normal alleles (Burghes 1997; Campbell et al. 1997; DiDonato et al. 1997a; Wirth et al. 1995b). Other studies found all SMA types represented in the population of patients with chimeric SMN genes (Devriendt et al. 1996; Hahnen et al. 1996; van der Steege et al. 1996; Velasco et al. 1996; Talbot et al. 1997a; Nishio et al. 1999). An analysis of SMN2 copy number as it correlates to severity in SMA patients with chimeric SMN genes has not previously been reported. We hypothesized that SMN2 copy number might account for the variation in phenotype observed within the population of chimeric SMA patients and tested this by comparing SMN2 copy number between type I and type III patients with chimeric SMN genes.

6.2 Results and discussion

Previous studies state that 5% of unaffected people homozygously lack SMN2 (Lefebvre et al. 1995). In this study, 14 out of 97 normal controls with at least two copies of SMN1 had zero copies of SMN2, indicating a much higher percentage (14.4%). The
difference between studies may be attributed to variation between populations. The
majority of normal individuals have one or two copies of SMN2 (Figure 6.1). It was also
found that 4% of this sample of the normal population had three copies of the SMN1
gene.

In order to determine if mild SMA patients have more SMN2 copies than severe
patients, we have performed a large-scale comparison of SMN2 copy number in SMA
patients with disparate ages of onset, including 52 type I cases that were sent for
molecular diagnosis before they were six months old and 90 patients who were at least 18
years of age and diagnosed as having type III SMA. The results are shown in Table 6.1
and clearly demonstrate that there are more copies of SMN2 in mild SMA cases
compared to severe cases based on a chi-square test (P<0.0001). Interestingly, 100% of
type III patients had at least three copies of SMN2 and 20/90 had four copies. This is in
contrast to only 3.8% of type I patients with three copies, while none had more than three
copies. No type III patients were identified with one or two copies of SMN2. This data
indicates that the presence of three or more copies of SMN2 is clearly correlated with a
milder phenotype. Based on this information, it can be said that the presence of one or
two copies of SMN2 predicts a severe phenotype, and three or more copies of SMN2 is a
good prognostic indicator.

This study sought to test whether the variation in severity between patients with
chimeric SMN genes was correlated with SMN2 copy number by analyzing the number of
SMN2 genes in one type I and 10 type III patients who lacked exon 7 and retained exon 8
of SMN1, while possessing exon 7 of SMN2. We observed that every chimeric type III
patient had at least three copies of SMN2, while the only type I chimeric patient tested had two copies. There were two additional chimeric patients who were 8 and 8.5 months at molecular diagnosis (possible type I) that had two copies of SMN2. If chimeric alleles were milder than normal SMN2 alleles, we might have expected to find some type III patients with two copies of SMN2. SMN2 copy number modifies the phenotype in deleted patients (McAndrew et al. 1997; Hahnen et al. 1996; Velasco et al. 1996; Campbell et al. 1997; Wirth et al. 1999; Hsieh-Li et al. 2000; Monani et al. 2000b; Mailman et al. 2001b) as well as in compound heterozygotes (Parsons et al. 1998a). This data supports the hypothesis that the variation in phenotype between patients with chimeric SMN genes is also correlated with SMN2 copy number and not with the fact that chimeric alleles are inherently mild. Our population does not contain a large proportion of type I chimeric patients. It would be interesting to analyze SMN2 copy number in a population with a larger number of such patients.

6.3 Conclusions

SMN2 copy number has clearly been shown to be elevated in type III patients compared with type I patients, and can provide prognostic information. Out of 90 type III patients, none had fewer than three copies of SMN2 and only 3.9% of type I patients had more than two copies of SMN2. Although we do not have data concerning the intermediate type II patients, we can say that it is very unlikely that a patient with one or
two copies of \textit{SMN2} will ever walk. Conversely, we can say that a patient with three or four copies of \textit{SMN2} should have onset of symptoms after at least six months of age and should at least be able to sit up.

We did not have access to sufficient patient samples to conclusively state that chimeric \textit{SMN} genes act in a functionally equivalent manner as normal \textit{SMN2} genes. The data that we have presented, however, is highly suggestive that \textit{SMN2} gene copy number is the most important factor in determining phenotype, regardless of whether or not the gene is chimeric. Patients that have chimeric genes as a result of gene conversion are likely to have more than two copies of \textit{SMN2}, because one copy of \textit{SMN1} is converted to \textit{SMN2}, adding to the \textit{SMN2} copies already present. This mechanism provides a reasonable explanation as to why patients with chimeric \textit{SMN} genes are more likely to have mild cases of SMA without relying on a model in which chimeric and normal \textit{SMN2} genes are not functionally equivalent.
Figure 6.1: Variation of SMN2 copy number in normal individuals with at least two copies of SMN1.
<table>
<thead>
<tr>
<th>Copy SMN2</th>
<th>Type I</th>
<th>Type III</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Copy SMN2</td>
<td>7 (13.5%)</td>
<td>0 (0%)</td>
<td>7 (4.9%)</td>
</tr>
<tr>
<td>2 Copy SMN2</td>
<td>43 (82.7%)</td>
<td>0 (0%)</td>
<td>43 (30.3%)</td>
</tr>
<tr>
<td>3 Copy SMN2</td>
<td>2 (3.9%)</td>
<td>70 (77.8%)</td>
<td>72 (50.7%)</td>
</tr>
<tr>
<td>4 Copy SMN2</td>
<td>0 (0%)</td>
<td>20 (22.2%)</td>
<td>20 (14.1%)</td>
</tr>
<tr>
<td>Total</td>
<td>52</td>
<td>90</td>
<td>142</td>
</tr>
</tbody>
</table>

Table 6.1: SMN2 copy number in type I compared with type III patients
6.5 Materials and methods

6.5.1 Patient clinical classification

We do not receive extensive clinical information about the majority of patients sent for diagnostic testing and almost never receive follow-up data. For these reasons, we were not able to reliably place all of our patients into SMA types I, II, and III. Lacking motor milestone information and dates of the death of these patients, we instead categorized 52 patient samples sent for molecular diagnosis between the ages of zero and six months as type I. Our group of 90 type III patients were taken from DNA samples of individuals classified as type III before enrollment in a gabapentin therapeutic drug trial. We realize that by analyzing the most severe and the most mild patients, we have chosen our data from the extreme ends of the diagnostic spectrum. This was the only way for us to ensure the segregation of these populations, and we feel that these extremes legitimately represent acute and chronic SMA cases, without providing information about intermediate presentation.

6.5.2 SMN2 dosage assay

Carrier analysis for the SMN1 and SMN2 genes was based on the protocol of McAndrew et al. (1997), however the same primers were fluorescently labeled with 6-FAM. When testing for SMN2 copy number in patients deleted for SMN1, controls with zero copies of SMN1 and with 1, 2, 3, or 4 copies of SMN2 were run on each gel. Amplification was carried out in a Perkin Elmer 9600 thermocycler (Applied Biosystems, Foster City, CA) using the following conditions: 5 min denaturation at 94°C followed by
21 cycles of (1 min at 95°C, 2 min at 55°C, 3 min at 72°C), and concluding with a 7 min extension at 72°C. Product was electrophoresed through a 5% LongRanger polyacrylamide gel (FMC Bioproducts, Rockland, ME) on an ABI Prism 377 DNA Sequencer (Applied Biosystems) at 3kV/h for 1.5 hours. Dosage ratios were analyzed using GeneScan software (Applied Biosystems).
CHAPTER 7

GENETICS OF PROXIMAL SMA UNASSOCIATED WITH SMN1

7.1 Introduction

Although we now have a highly sensitive and specific direct molecular test for 5q-proximal spinal muscular atrophy, the clinical diagnosis is not always as straightforward. In our experience, we have found that fewer than 43% of the patients sent to our molecular diagnostic laboratory are not deleted for the SMN1 gene (chapter 3). These patients share some of the clinical features of types I-III SMA including hypotonia, muscle weakness, and loss of ambulation. It is quite possible that some of these patients with a presentation resembling SMA have mutations in a different gene. In the field of muscular dystrophy, it was found that mutations in many of the genes coding for glycoproteins associated with dystrophin (the gene responsible for Duchenne muscular dystrophy) cause independent autosomal recessive muscular dystrophies. These include several forms of limb-girdle muscular dystrophy as well as Fukuyama-type congenital muscular dystrophy and North African Duchenne-like muscular dystrophy (Matsumura et al. 1992; Matsumura and Campbell 1993; Passos-Bueno et al. 1993; Ervasti and Campbell 1993; Matsumura et al. 1993a, 1993b; Matsumura and Campbell 1994:
Campbell 1995a, 1995b; Jung et al. 1996a, 1996b; Salih et al. 1996; Beckmann et al. 1996; Straub and Campbell 1997; Holt and Campbell 1998; Lim and Campbell 1998; Durbeej et al. 2000; Piccolo et al. 2000). We hypothesized that genes coding for proteins that associate with SMN would make good candidates that might be mutated in SMA-like patients. Yeast two-hybrid studies have demonstrated the interaction of several proteins with SMN, implying the possibility that such proteins act in a common biochemical pathway. These include Sm N (Liu et al. 1997), SIP1, fibrillarin (Liu and Dreyfuss 1996), Gemin 3 (Charroux et al. 1999), Gemin 4 (Charroux et al. 2000), profilin II (Giesemann et al. 1999), FUSE (Williams et al. 2000), papillomavirus E2 (Strasswimmer et al. 1999), and Htra2-B1 (Hofmann et al. 2000). The Sm N and SIP1 genes are expressed in neurons (Liu et al. 1997; Liu and Dreyfuss 1996) and so were considered by us to be good candidates of genes that might be mutated in patients with symptoms of SMA, but without SMN mutations.

7.2 Results and discussion

We hypothesized that interrupting the important biochemical pathway of SMN at a different point might result in a phenotype similar to SMA. The mechanism explaining why proximal spinal muscular atrophy is a disease specific to α-motor neurons, despite the fact that the SMN protein is ubiquitously expressed is unexplained. Furthermore, it is proposed that SMN plays an important role in pre-mRNA splicing. Such a function is
essential in every cell of the body, and it seems logical that ablating an important housekeeping process by mutation of the widely expressed SMN protein would affect other tissues as well.

Approximately 92-98% of recessive proximal spinal muscular atrophy cases are homozygously deleted for the \textit{SMN1} gene (Bussaglia et al. 1995; Cobben et al. 1995; Hahnen et al. 1995; Lefebvre et al. 1995; Rodrigues et al. 1995; Devriendt et al. 1996; Hahnen et al. 1996; Matthijs et al. 1996; van der Steege et al. 1996; Velasco et al. 1996; Burghes 1997; Campbell et al. 1997; DiDonato et al. 1997; Talbot et al. 1997a; Wirth et al. 1999; Wirth 2000). A further 2-5% of SMA patients are compound heterozygotes in the \textit{SMN1} gene. The molecular etiology of the remaining patients is unknown. We hypothesized that mutations in genes whose products associate with SMN might equate to interrupting a common biochemical pathway that could manifest in a similar phenotype. The \textit{SIP1} and \textit{Sm N} genes were chosen as candidates for mutational analysis in putative SMA patients lacking SMN1 mutations, because they are expressed in high levels in the \(\alpha\)-motor neurons.

\textbf{7.2.1 The Sm core protein Sm N}

snRNP (small nuclear ribonucleoprotein) molecules are an important part of the pre-mRNA splicing machinery (DeRobertis 1983; Fisher et al. 1985; Mattaj 1988; Feeney et al. 1989; Neuman de Vegvar and Dahlberg 1990; Zieve and Sauterer 1990). snRNPs are composed of snRNA molecules and Sm core proteins. Sm proteins bind to SMN amino acids 129-135 (Buhler et al. 1999) and form the link between SMN and
snRNPs. While it used to be thought that SMN and Sm proteins interacted only in the cytoplasm (Liu et al. 1997; Fischer et al. 1997; Buhler et al. 1999), it is now believed that they also interact in the nucleus (Meister et al. 2000). Although there are many different Sm proteins, Sm N expression is specific to neurons, and was, therefore, an attractive candidate gene for mutational analysis. When this study was undertaken, only the cDNA sequence of the Sm N gene was known (Genbank accession # X15892). In order to screen the entire gene for mutations via heteroduplex analysis, the intronic sequence at the boundaries of the exons had to be determined so that primers could be used to amplify each exon in its entirety. To elucidate the intronic sequence, we used a procedure called vectorette PCR (Riley et al. 1990), which was initially used to find the sequence at the end of YAC clones. Basically, the procedure allows for the exponential amplification of sequence when only one primer is known. It involves the ligation of random fragments of digested genomic DNA to an engineered oligonucleotide. PCR amplification between a primer on the vectorette oligonucleotide and a primer in the 5' part of the cDNA sequence yields a product that represents the sequence between the cDNA-specific primer and the first enzyme digestion site in the sequence (Figure 7.1). If that restriction site is in an intron, then sequencing of the product will identify the exon/intron border and a certain amount of intronic sequence. If the first product is entirely exonic, then the product from a different restriction library might be large enough to cross the intronic border. Otherwise, a step-wise procedure of choosing another cDNA primer downstream should eventually determine the exon/intron border. Oftentimes, the first amplification
between cDNA primer and vectorette primer (VECT224) was nonspecific, and a nested PCR reaction was required. Once a distinct band was produced, it was excised from an agarose gel (Figure 7.2) and sequenced.

The cDNA sequence along with the intronic sequence flanking the exonic borders is presented in Figure 7.3 (note: several years after this study was concluded, a different cDNA sequence was published—Genbank # XM_049809. I am not able to determine which is the correct sequence at this time). Intronic primers flanking each exon were used to screen patient samples for mutations by heteroduplex analysis. Patients were screened that had two copies of the SMN1 gene to avoid the possibility of testing compound heterozygotes. There were 203 such patients tested for abnormally migrating bands after heteroduplex analysis. No mutations were identified. A polymorphism of 462T→A, which does not alter an alanine residue, was found in a single patient.

7.2.2 SMN-interacting protein 1 (SIP1)

SIP1 is a 279 amino acid, 32 kDa protein found in the cytoplasm and in nuclear gems (Liu et al. 1997). SIP1 colocalizes with SMN in gems in immunohistochemical studies, while their expression only partially overlaps in the cytoplasm (Jablonka et al. 2001). Their complete colocalization in the nucleus implies that they participate in a shared pathway. This shared function likely is in recycling factors of the pre-mRNA splicing machinery (Figure 2.5). SIP1 specifically is thought to assemble Sm core proteins and snRNA molecules to form snRNPs (Fischer et al. 1997). In comparison with
other tissues, SIP1 (like SMN) is expressed at very high levels in the spinal cord. For the above reasons, SIP1 was a good candidate for a gene that is mutated in patients with symptoms resembling spinal muscular atrophy.

We went about screening this gene using the same strategy as for the SmN gene (7.2.1). Only the cDNA sequence was available for SIP1 at the time, and so we used vectorette PCR (Figure 7.1) to determine exon/intron borders before screening patients by heteroduplex analysis (Figure 7.4). Before finishing this experiment, another group published a study showing that the SIP1 gene was not mutated in SMA-like patients (Helmken et al. 2000). We decided not to finish the project, however, we did screen 203 patients for 27% of the coding sequence of SIP1 (Figure 7.5) and found no mutations.

7.3 Conclusions

We conclude from these studies that mutations in the SmN and SIP1 genes do not play an important role in the development of the spinal muscular atrophy phenotype. These genes are still viable candidate modifiers; mutations in these genes in patients with SMNI deletions may alter the phenotype of disease. Atypical SMA patients with SMNI mutations, such as those with arthrogryposis (Bürglen et al. 1996a; Bingham et al. 1997), might have mutations in such a modifying gene. Htra2-BI was thought to be a candidate gene that might be mutated in SMA patients without SMNI mutations (Hofmann et al. 2000), however this was later shown not to be the case (Helmken and Wirth 2000). There are several other proteins that interact with SMN, and all are candidates. It is
possible that the fewer than 5% of SMA cases without SMN1 mutations may have mutations in the 5' or 3'-untranslated region of SMN1 that are not typically screened.
7.4 Figures and tables

Figure 7.1: Vectorette PCR procedure
Three different digests (DraI, HpaI, RsaI) followed by first round of amplification. Second round of amplification. After nested amplification, if a discrete band is identified, the band was cut from an agarose gel, purified, and then sequenced.
Figure 7.3: Exon/intron borders of the *Sm* gene. cDNA sequence is in lowercase text and intronic sequence is in bold capital text.
Figure 7.4: Strategy for screening patients for mutations in the Sm N and SIP1 genes. 203 patients with two copies of SMNI were screened for abnormal heteroduplex results. No mutations were identified in either gene.
Figure 7.5: Sequence of *SIPI* cDNA (Genbank accession# AF027150) with partial intron border sequence. cDNA sequence in lowercase and intronic sequence in capital text.
7.5 Materials and methods

7.5.1 Creation of vectorette genomic libraries

Genomic DNA from a non-diseased individual was digested with one of three blunt-end cutting restriction enzymes to create three different genomic restriction fragment libraries. Digestion reactions were carried out overnight at 37 °C in 20 μL reaction volumes using 4 μg of DNA and 40 U of Dral, HpaI, or RsaI. The enzymes were then denatured for ten minutes at 55 °C. In a volume of 30 μL, 2 μg of digested genomic DNA was ligated to 100 ng of vectorette DNA (Riley et al. 1990) overnight at 16 °C in the presence of 3 U of T4 DNA ligase (Life Technologies).

7.5.2 PCR amplification

One μg of DNA was amplified in the presence of 150 ng of the VECT224 oligonucleotide primer (Riley et al. 1990) and of a cDNA-specific primer, along with 2.5 U of Taq polymerase (Life Technologies) in a final volume of 100 mmol/L Tris (pH 8.8), 16.6 mmol/L ammonium sulfate, 6.7 μmol/L EDTA, and 10 mmol/L 2-mercaptoethanol. The primer sequences were derived directly from the cDNA sequence (Genbank accession # AF027150). Amplification proceeded by a denaturation step of five minutes at 95 °C, followed by 30 amplification cycles of: (1 minute at 95 °C, 2 minutes at 55 °C, 3 minutes at 72 °C), and finished with an extension of 8 minutes at 72 °C (Ericomp thermocycler) (Monani and Burghes 1996; U. Monani personal communication). In order to achieve a specific product, it was often necessary to do a hemi-nested PCR.
reaction using another cDNA primer in conjunction with the VECT224 primer. PCR product was diluted 1:5000 and was amplified again at 35 cycles (Moynihan et al. 1996).

7.5.3 Heteroduplex analysis

Fifty μL of PCR product was incubated at 95 °C for five minutes to denature the DNA strands. It was then allowed to slowly ramp down to 37 °C for 30 minutes. Twenty μL of each mixture was then combined with 4 μL of 6X Triple Dye Loading Buffer (FMC BioProducts, Rockland, ME) and was electrophoresed through a 50 cM Hydrolink-MDE gel (AT Biochem) containing 150 g/L urea in TBE buffer (80.3 mmol/L tris base, 48.5 mmol/L boric acid, 1.53 mmol/L Na₂EDTA) for 15 hours at 1000 V. The gel was then stained in a solution of TBE containing 1 mg/L of ethidium bromide, and the DNA was photographed under UV light.

7.5.4 Sequence analysis

Sequencing was performed on samples that produced heteroduplexes on the MDE gel. The Wizard PCR Preps DNA Purification System (Promega, Madison, WI) was used to isolate longer ds-DNA strands away from the degraded DNA, which decreases the quality of the sequence. Double-stranded sequencing of the purified PCR product was performed using the ds-DNA Cycle Sequencing System (BRL, Gaithersburg, MD). Primers were end-labeled in a volume of 5 μL: 1 μL 5X kinase buffer (BRL), 2.5 μL (3 ng) primer, 1 μL [γ-^32P]-ATP, 0.5μL T4 polynucleotide kinase (BRL), and were incubated at 37 °C for 20 minutes, 55 °C for 10 minutes, and then placed on ice. The sequencing reaction proceeded in a volume of 37 μL: 5 mL end-labelled primer, 26 μL
ddH₂O, 4.5 μL 10X Taq sequencing buffer (BRL), 1 μL template DNA (from PCR amplification), 0.5 μL (5 U/μL Taq DNA polymerase. The cycle sequencing program consisted of: 6 minutes at 95 °C, 20 cycles: (30 seconds at 95 °C, 30 seconds at 55 °C, 30 seconds at 72 °C), followed by 10 cycles: (30 seconds at 95 °C, 1 minute at 68 °C), and finished with 8 minutes at 72 °C. The product from the sequencing reaction was analyzed using a 50 cm 5% denaturing polyacrylamide gel. Tumor DNA was compared to that of a normal control patient. Mutations were confirmed by repetition and by separately labeling both forward and reverse primers.
Spinal muscular atrophy is the leading inherited cause of infant death and is the second most common recessively inherited lethal disease after cystic fibrosis. This devastating neuromuscular disease cannot currently be cured. The research presented in this dissertation increases the sensitivity of diagnostic testing and allows for more accurate genetic counseling of families segregating the disease.

SMA is most often caused by homozygous mutation of the SMN1 gene. The symptoms of this disease, including hypotonia, muscle weakness, and loss of ambulation, are not specific to SMA. It is extremely important that molecular testing be performed on probands if subsequent family counseling is to be done. In the absence of a confirmed mutation, carrier studies are not informative. The importance of molecular testing for spinal muscular atrophy is exemplified by the fact that the SMN1 gene is intact in 57% of patients with symptoms resembling SMA. Since a carrier test for SMA became available in 1997, there has been a great interest in the testing of family members to determine their risk of having an affected child. This testing has uncovered the interesting finding
that 5% of parents possess two intact copies of \textit{SMN1} and require linkage analysis or monosomal hybrid studies for accurate risk assessment. With such a high rate of SMA-like patients testing negative for the common deletion, we recommend that carrier testing be reserved for at-risk relatives of patients who are positive for the \textit{SMN1} deletion. Our data indicate that if a molecular diagnosis of SMA is not made in a family, then carrier testing will be performed on the wrong gene or on a disease with no genetic etiology in more than half of the cases.

When sufficient family members can be recruited and when informative markers are present, \textit{de novo} mutations can be identified by haplotype analysis (Melki et al. 1994; Wirth et al. 1997a; Chen et al. 1999). This may not always be possible, however, if the rearrangement does not extend into the region containing the marker, or if families are not of sufficient size. What should happen if marker analysis is not informative? How should a parent with an affected child, two copies of \textit{SMN1}, and inconclusive marker results be counseled? This is a major dilemma in the field of SMA diagnostics. We have shown that through the analysis of individual chromosomes, \textit{SMN1} carrier status can be accurately determined and parents can be appropriately counseled. \textit{SMN1} deletions inherited from parents with two copies of \textit{SMN1} cannot be assumed to be \textit{de novo} mutation events. In this study, one parent was shown to carry a null chromosome and a two-copy \textit{SMN1} chromosome. Mechanistically, there should not be any reason why two-copy \textit{SMN1} chromosomes would not exist. Conversion events in which \textit{SMN1} sequence is replaced by \textit{SMN2} have been clearly demonstrated in SMA patients (Campbell et al. 1997; McAndrew et al. 1997; Taylor et al. 1998; Burghes 1997). Given the instability of
this genomic region, it would also seem likely that SMN2 sequence could convert to
SMNI. When copies of SMNI are lost as a result of unequal crossover, the reciprocal
event would be a duplication on the other chromosome. If de novo mutations represent
2% of these cases (Wirth et al. 1997), then SMA carrier parents with two-copy/zero-copy
chromosomes could be equally common. Such a parent has a recurrence risk that is the
same 25% as for all carriers of recessive mutations. A second parent had SMNI present
in all hybrids, indicating that a de novo deletion event was most likely, resulting in an
extremely low recurrence risk. Parents and other relatives of affected children who test
negative for carrier status may undergo somatic hybrid testing in the future to distinguish
these two possibilities.

Intragenic mutations are known to occur in 1-4% of spinal muscular atrophy
patients. These mutations are technically difficult to screen for in the diagnostic setting.
A cost-effective and rapid PCR-based intragenic mutation panel that identifies the seven
most common intragenic mutations found in compound heterozygotes. This assay could
be used to screen SMA-like patients who do not homozygously lack SMNI, as
determined by deletion analysis. Patients who test positive for an intragenic mutation
would then be tested by a SMNI dosage assay to determine if they are compound
heterozygotes. Because of the relatively low rate of actual 5q-SMA patients screened in
our diagnostic laboratory, we recommended that intragenic mutation screening be
reserved for patients for whom the diagnosis of 5q-linked spinal muscular atrophy is very confident. This panel will increase the diagnostic sensitivity of SMA testing by identification of affected patients who previously would have been missed.

This study has clearly demonstrated that SMN2 copy number is elevated in type III patients compared with type I patients and is the only known modifier of the SMA phenotype. Although it was known that more copies of SMN2 yields a less severe phenotype, a large-scale analysis of patients had not been previously carried out. The statistical significance of the difference in SMN2 copy number between patients with different types of SMA also had not previously been described. We compared SMN2 copy number in types I and III patients to determine how consistent the copy number was in each type. No type III patients had fewer than three copies of SMN2, and only 3.9% of type I patients had more than two copies of SMN2. This information yields important prognostic implications. Although we do not have data concerning the intermediate type II patients, we can state that it is very unlikely that a patient with one or two copies of SMN2 will be ambulant. Conversely, a patient with three or four copies of SMN2 should have onset of symptoms after at least six months of age and should at least be able to sit up.

We also examined SMN2 copy number in patients with chimeric SMN genes. These genes, which have exon 7 from SMN2 and exon 8 from SMN1, have been suggested to yield a milder phenotype. The data that we have presented is highly suggestive that the number of SMN2 gene copies is the most important factor in determining phenotype, regardless of whether or not the gene is chimeric. Patients with
chimeric genes resulting from gene conversion are more likely to have greater than two copies of \( SMN2 \), because one copy of \( SMNI \) is converted to \( SMN2 \), adding to the \( SMN2 \) copies already present. This mechanism provides a reasonable explanation as to why patients with chimeric \( SMN \) genes are more likely to have mild cases of SMA without relying on a model in which chimeric and normal \( SMN2 \) genes are not functionally equivalent.

Up to 5% of SMA cases cannot be explained by mutations in the \( SMNI \) gene. A possible explanation for this finding is that mutations in another gene cause a similar phenotype. We believed that a logical candidate gene would be one whose product interacts with SMN. We tested the \( S I P I \) and \( SmN \) genes for mutations in SMA-like patients, because their proteins interact with SMN and they are expressed at high levels in neurons. We conclude from these studies that mutations in the \( SmN \) and \( S I P I \) genes do not play an important role in the development of the spinal muscular atrophy phenotype. There are several other proteins that interact with SMN, and all are candidates. It is possible that SMA cases without \( SMNI \) mutations may be mutated in the 5' or 3'-untranslated region or in the introns of \( SMNI \) that are not typically screened.

This research has provided the most definitive evidence of the presence of carriers with two copies of \( SMNI \), described a valuable new tool for the diagnostic screening of compound heterozygotic patients, reported statistically conclusive data for the role of \( SMN2 \) in modification of the SMA phenotype, and ruled out the likelihood of mutations in two candidate genes playing a role in an SMA-like disease.
LIST OF REFERENCES


Dubowitz V (1964) Infantile muscular atrophy. A prospective study with particular reference to a slowly progressive variety. Brain 87:707-718


Gersappe A, Pintel DJ (1999) CA- and purine-rich elements form a novel bipartite exon enhancer which governs inclusion of the minute virus of mice NS2-specific exon in both singly and doubly spliced mRNAs. Mol Cell Biol 364-375


Hoffmann J (1892) Ueber Familiare progressive spinale muskelatrophie. Arch Psych (Berlin) 24:644-646


Lamond AI, Carmo-Fonseca (1993) The coiled body. TIB 3:198-204


Mailman MD, Heinz JW, Papp AC, Snyder PJ, Sedra MS, Wirth B. Burghes AHM. Prior TW (2001b) Molecular analysis of spinal muscular atrophy and modification of the phenotype by SMN2. In press


119


Monani UR, Lorson CL, Parsons DW, Prior TW, Androphy EJ, Burghes AHM. McPherson JD (1999a) A single nucleotide difference that alters splicing patterns distinguishes the SMA gene SMN\textsuperscript{I} from the copy gene SMN\textsuperscript{T}. Hum Mol Genet 8:1177-1183

Monani UR, McPherson JD, Burghes AHM (1999b) Promoter analysis of the human centromeric and telomeric survival motor neuron genes (SMN\textsuperscript{C} and SMN\textsuperscript{T}). Biochem Biophys Acta 1445:330-336


Müller B, Melki J, Burlet P, Clerget-Darpoux F (1992) Proximal spinal muscular atrophy types II and III in the same sibship are not caused by different alleles at the SMA locus on 5q. Am J Hum Genet 50:892-895


Noble SM, Guthrie C (1996b) Transcriptional pulse-chase analysis reveals a role for a novel snRNP-associated protein in the manufacture of spliceosomal snRNP's. EMBO J 15:4368-4379


Parsons DW, McAndrew PE, Monani UR, Mendell JR, Burghes AHM, Prior TW (1996) 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 11:1727-1732


Parsons DW, McAndrew PE, Allison PS, Parker WD Jr, Burghes AHM, Prior TW (1998b) Diagnosis of spinal muscular atrophy in an SMN non-deletion patient using a quantitative PCR screen and mutation analysis. J Med Genet 35:674-676


Pellizzoni L, Charroux B, Dreyfuss G (1999) SMN mutants of spinal muscular atrophy patients are defective in binding to snRNP proteins. Proc Natl Acad Sci, USA 96:11167-11172


127


