Genetic Investigation of both Complex and Mendelian Disorders: Neural Tube Defects and Native American Myopathy

Demetra Serena Stamm

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Curriculum of Genetics.

Chapel Hill
2007

Approved by:

Kari North

Cynthia Powell

Stephen Crews

Karen Mohlke

John Gilbert

ABSTRACT

Demetra Serena Stamm

Genetic Investigation of both Complex and Mendelian Disorders: Neural Tube Defects and Native American Myopathy

(Under the direction of Dr. Marcy Speer)

This dissertation seeks to map genes for two disorders: one complex and one Mendelian. The first project investigated the genetic basis of neural tube defects (NTDs) in one large multiplex family (8776). NTDs are among the most common debilitating birth defects occurring in ~1/1000 live births in the U.S. and are considered complex disorders with both genetic and environmental factors implicated. Linkage analysis using a genomewide single-nucleotide polymorphism (SNP) screen in family 8776 identified maximum LOD* scores of ~3.0 mapping to 2q33.1- q35 and 7p21.1-pter. Ascertainment of another branch of this family further supported linkage at these two loci providing a ~3.3 LOD* score, and decreased the linkage interval for 7p by 7.8 Mb. Further haplotype analyses narrowed the minimum candidate intervals, with the new regions mapping to 2q33.1-q35 (20.3 Mb) and to 7p21.1-21.3 (8.3 Mb). Within these regions, sequencing of exons in eighteen candidate genes did not identify a causative mutation and chromosomal abnormalities were undetected. Identification of an NTD

gene in this family may aid in a directed search for a disease gene for more typical NTD families.

The second project studied Native American myopathy (NAM), a rare congenital disorder observed in the Lumbee population. In this dissertation, shared clinical features are described among 14 individuals of Lumbee Indian descent with Native American myopathy. Notably, in these 14 subjects, we observed a 36% mortality rate and 29% had MH episodes. The Lumbee population is relatively isolated with evidence of consanguinity suggesting that homozygosity mapping methods may be employed for identifying the NAM disease locus. To this end, we conducted a genomewide SNP screen which demonstrated five regions of shared homozygosity (≥ 3 contiguous SNPs) in individuals with NAM. Genotyping microsatellites in these five regions lead to only one 5.6 Mb region of shared homozygosity amongst the NAM patients mapping to 12q13.13 to 12q14.1. Sequencing of four genes within this region did not identify a functional mutation.

Ultimately, the goal of these projects was to identify disease loci for NTD family 8776, and for NAM, to facilitate a targeted search of candidate genes for these debilitating congenital disorders.

DEDICATION

To Marcy Speer for her life contributions dedicated to furthering the understanding of congenital birth defects. To all the individuals whose lives have been affected by a congenital birth defect, and who through their altruistic contributions to these studies have helped to further the knowledge base of these disabling genetic diseases.

ACKNOWLEDGEMENTS

I would first like to acknowledge Marcy Speer, my advisor, for her overall generosity towards me as a student. She has provided me with many opportunities to pursue several avenues of research in order to broaden my knowledge base of human genetics. This built the foundation for a unique comprehensive training in the clinical, molecular and statistical facets of human genetics. I would also like to thank my committee members for their continuous support and insightful feedback on my projects. Cynthia Powell's guidance was essential for the Native American myopathy project -- I learned a tremendous amount from her in both clinical and academic realms and I always enjoyed discussing this project with her. I am very grateful for how dedicated and considerate she was towards me and our work together. Kari North, as my mentor at UNC-CH, was incredibly helpful and encouraging throughout my graduate studies; I thank her for her on-going academic and morale support and always making me feel capable and confident! I am very grateful to Karen Mohlke for all of her wonderful feedback during committee meeting and for my dissertation. Thank you to John Gilbert for providing feedback on molecular aspects of the projects. I am grateful to Eugene Orringer, Liz Garman, Bob Duronio, Cara Marlow and Sausyty Hermreck for their support during my graduate work. I appreciate Jessica Connelly and Kristie Deak for their time spent sharing their knowledge of human genetics with me. I would like to

thank Ian Sanderson, one of my most dear friends, who inspired me to be a physician scientist in the first place, and who continues to be a source of strength and inspiration for me. Renee Saville, my best friend, who I love like a sister, has seen me through many of life's hurdles and triumphs-- I thank her for always being there for me offering her gentle wisdom and infectious laughter. My family has been a constant source of support and love throughout my life, and I would like to thank them for always encouraging me to be passionate about my life choices and to remember that compassionate interaction is the most important ambition. I will miss Chay Kroner and Bill Hulme's late night discussions on philosophy, science and music, keeping my spirit alive during those long hours in the lab and graduate school in general. Most importantly, I would like to thank my husband, John Wallis, for his unwavering patience, unconditional love and kindness.

TABLE OF CONTENTS

Page

LIST OF TABLES xiii

LIST OF FIGURES xiv

LIST OF ABBREVIATIONS xv

CHAPTER 1 1

Introduction 1

1.1 Overview of dissertation projects 1

1.2 Neural Tube Defects are a Complex Group of Genetic Disorders 1

1.2.1 Neural Tube Defects: Clinical Impact 2

1.2.2 Epidemiology of Neural Tube Defects 4

1.2.3 Syndromic versus Nonsyndromic NTDs 6

1.2.4 Chromosomal Abnormalities in NTDs 7

1.2.5 Environmental Components to NTD Predisposition 8

1.2.6 Evidence for a Genetic Contribution to NTDs 10

1.2.7 Challenges to Identifying a Genetic Component 13

1.2.8 Why Investigate a Genetic Component for NTDs? 14

1.2.9 Normal Human Embryology of Neural Tube Development 15

1.2.10 Role of Folate in Development and Sufficient Maternal Folic Acid Consumption 18

1.2.11 Epigentic Effects and Parental Effects on Neural Tube Defects 22

1.2.12 Theories of Neural Tube Closure Sites 23

1.2.13 Key Molecular Process Underlying Neural Tube Closure 26

1.2.14 Translation of Mouse Models to Human NTDs 29

1.3 Approaches for Identifying Genes Related to Human Neural Tube Defects 31

1.3.1 Parametric Linkage Analysis 31

1.3.2 Nonparametric Linkage Analysis 36

1.3.3 Linkage Screen in 44 Multiplex NTD Families..... 39

1.3.4 Approaches to Candidate Gene Selection in Human NTDs 40

1.3.5 The Future for Gene Mapping Studies in Complex Disease 42

1.4 Native American Myopathy: Defining the Phenotype and Genetic Analysis.... 43

1.4.1 The Lumbee Population 44

1.4.2 Clinical Description of Native American Myopathy 47

1.4.3 Malignant Hyperthermia Susceptibility in the Lumbee Population 48

1.4.4 Clinical Complexity of NAM 49

1.4.5 Clinical Manifestations of Malignant Hyperthermia 50

1.4.6 Malignant Hyperthermia Susceptibility Loci and Associated Genes 51

1.4.7 *RYR1* Mutations and Associated MH and/or Congenital Myopathies..... 53

1.4.8 Homozygosity Mapping Approach to Identify the NAM Disease Locus... 55

1.4.9 Potential Pitfalls in Homozygosity Mapping 56

CHAPTER 2 58

SNP screen in a large multiplex neural tube defect family refines linkage to loci at 7p21-pter and 2q33.1-35 58

2.1 Abstract 59

2.2 Introduction 59

2.2.1 Neural Tube Defects: Clinical Impact 59

2.2.2 Evidence for genetic and environmental components 60

2.2.3 Initial genomic screen results..... 61

2.3 Materials and Methods 62

2.3.1 Clinical data collection and power analysis 62

2.3.2 Genotyping methods 63

2.3.3 Whole genome linkage analyses 64

2.3.4 Sequencing of Candidate Genes 65

2.4 Results 66

2.4.1 Linkage Results for Family 8776 66

2.4.2 Haplotyping 67

2.4.3 Candidate genes 68

2.5 Discussion 68

2.6 Acknowledgments 72

CHAPTER 3 76

Refinement of 2q and 7p loci and genomic convergence for candidate gene selection in a large multiplex neural tube defect (NTD) family 76

3.1 Abstract 76

3.2 Introduction 77

3.2.1 Human Neural Tube Defects 77

3.2.2 Previous NTD Gene Mapping Studies 78

3.3 Materials and Methods 79

3.3.1 Clinical Data Collection and Biological Samples 79

3.3.2 Genotyping Methods 80

3.3.3 Error Checking 81

3.3.4 Whole Genome Linkage Analyses 81

3.3.5 Marker Maps 81

3.3.6 Linkage Analysis 82

3.3.7 Fine mapping 83

3.3.8 Haplotype analysis 84

3.3.9 Chromosomal Aberration Detection Methods 84

3.3.10 Candidate Gene Prioritization 85

3.3.11 Expression Analysis 86

3.3.12 Sequencing of Candidate Genes 87

3.4 Results 88

3.4.1 Linkage Results for Family 8776 88

3.4.2 Haplotyping 89

3.4.3 Candidate Genes 90

3.4.4 Chromosomal Abnormalities 92

3.5 Discussion 92

3.6 Acknowledgments 97

CHAPTER 4 104

Native American myopathy: Congenital myopathy with cleft palate, skeletal anomalies, and susceptibility to malignant hyperthermia 104

4.1 Abstract 105

4.2 Introduction 105

4.2.1 Lumbee population 105

4.2.2 Native American myopathy: Clinical Description 106

4.2.3 Genetics of NAM 107

4.3 Materials and Methods 108

4.3.1 Clinical Ascertainment 108

4.4 Results 108

4.4.1 Natural history of Disease 108

4.4.2 Mode of Inheritance and NAM pedigrees 108

4.4.3 Cytogenetic Studies 109

4.4.4 Musculoskeletal and malignant hyperthermia 110

4.4.5 Dysmorphic features 111

4.4.6 Newborn and Toddler Period 111

4.4.7 Childhood 112

4.4.8 Adulthood 113

4.5 Discussion 114

4.6 Summary 118

4.7 Acknowledgements 119

CHAPTER 5 126

Genomewide SNP chip homozygosity mapping defines critical region in singleton Native American myopathy families 126

5.1 Abstract 126

5.2 Introduction 127

5.3 Materials and Methods 129

5.3.1 Clinical Description 129

5.3.2 Genotyping Methods 130

10K SNP Chip 130

5.3.3 Homozygosity Mapping and Copy Number Variant Analysis 130

5.3.4 Fine Mapping with Microsatellites 131

5.3.5 Sequencing of Candidate Genes 132

5.4 Results 133

5.4.1 Genetic Mapping 133

5.4.2 Candidate Genes 134

5.5 Discussion 135

5.6 Acknowledgements 139

CHAPTER 6 143

Discussion of Future Directions and Potential Limitations of Projects 143

6.1.1 Genetic Susceptibility to Neural Tube Defects in Family 8776 143

6.1.2 Environmental Exposures and Other Potential Risk Factors for NTD 148

6.1.3 Commentary on Future Genetic Studies of Human NTDs 151

6.2 Native American Myopathy Study 153

6.2.1 NAM: Clinical Evaluation and Natural History of Disease 154

6.2.2 Homozygosity Mapping in NAM 155

6.2.3 Identification of NAM gene 157

REFERENCES 160

LIST OF TABLES

Table 1.1 Responses to maternal nutrition in mouse models for NTDs. 31

Table 1.2 Genomewide Linkage Significance 35

Table 1.3 Differential diagnosis for NAM 50

Table 1.4 Malignant hyperthermia susceptibility loci 53

Table 2.1 Multipoint LOD scores (>2.5) for regions of interest on chromosomes 2 and 7. 73

Table 3.1 Genes Sequenced in NTD Family 8776 98

Table 3.2 Novel Variants Identified in Sequencing Screen of Family 8776 99

Table 3.3 Variants Sequenced in the Entire Family 8776 100

Table 4.1 Clinical description of patients with NAM 120

Table 5.1 Shared homozygous segments observed in affected individuals identified by SNP chip..... 140

Table 5.2 Information on microsatellites used to evaluate regions of shared SNP homozygosity. 141

Table 6.1 Cytogenetic abnormalities in chromosomal 2q and 7p regions associated with NTDs. 147

Table 6.2 Folic acid questionnaire results for family 8776 150

LIST OF FIGURES

Figure 1.1 Human neural tube closure and Carnegie Stages 10-13. 1

Figure 1.2 The folate and methionine metabolic cycles. 1

Figure 1.3 Regions of human and mouse neural tube closure sites. 1

Figure 1.4 PCP-dependent convergent extension movements..... 1

Figure 1.5 Logarithm of the Odds (LOD) score equation. 33

Figure 1.6 Genomic Convergence 42

Figure 1.7 Dysmorphic facies and skeletal abnormalities of NAM..... 1

Figure 2.1 Pedigree diagram of family 8776 74

Figure 2.2 Multipoint LOD score curves for regions of interest in family 8776..... 75

Figure 3.1 Pedigree diagram of NTD family 8776 with newly ascertained branch. 1

Figure 3.2 Multipoint LOD plots for chromosomes 2 and 7 in extended family 8776. . 102

Figure 3.3 The 2q and 7p disease-associated haplotypes for family 8776. 103

Figure 4.1 Observed clinical features of patients with NAM 121

Figure 5.1 The 12q13.13 to 12q14.1 haplotype for the five NAM families. 1

Figure 6.1 Pedigree plot of family 8776. 1

LIST OF ABBREVIATIONS

ABI2	abl interactor 2
AFP	alpha-fetoprotein
BMPR2	bone morphogenetic protein receptor, type II
CACNA1S	calcium channel, voltage-dependent, L type, alpha 1S subunit
CASP8	caspase 8
CASP10	caspase 10
CS	Carnegie stage
CGH	comparative genomic hybridization
CNV	copy number variants
CYP20A1	cytochrome P450, family 20, subfamily A, polypeptide 1
DGKB	diacylglycerol kinase, beta
DNA	deoxyribonucleotide
EEF1B2	eukaryotic translation elongation factor 1 beta 2
ETV1	ets variant gene 1
FZD5	frizzled homolog 5
FZD7	frizzled homolog 7
HSPD1	heat shock 60kDa protein 1
IBD	identical-by-descent
IBS	identical-by-state
ICA1	islet cell autoantigen 1

ICA1L	islet cell autoantigen 1,69kDa-like
IGFBP2	insulin-like growth factor binding protein 2
ITGA7	integrin α-7
IVCT	*in vitro* contracture test
LOD	logarithm of the odds
LOD*	nonparametric LOD score
MAP2	microtubule-associated protein 2
MEOX2	mesenchyme homeobox 2
MH	malignant hyperthermia
MHS	malignant hyperthermia susceptibility
MIM	Mendelian Inheritance of Man
MLC1sa	myosin alkali light chain 1 slow a
MoM	multiple of the median
MSAFP	maternal serum alpha-fetoprotein
MTHFR	5, 10-methylenetetrahydrofolate reductase
NAM	Native American myopathy
NRP2	neuropilin 2
NTD	neural tube defects
PCR	polymerase chain reaction
PDE1B	phosphodiesterase 1B, calmodulin-dependent
PECR	peroxisomal trans-2-enoyl-CoA reductase
PIP5K2C	phosphatidylinositol-4-phosphate 5-kinase, type II, gamma
PTPN11	protein tyrosine phosphatase, non-receptor type 11

RNA	ribonucleic acid
RYR1	ryanodine receptor 1
SAGE	serial analysis of gene expression
SCN4A	sodium channel, voltage-gated, type IV, alpha subunit
SNP	single nucleotide polymorphism
STK17B	serine/threonine kinase 17b
WNT6	wingless-type MMTV integration site family, member 6
WNT10A	wingless-type MMTV integration site family, member 10A
XRCC5	DNA repair protein XRCC5

CHAPTER 1

Introduction

1.1 Overview of dissertation projects

This dissertation includes two projects that both study congenital birth defects employing two different gene mapping approaches commonly applied to genetic epidemiological research. Neural tube defects (NTDs) are considered complex, involving many genes and environmental factors, and a component of my thesis focused on one large multiplex family that demonstrated a strong genetic effect. The other project studied a Mendelian disorder, Native American Myopathy (NAM), which is a rare congenital myopathy. **Chapter 2** discusses the linkage findings obtained from a SNP-based genomic screen of one large NTD multiplex family, 8776. Additional ascertainment in this family facilitated further linkage and haplotype studies presented in **Chapter 3**. **Chapter 4** describes the natural history and clinical observation in patients with NAM. Finally, **Chapter 5** discusses homozygosity mapping data for five individuals with NAM.

1.2 Neural Tube Defects are a Complex Group of Genetic Disorders

The neural tube is part of the central nervous system and is formed during embryonic development. The process of neural tube closure occurs through neurulation beginning in the third week of gestation. Failure of appropriate neural tube closure by the 4th week of gestation leads to a neural tube defect and these defects often arise before a woman is aware

of her pregnancy. The etiology of neural tube defects is considered complex with both genetic and environmental factors implicated.

1.2.1 Neural Tube Defects: Clinical Impact

Birth defects are the leading cause of death in neonates. The risk of an infant being born with a significant congenital abnormality in body structure or function is approximately 3-4%. NTDs occur at a rate of 1/1000 live births in the US [1], representing the second most common type of birth defects after congenital heart defects and the most common congenital malformation of the central nervous system. NTDs are phenotypically heterogeneous and are classified as either an open or closed defect based on either the absence or presence of a layer skin that covers the spinal defect. Open NTDs are brain and/or spinal tissue that are uncovered by skin tissue and therefore exposed at birth. In contrast, closed NTDs are covered by a layer of skin and consequently have less risk of infection.

The most common forms of NTDs are the open NTDs: anencephaly and spina bifida cystica. Anencephaly is the failure of the formation of the brain and back of the skull, and all infants with anencephaly are stillborn or die shortly after birth. Two types of spina bifida are commonly described, spina bifida occulta (SBO, closed) and spina bifida aperta (open). Spina bifida occulta is a vertebral defect without protrusion of any spinal contents. SBO patients are usually asymptomatic. The most common NTD, open spina bifida also called myelomeningocele, involves a defect in the vertebral arches leading to herniation of spinal tissue (spinal cord and meninges) at the base of the spine, resulting in deleterious neural deficits below the level of the lesion. Craniorachischisis, which is always incompatible with

life, is a rare severe form of open NTDs with fetuses having an entirely open neural tube (encompassing both the cranial and spinal regions).

For closed NTD defects, such as spina bifida occulta, often the only indication of the NTD is a cutaneous finding such as a skin tag, midline dimple, or hemangioma. Encephalocele is a closed NTD with a defect in the skull that leads to the protrusion of the brain outside of the skull into an enclosed sac of skin. Children affected with open spina bifida often receive surgery to close the defect within 48 hours after birth, and may require frequent surgical intervention for hydrocephalus. Surgical closure of an open neural tube is performed shortly after birth for the following two reasons: 1) to prevent possible infection of the neural tissues that could lead to meningitis; and 2) to help preserve neurological function of the tissues in close proximity to the open defect.

Hydrocephalus, which is an enlargement of the cerebral ventricles, is caused by an obstruction along the pathway of the cerebrospinal fluid (CSF) flow from the ventricles to the arachnoid villi. Ninety percent of children born with open spina bifida will have venticulomegaly, and 70% of these children will require shunt placement [2]. Interestingly, most patients with myelomeningocele also have Chiari II malformation with four potential sites of obstruction for the cerebral spinal fluid pathways including at the aqueduct of sylvius, cork in bottle, foramen of magendie, and venous outflow obstruction. This condition includes downward displacement of the medulla, fourth ventricle, and cerebellum into the cervical spinal canal, as well as elongation of the pons and fourth ventricle, probably due to a relatively small posterior fossa.

The Chiari II malformation is a complex congenital malformation of the brain and is nearly always associated with myelomeningocele. This co-occurrence of spina bifida and

Chiari II malformation may in part be explained by the disrupted cerebral spinal fluid pathways caused by the Chiari II malformation. For instance, the presence of disrupted CSF flow in the spine may have deleterious effects on appropriate neural tube formation and closure at the base of the spine, thereby leading to the associated spina bifida.

Patients affected with a spinal defect commonly have an impaired ability to walk and often need to use a wheelchair, and they may have little or no bowel and/or bladder control. The extent of impairment for the patient is dependent on the position of the spinal lesion. For example, a child with a cervical lesion may not be able to breathe on his or her own whereas a child with a lumbosacral lesion –the most common location for spina bifida—may have nerve damage affecting ambulation, bladder, and bowel function. In 1990, the estimated annual medical and surgical costs for persons with spina bifida exceeded $200 million [3]. However, the psychosocial and emotional toll for the family and child with NTD is immeasurable. These significant medical and personal costs have lead to several collaborative studies to investigate NTDs to better understand the genetic and environmental etiologies.

1.2.2 Epidemiology of Neural Tube Defects

Since a large percentage of fetuses affected with a NTD are likely to be spontaneously aborted before diagnosis, estimates of NTD incidence may be underestimated. Therefore, in a study of congenital birth defects, it is generally recommended to report prevalence estimates, often reported as live births, since there may be elected abortions and early fetal demise in the first few hours to days of life. Across the world, the prevalence of NTDs varies widely showing significant geographic and ethnic variation in NTD risk.

Prevalence studies and registries have been designed for epidemiological surveillance of birth defects in order to identify particular exposures in different geographical locations and populations which may either have protective mechanisms or potential teratogenic risks. In the US between the years of 1991-2004, the frequency of spina bifida was 19.56 per 100,000 live births and the frequency of anencephaly was 10.39 per 100,000 live births [4]. The highest US prevalence rates are noted in the South-East, particularly in the Appalachian region [2]. Using data reported to the National Birth Defects Prevention Network (NBDPN) for the years between 1999-2000, it was found that Hispanics had the highest prevalence of NTDs in the United States with ~5.8 per 10,000 live births, Caucasians were the next most frequent with 5 per 10,000 live births, followed by African Americans with ~3.8 per 10,000 live births [5]. More specifically, populations of Mexican descent along the Texas-Mexico border of the United States have a high prevalence of NTDs [6]. The high prevalence observed in this group may be due to any number of factors including genetic susceptibility, inadequate dietary folate levels, an occupational exposure such as pesticides on farms, and/or an environmental exposure such as fumonisin B1 (FB1). FB1 is a naturally occurring mycotoxin in corn—a dietary staple of this Hispanic population—and FB1 has been shown to reduce cellular folate uptake [6]. Moreover, an association between fumonisin exposure and NTD prevalence has been observed in women living along the Texas-Mexican border [7].

Other countries exhibit higher rates of spina bifida or anencephaly prevalence in comparison to the United States. For example, both Mexico and Guatemala have high prevalence rates of NTDs, rates that may also be attributed to FB1 exposure. The prevalence of NTDs in Mexico from the years of 1996-2000 was 13.5 per 10,000 live births for

anencephaly and 14.4 per 10,000 live births for spina bifida [8]. In India, a higher prevalence of NTDs are noted in the North (Punjab, Delhi) in comparison to Eastern cities [9]. For example, a retrospective study in Punjab, India reported an anencephaly prevalence of 36 per 10,000 live births [10]. Northern China's Shanxi Province has one of the highest prevalence of NTDs in the world, and during 2003 the birth frequency of NTDs was 138.7 per 10,000 live births [11]. Importantly, only 4.2% of mothers with children having NTDs used folic acid supplements during the periconceptional period, suggesting folic acid deficiency as a possible explanation for the high prevalence, although, other factors likely contribute to this high prevalence, for example teratogenic exposures and/or genetic effects.

Currently, it is unclear whether these observed differences in NTD risk are related to population-specific differences in genetic risk, environmental factors, or a combination of the two.

1.2.3 Syndromic versus Nonsyndromic NTDs

The etiology of NTDs is heterogeneous [12,13] both in phenotypic characteristics as well as the genetic predisposition. For instance, the majority of infants with NTDs are nonsyndromic (isolated cases without any other associated phenotypic features) and likely result from multifactorial etiologies. Different studies have reported that most cases of spina bifida and anencephaly (approximately 80-90% [14-17]) are nonsyndromic. Other patients with NTDs may have additional birth defects, and these NTDs are either considered part of a syndrome with a presumed cause (single gene mutation, chromosomal abnormality, or teratogen exposure) or if not part of syndrome, the child may be described as having multiple birth defects. The genetic mechanisms responsible for nonsyndromic NTDs versus multiple

birth defects and syndromic NTDs are likely to be different; however, it is also possible that there are pleiotropic effects of genes, or in the case of chromosomal abnormalities, that several genes may be altered leading to more tissues and organ systems to be adversely affected.

1.2.4 Chromosomal Abnormalities in NTDs

Chromosomal abnormalities are observed in both syndromic and nonsyndromic NTDs. Neural tube defects can be caused by chromosome rearrangements. Chromosomal abnormalities can include translocations, deletions and insertions/duplications, and the cytogenetic breakpoint may disrupt genes involved in neural tube development. Such chromosomal abnormalities are present in 5-17% of NTD cases; however, this estimate may not accurately predict live birth frequency since many of the studies were conducted prenatally and therefore also describe chromosomal abnormalities of stillbirths and spontaneous abortions [18-20]. For example, Sepulveda et al. conducted a prospective study of fetuses with open neural tube defects that underwent prenatal chromosome analysis [21], and found that trisomy 18 and 13 were the most common chromosomal abnormalities and both of these aneuploidies have high rates of fetal mortality. They collected cytogenetic data on 144 fetuses with open NTDs, and found that 10 fetuses (7%) had chromosomal abnormalities, 7 fetuses with trisomy 18 and 2 fetuses with trisomy 13. Characterizing the chromosomal abnormality in individuals affected with an NTD could be informative for identifying potential NTD candidate genes that either fall within the cytogenetic breakpoint or are in the deleted and/or amplified chromosomal segments.

The majority of NTDs are nonsyndromic and 2-16% of the individuals with nonsyndromic NTDs also carry chromosomal abnormalities [22]. However, if NTDs are observed with other congenital anomalies the frequency increases up to 24% [18]. The reported frequency may be even higher in spontaneously aborted fetuses, with approximately 53% of spontaneously aborted fetuses with an NTD having a cytogenetic abnormality [22]. Trisomy 18 is the most common aneuploidy associated with various forms of NTDs including spina bifida, anencephaly, craniorachischisis, encephalocele; trisomy 13 is also associated with spina bifida and encephalocele [22]. Other commonly associated chromosomal aberrations observed in NTDs include other trisomies and tetraploidy, duplication of 2p23-2pter, ring chromosomes 13 and 18 (formed by fusion of ends of chromosome), and deletions in chromosomes 13q and 22q, among many others.

There is variable presentation of NTDs with chromosomal abnormalities that may implicate a multifactorial nature of NTD manifestation even in syndromic cases, such that numerous genetic abnormalities in combination with particular environmental exposures may lead to expression of different NTD phenotypes.

1.2.5 Environmental Components to NTD Predisposition

The predisposition to and development of NTDs is complex, with both environmental and hereditary factors implicated. There are a vast number of exogenous risk factors correlated to the development of NTDs including maternal insulin-dependent diabetes mellitus, obesity, maternal use of anti-epileptic (valproic acid) and cholesterol (statins) medications, maternal elevated homocysteine levels, smoking, Agent Orange exposure, and mycotoxin fumonisin (a fungus toxin found on maize) [23-29].

The antiepileptic medications valproic acid (VPA) and carbamazepine have been associated with an increased risk for NTDs in both human studies and experimental animal models [24,30-36]. VPA has been associated with a variety of major and minor malformations, including neural tube defects, cleft lip and palate, cardiovascular abnormalities, genitourinary defects, developmental delay, endocrinological disorders, and limb defects. It has been estimated that 1-2% of all infants exposed to VPA during early pregnancy will have spina bifida aperta, a 20-fold increased prevalence over that observed in the general population [31]. VPA may cause these defects by interfering with retinoic acid metabolism [37]. With carmazepine, another anti-epileptic medication, the risk for exposed infants presenting with a NTDs ranges between 0.5% and 1% [25,32,38]. Lower blood folate concentrations in women treated for epilepsy have also been associated with increased congenital malformations [39]. However, folic acid supplementation in women with epilepsy is not always associated with protection against NTDs, suggesting maternal exposure of antiepileptic meds may be an underlying pharmacogenetic susceptibility[40].

A study of pregnant epileptic women revealed significant changes in drug metabolism enzymes with specific decreases in P450 1A2, xanthine oxidase and n-acetyltransferase activities, and a significant increase in hydroxylation activities during pregnancy [41]. These findings demonstrate that metabolism of antiepileptic drugs will be altered during pregnancy. And it is also possible that genetic polymorphisms in drug metabolizing genes in either the mother or infant can further change the pharmacokinetics of antiepileptics which may contribute to the 20-fold increased NTD risk in particular pregnancies.

Epigenetic factors may also influence susceptibility to NTDs, for example the maternal genetic background combined with environmental exposure may indirectly affect the development of the fetus in the intrauterine environment.

The main environmental intervention that has had a significant impact on preventing NTDs is maternal periconceptional supplementation with folic acid, which has been shown to reduce the recurrence risk for having a child with an NTDs by 70% [42] and estimated to reduce the risk in the general US population by 50% [43]. Fortification of the U.S. food supply with 140 mcg of folic acid per every gram of enriched grain products, since January 1998, has been estimated to reduce the prevalence of NTDs by 20-30% [44,45]. It was also shown that the post-fortification the newborn hospitalization rates for spina bifia decreased by 21% from the 1993-1997 to 1998-2002 which reflects the reduction in prevalence estimates [46]. Additionally, maternal serum alpha fetoprotein values have decreased by 32% since fortification, roughly mirroring the decreased prevalence estimates of NTDs [47]. However, the magnitude of folic acid's protective effect varies between different studies which may reflect differences in study design and/or genetic heterogeneity from differences in population-specific responsiveness to folic acid during pregnancy.

1.2.6 Evidence for a Genetic Contribution to NTDs

NTDs are likely caused by a complex interaction between environmental influences and genetic susceptibility. The significant reduction in NTDs with periconceptional supplementation with folic acid demonstrates that folate has a protective effect. This NTD protective effect may represent a maternal or fetal genetic susceptibility that can be modulated by gene-environment interactions. Several lines of evidence support a genetic

factor in NTDs [48]. Firstly, NTDs are associated with established genetic syndromes such as CHILD, Fraser, Jarcho-Levin, Meckel-Gruber, and Waardenburg syndromes [49]. As mentioned previously, NTDs are also associated with chromosomal rearrangements such as trisomies 13 and 18 [22]. Moreover, NTDs tend to breed true, meaning if there are multiple affected individuals within a family, they will tend to have the same phenotype [50-55] suggesting familial predisposition to particular phenotype may be due to similar genetic etiologies. Also, twin studies demonstrate a higher risk of NTD phenotype [56,57]. However, the twin studies compare like (monozygotic and dizygotic) versus unlike-sex (only dizygotic) instead of comparing the zygosity of the twin pairs (monozygotic versus dizyotic), which confounds the ability to estimate the degree of twin concordance. The limited twin data available, which is based on small sample sizes, claims concordance ranging from 3.7% to 18% [56]. Also, there are currently 155 mouse models with known genetic mutations causing an NTD phenotype, and these NTD mouse models provide further evidence in support of a genetic component underlying the formation of human NTDs [58].

Family history of more than one individual affected with an NTD [59-61] or familial clustering of NTDs is one of the strongest lines of evidence for a genetic contribution. In addition, there is a higher risk of recurrence of NTDs in offspring of parents with previous NTD pregnancies is higher, with a sibling recurrence risk of 2-5%, which is ~ 25-50 fold increase compared to what is observed in the general population [62,63].

Segregation analysis evaluates multiplex families with a particular disorder, such as NTDs, and attempts to fit a genetic model to the disease based on the transmission patterns observed in the families. In 1982, two formal segregation analyses were reported for NTDs. One family-based segregation analysis study suggested a genetic contribution by showing a

higher likelihood that the NTD phenotype was caused by a dominant gene with 75% penetrance than sporadic occurrence or recessive inheritance [64]. The other segregation analysis study in 298 nuclear NTD families observed results consistent with either a recessive major gene or environmental effects or a combination of both [65]. These differences in patterns of inheritance may reflect different family structures with likely different genes and modes of transmission. Since we believe NTDs to be multifactorial in etiology it is possible that many different forms of inheritance are contributing to disrupted neural tube closure during early embryogenesis.

Genetic gender differences in children born with anencephaly provide additional evidence for a genetic contribution to NTDs. In the United States it was observed that approximately 68% of anencephaly cases are female and 56% of open spina bifida cases are female [56]. Several theories have been proposed for the female preponderance of anencephaly cases such as there may be differences in spontaneous abortion rates in males more than females, sex-linked genetic effects that may be protective for females and/or deleterious for males, and that sex specific products, like hormones, might be affecting disease risk. The latter, seems less likely since neurulation precedes gonadal development. Cranial NTDs in mice have shown a female excess with approximately two-thirds of the exencephalic embryos being female in the SELH/Bc strain and in the *Cd*, *crn*, *ct*, *Efna5*, *Hipk1/Hipk2*, *Marcks*, and *xn* mouse mutants [58]. The molecular basis for the higher risk of females for cranial NTDs in both human and mice is currently unknown [58]. However, it has been proposed that the higher risk observed in female embryos may be explained by the higher demand of methyl groups required for appropriate cellular X-inactivation [66]. Further support of X-chromosome involvement in NTDs is from known X-linked conditions

associated with neural tube defects including CHILD syndrome, focal dermal hypoplasi, X-linked laterality sequence, X-linked midline defect, and X-linked neural tube defects[2].

While several lines of evidence suggest human NTDs have a genetic component, a major genetic factor has not been identified.

1.2.7 Challenges to Identifying a Genetic Component

The genes that contribute to inherited risk for having a child with a NTD are not well established. Due to the reduction in NTDs with folic acid supplementation, candidate genes in the folate metabolism pathway have been extensively investigated [67-69]. There have been conflicting data in human genetic studies on the folate metabolizing genes, specifically with regards to 766C-> T thermolabile isoform of (*MTHFR*), with some studies demonstrating an association and others not [68,69]. This discrepancy in human studies investigating association with the *MTHFR* polymorphism and NTD risk is likely caused by the differences in the study designs or statistical approaches, along with population specific differences in risk. For example, many of the studies showing significant association between *MTHFR* and NTDs are case-control studies, which are subject to sampling biases and population stratification [70-72]. Such confounding factors can lead to false positive associations; consequently, case-control studies need to be conducted with careful attention to possible bias and confounding by ethnicity and demographics. Family-based designs can be a reliable approach to analyzing the genetic contribution in human disease because they are robust to population stratification and the transmission of alleles from parents to proband can be traced.

Even though genes involved in folate metabolism are biologically plausible candidates, 30-50% of NTD cases are not responsive to nutritional folate [73] suggesting there may be other genetic mechanisms for folate-resistant NTDs. Undertaking the study of the genetic basis of human NTDs presents several complicating factors. First, multiplex families are relatively rare and traditional linkage studies using multiplex pedigrees to identify regions of interest based on identical by descent sharing have been limited. The majority of NTDs are sporadic, meaning that it is easier to ascertain simplex families (one affected individual in a family). Second, collection of affected individuals is challenging because of the increased spontaneous abortion and perinatal deaths associated with NTDs, low parent-to-child transmission due to decreased reproductive capacity in affected cases, and early termination of pregnancies detected prenatally [48,74,75].

1.2.8 Why Investigate a Genetic Component for NTDs?

Any insights into one or more genes predisposing to the development of NTDs would lend itself towards more accurate genetic counseling for families and ideally prevention of these frequently disabling birth defects. Three different prenatal diagnostic methods are currently available to assess a pregnant women's risk of having a child affected with an open NTD and these include MSAFP, amniocentesis to test AFP, and ultrasound. Typically, the first prenatal test for levels of maternal serum alpha-fetoprotein (MSAFP) takes place from 16-20 week gestation. Fetal AFP is a protein produced by fetal yolk sac, gastrointestinal tract and liver and a small amount can cross the placenta moving into the mother's blood stream. Fetal serum concentrations of AFP decline exponentially between the 14th to 32nd week gestation. MSAFP is a screening test for both open NTDs and Down Syndrome. An

increased risk of open NTDs is associated with high levels of MSAFP whereas an increased risk of Down Syndrome is associated with low level of MSAFP. An MSAFP cutoff level of 2.5 MoM or greater has a 73% detection rate for open spina bifida at 17 weeks gestation with false-positive rates of 1% when the AFP values are corrected for maternal age [2]. If a woman has an elevated MSAFP during pregnancy, an amniocentesis and ultrasound are suggested. Because the MSAFP is not a specific diagnostic tool but an associated metabolic marker, it can lead to false positive results. This can be anxiety provoking for the mother [76] and family and can lead to unnecessary medical procedures that increase risk for pregnancy loss. The identification of causative gene(s) for NTDs can enhance our prenatal testing capabilities generating more accurate screening tools. More importantly, preconceptional interventions may be implemented based on the familial genetic susceptibility, such that periconceptional folate supplementation may be more relevant for only a subset of pregnant women whereas other medical intervention(s) may be more efficacious for other individuals.

1.2.9 Normal Human Embryology of Neural Tube Development

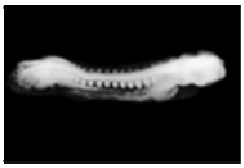
The process of neurulation leads to the formation of the neural tube and is a key morphogenetic event in human neurodevelopment. The neural tube gives rise to the brain and the spinal cord to form the central nervous system. The central nervous system (CNS) is the first adult organ system to appear during vertebrate development and formation of the human neural tube occurs in the very early stages of fetal development. Different regions of the neural tube form through two distinct processes: primary neurulation and secondary neurulation. Primary neurulation is the process by which a flat neural plate forms a cylindrical neural tube, resulting from the neural plate folding into the midline until the

neural folds (or edges of the neural plate) contact and fuse. Secondary neurulation is complete by the 8th week of gestation and is the process by which the neural tube forms at the most caudal end involving canalization of a dense mass of cells in the tail bud to form a hollow tube. Secondary neurulation is considered a mechanistically different process than primary neurulation since it occurs over several weeks as opposed to several days and does not require elevation and fusion of the neural plate.

As mentioned, the neural tube closure occurs early in the process of embryogenesis, and the early stages of fetal development during primary neurulation demonstrate distinctive morphological features [77]. Embryologists use these distinctive morphological findings in the human embryo to approximate the stage of development or embryonic Carnegie Stages (CS) because often it is difficult to determine the exact dates of conception [77]. This description of the Carnegie Stages focuses specifically on stages involving neural tube closure and are therefore postulated stages when anencephaly and spina bifida arise; these CS 10-13 are illustrated in Figure 1.1.

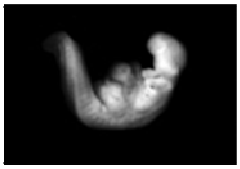
CS 10

At CS 10 (22 days gestation) 4-10 somites (giving rise to vertebrae) and the neural folds fuse in the midline.



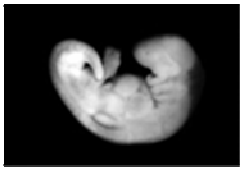
CS 11

At CS 11 (24 days gestation) both ends of the neural tube are open evidenced by rostral and caudal neuropores. Once 20 somites are present, the rostral neuropore closes within a few hours and failed closure may lead to anencephaly.



CS 12

By CS 12 (26 days gestation) 21-29 somites are present and the caudal neuropore closes taking about 24 hours for complete closure of the neural tube and any disruption of this process may lead to spina bifida.



CS 13

By CS 13 (28-32 days gestation) 30 somites are present and the neural tube has completely closed.

Figure 1.1 Human neural tube closure and Carnegie Stages 10-13.
Adapted from UNSW Embryology website by Mark Hill
http://embryology.med.unsw.edu.au/wwwhuman/Stages/CStages.htm

Closure of the neural tube prevents high levels of α –fetoproteins accumulating in amniotic fluid and thus forms the basis for MSAFP screening for open neural tube defects. Since closure takes place in the first few days of life, often well before a woman is even aware of her pregnancy, it is often impossible to take preventive measures, such as consuming folic acid. Closure of the caudal and rostral neuropores occurs at different stages of development and requires different amounts of time, suggesting there may be different underlying molecular mechanisms for each neuropore to close appropriately. These differences may result from spatial and temporal heterogeneity in molecular mechanisms involved in transformation of neural fold morphology to produce neural tube closure.

Since rostral and caudal neuropore closure takes place at different times, it seems possible that there may be different genetic risk factors involved with a cranial or spinal defect. To this end, a question commonly debated in the NTD community is whether

different NTD phenotypes have the same or different underlying genetic and/or environmental risk factors. Phenotypic classification of NTD mouse models have revealed that approximately 20% of mutants have either exencephaly or spina bifida aperta or both, 70% only have exencephaly, 5% have spina bifida aperta , and 5% have craniorachischsis [58]. These observations in mice suggest that particular genetic risk factors will only primarily affect a particular region of the neural tube whiles other genes may be involved with formation of the entire neural tube.

1.2.10 Role of Folate in Development and Sufficient Maternal Folic Acid Consumption

It is not currently understood why some NTDs are folate-sensitive while others are folate-resistant. Folic acid is involved in several biochemical pathways including nucleoside biosynthesis and production of methyl groups necessary for homocysteine metabolism [78]. Moreover, folate is crucial for the maintenance and production of cells, especially during periods of rapid cell division, as would be expected during neurodevelopment. Folate deficiency leads to impaired folate metabolism, and folic acid supplementation has beneficial effects on several developmental defects and diseases including neural tube defects (NTDs), cleft lip and cleft palate, omphalocele, congenital heart disease, cardiovascular disease, and cancer [5,78]. However, the exact process by which folic acid is protective against congenital birth defects is still unknown.

Folate derivates are involved in single-carbon transfer for many essential metabolic reactions. During early development many methylation reactions are taking place such that any impairment of the folate and methionine metabolic cycles involving the methylation

pathway may disrupt neurodevelopmental processes. For example, folate metabolism involves homocysteine (Hcy) remethylation to methionine integral for formation of Ado Met, a substrate for many transmethylation reactions including DNA methylation (shown in Figure 1.2). DNA methylation undergoes rapid changes during embryogenesis, especially during the first stages, so small delays or impairment in remethylation of genes could affect expression of important genes needed for closure of the neural tube.

Elevated plasma homocysteine levels have been observed in women with affected NTD children [79,80]; these elevated homocysteine levels have been shown to normalize when women are given adequate folic acid [80]. It is currently unknown why high homocysteine levels and low folate levels have been associated in women with NTD pregnancies. The elevated homocysteine may be a metabolic marker with a direct effect on NTD development or indirectly may reflect low folate status. Elevated homocysteine levels have been proposed to disrupt embryogenesis through several different mechanisms including: 1) toxic mechanisms to the fetus; or 2) disrupted methylation reactions in the fetus; or 3) increased oxidative stress in the fetus [81].

The widely studied *MTHFR* enzyme converts 5, 10- methylene THF to 5-methylTHF and consequently this enzyme regulates the balance between folate derivates being used for either DNA synthesis or methionine/adoMet synthesis involved in methylation reactions (shown in Figure 1.2). Many checks and balances are in place to ensure proper homeostasis between folate and methionine metabolic pathways; for example, adoMet inhibits the *MTHFR* enzyme.

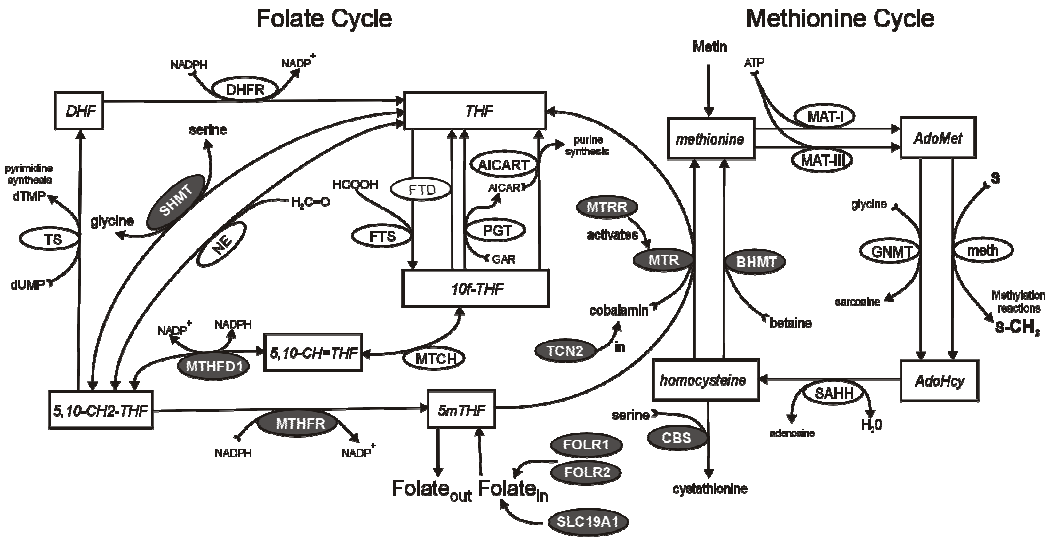


Figure 1.2 The folate and methionine metabolic cycles.
Substrates are shown in rectangular boxes while enzymes are shown in ellipses. Adapted from Nijhout, 2004; Reed, 2004 [82,83].

Folate is a water-soluble, B-complex vitamin acting as a coenzyme for single carbon transfers and nucleotide biosynthesis. In the United States, a woman may consume folate or folic acid in three ways: through food folate, folic acid from fortification, and/or through supplemental folic acid either from multivitamins or from a prescription of folic acid. The difference between folic acid and food folate is based on its bioavailability, and consumption of food folate on an empty stomach has approximately the same bioavailability as supplemental folic acid [84]. A way to estimate the degree of folate derivates consumed by women around the time of conception through the first month of pregnancy is by administering questionnaires that address their multivitamin use, folic acid supplementation, and specific diet. Unfortunately, these types of questionnaires are subject to recall bias. Ideally, to reduce the problem of recall bias, a study should be designed prospectively before a woman is pregnant to obtain a more accurate reflection of folate/folic acid intake. Further

quantification issues in estimating maternal folic acid intake comes from variability in folic acid dosage in different multivitamin formulations.

Despite the significant public health impact folic acid has had in the prevention of NTDs, the protective mechanism of action remains unclear [85]. Folic acid may be protective in the mother or the fetus or both. There may be varying reasons why certain women require a higher folic acid intake to protect against NTDs, such as some may have poor folate absorption, others may have increased excretion of folate, while others may have abnormal folate metabolism either from genetic mutations in the folate metabolism genes or from environmental disturbances. Studies have found that folate receptors have an important role in embryogenesis and NTD development [86,87].

Based on the observation that administering antiserum to folate receptors in pregnant rats caused developmental abnormalities [88], Rothberg and colleagues speculated that autoantibodies against folate receptors in women could also cause neural tube defects [89]. As they predicted, autoantibodies to folate receptors were found in 9 of the 12 women who had NTD-affected pregnancies and only in 2 of the 20 control subjects. The authors had two different explanations as to why supplemental folic acid may have a protective effect in relation to autoantibodies to folate receptors: 1) supplemental folic acid may bypass the folate receptors by being alternatively transported in the cell by the reduced folate carrier; and 2) according to their work, folic acid appeared to have a higher affinity to the folate receptor than the autoantibodies did, suggesting that high folic acid levels may displace the autoantibodies from the folate receptor. The sample size for this study was rather small, including only 12 women with a current or prior pregnancy affected with an NTD and 20 control women (pregnancy not affected by an NTD); therefore, replication of folate-

autoantibodies in a larger NTD dataset is warranted to determine the degree of reliability and generalizability of this finding.

1.2.11 Epigentic Effects and Parental Effects on Neural Tube Defects

Parental effects such as genomic imprinting may play a role in early developmental processes such as during neural tube closure. Epigenetic alterations acquired *in utero* can be inherited mitotically in somatic cells that change gene expression without altering the DNA sequence. The process of imprinting is re-established in each successive generation, and imprinted genes have varying temporal and spatial expression patterns. For example, imprinted genes may be monoallelically expressed from one parental chromosome in specific tissues and during particular developmental stages while having biallelic expression patterns in other parts of the body. Most imprinted genes are expressed in fetal and placental tissues [90,91], supporting the idea that monoallelic expression patterns may play a role in early developmental processes. Imprinting occurs through specific patterns of cytosine-phospho-guanine (CpG) dinucleotide sequence methylation that are established *de novo* by Dnmt3a (DNA-methyltransferase) and Dnmt3b in early development and then propagated during DNA replication by Dnmt1 [92].

Since one-carbon metabolism from the folate and methionine metabolism cycles provide methyl groups for biological methylation reactions, early human nutrition may influence DNA methylation patterns or epigenetic programming of the fetus *in utero* [93]. The large role maternal folic acid consumption has had in preventing NTDs suggests that there may be a maternal contribution to NTD susceptibility through the intrauterine environment.

A few human genetics studies have suggested a role for imprinting or parental sex effects in NTD risk. One study evaluated fifty families with individuals affected with nonsyndromic spina bifida [94] and observed twice as many unaffected female transmitting members than males, after excluding ascertainment bias, simple X-linked or mitochondrial inheritance and male infertility. Two other groups studying nonsyndromic NTDs in distant affected relatives also observed an increase in maternal transmission, and the authors discussed the possibility that maternally imprinted developmental genes may explain the female preponderance of normal transmitters [95,96]. Moreover, a recent study in 48 Irish NTD families observed that maternal first cousin pregnancies were more likely to end adversely (defined as premature children, stillbirths, and miscarriages) when compared to paternal first cousin pregnancies, and that excess risk with maternal line was associated to spina bifida occulta families [97].

1.2.12 Theories of Neural Tube Closure Sites

It has been postulated that the different forms of NTDs in both humans and mice could result from different fusion sites of the neural folds in the last stage of primary neurulation. Since the early 1990's different investigators have proposed varying theories on the number and location of these fusion sites in humans. The original "zipper theory" proposed that the first site of neural fold fusion occurred in the cervical region of the developing fetus with subsequent fusion of neural folds progressing bidirectionally in rostral and caudal directions [98]. This simplistic model was refuted by other later theories. For example, four sites of initiation of neural fold fusion were originally reported in the mouse by Sakai et al in 1989 [99] and later only three closure sites were corroborated by other groups

[100,101]. Based on this mouse work and observation of therapeutic abortuses and still born fetuses (19 embryos total) with different fusion defects along the neural tube, Van Allen et al in 1993 proposed that human embryos have five fusion sites with four corresponding neuropores [102]. The basis for this multi-site model was to explain the variety of NTDs that may result from failure of these closure sites [102,103]. Although Van Allen study did accurately predict the existence of multiple closure sites in humans, the five site model was not supported by later studies conducted in human embryos [104-106]. In 2000, Nakatsu et al evaluated 68 normal embryos of Carnegie stages 10-12 and 98 NTD embryos of Carnegie stages 11-23. Within these human embryos they observed three closure sites described as *A*, *B*, and *C*. The first closure site *A* was located at the upper cervical region and progressed both rostrally and caudally. The second closure site *B* occurred at the rhombenecephalon-mesencephalon junction; however, this closure site *B* only was observed in 10 embryos so it was not a widely accepted to be an actual closure site. And the third closure site *C* was observed occurring at the rostral tip of the neural groove. In 2002, O'Rahilly and Muller observed only two sites of initiation of neural fold fusion (α and β) in nighty-eight paraffin sections of Carnegie Stages 8-13 from the Carnegie collection of human embryos. Fusion site α was found in the rhombocephalon and fusion site β localized to the rostral tip of the prosenephalon [106] as illustrated in Figure 1.3. They also observed additional accessory sites of neural fold fusion but they were inconsistently observed and did not reveal a specific pattern of position. Thus, it is possible that closure site *B* observed by Nakatsu and colleagues (2000) was also an accessory site that is not always present. These accessory closure sites may reflect the complexity of molecular events involved in neural tube closure

and underlying genetic differences and environmental exposures that contribute to variable morphological events during neurodevelopment.

Mice only have three proposed sites of neural fold fusion localized to the cranial neural tube which further refuting the idea of the 5 site model proposed by Van Allen [107]. As illustrated in Figure 1.3 , mice have been proposed to have the following three neural tube closure sites: closure 1 occurs first and is initiated at the hindbrain/cervical boundary and then closure continues bidirectionally, closure 2 occurs at the forebrain/midbrain boundary, and finally, closure 3 occurs at the rostral end of the forebrain [107]. Interestingly, closure site 2 in mice appears to vary in location along the rostrocaudal axis and timing of closure and this polymorphic presentation was observed inbred mice, outbred mice, and backcrosses of *splotch* (sp^{2H}) mice to DBA2 mice and NZW mice [108]. Although the location of closure site 2 is polymorphic, the closure locations and timing of 1 and 3 appear to remain the same among different mice [107]. This morphological variation in closure site 2 is consistent with the two human embryo studies described whereby one study observed a closure site 2 while another study did not [105,106].

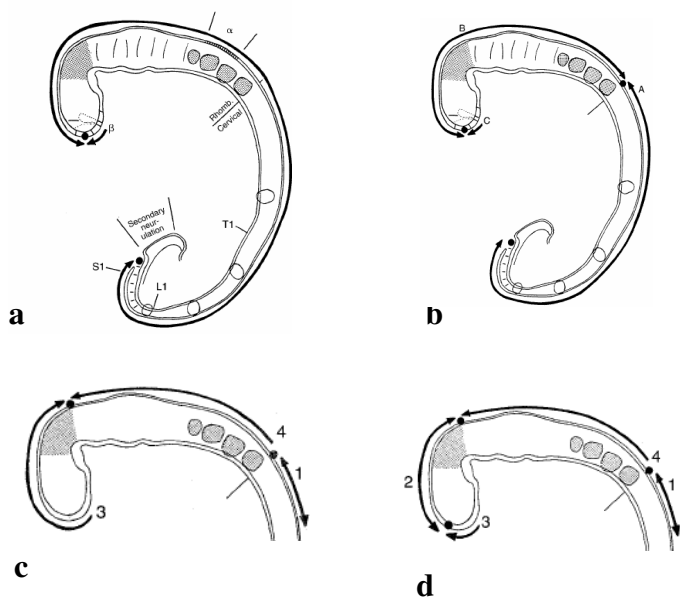


Figure 1.3 Regions of human and mouse neural tube closure sites.
The depiction uses an outline of an human embryo at Carnegie Stage 12 for observation in both mouse and human a) the human two-site closure model (designated by α and β) b) the human three site model (designated by A,B, and C) c) the mouse three closure site model, d) the mouse four closure site model. Adapted from O'Rahilly and Muller 2002 [106]

1.2.13 Key Molecular Process Underlying Neural Tube Closure

Neural tube closure is a dynamic process that requires a network of protein-protein and protein-DNA interactions in order to orchestrate the appropriate spatial and temporal expression of various factors involved in neurulation. Briefly, at the edge of the neural plate, neural folds form and then elevate and fuse in the dorsal midline to form the closed neural tube. Currently, the underlying processes thought to be involved with human neural tube closure have been elucidated from animal models, and the mouse model system is one of the most widely used animal models for studying NTDs. The plethora of mouse models have demonstrated dysregulation of numerous proteins that may result in a NTD, which is not

surprising given the incredible molecular complexity involved in the formation of the neural tube. Several of the gene products are transcriptional regulators and molecules involved in signal transduction. These products may regulate several different molecular pathways and could explain why some of the mouse models are not 100% penetrant.

The active processes known to be involved with neural tube closure include convergent extension cell movements, expansion of the cranial mesenchyme, contraction of actin microfilaments, regulated bending of the neural plate, balance of proliferation and programmed cell death of neuroepithelium and adhesion of neural folds [107]. A comprehensive evaluation of all the molecular mechanisms involved in neural tube closure is outside the scope of this dissertation, but since genes related to the planar cell polarity (PCP) pathway were preferentially selected as candidate genes for sequencing efforts described in Chapter 3, the PCP pathway is further described. The PCP pathway involving the highly conserved non-canonical Wnt-frizzled-dishevelled signaling pathway is required for initiation of neural tube closure. In *Xenopus*, it was first shown that an absence of PCP-dependent convergent extension causes neural folds to be widely spaced apart preventing apposition and fusion of the neural folds, and consequently producing embryos with short broad neural plates and failed neural tube closure [109,110]. Convergent extension movement is a process mediated through the planar cell polarity pathway through non-canonical Wnt signaling leading to inter-digitation of cells in the midline, in turn causing the neural plate to narrow and lengthen as shown in Figure 1.4. The PCP pathway requires binding of the Wnt ligand to the receptor Frizzled (Fz), triggering recruitment of cytoplasmic transducer Dishevelled (Dvl) to the plasma membrane to form a complex with

Fz; the PCP pathway signals are then transduced to the cytoskeleton through the activation of small Rho GTPase and JNK pathways [111]. *Drosophilia* genetics identified other signaling molecules for the PCP pathway whose interactions have not been identified and these include: Flamingo (Fmi), Strabismus (Stbm)/Van Gogh (Vangl), Prickle (Pk), Diego and Scribble (Scrbl)[111]. Moreover, PCP mutations in mice (Vangl2, Fzd3 and Fzd6 double mutants, Dvl1 and Dvl2 double mutants, Celsr1, and Scrbl) have been shown to result in widely spaced neural folds preventing neural tube closure (leading to craniorachischisis [112].) Overall, the PCP pathway appears to be an integral signaling cascade for appropriate neural tube closure during embryogenesis.

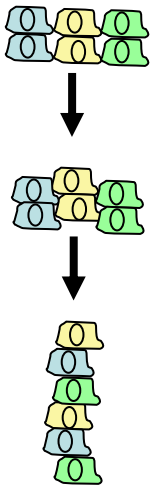


Figure 1.4 PCP-dependent convergent extension movements.

This process involves convergence of cells in the dorsal midline and mediolateral intercalations of cells whereby the tissue narrows in one axis and elongates in a perpendicular axis leading to an extension of the body axis. Adapted from De Marco et al 2006 [111].

1.2.14 Translation of Mouse Models to Human NTDs

The human homologues to several of the candidate genes gleaned from mouse models have been studied in human populations [58,69]. It is hypothesized that early developmental processes such as primary and secondary neurulation are conserved across vertebrate species, suggesting that genetic mutations in mice or other animal models may implicate important candidate genes for human NTDs. Although the number of mouse models with a known genetic mutation exhibiting NTD phenotypes is now at 155, the human homologs of these mouse genes have not, until recently, been particularly informative for human NTD studies [58,66,69,107]. Kibar and colleagues recently discovered three mutations in the *VANGL1* gene, with two mutations observed in two familial forms of spinal NTDs (V239I and R274Q) and one patient with a sporadic spinal NTD (M238T). The authors additionally demonstrated that the familial V239I mutation confers functional changes in VANGL1's protein-protein interactions with dishevelled-1,-2,-3 proteins [113]. Notably, the Lp mouse with a *Vangl2* mutation has craniorachischisis which is incomplete formation of both cranial and spinal neural tube [58]. Not only does this finding with *VANGL1* further support that there is a genetic contribution to NTDs but it also provides functional evidence of a rare genetic mutation associated with human familial NTDs.

Unfortunately, no other NTD mouse models have led to the discovery of human NTD genes, possibly because of the complex multifactorial etiology of NTDs. For example, phenotype disparity between mouse and humans manifests because the majority of NTD mouse models have exencephaly (which is equivalent to anencephaly phenotype in humans) whereas the majority of association analysis studies have samples from individuals with

lower open NTD lesion (myelomeningocele) with very little inclusion of anencephalic cases. Anencephaly is incompatible with life making collection of DNA samples for genetic analysis more challenging. Another factor that may complicate the mouse-human correlation is that most human NTDs are nonsyndromic and are likely due to multiple genes and environmental factors, while most NTD mouse models with null mutations are embryonic lethal having multiple other associated anomalies [114]. In order to emulate common complex human NTDs, an ideal mouse model of NTDs, would be nonsyndromic, have low penetrance of the NTD phenotype, and a reduced or increased NTD risk with environmental factors such as folic acid.

Only a few NTD mouse models exhibit similar multifactorial etiologies as is seen in human NTDs, with both genetic and environmental factors involved [58,66]. For example, the SELH/Bc mouse strain and curly tail mutation have oligogenic inheritance patterns. In SELH/Bc mice 5-30% have nonsyndromic forms of exencephaly and the genetic risk is caused by four loci (*Exen 1*, *2*, *3*, and *4*) acting additively with the frequencies of NTDs dependent on maternal diet as shown in Table 1.1 [58]. The curly tail (*ct*) mouse mutant may have a curled tail, spina bifida aperta, and exencephaly; three modifier genes have been shown to increase risk of having an NTD. Though SELH/Bc and curly tail (ct) have oligogenic inheritance patterns, neither appear to have decreased risk of NTDs with maternal folate supplementation [66]. Table 1.1 illustrates some mouse NTD models that have decreased frequency of NTDs with maternal nutritional supplementation of various factors including folic acid, folinic acid and methionine, among others, demonstrating a multifactorial etiology to these mouse models. Specifically, the folate-responsive NTD mouse mutants include Cart1, Crooked, Folbp1, and Splotch as shown in Table 1.1.

Table 1.1 Responses to maternal nutrition in mouse models for NTDs.

Mutant or strain	NTD	% NTD		Supplement
		Control	Treated	
Axial defects (Axd)	Spina bifida	30	15	Methionine
Cartilage homeo-protein 1 (Cart1)	Exencephaly/Facial Defect	100	20	Folic acid
Crooked (Cd)	Exencephaly	20	15	Folic acid
Curly tail (ct)	Spina bifida	15	5	Inositol
Folate binding protein 1 (Folbp1)	Exencephaly	100	10	Folinic acid
Splotch (Sp, Pax3)	Exencephaly/Spina bifida	100	60	Folic acid or thymidine
SELH/Bc strain	Exencephaly	25	10	Purina #5001 (vs #5015)

Adapted from Juriloff 2000 [66].

1.3 Approaches for Identifying Genes Related to Human Neural Tube Defects

Linkage mapping can be used in complex diseases such as NTDs to identify underlying genes only with prior knowledge that the trait is inherited. Linkage is a powerful tool to help identify disease genes without having any prior information about the molecular mechanisms or biological properties of the disease. One of my dissertation projects involved the genetic basis for the neural tube defects and the linkage techniques described in this section focus on the linkage methods used for NTD family 8776 (Chapters 2 and 3).

1.3.1 Parametric Linkage Analysis

Linkage analysis is a traditional approach to mapping genes in families with more than one individual affected with a Mendelian disease. Linkage is a gene mapping tool used to evaluate co-segregation of a locus and a particular trait within families. Genetic linkage

measures the tendency for two or more genetic loci that are usually in close physical proximity on a chromosome to segregate together in a family. In 1955, Morton originally described likelihood-based methods for scoring recombinants and non-recombinants to determine linkage using a parametric or model-based approach. Using this method, the degree of linkage can be measured by the number of recombination events between two loci based on the total observed meioses, and with a recombination fraction designated as θ. Complete linkage occurs when there is no recombination at a $\theta = 0$, and a lesser degree of linkage between the loci when the number of recombinants has a $\theta < 0.5$. The evidence of no linkage is when the observed recombination fraction is 50% ($\theta = 0.5$). These recombination events take place during meioses when homologous chromosomes pair and crossover. The frequency by which the alleles segregate together is a factor of the distance between them along a chromosome. The proximity of alleles to each other can be described by genetic linkage which is measured using centiMorgan (cM), where 1 cM represents 1% recombination. The marker maps used for linkage analysis are derived from mathematical formulas to determine genetic map distance developed by Haldane in 1991 [115] and Kosambi in 1944 [116]. These genetic maps are based on the frequencies of recombination events between marker alleles. The more commonly used Kosambi map function accounts for interference, a phenomenon whereby crossovers of homologous chromosomes at a specific chromosomal location interferes with crossover events within the same location.

The LOD score serves as a test of the null hypothesis of no linkage versus the alternative hypothesis of linkage. The equation for the parametric LOD score is illustrated in Figure 1.5.

$$LOD = \log_{10} \frac{(1-\theta)^{NR}\,\theta^{R}}{(1/2)^{NR+R}} = \log_{10} \frac{(Likelihood\ \ for\ linkage)}{(Likelihood\ \ for\ no\ linkage)}$$

Figure 1.5 Logarithm of the Odds (LOD) score equation.
NR representing non-recombinant events and R indicating recombinant events.

To score recombinant and non-recombinant individuals within a pedigree, the genetic mode of inheritance is used to assign the disease locus based on individuals' affection status, and then phase of the marker and disease alleles can be determined. If phase is unknown within a pedigree and there is no a priori expectation to assign phase to a particular phase arrangement, then the LOD score will be estimated over the various possible phases. The ability to establish phase is often dependent on whether both parents are genotyped and if they have different heterozygous alleles at each marker. Linkage analysis using parametric linkage methods assumes a genetic model based on Mendelian inheritance patterns, such as specifying whether the disease is autosomal dominant, recessive, or an X-linked. Model-based linkage methods also require knowledge of disease allele frequency, marker allele frequency, and penetrance of trait of interest. Allele frequencies are commonly calculated from the existing dataset being tested for linkage; however, these frequencies may not reflect true population frequencies, particularly when allele frequencies are estimated using a small set of families. Misspecification of any of these parameters, especially the genetic model can decrease the power to detect linkage [117-120].

Full pedigree analysis includes all available phenotypic and genotypic data and also allows specification of a penetrance model. In complex diseases, like neural tube defects, often an affecteds-only or low-penetrance approach is employed. In this affecteds-only approach the phenotypic information for unaffected individuals is not used because they may

carry the disease locus but not exhibit the phenotype due to a myriad of potential causes such as incomplete penetrance of the gene, absence of another susceptibility gene and/or environmental exposure. Therefore, an affecteds-only linkage approach only uses the phenotypic information from those individuals who express the trait of interest. In this model, the genotypic data from the unaffected individuals can be used to infer missing genotypes and establish phase, although phenotypic data from unaffecteds will not directly contribute to the LOD score. This type of an analysis can be more powerful for detecting linkage when reduced penetrance or subtle non-specific phenotypic findings may lead to misdiagnosis. In addition, phenocopies -- when a phenotype mimics disease but has nongenetic etiology -- may also decrease the power to detect linkage, and may be accounted for by using liability classes based on estimated phenocopy rates.

Two-point linkage analysis compares one marker locus to the disease locus, and estimates the recombination fraction one marker at a time. Multipoint linkage analysis is an extension of two-point analysis in which linkage of a disease locus to a marker is tested over several markers at a time. In a single pedigree, the multipoint results will show co-segregation of many marker haplotypes with the trait of interest through the pedigree. Multipoint analysis can be more informative in situations when there is an uninformative marker or missing genotypic data that can be inferred from the haplotype structure of other family members. Multipoint LOD scores can be helpful for gene mapping when an affected individual is recombinant for the region of interest because this crossover event will provide a negative LOD score along the curve, decreasing the evidence in favor of linkage at that particular location.

Genomewide screens are used to identify novel loci for both Mendelian and complex diseases, and with any statistical test, there is the possibility for false positive results and thresholds for significance. Morton defined a LOD score of $\geq$ 3.0 (odds are 1000:1 that loci are linked at θ rather than unlinked) to be considered a "significant" linkage result, and a LOD score of 2 to 3 was "suggestive" of linkage, whereas a LOD score $\leq$ -2 provides evidence to exclude linkage. Scores between -2.0 and 3.0 are often considered inconclusive linkage findings suggesting that additional informative family members or families and/or more informative markers need to be included. Lander and Kruglyak later defined genomewide significance for linkage studies using more conservative estimates. Their new estimates require a parametric LOD score of 3.3 and nonparametric, allele sharing methods, to have a LOD score of 3.6. The underlying assumption of these significance thresholds is that there will be one false positive outcome in 20 genome scans for a type I error rate of 5%. These criteria for genomewide scan significance levels are shown in Table 1.2.

Table 1.2 Genomewide Linkage Significance

Mapping Method	LOD Score	Level of Significance
Parametric LOD score analysis	3.3	Significant
	1.9	Suggestive
Nonparametric Allele sharing method	3.6	Significant
	2.2	Suggestive

Estimates based on work by Lander and Kruglyak in 1995 [121]

1.3.2 Nonparametric Linkage Analysis

Many common diseases are heritable but do not have a classic Mendelian pattern of inheritance. These diseases are often complex with multiple genetic and environmental components involved, making it difficult to model the mode of inheritance appropriately. Nonparametric linkage analysis methodologies, also known as model-free approaches, do not require prior knowledge of the mode of inheritance or the disease model. Since misspecification of the mode of inheritance in parametric linkage methods can lead in a decrease in power to detect linkage, these model-free approaches are often used for complex diseases for which the mode of inheritance is unknown. Methods designed to map complex human diseases using a nonparametric approach with relative pairs will quantify the degree to which related individuals share alleles at the marker loci. These relative pair allele-sharing methods either evaluate alleles that are identical by state (IBS) or alleles that are identical by descent (IBD). Specifying alleles as IBS or IBD between relative pairs can be established through examining the inheritance of markers in families. Alleles that are IBD can be shown in pedigrees to be inherited from a common ancestor; in contrast, alleles that are IBS are the same alleles but cannot be conclusively proven to be inherited IBD. Nonparametric methods evaluate whether affected relatives share alleles identical by descent at a particular locus more frequently than expected by chance alone to localize the disease locus. One of the first nonparametric approaches was the affected sib pair (ASP) method that calculates the sharing of alleles identical by descent between sibling pairs [122,123]. Each relative pair will be expected to share a certain proportion of their alleles IBD. For example the Mendelian expectation is that siblings will share 0 alleles IBD with a probability of 0.25, 1 allele with probability 0.5, or 2 copies of an allele IBD at an expected frequency of 0.25. In ASP, if two

affected siblings share particular alleles IBD greater than expected by chance, then this finding suggests that the marker allele may be linked to the disease trait of the siblings. Although the ASP methods are useful and easy to calculate they did not account for other relative pairs within extended multiplex families. Later, an affected relative pair (ARP) method which maximizes information over all affected family members and uses the nonparametric linkage analysis (NPL) statistic was established [124,125] and implemented in the Genehunter, Allegro and Merlin linkage programs.

The NPL statistic calculates the inheritance vector by specifying the founder alleles inherited by each of the individuals that are non-founders in the pedigree. In some cases, the inheritance vector cannot be estimated perfectly since there may be missing genotypic data or there may be uninformative markers with low heterozygosities. To address the issue of inheritance uncertainty, the NPL statistic calculates a probability distribution that assigns a prior and posterior probability to estimate the probability of inheritance vectors. The inheritance distribution is calculated using the hidden Markov models (HMMs) that are implemented in the Lander-Green algorithm [124]. The prior probability of each inheritance vector is equally as likely without genotypic data; however, as genotype data is added the inheritance distribution selects more likely inheritance vectors. Once the inheritance vector is estimated, the program calculates linkage scores for both parametric and nonparametric methods. The NPL scores are calculated using the scoring functions S_{pairs} or $S_{all.}$ [126], which compare the degree of IBD sharing to expected IBD estimates for the relationship pair. S_{pairs} counts the degree of alleles shared IBD by each affected relative pair and compares this to the expected degree of IBD sharing. On the other hand, the S_{all} statistic evaluates permutations of the inheritance vectors of each allele to count the number of identical shared

permutations for the affected relative pairs, and this statistic gives increasing weight as the number of affected relative pairs share more alleles IBD. Generally speaking, S_{all} works best for additive and dominant diseases, whereas S_{pairs} is best for recessive diseases. The normalized S_{pairs} or S_{all} score is described as the Z score or the NPL. Of note, NPL is a conservative test that is more towards the null hypothesis of no linkage due to an overestimation of the variance from the assumption that all markers are completely informative. Also a likelihood-based measure of linkage is available, referred to as the LOD* statistic that incorporates the NPL score as part of its estimate and is considered comparable to a standard LOD score [127]. Unlike the traditional LOD score, the LOD* score is always a non-negative result and so may not be informative for exclusion of linkage.

Two commonly used computing algorithms for linkage analysis include the Elston-Stewart and Lander-Green algorithms. The Elston-Stewart algorithm, which was introduced in 1971 [128], does not require phase to be completely known for each pedigree, because this algorithm can estimate phase using maximum likelihood estimates; additionally, this algorithm allows for computational speed linear with an increasing number of individuals, but the algorithm becomes exponentially more time consuming with more markers. The Elson-Stewart Algorithm is implemented in such programs as Linkage, Mendel and Vitesse [129-132]. The Lander-Green algorithm [124] uses inheritance vectors (described in the nonparametric linkage methods section) in a family to estimate linkage. For example, at every marker, the observed marker phenotypes and genotypes are used to assign probabilities for each possible inheritance vector. The Lander-Green algorithm allows for computational speed linear with an increasing number of markers and is exponential with the number of individuals in the pedigree. With the advent of denser genotyping platforms, linkage

programs with the Lander-Green algorithm are more commonly used since the computational time is linear with increasingly more markers. The Lander-Green algorithm is implemented in Genehunter and Allegro linkage programs [125,133]. Allegro, which was developed after Genehunter, allows for more extended pedigrees (20-30% larger) with a bit size of 24 [133]. The bit size is calculated by $2n\text{-}f$ with n representing the number of non-founders and f is the number of founders in the pedigree. Often trimming of large extended pedigrees is performed in which the most informative individuals are retrained within the pedigree. For example, trimming was required for the linkage analysis conducted in the large multiplex NTD family 8776, and the individuals included in linkage analysis are described in Chapters 2 and 3.

1.3.3 Linkage Screen in 44 Multiplex NTD Families

In 2005, one of the first linkage screens of NTDs was reported in 44 multiplex families [134]. The linkage screen included 402 microsatellite markers spaced on average ~ 10 cM apart and the majority of the families were Caucasian (n=41). A total of 292 individuals were included in this genomic screen with a total of 89 sampled affected, of which 50 had lumbosacral myelomeningocele. The other affected individuals had other types of NTDs including anencephaly, cervical myelomeningocele, craniorachischisis, encephalocele, lipomyelomeningocele, rachischisis, and thoracic myelomeningocele. To control for potential genetic heterogeneity underlying different NTDs, the phenotype for linkage analysis was first narrowly defined as only those individuals with lumbosacral myelomeningocele, and also a broader phenotypic definition was used leading to inclusion of all individuals affected with an NTD. The study population was composed of 21 affected

sibling pairs, 12 affected avuncular pairs, and 35 other affected relative pairs. Multipoint parametric and nonparametric analyses using Allegro identified regions of interest on chromosomes 7 and 10. The highest LOD score of 2.45 was on chromosome 7 under the parametric model, and the linked region spanned a 24 cM interval between D7S3056 and D7S3051. Further investigation of the genomic screen data identified a single large family, 8776, as primarily driving the linkage results on chromosome 7; on its own generating a parametric LOD score of 2.40. This screen revealed that family 8776 clearly demonstrates a strong genetic effect, and additional gene mapping analyses in family 8776 are described in Chapters 2 and 3.

1.3.4 Approaches to Candidate Gene Selection in Human NTDs

After a region of interest has been defined by linkage studies, candidate genes can be prioritized for further evaluation. The human genome project has provided a rich sequence dataset for candidate gene selection. Several bioinformatics sites compile extensive information on gene annotation, expression profiles in various tissues, and protein structure and function; commonly used sources for human candidate gene evaluation studies include: UCSC Genome Bioinformatic Site, National Center for Biotechnology Information (NCBI), and Ensembl Genome Browser. These complementary bioinformatics websites can be used to compile a comprehensive list of genes residing in positional target regions.

Candidate gene studies can select genes either that map to regions identified by positional cloning methods or based solely on the gene's biological plausibility without any gene mapping information. Alternatively, candidate gene selection can also preferentially use many converging lines of evidence to choose a particular gene; this approach has been

previously termed “genomic convergence” [135-137]. For example, studying neural tube defects may use any of the following three converging lines of evidence for selection of candidate genes: 1) positional candidates from promising linkage or whole genome association analysis results; 2) known biological mechanisms that may affect neural tube closure such as genes involved in folate metabolism and early developmental neurobiology and genes from NTD mouse models; and 3) gene expressed in tissue of interest. The genomic convergence model is illustrated in Figure 1.6; this approach was applied for candidate gene selection in family 8776 as described in Chapter 3.

Once a candidate gene is selected, a way to interrogate the gene may include use of known variants (such as SNPs) and genotype using either Taqman SNP allelic discrimination methods or SNP arrays designed by such biotechnology companies as Illumina and Affymetrix that can be used for association analysis. To capture additional variation in the population of interest, resequencing may be done to identify novel variants that may reveal functional variants and/or other variants that may also be used for association analysis or tested for functional characteristics.

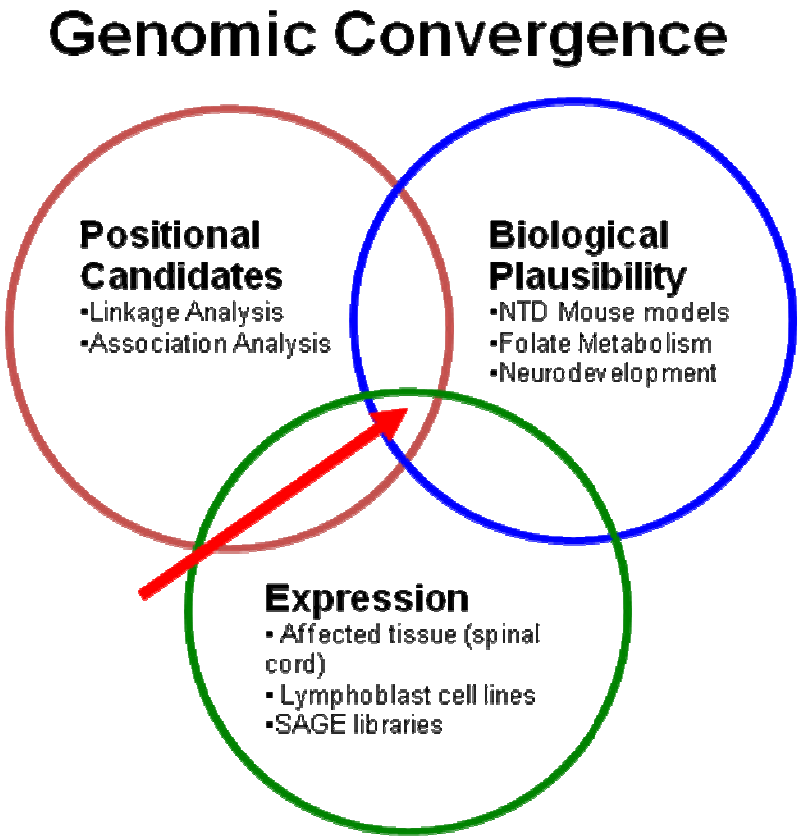


Figure 1.6 Genomic Convergence

Combination of multiple sources of information for candidate gene selection. The genes of highest priority to follow-up are those that fall into the region defined by the intersecting lines of evidence as illustrated by the red arrow; NTD candidate gene selection is used as an example.

1.3.5 The Future for Gene Mapping Studies in Complex Disease

Linkage analysis studies depend on the number of informative meisoses from families to estimate recombination. The informativeness of the marker set used in whole genome linkage analysis will be based on the heterozygosity of the genetic markers and the spacing of the markers across the genome; however, even dense marker sets will still lead to large intervals of interest spanning ~10-20 Mb. In humans, single nucleotide polymorphisms are the most common type of genetic variation and several million SNPs have been documented in the dbSNP database [138]. The development of efficient high-throughput SNP platforms

have catapulted the world of human genetics into a whole genome association analysis paradigm, where ideally the disease loci and/or variants can be mapped at a finer scale than linkage methods. In addition, these dense oligonucleotide microarrays, along with array-CGH, have identified copy number variants in the presumed normal population [139-144]. These recently identified known copy number variants (CNVs) leading to amplification or deletion of portions of chromosomes may confer risk to common complex disease. As each year passes, the density of arrays increases, now dense 500K or 1M SNP chip arrays not only allow the potential for whole genome association analysis but also allow for fine resolution detection for either known or novel copy number variants. However, many of the characteristics of the CNVs in the human population are currently unknown including the total number, expected frequencies for particular populations, genomic content, gene content, and patterns of linkage disequilibrium [145]. Perhaps in the future CNVs will be informative for the study of NTDs, and Chapter 6 comments on the feasibility and rationale behind WGA in the study of NTDs.

1.4 Native American Myopathy: Defining the Phenotype and Genetic Analysis

Native American Myopathy (NAM) is a congenital myopathy [MIM 255995] first described in the Lumbee American Indian population a multiracial group indigenous to Robeson County, south-central North Carolina [146]. NAM is a congenital myopathy characterized by cleft palate and a susceptibility to malignant hyperthermia provoked by anesthesia. NAM likely segregates in an autosomal recessive manner. The Lumbee people are a relatively isolated population with a high rate of consanguinity suggesting that autosomal recessive diseases in this population may be inherited through a common ancestor.

Despite the high mortality and morbidity seen in NAM, little is known about the clinical presentation or the genetic mutation causing the disease.

1.4.1 The Lumbee Population

The Lumbee people are a Native American Indian tribe from south-central North Carolina. The "Lumbee" name is derived from the Lumber River that passes through Robeson County, the region in which the Lumbee population is concentrated. Currently, the Lumbee Tribe of North Carolina is estimated to have more than 45,000 members making them the largest tribe east of the Mississippi River [147]. The Lumbee population is considered a relatively isolated population with three different ethnic backgrounds including a mixture of American Indian, European, and African American ancestry [148].

The origin of the admixed Lumbee population has many theories. One commonly repeated theory, proposed by Hamilton McMillian in 1885, is that the Lumbee people are descendents of Sir Walter Raleigh's "Lost Colony" who mixed with the Croatan Indians, an Algonquin-speaking tribe from the North Carolina coast [149]. The "Lost Colony" saga commenced in 1587, when under the authority of Sir Walter Raleigh, John White and 120 men and women embarked on expedition to establish a colony on Roanoke Island [148]. Once the colony was established, John White returned to England for additional supplies, and upon his return in 1591 the colony had disappeared. It is unclear what happened to the colony, they may have died of starvation or disease or perhaps intermarried with the Croatan Indians. Interestingly, in the 1730's when the first English and Scotch settlers arrived in Robeson County, they described encounters with individuals of mixed race and speaking an English dialect [148]. Two other theories of the Indian origin of the Lumbee people exist.

One is that the Indian ancestry is derived from the Cherokee, a powerful Iroquois-speaking tribe who dominated the Western Carolinas prior to European settlement, and the other theory is the Indian ancestry may be from the Catawaba, a Siouan-speaking tribe, from the Piedmont Carolinas [150].

In the 1960's, William Pollitzer, an anthropologist at the University of North Carolina at Chapel Hill, studied the origins of the Lumbee people. He described the Lumbee population as a "tri-racial isolate" that dates back as far as the American Revolution [151]. He further reported: "Considerable phenotypic variation is found within the isolate today, with extremes of skin color from light to dark and of hair from very curly to straight. The morphology of the face also suggests broad racial backgrounds" [150]. In order to estimate the degree of admixture in the "tri-racial isolate", Pollitzer investigated the type of blood factors and hemoglobin within 1273 individuals in the Lumbee population. He estimated, based on three loci (ABO genes, Rh, and hemoglobin pattern), that the racial composition of the Lumbee was ~34% Caucasian, ~53% African American, and ~13% American Indian. These estimates were suggestive of a small amount of American Indian ancestry, and not surprisingly, the Lumbee tribe, who identify primarily with their Native American ancestry, did not appreciate the implications of this study (Cynthia Powell, personal observation).

Whatever the original pattern of ancestry, the Lumbee people are now a relatively isolated population [148]. Indeed, Pollitzer speculated that the large number of individuals with the same last name suggests a pattern of historical inbreeding. For example, Pollitzer described the marriage records from 1800 to 1860's of 132 men who had 12 surnames common to the current day tribe, and of these 132 men, 23 married women with the same surname. More current observations of common surnames and consanguinity within families

also support Pollitzer's previous studies (Cynthia Powell, personal observation). Pollitzer believed the Lumbee's social isolation may have resulted from members of the population not wanting to marry outside of their tribe [151]. This social isolation was probably heightened by low socioeconomic status within this group limiting their ability to migrate to other regions and to marry outside of their group.

The Lumbee people have a very active tribal community evidenced by many tribal organizations, community events celebrating their Native American Indian ancestry, and Lumbee websites dedicated to preserving their history and describing their struggle for federal recognition as a bona fide Native American tribe [152]. The Lumbee people, who identify strongly with their Native American ancestry, have been attempting to obtain federal recognition as an Indian tribe to request benefits since 1888. In 1885, the State of North Carolina first recognized the population to be Croatan Indians (the former name for the Lumbees). This state recognition allowed the Croatan Indian to petition for a school system for the exclusive use of tribal members. Much later, in 1956, the United States House of Representatives passed the Lumbee Act, that recognized the Lumbee people as a Native American tribe. However, the Lumbee Act did not provide full federal recognition as an American Indian tribe, thus preventing the tribe from full tribal benefits such as for health care and housing from the Federal Government. In 1987, the Lumbees petitioned the U.S. Department of the Interior for federal recognition to gain federal monetary benefits reserved for American Indian tribes, and this petition was denied.

1.4.2 Clinical Description of Native American Myopathy

In 1987, Bailey and Bloch first described Native American Myopathy in a 3-month old of Lumbee descent with multiple congenital anomalies and a malignant hyperthermia episode under general anesthesia for placement of a gastrostomy tube [153]. In 1988, Stewart et al described six Lumbee children affected with congenital myopathy, kyphoscoliosis, and cleft palate. Two of these children had documented malignant hyperthermia during cleft palate repair [146]. The 6 children ranged in age from 5 to 13 years old and two were siblings, and one of the patients was a cousin to this sib-pair. All children had normal baseline creatine kinase levels and three of the six muscle biopsies had nonspecific myopathic changes. The one male enrolled in the study had undescended testes (cryptorchidism). The typical clinical presentation of NAM (abnormal facies with ptosis and low-set ears, muscle wasting and kyphoscoliosis) is illustrated in Figure 1.7. Overall, clinical features described for NAM include short stature, cleft palate, high arching palate, myopathic facies (malar hypoplasia, ptosis), scoliosis, congenital joint contractures, clubbed feet, normal baseline creatine kinase,and susceptibility to malignant hyperthermia [146,153-156]. These descriptions highlight both shared NAM characteristics as well as the phenotypic heterogeneity in the presentation of NAM. The phenotypic heterogeneity seen in NAM patients may be explained by several different mechanisms including different environment exposures, pleiotropic effects of the NAM gene, epigenetic influences on disease progression and expression, and/or potential misdiagnosis. Chapter 4 provides a more in depth description of the clinical presentation of 14 patients with NAM.

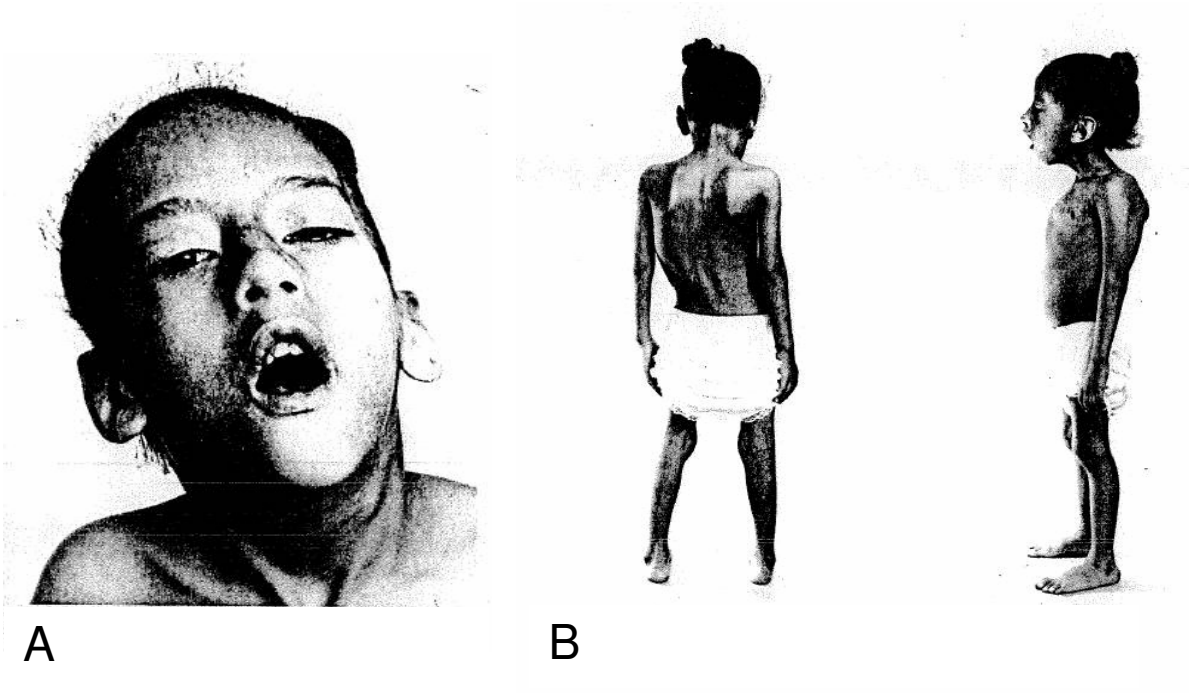


Figure 1.7 Dysmorphic facies and skeletal abnormalities of NAM.

Panel A: Dysmorphic facies with low-set ears and ptosis. Panel B: Kyphoscoliosis and cachexia with muscle wasting. Reprinted from Pediatric Neurology, 4(6), Stewart CR, Kahler SG, Gilchrist JM, Congenital Myopathy with Cleft Palate and Increased Susceptibility to Malignant Hyperthermia: King Syndrome?, 371-374., Copyright (1988), with permission from Elsevier.

1.4.3 Malignant Hyperthermia Susceptibility in the Lumbee Population

Susceptibility to malignant hyperthermia is common in Lumbee people even in the absence of congenital myopathy [154]. The prevalence of MH in the Lumbee population is likely to be much higher than ~1:30,000 observed in the general population [157]. Because of the known MH susceptibility risk within the Lumbee people, clinicians in the Lumberton region and surrounding areas will often use non-halogenated non-volatile anesthetics in patients of Lumbee descent. Several mechanisms may be causing MH to be more frequent in

the Lumbee people; for instance, the presence of MH without congenital myopathy may be associated with individuals that are heterozygous for the NAM genetic mutation whereas individual with all the features of NAM may be homozygous for the genetic mutation. Another possibility is that some individuals homozygous for the NAM genetic mutation may have reduced penetrance leading to a milder phenotype such as MHS only. Alternatively, another genetic mutation may be segregating though the population that confers predisposition to MH only. Particular gene-gene and gene-environment interactions may also determine whether NAM and/or MHS will be expressed.

1.4.4 Clinical Complexity of NAM

The phenotypic presentation of NAM is not well characterized evidenced by a lack of clinical information in the literature. Dissecting the unique clinical characteristics of NAM has been complicated by the phenotypic heterogeneity in the presentation of NAM and by King-Denborough Syndrome (KDS) and Noonan Syndrome (NS) that have similar phenotypic characteristics to NAM. A key distinction is that both KDS and NS tend to be inherited as autosomal dominant disorders whereas NAM appears to be inherited in an autosomal recessive manner. KDS is a disorder that appears markedly similar to NAM, with both features of myopathy and susceptibility to malignant hyperthermia. However, we believe NAM to be clinically distinct from KDS because patients with KDS have increased baseline CK levels and they lack cleft palate, contractures, and clubbed feet. In addition, Noonan syndrome affected individuals have short stature and dysmorphic features that resemble NAM. Such clinical resemblance led one patient with NAM to be initially diagnosed with Noonan Syndrome (S. Kahler, personal communication). However, NS can be clinically distinguished from NAM, because patients with NS tend to have structural

cardiac defects and they do not have congenital myopathy, cleft palate or congenital joint contractures as seen in NAM. In addition, Froster-Iskenius and colleagues described an autosomal recessive form of congenital contractures with similar features to NAM including myopathic face, arthrogryposis, ptosis, cleft palate, and susceptibility to malignant hyperthermia, but these patients also had congenital torticollis and neck and axillary webbing, which are not observed in NAM [158]. Thus, although these other diseases may resemble NAM, the differences in phenotype suggest NAM to be clinically distinct as shown in Table 1.3.

Table 1.3 Differential diagnosis for NAM

Disease	Inheritance pattern	Disease Locus	Gene	Traits common to NAM	Dissimilar to NAM
King-Denborough Syndrome	majority autosomal dominant; Some cases autosomal recessive	19q13.1	*RYR1*	congenital myopathy, short stature, scoliosis, cryptorchidism, myopathic facies, MH	below average intelligence, high CK; No cleft palate or arthrogryposis
Noonan Syndrome	Autosomal dominant	12q24.1	*PTPN11*	short stature, dysmorphic facies with ptosis, feeding difficulties as a newborn	cardiac defects; No MH, cleft palate or arthrogryposis
Froster-Iskenius	Autosomal recessive	Unknown	Unknown	myopathic face, arthrogryposis, ptosis, cleft palate, MH	congenital torticollis, neck and axillary webbing
Gordon Syndrome	Autosomal dominant	Unknown	Unknown	short stature, cleft palate, equanovarus deformities, camptodactyly	pterygium colli, and vertebral anomalies; No MH

This a partial list of the diseases and known loci presenting with similar features to NAM.

1.4.5 Clinical Manifestations of Malignant Hyperthermia

Malignant hyperthermia (MH [MIM 145600]) is a potentially life-threatening pharmacogenetic disorder of the skeletal muscle that is triggered in genetically susceptible individuals exposed to inhaled anesthetics and/or depolarizing muscle relaxants. MH crisis is a hypermetabolic response to potent inhaled anesthetic agents that may result in any of the following symptoms: tachypnea, tachycardia, unstable arterial blood pressure, increased carbon dioxide production, increased oxygen consumption, acidosis, muscle rigidity,

rhabdomyolysis, hyperthermia, elevated creatine kinase and hyperkalemia [159]. If MH crisis is not reversed, severe tissue damage and death may result. Treatment of MH involves discontinuing the eliciting anesthetic agent, hyperventilation with 100% oxygen, and administration of dantrolene sodium, a muscle relaxant acting to depress excitation-contraction coupling of skeletal muscle [160,161]. Dantrolene is the most effective pharmacologic agent used to halt the progression of MH [160,161]. The use of dantrolene as a main treatment intervention of MH has decreased the risk of mortality from 80% in the 1960s to less than 10% of cases today [162]. MH susceptibility in patients can be determined by applying the standardized European *in vitro* contracture test (IVCT) that evaluates a muscle biopsy specimen for the degree of muscle tension in response to halothane and caffeine [163]. However, the IVCT is invasive, requiring skeletal muscle biopsy and consequently is not widely available or performed [164].

1.4.6 Malignant Hyperthermia Susceptibility Loci and Associated Genes

Predisposition to malignant hyperthermia is commonly inherited as an autosomal dominant trait and has been reported to be genetically heterogeneous with six different malignant hyperthermia susceptibility (MHS) loci identified: MHS1-6 (Shown in Table 1.4). Two MHS genes have been identified, including the ryanodine receptor type 1 (*RYR1*) , which is a calcium release channel of skeletal muscle sarcoplasmic reticulum; *CACNA1S,* which encodes the α1-subunit of the human skeletal muscle dihydropyridine-sensitive L-type voltage-dependent calcium channel (VDCC) (Table 1.4). The *SCN4A* gene which encodes the sodium channel voltage-gated type IV alpha is a potential MHS gene but only linkage to this gene has been observed thus far [165,166]. MHS1 [MIM 145600] maps to 19q13.1 and

may be caused predominantly by dominant mutations in the *RYR1* gene [167-169]. King-Denborough Syndrome is also included as part of MHS1. MHS2 [MIM 154275] maps to 17q11.2 - q24 [170], and two studies provided marginal evidence for linkage of MHS2 to the *SCN4A* gene but no genetic mutation has been identified [165,171]. MHS3 [MIM154276] maps to 7q21-22 near the *CACNA2D1* gene and evaluation of coding regions in affected patients did not identify any pathogenic mutations [172,173], MHS4 [MIM 600467] maps to 3q13.1 [174], and MHS6 [MIM 601888] maps to 5p [175]. MHS5 [MIM 601887] was shown in three European families to map to 1q32 locus [175], and a 3333G-A transition in the *CACNA1S* gene, resulting in an Arg1086His amino acid substitution in the protein, was identified in a single large French family [176]. Later, mutations in the *CACNA1S* gene were also observed in a large North American family [177]. However, Stewart and colleagues observed that only 1% of the North American MHS population of 112 people had the Arg1086 His mutation [177]. Mutations in *CACNA1S* gene are associated with either malignant hyperthermia susceptibility or hypokalemic periodic paralysis; however, the *CACNA1S* mutation for MHS is localized to a different part of the alpha-1-subunit of the skeletal muscle voltage-dependent calcium channel compared with mutations for hypokalemic periodic paralysis [176]. Notably, the *CACNA1S* and *CACNA2D1* genes encode subunits of the dihydropyridine receptor which is functionally coupled to the ryanodine receptor. Overall, the *CACNA1S* gene is considered to be rarely associated with MHS and instead the majority of mutations identified in families with malignant hyperthermia have been in the *RYR1* gene. Although the majority of *RYR1* mutations are autosomal dominant, a few autosomal recessive mutations have recently been identified [169]. Given the allelic and locus heterogeneity of MH, susceptibility to malignant

hyperthermia may be dependent upon the effects of more than one gene. For example, a small association analysis study (77 UK nuclear families) of several MH candidate loci suggested that the MH phenotype is influenced by *RYR1* as the major locus and also by additional modifier genes [178].

Table 1.4 Malignant hyperthermia susceptibility loci

MHS locus	Location	Associated Gene	Mutation
MHS1	19q13.1	*RYR1*	Yes, multiple
MHS 2	17q11.2-q24	*SCN4A*	No
MHS 3	7q21-22	*CACNA2D1*	No
MHS 4	3q13.1	-	-
MHS 5	1q32	*CACNA1S*	Yes, rare
MHS 6	5p	-	-

1.4.7 *RYR1* Mutations and Associated MH and/or Congenital Myopathies

The mechanism of MH involves increased intracellular calcium ions in skeletal muscle resulting in protracted muscular contraction and hypermetabolism [179]. Porcine animal models susceptible to MH have been useful for characterizing the genetics of MH; these animal models exhibit an autosomal recessive form of disease and sequencing of the *RYR1* gene in these pigs revealed a missense mutation (Gly614Arg) [180]. Identifying this mutation in pigs stimulated a search for *RYR1* mutations in MH-susceptible human families, and both missense mutations and deletions [181] have also been reported in humans affected with MH. MH-susceptible porcine breeds also demonstrated that mutated RYR1 receptors allowed for greater efflux of calcium ions [182].

Susceptibility to MH appears to be from abnormal calcium ion balance in skeletal muscle, and many cases have been shown to result from mutations in the *RYR1* gene product in the sarcoplasmic reticulum. The skeletal muscle ryanodine receptor (*RYR1*) gene comprises 106 exons and encodes a protein of 5,038 amino acids. The RYR1 protein functions as the main calcium opening channel of the sarcoplasmic reticulum involved in excitation-contraction coupling in skeletal muscle, whereby the opening of RYR1 leads to an efflux of calcium ions into the myoplasm. The released calcium ions activate muscle contraction by attenuation of troponin C inhibition of the contractile proteins myosin and actin.

Congenital myopathies, sometimes associated with malignant hyperthermia, are a phenotypically heterogenous group of neuromuscular disorders that are characterized by specific histopathologic findings in skeletal muscle from muscle biopsy. The key histological feature seen in skeletal muscle samples has been used in the naming of these myopathies, which include: central core disease, nemaline myopathy, multiminicore disease, and centronuclear myopathy. These congenital myopathies overlap in phenotypic, genotypic, and histopathologic characteristics [183] suggesting that different forms of the same myopathic disease process may exist. For example, malignant hyperthermia susceptibility (and King-Denborough disease), central core disease, and multi-minicore disease have all been associated with *RYR1* mutations [184]. The majority of *RYR1* mutations in CCD and MHS cause the RYR channel to be leaky leading to depletion of calcium stores in the sarcoplasmic reticulum and elevated levels of calcium in the cytosol [185]. Over 100 mutations in *RYR1* have been described [157] with only 9 of these reported in multimincore disease, and the majority of the *RYR1* mutations are present in MHS and CDC. General patterns in the

phenotypic effects of *RYR1* mutations are becoming more clear; for instance, dominant *RYR1* mutations disrupting either the cytoplasmic -terminal or the central domains of the RYR1 protein predominantly lead to a MHS phenotype, while dominant mutations in the C-terminal region of the RYR1 protein cause the CDC phenotype [169]. Even though more *RYR1* mutations have been elucidated and associations between these mutations and the development of congenital myopathies and/or malignant hyperthermia have been well documented, the exact molecular mechanism underlying the various associated phenotypes still remains undetermined.

1.4.8 Homozygosity Mapping Approach to Identify the NAM Disease Locus

In 1987, Lander and Bostein proposed using a homozygosity mapping approach to identify rare sharing of homozygous regions for recessive diseases born from consanguineous matings [186]. Homozygosity mapping identifies genomic segments shared identical-by-descent inherited from a common ancestor (also referred to as autozygosity) by patients with recessive diseases from populations with limited migration into or out from the population and/or with high rates of consanguinity, allowing characterization of a genomic region that harbors the causative genetic mutation [186,187]. Moreover, homozygosity mapping is a form of linkage disequilibrium (LD) mapping whereby the disease gene is in LD with a conserved haplotype that arises from founder effects [187]. However, these conserved homoyzygous haplotypes can be large segments complicating gene discovery, but recombination events over several years within the population may lead to smaller segments of the original founder chromosome facilitating more productive homozygosity mapping studies. While individuals in the general population can share small regions of homozygosity

that are either identical by state (IBS) or identical by descent (IBD), inbred individuals are more likely to share longer stretches of homozygous alleles inherited IBD from a common ancestral haplotype. However, contrary to this assumption, it has recently been shown that presumable outbred populations can also have long stretches of homozygous segments that are considered ancestral haplotypes with alleles inherited identical-by-descent [144,188-190].

Homozygosity mapping is typically performed using linkage analysis methods in multiplex consanguineous families and has facilitated the identification of causal genetic variants of recessive diseases from consanguineous populations [191-193]. Similarly, the genetic basis of NAM could be investigated by utilizing homozygous mapping methods based on the observation that Lumbee people are a relatively isolated population with a high degree of consanguinity (Cynthia Powell, personal observation). Our hypothesis underlying the genetic analysis of NAM was that affected individuals with NAM would share a region of homozygosity flanking the NAM disease gene that was inherited IBD from a common ancestor. Genetic mapping results using homozygosity mapping methods to map the NAM disease locus are described in Chapter 5.

1.4.9 Potential Pitfalls in Homozygosity Mapping

In a highly inbred population like the Lumbee population, one may observe several autozygous chromososmal regions, such that identifying the true disease-associated haplotype may be difficult. Moreover, even though homozygosity mapping may be a powerful tool to map novel loci in rare autosomal diseases, the actual disease locus may not be detected if any degree of allelic heterogeneity is present within the population. Exemplifying this potential issue, two studies of rare autosomal recessive disease in isolated

consanguineous populations identified different mutations in the same disease gene, such that a few individuals were either heterozygotes for a mutation or compound heterozygotes [194-196]. This observed allelic heterogeneity within consanguineous kindreds then obscured the actual disease locus due to loss of homozygous markers flanking the disease mutation [194-196]. Additionally, when using linkage methods for homozygosity detection, the presence of hidden consanguinity in pedigree founders may inflate the likelihood estimate, which may lead to identifying spurious homozygous segments [195]. These potential issues in the genetic analysis of NAM are further addressed in Chapter 6.

CHAPTER 2

SNP screen in a large multiplex neural tube defect family refines linkage to loci at 7p21-pter and 2q33.1-35

The following material was published as a manuscript in the *Birth Defects Research Part A: Clinical and Molecular Teratology* under the following citation:

Stamm DS, Rampersaud E, Slifer SH, Mehltretter L, Siegel DG, Xie J, Hu-Lince D, Craig DW, Stephan DA, George TM, Gilbert JR, Speer MC. High-density single nucleotide polymorphism screen in a large multiplex neural tube defect family refines linkage to loci at 7p21.1-pter and 2q33.1-q35. Birth Defects Res A Clin Mol Teratol 2006; 76(6):499-505. Copyright © (2006 and Wiley-Liss Inc.)

This material is copyrighted by and reproduced with the permission of Wiley-Liss Inc. a subsidiary of John Wiley & Sons, Inc.

2.1 Abstract

Neural tube defects (NTDs) are considered complex with both genetic and environmental factors implicated. To date, no major causative genes have been identified in humans despite several investigations. The first genomewide screen in NTDs (Rampersaud et al. 2005) demonstrated evidence of linkage to chromosomes 7 and 10. This screen included forty-four multiplex families and consisted of 402 microsatellite markers spaced approximately 10 cM apart. Further investigation of the genomic screen data identified a single large multiplex family, pedigree 8776, as primarily driving the linkage results on chromosome 7. To investigate this family more thoroughly, a high-density single nucleotide polymorphism (SNP) screen was performed. Two-point and multipoint linkage analyses were performed using both parametric and nonparametric methods. For both the microsatellite and SNP markers, linkage analysis suggested the involvement of a locus or loci proximal to the telomeric regions of chromosomes 2q and 7p, with both regions having nonparametric LOD* scores of ~3.0, yielding very similar evidence in favor of linkage. In addition to mutations and/or variants in a major gene, these loci may harbor a microdeletion and/or translocation; potentially, polygenic factors may also be involved. This single family may be promising for narrowing the search for NTD susceptibility genes.

2.2 Introduction

2.2.1 Neural Tube Defects: Clinical Impact

Birth defects are the leading cause of death for neonates in the United States; the risk of an infant being born with significant congenital abnormality in body structure or function

is approximately 3-4%. Neural tube defects (NTDs) occur at a rate of 1 per 1000 live births in the United States [1,197], representing both the most common congenital malformation of the central nervous system and the second most common type of birth defect. Neural tube closure occurs during the first three to four weeks of development. NTDs are due to failure of closure of the neural tube during early embryogenesis, typically before a woman is aware of her pregnancy. NTDs can present clinically with large phenotypic variation. The most common presentations are spina bifida (spina bifida cystica, open spina bifida) and anencephaly. These phenotypes represent the classic definition of NTDs and whether they are caused by the same underlying genetic basis remains controversial.

Neural tube defects can frequently be disabling. Children affected with spina bifida often undergo surgery to close the defect within 48 hours after birth, and may require frequent surgical intervention for complications such as hydrocephalus. Affected patients commonly have an impaired ability to walk and often need to use a wheelchair, and they may have little or no bowel and/or bladder control. Severely affected NTD cases may require aggressive surgical, medical, and rehabilitative care. Thus, for a child and their family affected with an NTD, the psychosocial and emotional tolls are immeasurable. These significant medical and personal costs have lead to several collaborative studies to investigate NTDs to better understand the genetic and environmental etiologies.

2.2.2 Evidence for genetic and environmental components

NTDs are caused by a complex interaction between genetic and environmental risk factors. Several exogenous risk factors are correlated with the development of NTDs, including maternal insulin-dependent diabetes mellitus, obesity, maternal use of anti-

epileptic (valproic acid), and and maternal consumption of the mycotoxin fumonisin [23-29]. An important environmental intervention for NTDs is periconceptional supplementation with folic acid, which has been shown to reduce the recurrence risk for NTDs by reducing the recurrence risk for NTDs by 50-70% [43,85,198]. However, folate supplementation has not entirely eliminated the risk of NTDs, suggesting that there are additional underlying genetic factors that may contribute to the development of NTDs.

Several lines of evidence support a genetic component to NTDs [48]. The most compelling evidence for a genetic factor is the higher recurrence risk for NTDs among offspring from parents with previous NTD pregnancies [62,63]. Notably, the NTD recurrence risk for siblings is 2-5%, representing a 25-50 fold increased risk above that observed in the general population [199]. NTDs are also associated with established genetic syndromes including Meckel-Gruber syndrome, anterior sacral meningomyelocele, anal stenosis, trisomies 13 and 18, and other chromosomal aberrations [18,19,22]. Furthermore, a plethora of mouse models with mutations leading to NTDs have been identified [200]. Overall, these observations support a genetic component to the development of NTDs.

2.2.3 Initial genomic screen results

Results from the first genome screen of NTDs were recently reported [201]. This screen included 44 multiplex families and considered 402 microsatellite markers with approximate 10 cM spacing. The genomic screen identified regions of interest on chromosomes 7 and 10, with the highest LOD score of 2.45 on chromosome 7 using parametric linkage analysis. Further investigations of the genomic screen data identified a single large family 8776, a multigenerational Caucasian family with lumbosacral

myelomeningocele segregating, as primarily driving the linkage results on chromosome 7. Additionally, this family revealed a region on chromosome 2 with evidence in favor of linkage (unpublished data).

In this study of family 8776, the microsatellite genomic screen was followed up with a whole-genome high-density SNP screen to confirm the previously identified region on chromosomes 7 and 2 and narrow the minimum candidate interval for each. In addition, the SNP screen could identify previously unidentified regions of linkage. The utility of SNPs over microsatellites for genomewide screens has been validated, with studies demonstrating increases in information content with SNP screens being dependent on family structure, previous marker density, and other factors [202-205]. However, this increased density comes at a cost: simulation-based studies have lead to concern regarding inter-marker linkage disequilibrium (LD) in dense SNP-genomic screens. These simulations demonstrated that dense SNP data can cause false positive elevation of multipoint LOD scores, especially in the presence of missing parental genotypes [206,207].

2.3 Materials and Methods

2.3.1 Clinical data collection and power analysis

NTD family 8776 was identified by self-referral through the NTD Collaborative Study. First-degree relatives and relatives connecting related affected individuals were ascertained in this extended family. A detailed family history [208] was obtained and medical records including operative reports and pre-surgical x-ray films were collected for review of diagnosis by a neurosurgeon (TMG). Of the affected individuals in the pedigree,

three were diagnosed with lumbosacral myelomeningocele and a fourth individual was diagnosed with congenital dermoid cyst with tethered cord.

Certified phlebotomists obtained blood samples from affected individuals and related family members at medical centers, clinics, or by visiting participants' homes. Some participants sent in mailer kits with their enclosed samples. This study was conducted under the oversight of Duke University Medical Center Institutional Review Board and informed consent was obtained from all participants.

Power analysis of family 8776, under the broad phenotypic classification, was performed for the microsatellites screen by using the program SIMLINK [209,210]. The assumptions for the power studies included an autosomal dominant inheritance pattern with low penetrance (affecteds-only) model, a disease allele frequency of 0.001, and a marker with 75% heterozygosity linked at 5% recombination with the disease allele. These power studies generated an average parameric LOD score for family 8776 at a recombination fraction $\theta = 0.05$ of 1.063 ± 0.02 and a maximum LOD score of 2.515. These power estimations are based on two-point parametric analysis with a specified dominant mode of inheritance; therefore this analysis may underestimate power to detect linkage when using a high density SNP approach.

2.3.2 Genotyping methods

Blood samples were collected on four affected individuals and 27 unaffected individuals in family 8776. DNA was extracted from whole blood using the Puregene system (Gentra Systems, Minneapolis, MN). Whole genome SNP typing was performed using the Affymetrix 10K SNP Chip (Affymetrix, Santa Clara, CA). In brief, the 10K SNP

Chip requires only 250 ng of genomic DNA and yields genotypes for approximately 10,644 SNPs from the SNP Consortium database with an average heterozygosity of 0.37, and covering on average approximately one SNP every 210 kb. Microarrays were scanned by using the GeneChip Scanner 3000 according to the manufacturer's protocol (Affymetrix). Data acquisition was performed by using the GeneChip GCOS software.

2.3.3 Whole genome linkage analyses

The integrity of the family's pedigree structure was confirmed as previously reported [211] using data generated from the genomewide microsatellite screen with the program RELPAIR [212,213].

All marker information from the genomic screen was databased using the PEDIGENE system [214]. Data were evaluated for Mendelian inconsistencies using PedCheck [215]. Markers were retained in the map files for linkage analysis if they demonstrated at least 85% efficiency. Allele frequencies for the SNP markers were estimated based on 126 unrelated Caucasian individuals that were genotyped on the 10K SNP chip. Map order for the SNP markers was generated from sequence data provided by Translational Genomics (Phoenix, AZ). To account for potential inter-marker LD in the dense SNP screen, LD Select was used [216] with pairwise $r^2 > 0.16$ (Boyles et al. 2005) to generate map files that contained only tagging SNPs.

Linkage analysis was performed using the program Allegro [133] for both two-point and multipoint analysis. Because the exact inheritance pattern of NTDs is unknown, different analytic approaches were applied to these data, including parametric (assuming dominant inheritance and a disease allele frequency of 0.001) and nonparametric analysis.

For the nonparametric analysis, an identity-by-descent relative pair sharing method LOD* was assessed between all pairs of affected individuals using S_{pairs} [126] and the exponential model [127] as implemented in Allegro. Since the unaffected individuals in the pedigree 8776 could represent asymptomatic gene carriers or have undiagnosed NTDs, a low penetrance affecteds-only parametric model was applied such that unaffected individuals contributed no information to the LOD score other than to infer missing genotypic data on unsampled individuals. Size limitations in Allegro's computational capability allowed inclusion of only 24 individuals from family 8776 in the linkage analyses, so the least informative individuals were eliminated from the analysis. Phenotypic criteria included both broad (all 6 affected individuals) and narrow (only 5 affected individuals with lumbosacral myelomeningocele), with the narrow phenotypic definition excluding individual 100, who has a congenital dermoid cyst. The two affected individuals 102 and 1075 who were not sampled were included in the analyses, because the genotypic information on their siblings and/or parents using maximum-likelihood estimation techniques can help to infer the affected individual's missing genotypes.

To confirm segregation of a disease-linked haplotypes within regions of interest, haplotype analysis was performed using both visual inspection and a statistical approach with Simwalk [217].

2.3.4 Sequencing of Candidate Genes

Several genes mapping to the regions of interest were identified using an internally developed bioinformatics tool Genominator2 that operates via Ensembl [218]. Primers were designed using Primer 3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) on

genomic sequence from UCSC's database (http://genome.ucsc.edu/). Primers were designed to flank the exons and include the splicing region to account for splice variants. PCR products covered all the coding exons and respective intronic boundaries. PCR product was purified using Sephadex G-50 fine (Amersham Biosciences, Piscataway, NJ) and QuickStep2 SOPE Resin (Edge BioSystems, Gaithersburg, MD). Samples were sequenced using the BigDye Terminator v3.1 Cycle Sequencing Ready Reaction kit (Applied Biosystems, Foster City, CA) and purified using Sephadex G-50 fine. Both forward and reverse strands were sequenced. The sequence data was analyzed to evaluate for putative polymorphisms using ABI 3100 Data Collection Software Version 1.01 and ABI Sequencing Analysis 3.7. All sequences for candidate genes *MEOX-2* and *FZD7* were visually inspected with this software to evaluate noteworthy variants such as a premature termination codon, a frameshift mutation(s) or a missense mutation(s). The sequence data were compared to unaffected parents to ascertain if a variant was transmitted or a de novo mutation.

2.4 Results

2.4.1 Linkage Results for Family 8776

The individuals included in the linkage analyses using Allegro are indicated in Figure 2.1. The regions of interest were defined by a one LOD unit drop support interval from the peak LOD to approximate 95% confidence intervals. Table 2.1 summarizes the linkage findings for the SNP data based on the results from Allegro multipoint LOD scores > 2.5 and with two-point LOD scores $\geq$ 1.8; plots for these regions are shown in Figure 2.2. The strongest evidence for linkage occurred near the telomeres of chromosomes 7 (7p21.1-pter) and 2 (2q33.1-q35). Specifically, the maximum multipoint nonparametric LOD* score was

3.01 on chromosome 2 in the region spanning rs1379007-rs1992235. For chromosome 7, the maximum multipoint nonparametric LOD* score was 2.99 in the region spanning rs2353788-rs1368215. These results provide an approximate 10% increase in information content over the maximum multipoint scores in the microsatellite genomic screen.

When individual 100, who has a closed NTD, was included as affected (under the broad phenotype), her genotypic information increased the overall LOD score under both parametric and nonparametric analyses.

2.4.2 Haplotyping

Preliminary haplotyping using Simwalk [217] and hand-haplotyping confirmed the linkage regions identified in Table 2.1. The haplotype results are consistent with autosomal dominant transmission with reduced penetrance. Individual 100, who has the closed NTD variant of congenital dermoid cyst, also carried the disease-associated haplotypes on chromosomes 2 and 7. Visual inspection of marker data for chromosome 2 expanded the candidate interval to span markers D2S1776 to D2S1363 approximating a 55.2 cM region. For chromosome 7, marker coverage of 7pter extended only to the middle of 7p22.3; as a result, the flanking marker for the 7pter end of the disease-associated haplotype could not be determined. Thus, the minimum candidate interval for chromosome 7 spans from D7S3051 to 7pter approximating a 32.3 cM region. The presumed unaffected individual 126 carried the chromosome 7 disease-associated haplotype and showed a crossover between D7S2201 and D7S641. If individual 126 is truly “unaffected”, the crossover could reduce the minimum candidate interval to 17.4 cM.

2.4.3 Candidate genes

Several biologically plausible candidate genes reside in the regions of interest defined by haplotyping results. Using Ensembl (http://www.ensembl.org/index.html), a list of candidate genes for chromosomes 2 and 7 was assembled for mutational analysis. For chromosome 2, the frizzled 7 receptor (*FZD7*) gene that maps to 2q33.1 was sequenced. Members of the frizzled gene family encode 7-transmembrane domain proteins that are receptors for Wnt signaling proteins. The highly conserved Wnt-frizzled signal non-canonical transduction pathways have been shown in several animal models to be involved in convergent extension, the process that leads to elongation of the neural plate during neural tube development [107,110,219]. Human mesenchyme homeobox 2 (*MEOX 2*), located at 7p21.3-p22.1, is a homeobox gene involved in formation, patterning, and differentiation of somites. The mouse homolog *Meox-2* is required for migration of limb bud and null *Meox-2* mice have decreased limb muscle mass and demonstrate downregulation of *Pax3* [220]. Both genes evaluated (*FZD7* and *MEOX-2*) did not show any novel genetic polymorphisms within the coding exons.

2.5 Discussion

For some complex diseases, traditional linkage analysis approaches in rare families with Mendelian or near-Mendelian inheritance patterns has identified susceptibility loci. For example, investigators who studied rare multigenerational families with early-onset fully penetrant Alzheimer's disease identified linkage to chromosome 14 and subsequently identified the amyloid precursor protein (*APP*) gene [221]. Later investigations identified a second locus and the genetic risk gene presenilin 1 (*PSEN 1*) [221]. Other examples using

single large families to identify linkage regions that influence disease susceptibility include familial focal segmental glomerulosclerosis [222], coronary artery disease [223], dementia [224], tonic-clonic seizure [225] sensoineural hearing loss [226,227], among others. This focused approach of genetic analysis is advantageous for minimizing genetic heterogeneity, thereby increasing the investigator's ability to identify areas of allele sharing in large multiplex families.

Using a high-density SNP genomic screen, linkage of Family 8776 was confirmed for the previously reported locus at 7p21.1- pter. Additionally, a locus at 2q33.1-q35 yielded increased evidence in favor of linkage over that from the earlier microsatellite genomic screen [228]. These two regions of interest exhibit similar statistical support for an NTD susceptibility locus, which may reflect multiple distinct genetic effects. Haplotype analysis verified these regions of interest and demonstrated that the high-risk haplotypes segregate in a pattern consistent with autosomal dominant inheritance and reduced penetrance, or alternatively, an environmental factor that triggers disease presentation. As with any genetic linkage study, these results could represent a spurious finding.

An avenue for further investigation involves more precise phenotypic definition in the presumptively unaffected individuals. For instance, determining whether these unaffected individuals have subtle dysraphic changes such as spina bifida occulta may stimulate broadening of the NTD phenotype, increase the power for future linkage analyses, enhance the ability to minimize a candidate interval, and identify individuals at high risk to conceive a pregnancy with an NTD. Alternatively, the failure to identify radiographic changes in presumptively unaffected "transmitting" individuals may reflect a non-penetrant allele or a susceptibility locus in the absence of an eliciting environmental insult. Studies to elucidate

the phenotype involve radiographic confirmation by x-ray and/or magnetic resonance imaging and may be difficult to obtain in a research setting due to increased risk to the subject. Any detected differences in affected individuals' NTD phenotype from family 8776 may potentially be explained by pleiotropic effects of a common underlying gene, different genetic backgrounds containing a modifier gene, and/or different environmental exposures at a critical developmental time point. Two individuals in family 8776 demonstrate the importance of accurate phenotypic definition.

For instance, when individual 100, who has an NTD variant of congenital dermoid cyst, was included in the analyses as affected, the LOD scores for the 7p and 2q regions were increased. Individual 100 also carries the same high-risk haplotype as other affected individuals with lumbosacral myelomeningocele in the family.

And, haplotype analysis also revealed that individual 126, who appears phenotypically unaffected, to be recombinant for the disease-associated chromosome 7 haplotype. Since individual 126 is a minor, radiographic studies are not available for clarification of the phenotype. If individual 126 demonstrated no evidence for a radiographic abnormality, the 7p interval could potentially be narrowed by 14.9 cM to a 17.4 cM region between D7S641-7pter. However, this crossover for the unaffected individual should be interrupted cautiously because she may harbor a subtle dysraphic variant which would lead to a change in her affection status, thereby modifying the minimum candidate interval.

Several promising candidate genes reside in the regions of interest for chromosomes 2 and 7. *FZD7*, *CASP8, CASP10*, *PAX3*, and a cluster of *HOXD* genes map to the region of interest on 2q. *MEOX2* and *TWIST1* genes map to 7p. Sequence data for *FZD7* and *MEOX2* genes did not reveal any potential causal variants for this family; however, the promoter

regions for these genes have yet to be evaluated. In addition to evaluation of candidate genes, family 8776 is also under investigation for potential chromosomal abnormalities using both comparative genomic hybridization (CGH)-arrays and fluorescent *in situ* hybridization (FISH).

Overall, the highest linkage results for family 8776 map near the telomeric regions of chromosomes 2 and 7. Telomeric regions typically are difficult to map due to limited marker density, and in fact, the chromosome 7 region of interest could be narrowed with an increased marker density. Regions proximal to the telomeres tend to have a high degree of interchromosomal recombination, so it is plausible that a microdeletion and/or translocation event has taken place in one or both of these regions. Chromosomal abnormalities are associated with NTDs, and reports suggest 2-16% of nonsyndromic NTDs have detectable cytogenetic abnormalities [22]. The 2q region of interest for family 8776 includes the *PAX3* gene, for which an interstitial de novo microdeletion of 2q35-q36.2 has been previously reported for two patients with Waardenburg syndrome type III and myelomeningocele (OMIM 148820) [229]. *PAX3* is a member of the Paired Box transcription factor gene family that has been extensively studied for its role in early development. *Pax3* is expressed in the mouse neural tube and mutant *Pax3* mice (*Splotch)* embryos have neural tube defects, including spina bifida and exencephaly, (the equivalent of human anencephaly).

Focusing on rare Mendelian-like families for linkage studies can identify susceptibility loci that may later be shown to alter risk in the non-Mendelian forms of the disease. Family 8776, as one of the largest reported NTD families, may represent an important resource for narrowing the search for NTD candidate genes since it provides a unique opportunity for identification of a single major locus. However, evaluation of this

one family may or may not be generalizable to all NTDs, so any putative high-risk haplotypes, genetic variant(s) and/or chromosomal abnormalities may or may not be unique to this family. Consequently, any results identified in family 8776 must be characterized in other families to determine generalizability.

2.6 Acknowledgments

We thank all of the patients and family members who participated in this study for their generous contributions of time, energy, and biological samples. We also thank Bei Zhao, Kristen Deak, and Carol Haynes for their contribution to this project.

Table 2.1 Multipoint LOD scores (>2.5) for regions of interest on chromosomes 2 and 7.

				Parametric		Nonparametric	
Cytogenetic Band	**Marker**	**Region Spans**	**One LOD unit (CI) (Total distance, cM)**	Broad n=6	Narrow n=5	Broad n=6	Narrow n=5
Chrom. 2 **2q33.1-q35**	SNPs	rs1437872-rs2888386	196.5 – 213.6 (17.1 cM)	2.69	2.39	3.01	2.41
	Microsat.	D2S2978-D2S1363	186.0 – 229.2 (43.2 cM)	2.32	2.02	2.64	2.04
Chrom. 7 **7p21.1-pter**	SNPs	rs2389831-7pter	0 – 31.68 (31.7 cM)	2.66	2.36	2.99	2.38
	Microsat.	D7S3051-7pter	0 – 32.34 (32.3 cM)	2.31	2.02	2.63	2.03

n represents the number of affected individuals included in each analysis.

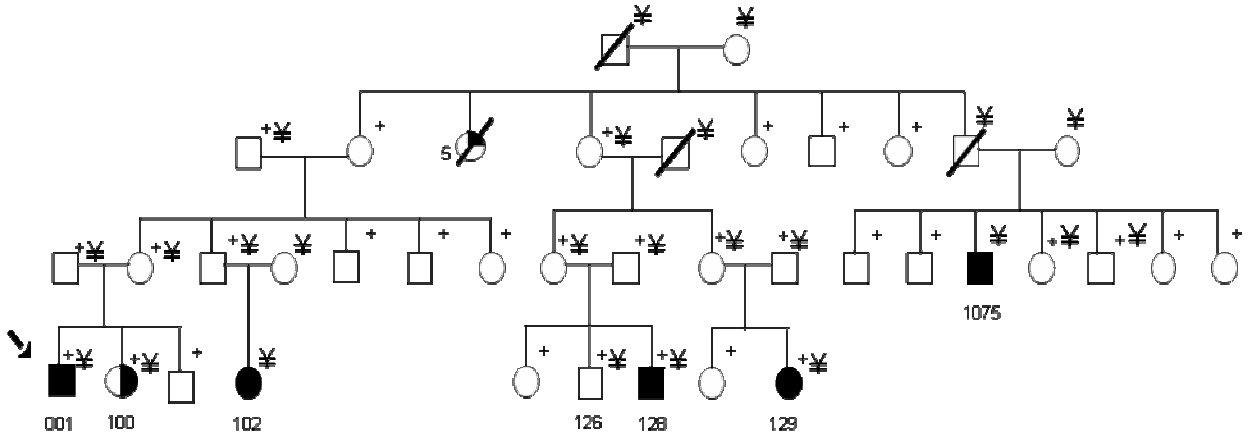
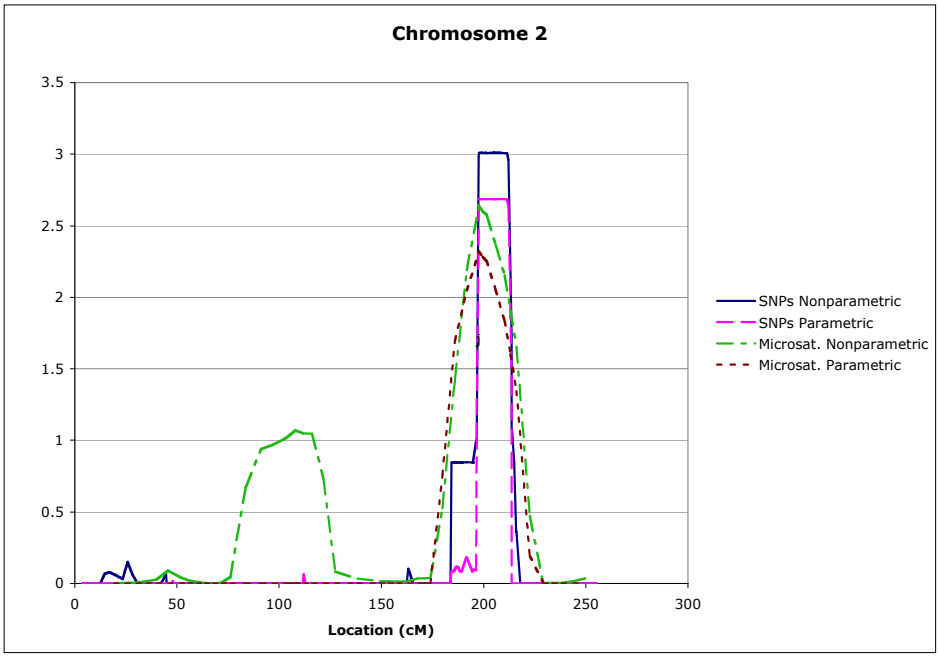


Figure 2.1 Pedigree diagram of family 8776

Condensed version of extended family 8776. Fully shaded symbols represent individuals affected with lumbosacral myelomeningocele. The half shaded circle for individual 100 represents congenital dermoid cyst. The quarter shaded circle denotes the 5 individuals who are affected by history alone. + indicates sampled (DNA) individuals. § refers to individuals included in linkage analysis. Individuals 102 and 1075 are reported to have lumbosacral myelomeningocele; however, the medical records and radiographic studies are not currently available to confirm this diagnosis.

A



B

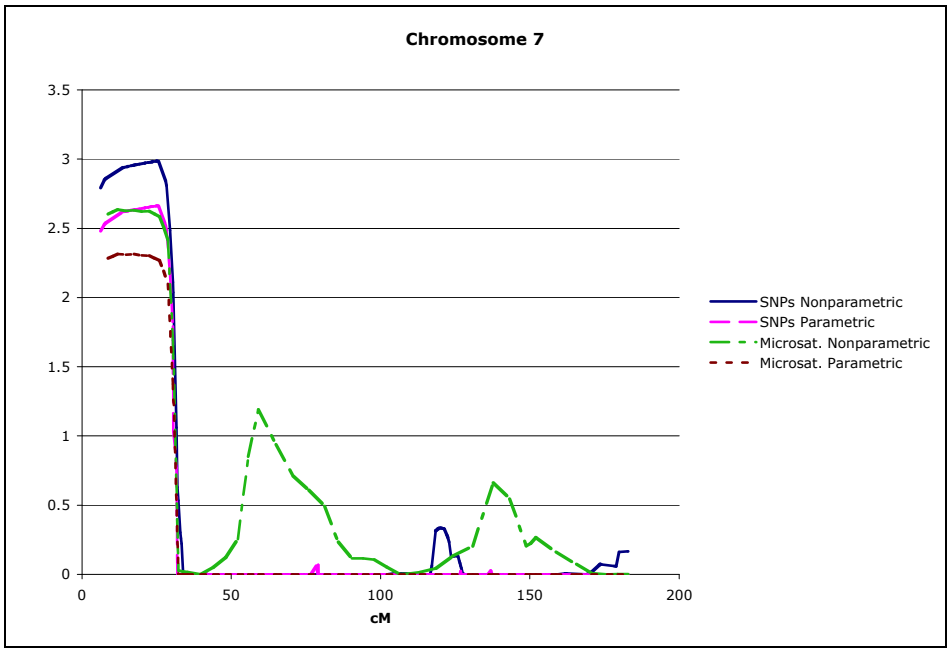


Figure 2.2 Multipoint LOD score curves for regions of interest in family 8776

Both SNP and microsatellite marker linkage results using parametric and nonparametric linkage methods are plotted for chromosomes 2(A) and 7 (B). The total genetic distance covered by the analyzed markers is given on the X-axis in cM.

CHAPTER 3

Refinement of 2q and 7p loci and genomic convergence for candidate gene selection in a large multiplex neural tube defect (NTD) family

3.1 Abstract

Neural tube defects (NTDs) are considered complex disorders that arise from an interaction between genetic and environmental factors. NTD family 8776 is a large multigenerational Caucasian family that provides a unique resource for the genetic analysis of NTDs. Linkage analysis using a genomewide SNP screen in family 8776 and multipoint nonparametric methods mapping previously identified maximum LOD* scores of ~3.0 mapping to 2q33.1- q35 and 7p21.1-pter. We ascertained an additional nuclear branch of 8776 and conducted additional linkage analysis, fine mapping and haplotyping
We observed increased evidence for linkage with LOD* scores of ~3.3 for both regions. Haplotype analyses narrowed the minimum candidate intervals to a 20.3 Mb region in 2q33.1-q35 between markers rs1050347 and D2S434, and a 8.3 Mb region in 7p21.1-21.3 between a novel marker 7M0547 and rs28177. Within these candidate regions, 16 genes were screened for mutations; however, no obvious causative NTD mutation was identified. Evaluation of chromosomal aberrations using array-CGH, subtelomeric fluorescent *in situ* hybridization, and copy number variant detection techniques within the 2q and 7p regions did not detect any novel chromosomal abnormalities. This large NTD family has identified two

genomic regions that may harbor NTD susceptibility genes. Ascertainment of another branch of family 8776 and additional fine mapping permitted a 9.1 Mb reduction of the NTD candidate interval on chromosome 7 and 37.3 Mb on chromosome 2 from previously published data. Identification of one or more NTD susceptibility genes in this family could provide insight into genes that may affect other NTD families.

3.2 Introduction

3.2.1 Human Neural Tube Defects

Neural tube defects (NTDs) are among the most serious congenital birth defects occurring at a rate of 1 per 1000 live births in American Caucasians [1,197]. NTDs result from failure of closure of the neural tube during the first three to four weeks of fetal development. NTDs are phenotypically heterogeneous disease with a large spectrum in the clinical presentation and degree of impairment. The most common NTD presentations are anencephaly and lumbosacral myelomenginocele (also known as spina bifida). Anencephaly occurs when the rostral neural tube fails to form and is associated with 100% mortality of anencephalic fetuses whereas spina bifida occurs when the caudal neural tube fails to close appropriately. The majority of NTDs are non-syndromic [15,16] that is, the failed closure of the neural tube in the cranial or spinal location, or both, is the only defect present at birth.

Non-syndromic NTDs are thought to have a multifactorial etiology with a complex interplay of several genetic and environmental factors [230]. A number of these environmental risk factors have been associated with maternal risk of having a child affected with an NTD. Most significant among these has been the impact that periconceptional folate supplementation has made a on NTD prevalence in the US population, decreasing the risk of

having a child with an NTDs by 50-70% [43,85,231]. Approximately 30-50% of NTDs are not prevented by folic acid supplementation, suggesting that additional factors including genetic risk factors, may contribute to the development of NTDs. The most compelling evidence for a genetic factor is a 2-5 % recurrence risk for NTDs among offspring from parents with previous NTD pregnancies [199]. Interestingly, syndromic NTDs are also associated with established genetic syndromes including Meckel-Gruber syndrome, anterior sacral meningomyelocele, anal stenosis, trisomies 13 and 18, and other chromosomal aberrations [18,19,22]. Furthermore, 155 mouse models with known genetic mutations leading to NTDs have been characterized -- one of these is the Lp mouse with craniorachischisis caused by a *Vangl2* mutation [58]. Recently, Kibar and colleagues discovered two mutations in the *VANGL1* gene in two patients with familial forms of non-syndromic NTDs (V239I and R274Q) and the V239I mutation abrogated protein-protein interactions with disheveled binding partners [113]. Not only does the *VANGL1* finding support a genetic contribution to NTDs, but it also provides functional evidence of a rare genetic mutation associated with human familial NTDs. This work demonstrates that family-based genetic studies can illuminate rare genetic effects and further our understanding of the complex molecular mechanisms underpinning human neural tube closure.

3.2.2 Previous NTD Gene Mapping Studies

A whole genome screen performed on 44 multiplex families using 402 microsatellite markers at ~ 10 cM spacing identified linkage peaks on chromosomes 7 and 10 with parametric LOD scores of 2.45 and 2.08, respectively [134]. One large family, 8776, was found to be primarily responsible for the chromosome 7p linkage peak.

Further genetic analysis of family 8776 alone revealed similar evidence of linkage on chromosome 2, with both linkage regions in 7p21.1-pter and 2q33.1-q35 having a nonparametric S_{pairs} LOD* scores ~3.0 assuming a broad phenotypic definition for NTDs in which all affected individuals with a broad spectrum of NTDs were considered affected [61]. The 2q and 7p linkage regions were confirmed by haplotype analysis, and we also determined that the 2q and 7p NTD associated haplotypes segregated in a pattern consistent with autosomal dominant inheritance with reduced penetrance.

In this current genetic analysis, we have ascertained a new branch of the family 8776 that led to the refinement of the previous linkage interval on 7p. Additional fine mapping and haplotype analysis further narrowed the minimum candidate gene interval for both the 2q and 7p loci. We have used convergent methodologies to identify to prioritize candidate genes for mutational screening and have investigated the possibility of chromosomal abnormalities within family 8776 as a mechanism for NTD development. As family 8776 is one of the largest known multigenerational NTD families showing evidence of a strong genetic effect, it provides a unique opportunity for identification of a novel NTD gene.

3.3 Materials and Methods

3.3.1 Clinical Data Collection and Biological Samples

NTD family 8776 was identified by self-referral through the NTD Collaborative Study. First-degree relatives and relatives connecting related affected individuals were ascertained in this extended family. A detailed family history [208] was obtained, and medical records including operative reports and pre-surgical x-ray films were collected for review of diagnosis by a neurosurgeon. A total of 5 affected and 29 unaffected individuals

were available for genotyping. Of the sampled affected individuals with non-syndromic NTDs in the pedigree, four were diagnosed with lumbosacral myelomeningocele and individual 100 was diagnosed with congenital dermoid cyst with tethered cord. A sixth affected individual (1075) is reported by family history to be affected with lumbosacral myelomeningocele.

Certified phlebotomists obtained blood samples from affected individuals and related family members at medical centers, clinics, or by visiting participants' homes. Some participants sent in mailer kits with their enclosed samples. This study was conducted under the oversight of Duke University Medical Center Institutional Review Board and informed consent was obtained from all participants.

Power analysis of family 8776 for the SNP screen was performed using the program SIMLINK [209,210]. The five newly sampled individuals affected with an NTD were included in the power analysis. The assumptions for the power studies included an autosomal dominant inheritance pattern with low penetrance (affecteds-only) model, a disease allele frequency of 0.001, and a SNP marker with 50% heterozygosity linked at 5% recombination with the disease allele. These power studies generated a maximum parametric LOD score of 2.07 at $\theta = 0.05$. SIMLINK assumes an underlying Mendelian inheritance pattern, which may not accurately represent the complexity of NTDs.

3.3.2 Genotyping Methods

Since the previous publication [61] blood samples were collected on one nuclear branch in family 8776 including affected individual (102) and 2 unaffected individuals (1006, 103) . DNA was extracted from whole blood using the Puregene system (Gentra

Systems, Minneapolis, MN). Individual genotyping was performed on the Affymetrix 10K version 2 chip (Affymetrix, Santa Clara, CA) using the standard Affymetrix protocol. This SNP chip provides genotypes for approximately 11,500 SNPs that have an average heterozygosity of 0.37 and average spacing of one SNP every 210 Kb. Microarrays handling and data acquisition methods are the same as previously reported [61]. All marker information from the genomic screen was databased using the PEDIGENE system [214].

3.3.3 Error Checking

Mendelian pedigree inconsistencies present within the SNP data were identified using PedCheck [232] and were removed prior to genetic analysis. Using 685 SNP markers genotyped across the genome, intra-familial relationships for family 8776 were inferred using a maximum likelihood approach implemented in RELPAIR [233,234] in order to verify the genetic relationships for the newly ascertained 8776 branch containing individuals 1005, 1006, 102, and 103.

3.3.4 Whole Genome Linkage Analyses

3.3.5 Marker Maps

Map order for the SNP markers was generated from sequence data provided by Translational Genomics (Phoenix, AZ). SNPs with at least 85% genotyping efficiency for family 8776 were retained for linkage analysis. Allele frequencies for the SNP markers were estimated based on 126 unrelated Caucasian individuals' DNA also genotyped on the 10K Affymetrix SNP chip. To account for potential inter-marker LD in the 10K SNP screen, LD

Select was used [216] with a pairwise $r^2 > 0.16$ (Boyles et al. 2005) to create map files with only tagging SNPs. After efficiency and inter-marker linkage disequilibrium (LD) were incorporated, the markers maps used for linkage analysis included 5871 SNPs. The rationale behind accounting for inter-marker LD in multipoint linkage analysis is based on simulation work demonstrating that dense SNP data can lead to false positive elevation of multipoint LOD scores, particularly when there are missing parental genotypes [206,207].

3.3.6 Linkage Analysis

We used similar linkage approaches as previously reported for this family [61]. Briefly, linkage analysis was performed using the program Allegro [133] for both two point and multipoint analysis. Since the inheritance pattern of NTDs is unknown, different analytic approaches were applied to these data, including parametric (assuming dominant inheritance and a disease allele frequency of 0.001) and nonparametric analysis (model-free). It is possible that unaffected individuals in the pedigree 8776 could represent asymptomatic gene carriers or have undiagnosed NTDs; thus, a low penetrance affecteds-only parametric model was applied such that unaffected individuals did not affect the LOD score other than to help infer missing genotypic data on unsampled individuals. Phenotypic criteria included as affected all six individuals with lumbosacral myelomeningocele, as well as individual 100 who has an NTD variant of congenital dermoid cyst. Previous linkage findings demonstrated that individual 100 contributed to the LOD score for both 2q and 7p. Individual 1075, who is affected by historical account only and not sampled, was also included in the analyses, because the genotypic information from siblings and/or parents can help to infer the affected individual's missing genotypes. For the nonparametric analysis, an identity-by-descent

relative pair sharing method LOD* was assessed between all pairs of affected individuals using S_{pairs} and S_{all} [126] and the exponential model [127] as implemented in Allegro. The regions of interest were defined by a 1 LOD/LOD* unit support interval from the peak LOD/LOD* to approximate 95% confidence intervals. The LOD score reflects the parametric linkage results and the LOD* is the linkage results from the nonparametric methods.

The same 24 individuals were retained in the current linkage analysis as previously reported [61]. The nuclear branch with individuals 1006, 1005, and 102 was included in this original linkage set based on the rationale that the genotypic data from 1005 could help to infer affected individual 102's potential genotypes. In this current study, the new genotypic data for individual 102, who is affected with lumbosacral myelomeningocele, provided additional power (estimated increase in parametric LOD score by 0.3) to detect linkage.

3.3.7 Fine mapping

An additional 19 microsatellite repeat markers in 2q and 11 (including two novel variants) microsatellite repeat markers in 7p were genotyped to further refine the linkage regions. Microsatellite markers were prioritized based on high heterozygosity and spacing (~ 2 Mb) within the region. We also used UCSC Bioinformatics Table Browser (http://genome.ucsc.edu/cgi-bin/hgTables?command=start) to identify additional repetitive regions for the development of novel genetic markers as the chromosome 7 interval had a paucity of known microsatellites. Primers were designed in Primer 3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) and the novel microsatellite markers were named 7M0546 and 7M0547. 7M0546 maps to chr7: 14043894-14044124 (NCBI Build 35); the forward primer is 5'CATTGTTGGCTGACGAATTG 3' and the

reverse primer is 5' CTAGGGGCACAGGTGAGAAG 3'. 7M0547 maps to chr7: 16053379-16053552 and a forward primer of 5'GAATGGAACAGGGATTTGGA 3' and the reverse primer is 5' CACTGCATCCATATTGGGTTT 3'. Microsatellite genotyping methods are the same as previously published [61].

3.3.8 Haplotype analysis

To confirm segregation of haplotypes with disease status and to further evaluate haplotypes from fine mapping, haplotype analysis was performed using both visual inspection and a statistical prediction using Simwalk [217],

3.3.9 Chromosomal Aberration Detection Methods

We carried out genomewide copy number profiling to identify genomic gain or loss that could be associated with NTD within family 8776. as previously described [235]. Briefly, PCR representations of 30,000 overlapping 100-200 Kb clones generated from the Human Genome Project are used as array probes. NTD DNA and sex-mismatched control DNA were differentially labeled with Cy3 or Cy5 and co-hybridized to the genomic tilepath array [236]. Hybridization signals were visualized using an Axon 4100A dual laser scanner and output data uploaded into BlueFuse (www.cambridgebluegnome.com/) for signal normalization and clone filtering prior to log2 ratio calculation (as previously reported by [235]). Array-CGH was conducted on DNA of 8 individuals from family 8776; specifically, CGH arrays were run on 4 sampled affected individuals (001, 100, 128, 129), 2 mothers transmitting the 7p and 2q high-risk haplotypes (1001, 1023), an unaffected not carrying the

high-risk haplotypes (131), and an unaffected individual (126), who carries the 2q and 7p haplotype.

SNP data available from the Affymetrix 10K SNP Chip was also used to test for potential copy number variants (CNV) and loss of heterozygosity (LOH). More specifically, the .cel files which provide the intensity information for a given probe on an array were utilized in two different analysis tools that both compare cases versus controls for copy number changes and loss of heterozygosity. The first was DChip (http://biosun1.harvard.edu/complab/dchip/) and the second tool was CNAT 4.0 (http://www.affymetrix.com/products/software/specific/cnat.affx). In each case the SNPs were first viewed in groups of 5 in both a smoothed median and a Hidden Markov model. The SNP threshold was then gradually lowered to a single SNP using both algorithms.

Subtelomeric FISH (fluorescent *in situ* hybridization) analysis on family 8776 was performed using ToTelVysion™ Multicolor DNA probe mixtures (Abbott Laboratories, Abbott Park, IL) per the manufacturers recommended protocol. FISH analysis was performed on two nuclear, branches of family 8776; more specifically, in individuals 1000, 1001, 001, 100, 1024, 1023, 126, 127, and 128.

3.3.10 Candidate Gene Prioritization

We identified 89 and 9 RefSeq genes mapping to the 2q and 7p disease-associated haplotype regions, respectively, using the UCSC genome browser (http://genome.ucsc.edu/). To focus our mutational sequencing efforts we prioritized candidate genes based on each of the gene's following characteristics: 1) biological plausibility (e.g. NTD mouse model, involved in folate or glucose metabolism); 2) expressed in brain or fetus; and 3) differential

expression in immortalized lymphoblast cells lines from three NTD cases of family 8776 (individuals 1, 100,128) compared to 9 Coriell control cell lines (GM03797, GM07535, GM03714, GM03798, GM12497, GM13072, GM08336, GM10958 , GM01207; Camden, NJ). The control cell lines were selected to roughly age match (the age of the subject when cell lines were immortalized) for the cases and controls—the cases had an mean age of 4.2 years old $\pm$ 2.8 (SE) at age of immortalization and the control cell lines were 7.9 years old $\pm$ 1.5 (SE).

3.3.11 Expression Analysis

The whole genome expression work was performed on the Illumina's Sentrix® Human-6 version 1 Expression BeadChip (San Diego, CA) using the standard Illumina protocol for these RNA chips. These chips allow for assessment of six independent RNA samples in parallel on a single chip containing probes for 46,000 transcripts with 24,000 being annotated genes. Each transcript has a full-length 50mer oligo probe that is present on average 30 times in order to increase the reproducibility and reliability of the expression data generated. Illumina's Bead Studio 2.3 was used for expression analysis. The rank invariant normalization method was performed (rank invariant genes have expression values with consistent order in different samples when compared to other genes on the expression panel). The Illumina custom error model was applied to the data; this model assumes the target signal intensity is normally distributed among replicates and takes into consideration three components of variation: sequence biological variation, nonspecific biological variation, and technical error. Differentially expressed genes with p values $\leq$ 0.05 residing in the 2q and 7p haplotypes were prioritized for candidate gene selection.

3.3.12 Sequencing of Candidate Genes

Exon locator and eXtractor and Resequencing (ELXR) were used to select primer sets for resequencing (http://mutation.swmed.edu/ex-lax/index.html). If primers failed to produce PCR products, then primers were redesigned using Primer 3 on genomic sequence from the UCSC genome browser. PCR amplicons covered all exons and respective intronic boundaries. First, we conducted a screen for potential variants by sequencing eight individuals (four affected individuals: 001, 100, 128, 129; two parents: 1000, 1001; one unaffected individual 126, who carries 2q and 7p haplotype; and a sample from the Centre d'Etude du Polymorphisme Humain (CEPH) genotyping database (Family 1331, individual 2). PCR products and corresponding primers were sent to Agencourt's Genomic Service (Beverley, MA) for sequencing. In the initial screen, only the forward direction was sequenced. Subsequent sequencing was performed in both forward and reverse directions to contiguate sequence from several overlapping PCR products. Sequence data were visually inspected to identify nucleotide variants using ABI 3100 Data Collection Software Version 1.01 and ABI Sequencing Analysis 3.7. The sequence data were compared to unaffected parents to ascertain if a variant was either transmitted to their two affected individuals (001 and 100) or a de novo mutation.

Variants observed in the mutational screen were further prioritized for resequencing in the remaining individuals of NTD family 8776. The decision for additional re-sequencing efforts in family 8776 was based on the following features of each variant identified in the screen: 1) whether all four affecteds share the same variant; 2) potential functional significance of variant (coding or UTR); 3) degree of conservation; 4) whether novel or

known variant; and 5) when available, allele frequencies. Cross-species conservation predictions were based on the UCSC genome browser's conservation track. Sequencing the entire family included DNA from 34 NTD 8776 family members and 2 CEPH controls.

All novel mutations identified in this study were named according to the recommendations of the Human Genome Variation Society (HGVS) described by [237] with the ATG translation start codon assigned +1. Gene names were assigned based on the HUGO-approved gene symbol (http://www.gene.ucl.ac.uk/cgi-bin/nomenclature/).

3.4 Results

3.4.1 Linkage Results for Family 8776

Intra-family relationships for NTD family 8776 were verified using RELPAIR and the pedigree structure of 8776 is shown in Figure 3.1. The individuals included in the current linkage studies (Figure 3.1) are the same set as previously reported [61], but with the addition of genotypic data for individuals 1006 and 102. Figure 3.2 shows LOD plots for the regions with the strongest evidence of linkage mapping to chromosomes 2 (2q33.1-q35) and 7 (7p21.1-7p21.3). The maximum multipoint parametric LOD and nonparametric LOD* scores for chromosome 2 were 3.31 and 2.99 spanning the intervals between rs1961468 - rs956133 and rs1379007 - rs1912181, respectively; for chromosome 7, they were 3.29 and 2.96 at rs2215949 - rs763534 and rs719680 - rs1368215, respectively. Specifically, the 1 LOD support interval for chromosome 2 is rs1437872 to rs2888386, the same as previously published, whereas for chromosome 7 the region is flanked by markers rs2389831 to rs28177, a 9.7 Mb interval, leading to a narrowing of the 7p linkage region by 7.8 Mb (15.5 cM). Notably, both regions of interest had an increase in the LOD scores by 0.3 with the

additional genotypic data for individual 102. The S_{pairs} and S_{all} linkage results were the same within the 1 LOD* support interval.

3.4.2 Haplotyping

We previously identified regions of linkage to chromosomes 2 and 7 within NTD family 8776 [61]. To further refine these loci we genotyped an additional 19 and 11 microsatellite repeat markers on 2q and 7p, respectively, with inclusion of a new branch of family 8776 (Figure 3.3). NTD-associated haplotype analysis of the additional markers substantially decreased the minimal critical regions of interest on both chromosome 2 and 7.

For chromosome 2, the haplotype previously was reported to span D2S1776 to D2S1363, a 57.6 Mb interval mapping to 2q24.3-q36.3 [61]. Genotype analysis of the 19 microsatellites on chromosome 2 substantially reduced the region to a 23.5 Mb interval between D2S425 to D2S434. We further refined this interval by sequencing exon 2 of the 'heat shock 60kDa protein 1' gene (HSPD1 [MIM 118190]) within all family members of 8776. Sequence analysis revealed a recombination event within individual 129 for a coding synonymous SNP within HSPD1 (rs1050347) that reduced the chromosome 2 by 3.2 Mb, redefining the minimal critical region to 20.3 Mb between rs1050347 - D2S434 mapping to 2q33.1-q35. Two presumptively unaffected individuals 103 and 126 also have the 2q high-risk haplotype as shown in Figure 3.3.

Our previously reported chromosome 7 NTD haplotype spanned a 17.4 Mb interval between D7S2508 to 7pter [61]. Genotype analysis of the additional 11 microsatellite markers within the additional branch of family 8776 identified cross over events within individual 1005, the father of 102 between novel markers 7M0546 and 7M0547 and affected

individual 102 between SNP markers rs28177 and rs2349775. The new data shows a conserved 7p haplotype segregating with affection status mapping to an 8.3 Mb interval between rs28177-7M0547 within 7p21.1-21.3 (Figure 3.3), a decrease of 9.1 Mb from our previous data [61]. Haplotype analyses also revealed two unaffected individuals, 103 and 126, carry the 2q haplotype and are recombinant for the 7p haplotype. Specifically, individual 103 is recombinant between D7S2547 and rs2215949 and individual 126 between rs725934 and D7S2514 (Figure 3.3).

3.4.3 Candidate Genes

Several biologically plausible candidate genes reside in the regions of interest defined by haplotyping results. Using UCSC, a list of candidate genes for chromosomes 2 and 7 was assembled. Approximately 89 and 9 annotated genes for 2q and 7p minimum candidate intervals, respectively. Our decision to sequence genes was prioritized first on their biological plausibility based on what is known from animal models on neural tube closure. For instance, biologically plausible candidate genes include those involved in folate metabolism [58,68,69], in planar cell polarity pathway [109,110], and in apoptosis [58,200]. Thus, the genes *FZD5*, *WNT10A* and *WNT6* were preferentially evaluated for their role in the planar cell polarity pathway; of note, only two exons were successfully sequenced in the WNT genes since the other exonic regions had high homology to many other genomic locations complicating primer design. Three genes involved in apoptosis were sequenced in family 8776 including *CASP8*, *CASP10*, and *STK17B*.

Table 3.1 shows the genes that were selected for mutational sequencing and describes for each gene the particular variant evaluated with the corresponding total number of exons

successfully sequenced. The majority of exons were sequenced but for those exons that were not sequenced it was often due to the sequence being highly homologous to other regions of the genome making primer design challenging. There were 1372 differentially expressed transcripts in the 3 case to 9 control comparison, 4 of which mapped to the 2q and 7p haplotype intervals, and three were selected for sequence analysis; *HSPD1*, *PECR* and *IGFBP2* (Table 3.1), and the other differentially expressed gene *ICA-1* was not sequenced because it mapped near the end of the haplotype (13.6 kb from the flanking marker rs28177). Table 3.1 indicates of these 3 genes sequenced whether they had increased or decreased patterns of expression.

All eight novel variants identified in our mutational screen by re-sequencing were non-coding (intronic or untranslated) polymorphisms (Table 3.2). Variants were prioritized for additional sequencing in the entire family based on whether $\geq$ 3 of the 4 affected individuals shared the same variant since we would expect the same variant to be segregating with disease. Only two novel variants (NM_003468.2:c.-164 G>A and NM_177538.1: c122 + 129 G > A) were observed. We did not carry out further sequence analysis of NM_177538.1: c122 + 129 G > A in *CYP20A1* since it was observed in all 8 individuals sequenced including the CEPH sample and the father 1000 who does not carry the 2q or 7p haplotype. Twelve variants detected in the screen were also sequenced in the entire family 8776 (Table 3.3). Following evaluation, no variant appeared to be specifically associated with disease status. Hence, it was interpreted from this finding that these mutations are unlikely to be NTD susceptibility variants.

3.4.4 Chromosomal Abnormalities

Copy number analysis by tilepath array-CGH of 8 members of family 8776 failed to identify copy number variation associated with NTD within the two regions of interest in 2q and 7p. Analysis of SNP data across the entire genome of these individuals also failed to identify copy number changes that may be associated with NTD. Subtelomeric *FISH* analysis for probes to 7p and 2q on two triads of family 8776 did not identify any chromosomal aberrations. Together, these findings provide evidence supporting normal chromosomal structure in family 8776 for the 2q and 7p regions of interest.

3.5 Discussion

In this study, fine mapping, haplotype analysis and pedigree expansion of NTD family 8776 lead to a substantial refinement of regions of interest previously identified at 2q and 7p. Pedigree expansion resulted in an increased LOD* from ~3.0 to ~3.3 for both 2q and 7p linkage peaks and decreased the 7p linkage region to an interval that is 7.8 Mb (15.5 cM) smaller than previously described [61]. Moreover**,** fine mapping and haplotype analysis of these two regions led to a decrease in the region of interest in 2q by 37.3 Mb to a 20.3 Mb region and in 7p there was a 9.1 Mb decrease to a 8.3 Mb interval.

Further narrowing of this 7p haplotype could be achieved if we were able to determine whether one of the presumed unaffected individuals 103 and 126 were truly not affected by an NTD. When haplotypes of individuals 103 and 126 are combined they completely overlap with the 7p NTD-associated haplotype, so if both individuals were found to be definitively unaffected this region may be excluded altogether. However, the

recombinant haplotypes for these two unaffected individuals should be interpreted carefully because these individuals may have a subtle NTD variant that has not been diagnosed. Radiographic studies (X-rays and MRI) used to diagnose NTDs are not currently available for either of these individuals, and even had such studies been available the utility of a negative imaging result is unclear. Of note, an X-ray on transmitting relative 1001 did not identify any bony defects.

The chromosome 7 linkage interval localizes to the subtelomeric region of chromosome 7p. Telomeres have previously been shown to have a high degree of inter-chromosomal recombination [238]. It is plausible, therefore, that a microdeletion and/or translocation event may have taken place in one or both of these regions and 2-16% of non-syndromic NTDs have detectable cytogenetic abnormalities [22]. We evaluated family 8776 for chromosomal abnormalities in the 2q and 7p regions and detected no evidence of abnormality using three independent techniques: 1) array-CGH; 2) subtelomeric *FISH*; and 3) CNV analysis of the SNP mapping array data. Although none of these approaches identified loss or gain of genomic material, it is possible that a small chromosomal aberration(s) exists within these regions that may only be detected using denser genotyping chips or tiling path oligo-arrays.

The chromosome 2 and 7 linkage regions exhibit similar statistical support for an NTD susceptibility locus which may reflect multiple distinct genetic effects or epistatic interactions associated with the disease. As with all gene mapping studies, there still remains the possibility that one or both of these haplotypes are shared by chance and do not harbor an

NTD high-risk variant. However, this seems unlikely given the high evidence in favor of linkage for this family.

NTDs are most likely caused by polygenic factors combined with a multitude of environmental factors. A multifactorial threshold model can be applied to diseases like NTDs with complex disease inheritance [239]. Within this model the presence of the complex disease phenotype is detected once the combined additive genetic and environmental effects coalesce to a particular liability threshold. In family 8776, the liability threshold may not have been reached for the unaffected individuals that carry either the 2q or 7p or both haplotypes, such that certain genetic variants or environmental stimulus may not be present to reach the additive gene and environment threshold. Interestingly, individual 100, who carries the 2q and 7p haplotypes, has a different clinical presentation (congenital dermoid cyst) from the other 5 individuals affected with lumbosacral myelomeningocele. This different phenotype or lack of phenotype in "unaffecteds" carrying the high-risk haplotype(s) may be due to any of the following factors: 1) pleiotropic effects of the same susceptibility gene(s), 2) differences in genetic backgrounds with modifier gene(s), and/or 3) differences in-utero exposures at time of neural tube closure.

Epigenetic mechanisms may also play a role in normal and dysregulated neural tube closure. Methylation of cytosine nucleotides is an epigenetic modification of DNA that occurs during the early development of germ cells and pre-implanted embryos. NTDs have been shown to develop in mutant mice that are null for DNA (cytosine-5)-methyltransferase 3B (*Dnmt3b*) and DNA (cytosine-5)-methyltransferase 3L (*Dnmt3l*) [58], genes that are known to establish methylation of maternally imprinted genes throughout mouse

development [240]. These data suggest that aberrant methylation may play a role in neural tube defects within humans. Interestingly, methylation patterning during fetal development has been shown to be affected by dietary methyl donors (methionine and choline) and critical co-factors (folic acid, B_{12}, pyridoxal phosphate) [92] . Thus, the nutritional status of the mother may affect essential methylation reactions such as establishment and maintenance of methyl marks on imprinted genes.

In NTD multiplex family 8776, aberrant maternal imprinting may implicated in the disorder because the sex ratios for NTD transmitting unaffected members of the family to affected children is 5:2 female to male ratio, suggesting a stronger maternal transmission pattern. Although this observation is not statistically significant and may reflect ascertainment biases toward maternal interviews, it illustrates that maternal imprinting in family 8776 and/or general DNA methylation patterns may increase the risk of having a child affected with an NTD.

We selected 16 genes promising candidate genes from the refined linkage regions on chromosome 2 and 7 for mutational sequence analysis in family 8776. Twelve variants identified within these genes were evaluated in the entire family, however, they were not specifically associated with disease status; therefore, it is our belief that the actual NTD susceptibility variant(s) from these intervals remain to be identified. Despite this negative outcome, sequencing efforts for family 8776 are still underway. However, other avenues of research may also prove useful in identifying candidate genes. For instance, preliminary work in our laboratory evaluating expression patterns of human fetal embryos from SAGE (serial analysis of gene expression) libraries of human fetal neural tube tissue from Carnegie

Stage C12 (neural tube closing) and Carnegie Stage 13 (neural tube closed) may be a valuable tool for selecting additional candidate genes [241].

There are likely many rare and common variant interactions coupled with environmental milieu that together predispose a child to an NTD. Unraveling the separate components of this genetic complexity often requires focused study design and mapping approaches. Focusing on genetic mapping in rare Mendelian-like families, like family 8776, can help identify susceptibility loci and/or genes that may later be shown to confer risk in the non-Mendelian forms of the disease. Specifically, for this family, identification of the causal NTD gene(s) may allow for identification of individuals who are at an elevated risk of conceiving offspring with an NTD. However, such a causal variant(s) may or may not be generalizable to most subjects with NTDs, and therefore any promising variant(s) identified in this family will also need to be examined in a larger subset of families. Even if a causal variant(s) is identified in this family and fails to be found in other NTD cases, the genetic information gleaned from 8776 is still valuable because the gene(s) identified in 8776 may provide insight into the pathogenesis of human NTDs and of early neurodevelopment in general.

In summary, we have used one of the largest known multigenerational NTD families that show evidence of a strong genetic effect to substantially refine two regions of genetic linkage on chromosome 2 and 7. These data provides a valuable resource for identifying novel candidate genes implicated in the molecular etiology of NTDs.

3.6 Acknowledgments

We are grateful to the on-going participation of the patients and families for without their generation contributions this study would not be possible. We also would like to thank Putting Xu and Kristen Deak for input on laboratory techniques, Carol Haynes for assistance with database management, Andrew Dellinger for input on SAGE libraries and Bei Zhao for technical assistance.

Table 3.1 Genes Sequenced in NTD Family 8776

gene	chr	molecular mechanism	DE	#E	ES	gene name
ABI2	2	arg protein tryosine kinase binding protein		10	2-10	NM_005759.3| abl interactor 2
CASP10	2	apoptosis		10	10	NM_032977.2| caspase 10, apoptosis-related cysteine peptidase
CASP8	2	apoptosis		9	9	NM_033355.2| caspase 8, apoptosis-related cysteine peptidase
CYP20A1	2	drug metabolism; cholesterol synthesis		13	2-5, 7-13	NM_177538.1| cytochrome P450, family 20, subfamily A, polypeptide 1
EEF1B2	2	translation elongation factor		6	6	NM_001959.3| eukaryotic translation elongation factor 1 beta 2
FZD5	2	signaling; planar cell polarity pathway		2	2	NM_003468.2| frizzled homolog 5 (Drosophila)
HSPD1	2	heat shock protein	UP	12	12	NM_002156.4| heat shock 60kDa protein 1
IGFBP2	2	Insulin growth factor	DOWN	4	2, 3	NM_000597.2| insulin-like growth factor binding protein 2
MAP2	2	microtubule-associated protein		17	17	NM_031846.1| microtubule-associated protein 2
NRP2	2	vascular endothelial growth factor; axon guidance		16	16	NM_018534.3| neuropilin 2, transcript variant 4
PECR	2	peroxisomal trans-2-enoyl-CoA reductase	UP	8	8	NM_018441.2| peroxisomal trans-2-enoyl-CoA reductase
STK17B	2	apoptosis		8	1-7	NM_004226.2| serine/threonine kinase 17b (apoptosis-inducing)
WNT10A	2	signaling; planar cell polarity pathway		4	2,3*	NM_025216.2| wingless-type MMTV integration site family, member 10A
WNT6	2	signaling; planar cell polarity pathway		4	2,4*	NM_006522.3|wingless-type MMTV integration site family, member 6
DGKB	7	regulate intracellular levels of diacylglycerol (DAG)		24	24	NM_004080.1| diacylglycerol kinase, beta 90kDa
ETV1	7	transcription factor		12	12	NM_004956.3| ets variant gene 1

Key legend:
DE =Differential expression of genes from lymphoblast cell lines comparing 3 NTD cases to 9 controls. UP denotes the gene was found to have increased expression levels in the cases and DOWN denotes decreased expression levels of the gene.
#E = Number of exons present in the gene sequenced
ES=Exons successfully sequenced
*Only two exons were sequenced in the WNT genes because the other exonic regions had high homology to many other genomic locations complicating primer design.

Table 3.2 Novel Variants Identified in Sequencing Screen of Family 8776

Gene	Variation	NCBI ss#	Type of Variant	Individual(s) with variant
HSPD1	NM_002156.4:c.606 + 42 G>T	ss77831713	intronic SNP	1,100, 126, 128, 1000,1001, CEPH
HSPD1	NM_002156.4:c.606 + 53 delGTTTGAGTTATCTGATGATC AAAA	ss77831714	intronic deletion (-GTTTGAGTTATCTG ATGATCAAAA)	1, 100, 126, 128, 1001, CEPH
FZD5	NM_003468.2:c.-164 G>A	ss77831715	5' untranslated SNP	126,128,1000,1001
STK17B	NM_004226.2:c-130 C >T	ss77831716	5' untranslated SNP	1, 129, 1000
CASP8	NM_033355.2:c.550 + 63 G > A	ss77831717	intronic SNP	1001
CYP20A1	NM_177538.1: c122 + 129 G > A	ss77831718	intronic SNP	1, 100, 126, 128, 129, 1000, 1001, CEPH
DGKB	NM_004080.1:c516 + 155 G > C	ss77831719	intronic SNP	1,100,1000
NRP2	NM_018534.3:c1786 + 122 C>A	ss77831720	intronic SNP	001, 1000

Table 3.3 Variants Sequenced in the Entire Family 8776

Gene	Variant	Type of Variant	Known/ Novel Variant	Conserved	Allele frequencies (Caucasian)
CYP20A1	rs13015873	intronic SNP (C/T)	known	mild	***C=0.985; T=0.015*** [d]
CYP20A1	rs10586509	intronic deletion (-/TTAA)	known	yes	***Del=0.985; no Del=0.015*** [d]
CYP20A1	rs2351771	intronic SNP (T/G)	known	yes	***G= 0.986; T =0.014*** [d]
EEF1B2	rs11546386	coding nonsynonymous SNP (G/A)	known	yes	***G= 0.917; A =0.083*** [d]
HSPD1	rs1050347	coding synonymous SNP (C/T)	known	yes	C= 0.08; T= 0.920 [b]
NRP2	rs849526	coding synonymous SNP (A/G)	known	yes	A=0.5;G=0.5 [b]
NRP2	rs2041808	intronic SNP (A/G)	known	mild	***C= 0.672; A =0.328*** [d]
STK17B	rs12478353	coding synonymous SNP (A/T)	known	yes	A=0.658; T= 0.342 [a]
STK17B	rs1054537	coding synonymous (C/T)	known	yes	C=0.62; T= 0.38 [b]
CASP8	rs2141331	3' untranslated SNP (A/G)	known	no	G=0.919; A=0.081 [c]
ETV1	rs9639168	coding nonsynonymous SNP (C/T)	known	yes	C=0.267; T= 0.733 [a]
HSPD1	NM_002156.4 :c.606 + 53 delGTTTGAG TTATCTGAT GATCAAAA	intronic deletion (-GTTTGAGTTATCTGA TGATCAAAA)	novel	no	***Del=0.371; no Del=0.629*** [d]

Allele frequencies are based on the following populations: a = HapMap CEU; b= CEPH; c= CAUC1; d= No publically available allele frequencies; allele frequencies are calculated from our sequencing efforts and are shown in bold italics. Del= deletions. Conserved: indicates regions of sequence conserved by alignment of 28 different species using USCS Genome Browser.

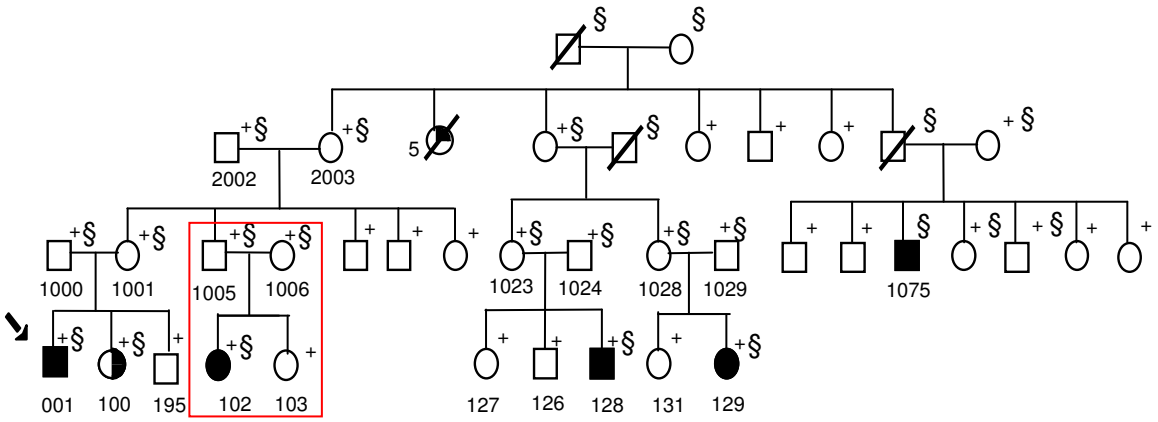
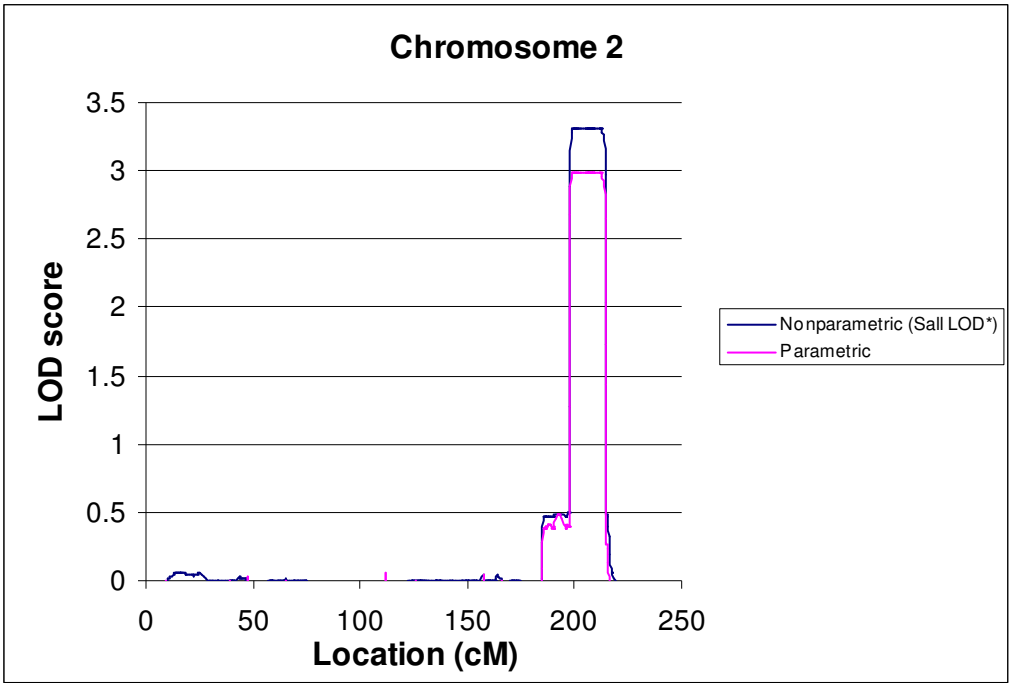


Figure 3.1 Pedigree diagram of NTD family 8776 with newly ascertained branch. Condensed pedigree diagram of the extended multiplex Caucasian NTD family 8776. Fully shaded symbols represent individuals affected with lumbosacral myelomeningocele. The half shaded circle for individual 100 represents congenital dermoid cyst. The quarter shaded circle denotes the 5 individuals who are affected with spina bifida by history alone, all of whom died in the first few days of life. + indicates individuals with DNA samples. § refers to individuals included in linkage analysis. Individual 1075 is reported to have lumbosacral myelomeningocele; however, the medical records and radiographic studies are not currently available to confirm this diagnosis. The red box denotes the recently ascertained branch. Individuals 128 and 129 are reported to have hydrocephalus consistent with Chiari II malformation.

A.



B.

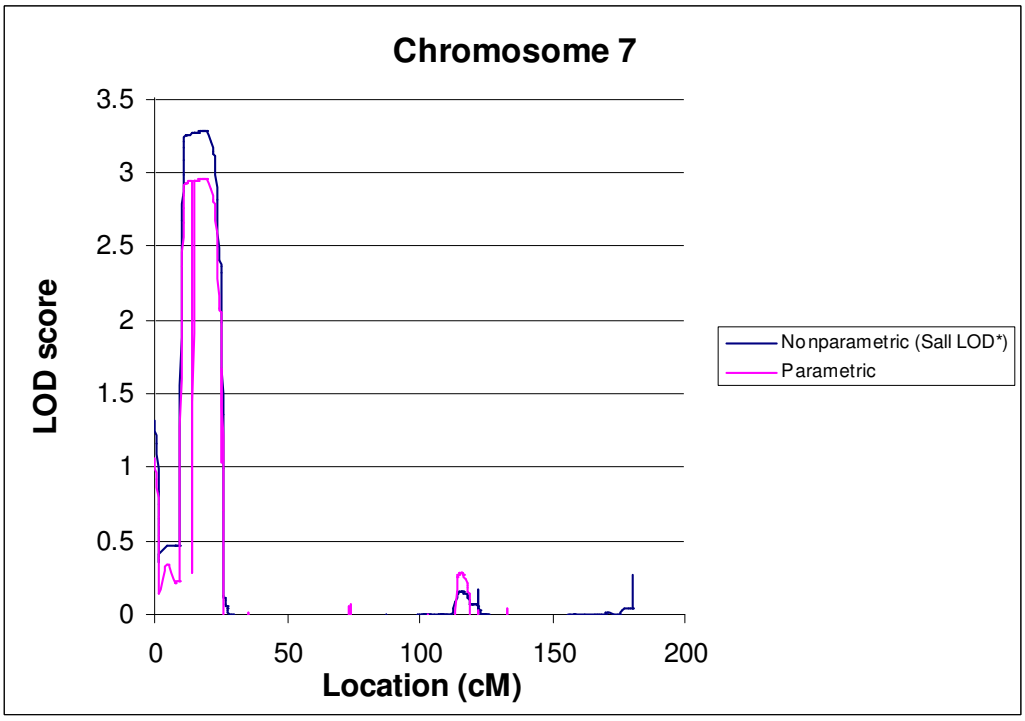
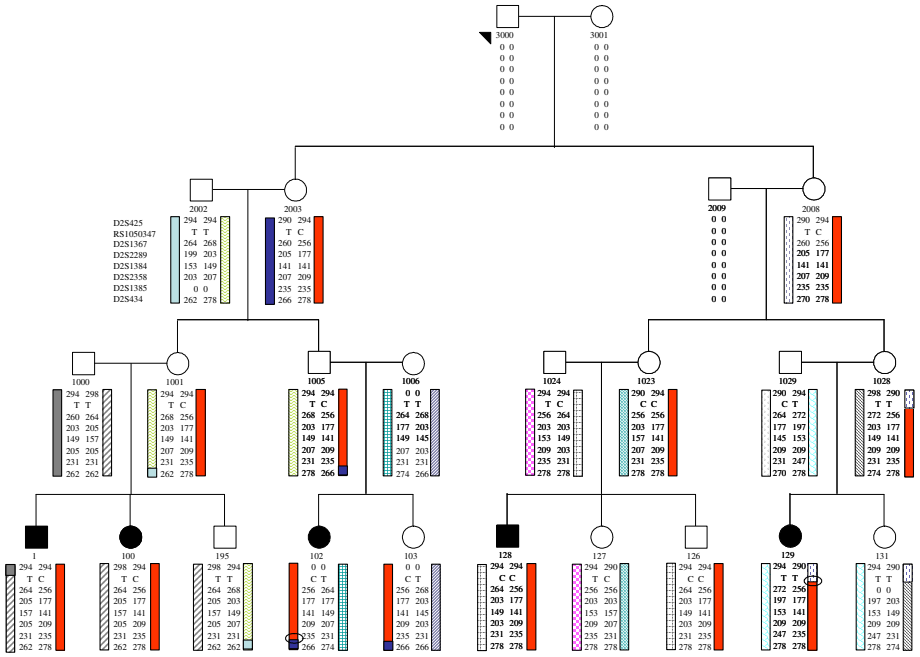


Figure 3.2 Multipoint LOD plots for chromosomes 2 and 7 in extended family 8776. Multipoint logarithm of odds (LOD) plots for chromosomes 2 and 7 in extended family 8776. Linkage results for chromosomes 2 (A) and chromosome 7 (B). SNP linkage results using parametric and nonparametric linkage methods are plotted. The genetic distance for the analyzed markers is given on the X-axis in cM.

A. Chromosome 2 Haplotype



B. Chromosome 7 Haplotype

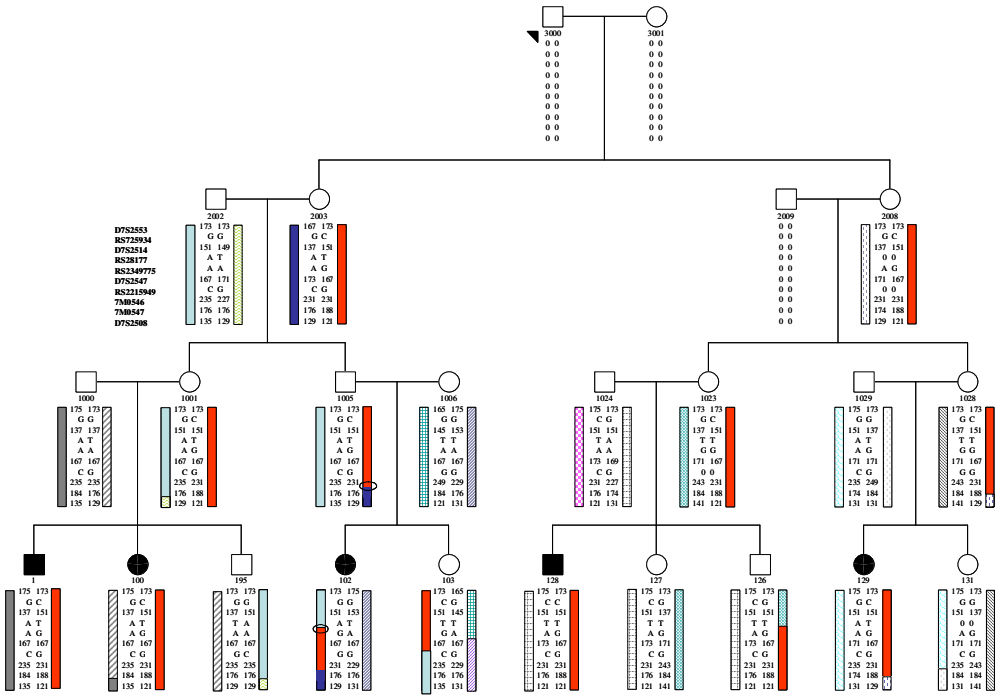


Figure 3.3 The 2q and 7p disease-associated haplotypes for family 8776.
The haplotype results are consistent with autosomal dominant transmission with reduced penetrance as shown with the red bar segregating with affection status. Only a subset of the markers is shown in these haplotypes. The circles surrounding the haplotype bars indicates where a critical recombinant haplotype resides that allowed narrowing of these minimum candidate intervals.

CHAPTER 4

Native American myopathy: Congenital myopathy with cleft palate, skeletal anomalies, and susceptibility to malignant hyperthermia

Demetra S. Stamm [1,2], Arthur S. Aylsworth [1], Jeffrey M. Stajich [2], Stephen G. Kahler[3], Marcy C. Speer [2], Cynthia M. Powell [1]

[1]University of North Carolina, Chapel Hill, NC; [2]Center for Human Genetics, Duke University, Durham, NC; [3] University of Arkansas for Medical Sciences, Little Rock, AR

Submitted to *American Journal of Medical Genetics*

4.1 Abstract

Native American myopathy (NAM) [MIM 255995], a putative autosomal recessive disorder, was first reported in the Lumbee Indians of North Carolina. NAM features include congenital weakness and arthrogryposis, cleft palate, ptosis, short stature, kyphoscoliosis, talipes deformities, and susceptibility to malignant hyperthermia (MH) provoked by anesthesia. This report documents the phenotypic complexity and natural history of this rare congenital disorder in fourteen individuals with NAM. Findings include a previously unreported 36% mortality by age 18. Based on this study, our conservative estimate for prevalence of NAM within the Lumbee population is ~2:10,000; however, birth incidence remains unknown.

4.2 Introduction

4.2.1 Lumbee population

The Lumbee people are a Native American group from the Lumber River region of south-central North Carolina [146]. The Lumbee have been culturally isolated for the past two centuries leading to frequent intra-tribe marriages. Currently, the Lumbee Tribe of North Carolina is estimated to have more than 45,000 members making them the largest tribe east of the Mississippi River [147]. The origin of the Lumbee people is not well documented. One theory is that the Lumbee tribe descended from Sir Walter Raleigh's Lost Colony of Roanoke Island [148].

4.2.2 Native American myopathy: Clinical Description

In 1987, Bailey and Bloch first described Native American myopathy [MIM 255995] in a 3 month old American Indian infant of Lumbee descent with multiple congenital anomalies including cleft palate, micrognathia, talipes equinus and arthrogryposis. The infant had an episode of malignant hyperthermia (MH) while under general anesthesia for placement of a gastrostomy tube [153]. Later, in 1988, Stewart, Kahler, and Gilchrist described six Lumbee children affected with congenital myopathy, skeletal anomalies, and susceptibility to malignant hyperthermia [146]. These 6 children ranged in age from 5 to 13 years old and two were sibs with one of the patients a cousin to this sib pair. They had normal baseline creatine kinase levels. Two of the children developed malignant hyperthermia during cleft palate repair. Clinical features of this disorder described in the literature include short stature, cleft palate, high- arched palate, myopathic facies (malar hypoplasia, ptosis), scoliosis, congenital joint contractures, and talipes deformities [146,153-156].

One of the greatest risks for patients with NAM is from MH crises during exposure to certain anesthetics. MH is a potentially life threatening pharmacogenetic disorder of the skeletal muscle that is triggered in genetically predisposed individuals exposed to inhaled halogenated anesthetics and/or depolarizing muscle relaxants. Symptoms of MH can include hyperthermia, muscle hyperactivity/hypermetabolism, rhabdomyolysis, tachycardia, tachypnea, rigidity, and metabolic acidosis. Succinylcoline and volatile anesthetics are often MH-triggering drugs in susceptible individuals [160]. If MH is not reversed, it can lead to severe tissue damage and death. Despite the high rate

of morbidity and mortality associated with NAM, little is known about the etiology and natural history of this condition.

4.2.3 Genetics of NAM

To date, the genetic etiology of NAM remains unknown. The Lumbee population is considered inbred based on a few common surnames and evidence of consanguinity within families (Powell, personal observation). Given the cultural isolation and high-rate of consanguinity in the Lumbee population, it is likely that the genetic basis of NAM is due to a founder effect. Founder effects can lead to a high frequency of recessive genetic disorders in an isolated or inbred population when members of the population harbor a disease-causing mutation inherited from a common ancestor. Because offspring of related parents from an isolated community share more alleles identical by descent, they will have a higher proportion of homozygous loci than the offspring of unrelated parents. These homozygous loci may harbor genetic mutations causing recessive diseases. Two examples of recessive disease in individuals of Lumbee ancestry described in the literature are glutaric academia type 1 [242] and hyper IgM syndrome [243]. Stewart and colleagues also mentioned that the mode of inheritance for NAM appeared to be autosomal recessive [146].

In this study, our aim is to further characterize both the natural history of the disease and the mode of inheritance for NAM.

4.3 Materials and Methods

4.3.1 Clinical Ascertainment

This study evaluated patients of Lumbee Indian descent affected with congenital muscle weakness who demonstrate two or more of the following phenotypic characteristics: myopathic facies, susceptibility to malignant hyperthermia, kyphoscoliosis and cleft palate. Subjects included in this study were diagnosed with NAM by clinical geneticists at Duke University or at the University of North Carolina at Chapel Hill; these two large medical centers are the primary providers of medical genetics services to the Lumbee people. Standardized family histories were conducted to provide information about potential inheritance patterns and whether consanguinity was present in the families. Clinical data describing the parents were obtained through recorded family history and clinical observations. Data on patients were gathered from clinical evaluations and medical record review from The University of North Carolina at Chapel Hill and Duke University. Data were collected in compliance with Duke University and the University of North Carolina Institutional Review Boards.

4.4 Results

4.4.1 Natural history of Disease

4.4.2 Mode of Inheritance and NAM pedigrees

Fourteen Lumbee individuals affected with NAM were included in this study, denoted as individuals 1-14 in Table 4.1. There was no gender discrepancy. The study consisted of 7 females and 7 males, supporting an autosomal mode of inheritance for

NAM. Ten detailed family histories were available and pedigree analysis revealed consanguinity within three families for individuals 4, 8, and 11.Individual 4's parents are 3rd cousins, individual 8's parents are 2nd cousins, and individual 11's parents are first cousins once removed. Individuals 3 and 4 are half –sibs with the same mother. The parents with clinical information available showed no evidence of NAM features (cleft palate, significant muscle weakness, talipes deformities, contractures). All NAM pedigrees show two unaffected parents of Lumbee descent having a child affected with NAM without any evidence of X-linked or dominant inheritance. Therefore, these pedigrees provide further evidence for NAM being an autosomal recessive condition, whereby the parents are carriers of a single copy of the recessive NAM mutation.

4.4.3 Cytogenetic Studies

Twelve patients had G-banded chromosome analysis of peripheral blood. Two patients (individuals 5 and 6) had a balanced interstitial paracentric inversion of chromosome 1 (q43q44). Individual 5's mother who is unaffected also was heterozygous for 46,XY inv(1)(q43q44). The ten other patients who had chromosome analysis performed were normal. As children with cleft palate or a submucosal cleft are often evaluated for 22q deletion syndrome [MIM 606232], four patients (individuals 3, 5, 10, 14) had testing for 22q deletion using fluorescence in situ hybridization *(FISH)* analysis. No deletion was present in any of the four tested.

4.4.4 Musculoskeletal and malignant hyperthermia

All the patients were reported to have congenital non-progressive myopathy with myopathic facies. Four out of the fourteen patients had MH episodes giving a 29% frequency of MH among this study group (see Table 4.1). These MH reactions took place during surgery for talipes deformities or cleft palate repair. In all four patients presenting with MH, the medical team managed them by discontinuing the provoking anesthetic, aborting the surgery, and administering dantrolene, and all had a positive outcome with 100% survival. However, individual 6's sister, who was not included in this current study for lack of documented clinical history, died at age 2 years from a malignant hyperthermia crisis during surgery for talipes deformity repair.

Muscle biopsies were performed on six patients to characterize the muscle pathology for diagnostic purposes. These biopsies showed non-specific myopathic changes with no evidence of mitochondrial abnormalities or central core disease. Additional diagnostic studies were done to evaluate muscle weakness including electromyography (EMG) and nerve conduction studies. EMG was performed on four patients: 3 were normal, and individual 6 had prominent distal myopathy which may be attributable to a motor vehicle accident injury. Nerve conduction studies performed on two patients were normal.

Scoliosis was observed in 9 out of the 14 patients (64%) as shown in Table 4.1; some individuals wore braces and others had surgical rod placement. Scoliosis progressed during both childhood and adult life causing increasing impairment of both lung function and mobility over time. Patients with braces had scoliosis that progressed despite this intervention. Talipes deformities were a common presentation in the NAM patients with

10/14 (71%) having talipes equinovarus as reported in Table 4.1. Congenital joint contractures were observed in 12 of the 14 patients (86%; see Table 4.1).

4.4.5 Dysmorphic features

All patients observed in this study had myopathic facies. Figure 4.1 shows some of those common myopathic features. A myopathic face appears expressionless, consisting of bilateral ptosis, down-turned corners of the mouth with inability to elevate the corners of the mouth, and open-mouthed posture. All of the patients had unusual eye morphology with a variety of manifestations: telecanthus, ptosis, short palpebral fissures, and/or downslanted palpebral fissures. Eight of the fourteen (57%) patients had ptosis (see Table 4.1). Two of the patients had posteriorly rotated ears. Seven patients had micrognathia as reported in Table 4.1. Four of the patients had facial hemangiomas in various sites including the forehead, glabella, nasal tip, eyelids, and upper lip. Palate abnormalities were present in all 14 patients evaluated. Nine out of the 14 (64%) children had posterior cleft palate, five had a high-arched palate, and three had probable submucosal cleft palate (Table 4.1). Older patients develop a long and narrow face as shown in Figure 4.1 (F).

4.4.6 Newborn and Toddler Period

Three of the 14 patients died in the first year of life (21% mortality rate). Causes of death included severe pulmonary hypoplasia (patient 1) and apnea secondary to acute fulminant enterococcal pneumonia (patient 3); the cause of death for patient 5 is

unknown. Of the eleven patients seen in the newborn period, one was macrocephalic with a head circumference at or above the 97th centile and three of the patients had a head circumference greater or equal to the 90th centile (Table 4.1). Three of seven patients who had brain MRIs showed evidence of enlarged ventricles. Gastrointestinal issues were common in the first years of life including feeding problems in 10/14 of the patients (71%), aspiration, and gastroesophageal reflux in 5 patients (36%) (Table 4.1). Feeding problems were secondary to oral hypotonia, cleft palate and micrognathia and often caused poor postnatal weight gain. Cryptorchidism was present in 5 of the 7 males (71%). All patients who survived the first few months of life had delays in motor development, likely due to their muscle weakness as shown in Figure 4.1C, where individual 2 requires support for ambulation at 2 years of age. Distal contractures and equinovarus also contribute to delayed ambulation. All patients seen in the newborn period had a myopathic face. Figure 4.1A and 1B show Individuals 9 and 2 at three months of age having myopathic facies with ptosis and open, down-turned mouth.

4.4.7 Childhood

Despite motor delays in the infant and toddler period, all patients seen later in childhood had normal intelligence. Expressive language delay and hypernasal speech attributable to cleft palate, residual palatal fistulae, velopharyngeal insufficiency and oral hypotonia was still present in some children. By early childhood, head circumferences were closer to the mean than during the newborn period.

4.4.8 Adulthood

Two adults with NAM (individuals 6 and 13) had short stature with heights below the 3rd centile. These two adults had normal intelligence and both completed high school. Most patients during later childhood and adulthood were cachectic with significant muscle wasting. In adulthood, the face is long with a prominent chin as illustrated in one adult patient shown in Figure 4.1 F. Despite individual 6's severe development delay as a child (walked at 5 years of age and had language delay), as an adult he maintained a job as a barber for many years. This patient, who is currently 40 years old, is the oldest patient in this study. He has mild scoliosis. His back pain has become worse over the past few years; he has a foot drop, and now requires a walker to ambulate. It should be noted that in addition to having features of NAM, this patient was involved in a motor vehicle accident during adulthood. It is unclear whether his chronic pain and foot drop are secondary to the accident or his underlying NAM. Patient 13 is now 30 years old, has chronic respiratory failure with a permanent tracheostomy, and requires overnight continuous mechanical ventilatory support. Even though individual 13 has respiratory compromise, she delivered a normal healthy baby with low forceps delivery in 2001[155].

Progressive scoliosis appears to be a major cause of morbidity in adolescents and adults with NAM, contributing to restrictive lung disease and pulmonary insufficiency. The only individual to die after childhood was individual 11. This was due to respiratory compromise with decreased lung volume from severe kyphoscoliosis and exacerbated by tobacco use. Two other individuals 4 and 5 died of unknown causes. Overall, these fourteen individuals with NAM had by age 18 a 36% mortality rate (5/14 patients).

4.5 Discussion

In this study, we report shared clinical features among 14 individuals of Lumbee Indian descent with Native American myopathy. Many of the features of NAM are likely due to underlying muscle weakness in utero and postnatally. For example, the neonatal arthrogryposis observed in patients with NAM may be caused by hypotonia and hyopokinesia in utero. The neonatal respiratory failure is likely to be secondary to one or all of the following: hypotonia, micrognathia, glossoptosis, gastroesophageal reflux, and aspiration. The feeding problems and failure to thrive may be secondary to the patient's oral hypotonia, micrognathia and cleft palate. Feeding problems may be ameliorated in patients with special cleft palate nipples and bottles, feeding therapy, and treatment of gastroesophageal reflux. The chronic respiratory compromise seen in adults is likely due to their myopathy and severe scoliosis that leads to restrictive lung disease.

MH is a well known risk factor among the Lumbee Indian population [146,153,154]. Within our study, we found that four out of 14 patients (29%) had MH episodes, a similar frequency to two out of six patients (33%) observed in 1988 by Stewart and colleagues [146]. However, the true frequency of susceptibility to MH among these patients is likely to be much higher because the patients without episodes of MH either had no history of surgery or did not receive a MH-provoking anesthetic.

Anecdotal evidence suggests an increased frequency of congenital musculoskeletal anomalies [154] and a higher susceptibility to malignant hyperthermia in Native Americans of Lumbee background [153]. For example, in 1989 Meluch and colleagues reported a case study of an American Indian male of Lumbee descent who

underwent frontal craniotomy for resection of a pilocytic astrocytoma [154]. Subsequent surgeries for placement of ventriculoperitoneal shunts with halothane and isoflurane anesthesia triggered malignant hyperthermia crises. To address the MH susceptibility risk within the Lumbee people, clinicians in the Lumberton region and surrounding areas often will use non-halogenated non-volatile anesthetic management for their Lumbee patients (personal communication). Furthermore, the above case report of a Lumbee male did not describe the patient as having any of the common features of NAM such as muscle weakness, skeletal abnormalities, or dysmorphic facies. This finding suggests that the Lumbee people may have an isolated risk of MH without the presence of congenital myopathy. Two possible explanations for MH without congenital myopathy seen in the Lumbee population are: 1) individuals with isolated MH are heterozygous for the NAM genetic mutation whereas being homozygous for this mutation leads to both the congenital syndrome and MH; or 2) susceptibility to malignant hyperthermia is due to another genetic variant that is present in this population.

The occurrence of both malignant hyperthermia and myopathy is not unique to Native American myopathy. MH crises have been documented in patients with other myopathies such as myotonia fluctuans, Duchenne and Becker dystrophy, myotonia congenital, and myotonic dystrophy [244]. Malignant hyperthermia in patients with other musculoskeletal anomalies is not uncommon. A review by Strazis and Fox of 503 published MH cases identified a much higher frequency of congenital anomalies such as cleft palate, club foot, scoliosis, ptosis, and cryptorchidism compared to the general surgical population [245].

The frequency of NAM within the Lumbee Indian population is unknown. We can make a minimum estimate of the current NAM prevalence based on our collected population, if we consider that at least 9 individuals affected with NAM are alive and the Lumbee tribe is estimated to be ~ 45,000 [147]. This estimated minimum prevalence of NAM within the Lumbee population would be 9/45,000 or 2/10,000. The true prevalence of NAM, however, is likely to be much higher since many patients with NAM may not have been identified through our study. Conversely, we may have underestimated the population size of the Lumbee tribe leading to a marginal inflation of estimated prevalence. The high degree of infant mortality (21%) suggests that the birth incidence is likely much higher than the Lumbee population prevalence. Underestimation of prevalence may also be due to phenotypic heterogeneity of NAM leading to non-inclusion in the study of individuals who did not meet our phenotypic criteria.

The prevalence of NAM may also have been underestimated if affected individuals are diagnosed with other conditions. Since many of the clinical features of NAM resemble other congenital myopathies, it is likely that patients with NAM may have been diagnosed with another similar disorder such as King Syndrome, Noonan syndrome, or other myopathies. King Syndrome (KS) [MIM 145600] is an autosomal dominant disorder characterized by dysmorphic facies (low-set ears), cryptorchidism, short stature, and susceptibility to MH. King Syndrome and NAM are both phenotypically heterogeneous with a myriad of clinical presentations, but the classic features for both are the congenital myopathy and malignant hyperthermia. While patients with KS have many similar clinical features of NAM, they rarely have cleft palate, contractures, or clubbed feet [146]. Noonan syndrome [MIM 163950] is an

autosomal dominant dysmorphic syndrome characterized by hypertelorism, ptosis, low-set posteriorly rotated ears with thickened helices, and a small pointed chin. While individuals with Noonan syndrome also have short stature and dysmorphic features that resemble NAM they frequently have structural cardiac defects not seen in NAM and typically do not have congenital myopathy, cleft palate or congenital joint contractures as seen in NAM. Froster-Iskenius and colleagues described four children with similar features to NAM such as myopathic face, congenital contractures, ptosis, cleft palate, and susceptibility to malignant hyperthermia, but these patients also had congenital torticollis and neck and axillary webbing, which are not observed in NAM [158]. Some of the other diagnoses considered by physicians during evaluation of the patients in our series included: central core disease, neurogenic arthrogryposis, Freeman-Sheldon syndrome, amyoplasia, Trisomy 18, fetal akinesia sequence, Christian adducted thumb syndrome, and Gordon Syndrome ([MIM 114300] autosomal dominant distal arthrogryposis with cleft palate and club foot) [246].

Cytogenetic studies were conducted for most patients in this study, in order to identify other potential causes for their multiple congenital anomalies. FISH analysis for 22q deletion syndrome on four patients revealed no deletion. Two patients showed apparently balanced interstitial inversions of chromosome 1 (q43q44). The ten other patients who had chromosome analysis performed were normal. This inversion could be interpreted either as an unrelated finding or disrupting the causative gene region. Based on the findings that individual 5's unaffected mother is also heterozygous for the inversion and the lack of this inversion in other patients with NAM, it seems more plausible that the balanced inversion observed in two patients is not responsible for the

NAM phenotype and is most likely a benign polymorphism in the Lumbee population. Alternatively, NAM may result if this inversion causes a breakpoint in the putative NAM gene, and the other chromosome carries a point mutation. If this were the case, the other ten individuals with NAM who do not carry the inversion would require two copies of the other point mutation.

Native American myopathy is a putative autosomal recessive disorder. To clarify the inheritance patterns for NAM we examined several NAM pedigrees that all support autosomal recessive inheritance by the following findings: 1) None of the affected patients has an affected parent; 2) The majority of affected patients' sibs are unaffected 3) equal ratio of affected females to males; 4) evidence of known consanguinity between the parents for some patients; and 5) all patients have parents of Lumbee ancestry.

The high degree of consanguinity present within the population means it is plausible that any two Lumbee individuals affected with a recessive disease share genetic material from a common ancestor. Thus, this condition is an excellent candidate for homozygosity mapping to identify the NAM loci or gene [186,187]. Overall, the identification of the NAM disease locus or gene mutation could lead to understanding the molecular mechanisms behind the disease and provide potential genetic testing for Lumbee families.

4.6 Summary

In this study, we report shared clinical features among 14 individuals of Lumbee Indian descent with Native American myopathy. All of the patients exhibited dysmorphic features including unusual eye morphology with ptosis, and/or downslanting

palpebral fissures. During the newborn period, numerous clinical features were commonly observed including congenital joint contractures, equinovarus, cleft palate, micrognathia, significant feeding difficulties, and cryptorchidism. Intellectual development was normal but motor development delayed. In childhood, additional clinical features were documented such as MH-crises, and severe scoliosis. During the newborn period the mortality rate was documented to be 21%, and if extended to adulthood there was an overall 36% mortality rate. Pedigree inspection provided evidence in support of an autosomal recessive inheritance pattern for this condition, and the extensive consanguinity observed in these pedigrees suggests that a homozygosity mapping approach may be productive for gene identification studies.

4.7 Acknowledgements

The authors would like to thank the patients and their family members. This work would not be possible without their participation. We gratefully acknowledge support from Muscular Dystrophy Association.

Table 4.1 Clinical description of patients with NAM

Patient number	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
Natural history														
sex	F	M	F	M	M	M	F	F	F	M	M	M	F	F
age at diagnosis	1 d	2y 4m	84 d	2 d	1 d	30 y	22 d	1 d	6 d	1 d	14 y	1 d	NK	9 d
birth weight*	3-10%	<3%	10%	10%	<3%	50%	3-10%	<3%	3-10%	3-10%	50%	3-10%	NK	<3%
birth length*†	NK†	3-10%*	10%*	10-25%*	<3%*	NK	10-25%*	75%*	50%*	75%*	NK	NK†	NK	3-10%*
head circumference*‡	75%*	50%*	25%*	90-97%*	10%*	90%‡	25%*	50%*	90%*	90%*	90%‡	>97%*	25%‡	<3%*
delayed motor development	NK	+	NK	+	NK	+	+	+	+	+	+	NK	NK	+
feeding problems	NK	+	+	NK	NK	+	+	-	+	+	+	+	+	+
GE Reflux	NK	-	-	NK	+	+	+	NK	NK	+	NK	NK	+	NK
malignant hyperthermia	-	+	-	+	-	+	-	-	-	-	-	-	+	-
deceased (age)	1d	No	5 m	< 6 y	1 m	No	No	No	No	No	18 y	No	No	No
Dysmorphology														
Myopathic face	+	+	+	+	+	+	+	+	+	+	+	+	+	+
ptosis	+	+	-	+	NK	+	+	+	+	NK	+	NK	NK	-
downturned mouth	-	+	+	NK	NK	NK	+	+	+	NK	NK	NK	+	NK
Cleft palate	+	HAP, SMC?	+	+	+	HAP	HAP, SMC?	+	+	+	HAP, SMC?	HAP	+	+
micrognathia	-	+	+	-	+	-	-	+	-	+	-	+	-	+
Muskoskeletal features														
congenital joint contractures	+	+	+	+	+	NK	+	+	+	+	+	+	NK	+
equinovarus	+	+	+	-	+	+	+	+	+	+	-	+	NK	-
congenital scoliosis	NK	+	NK	+	NK	+	+	+	+	NK	+	+	+	-

NK (Not Known)

* weight, length, and head circumference based on fetal-infant growth chart for preterm infants and reflect percentiles in the newborn period.

† length measurement could not be obtained because feet were at her ears due to severe hip flexion.

‡ represents head circumference percentiles based on Nelhaus, G., Composite International and Interracial Graphs, PEDIATRICS 41: 106, 1968.

CP (Cleft palate); SMC? (Probable Submucousal Cleft); HAP (High-arched palate)

Figure 4.1 Observed clinical features of patients with NAM

A. Patient 9 at three months of age. Note ptosis, myopathic face, and cleft palate.

B. Patient 2 at three months of age. Note myopathic face, forehead, nasal and philtrum hemangioma, and congenital talipes equinovarus.

C. Patient 2 at age 2 years. Note myopathic face, down-turned open-mouthed posture, and inability to stand unassisted.

D. Patient 2 at age fourteen showing cachectic appearance with muscle wasting and camptodactyly demonstrated with attempt at finger extension.

E. Patient 8 at age ten years of age. Note ptosis, myopathic facies and camptodactyly.

F. Patient 6 at age thirty years old. Note long face and prominent mandible.

CHAPTER 5

Genomewide SNP chip homozygosity mapping defines critical region in singleton Native American myopathy families

5.1 Abstract

Native American myopathy (NAM) is an autosomal recessive congenital myopathy first reported in the Lumbee Indian people. Features of NAM include congenital weakness, cleft palate, ptosis, short stature and susceptibility to malignant hyperthermia. We identified five individuals with NAM from the Lumbee population, and hypothesized that these affected individuals have disease alleles shared identical-by-descent (IBD) inherited from common ancestry. To identify a NAM disease locus, homozygosity mapping methods were employed on a genomewide 10K single-nucleotide polymorphism (SNP) screen. To confirm regions of homozygosity identified in the SNP screen, microsatellite repeat markers were genotyped within those homozygous segments. The SNP data demonstrated five regions of shared homozygosity in individuals with NAM. The additional microsatellite genotyping data narrowed the region to one common segment of homozygosity spanning D12S398 to rs3842936 (5.6 Mb) mapping to 12q13.13-14.1. This study reports the first gene mapping of NAM using SNP-based homozygosity mapping in only a few affected individuals from simplex families and identified a putative NAM locus. Identifying the genetic basis of NAM may suggest new

genetic etiologies for other more common conditions such as cleft palate and malignant hyperthermia.

5.2 Introduction

The Native American myopathy [MIM 255995] is a congenital myopathy first described in the Lumbee people, a multi-racial group indigenous to Robeson County in south-central North Carolina. In 1987, Bailey and Bloch first described Native American myopathy in a 3-month old American Indian infant of Lumbee descent with multiple congenital anomalies including cleft palate, micrognathia, talipes equinus and arthrogryposis [153]. Later, Stewart et al described NAM in six Lumbee children and highlighted malignant hyperthermia as a feature of NAM that warrants medical concern in this myopathy [146]. Malignant hyperthermia (MH) is a potentially life-threatening pharmacogenetic disorder of the skeletal muscle that is triggered in genetically susceptible individuals exposed to inhaled anesthetics and/or depolarizing muscle relaxants. Though clinical presentation of NAM is relatively heterogenous, some other typical features of NAM include muscle wasting and kyphoscoliosis secondary to congenital weakness, contractures, cleft palate, abnormal facies with ptosis and low-set ears, and talipes deformity [146,153-156].

The Lumbee population is a relatively isolated population that is considered inbred based on a few common surnames and evidence of consanguinity within families (personal observation, C.P.). NAM pedigrees suggest an autosomal recessive pattern of inheritance, and clinical documentation in a few families shows unaffected parents having more than one child with affected with NAM (personal observation, C.P.). The

high degree of speculated consanguinity within this isolated population suggests that Lumbee individuals affected with NAM inherited identical genetic material containing the disease gene from a common ancestor. This phenomenon, whereby a couple transmits the same ancestral chromosomal segment to one of their offspring, is termed autozygosity. Based on this rationale, homozygosity mapping may identify regions of autozygosity within the NAM families and facilitate mapping the NAM locus.

Homozygosity mapping identifies genomic segments shared identical-by-descent among patients with recessive diseases from populations with limited migration into or out of the population and/or with high rates of consanguinity, allowing characterization of a genomic region that harbors the causative genetic mutation [186,187]. Homozygosity mapping is a form of linkage disequilibrium (LD) mapping whereby the disease-causing gene is in LD with a conserved haplotype that arises from founder effects [187]. Founder effects occur when only a small number of individuals bring a chromosome with a disease mutation into a small isolated population and a subsequent increase in this population size will lead to a large proportion of individuals carrying this founder chromosome. Recombination events in this population over time lead to smaller segments of the original founder chromosome containing the disease gene, thereby facilitating productive homozygosity mapping studies. Since the development of dense whole genome SNP arrays, several rare recessive Mendelian loci and genes have been mapped [192,193,247] using homozygosity mapping approaches. We reasoned that whole genome SNP homozygosity mapping might facilitate identification of the NAM locus.

In this study, we conducted homozygosity mapping in NAM to identify a disease-associated haplotype in which only simplex NAM families (as opposed to multiplex families) are available.

5.3 Materials and Methods

5.3.1 Clinical Description

The typical clinical presentation of NAM includes abnormal facies with ptosis and low-set ears, muscle wasting and kyphoscoliosis. NAM subjects were of Lumbee Indian descent with congenital muscle weakness who demonstrate two or more of the following phenotypic characteristics: myopathic facies, susceptibility to malignant hyperthermia, kyphoscoliosis, and cleft palate. Based on these criteria, five Lumbee families were included in this study. These families are from North Carolina and the probands were ascertained from the University of North Carolina at Chapel Hill's Genetics Clinic and from Duke University's Muscular Dystrophy Clinic. From these five families, 17 DNA samples were available. All data and samples were collected following informed consent of subjects; this study was approved by the Duke University Medical Center Institutional Review Board. All families include at least an affected child and unaffected mother (family 2281). Family 3012 also included a sampled unaffected father of the case, and families 2118, 3008, and 3013 include a sampled unaffected father and an unaffected sibling of the proband. These families report no connection to one another within three degrees of relationship to the proband.

5.3.2 Genotyping Methods

10K SNP Chip

DNA was extracted from whole blood using the Puregene system (Gentra Systems, Minneapolis, MN). Genotyping was performed on the Affymetrix 10K version 2 Chip using the standard Affymetrix protocol (Affymetrix, Santa Clara, CA) on all 17 individuals from the five families. Subjects' DNA was treated with standard GeneChip Mapping 10K *Xba*I Assay protocol as described previously [191]. Data acquisition was performed by using the GeneChip GCOS software. Data were checked for Mendelian errors and an inheritance check was performed using the software GeneSpring GT v1.0 (Silicon Genetics). SNP data were evaluated for Mendelian inconsistencies using PedCheck [232] and SNPs demonstrating non-Mendelian inheritance were removed from the analyses.

5.3.3 Homozygosity Mapping and Copy Number Variant Analysis

The 10K SNP genotyping data for the families were imported into an Oracle database. Once the data were merged into the database, SQL scripts were developed to query the Oracle database for homozygous SNPs shared amongst the individuals with NAM. Regions with ≥ 3 consecutive homozygous SNPs were selected for follow-up with microsatellite repeat markers.

SNP data available from the Affymetrix 10K SNP Chip was also used to evaluate for potential copy number variants (CNV) and loss of heterozygosity (LOH). The .cel files that provide intensity information for a given probe on an array were utilized in two different analysis tools (DChip (http://biosun1.harvard.edu/complab/dchip/) and

CNAT4.0 (http://www.affymetrix.com/products/software/specific/cnat.affx)), each of which makes two comparisons: comparing copy number changes between cases and controls and characterizing loss of heterozygosity. In each case the SNPs were first viewed in groups of 5 in both a smoothed median and a Hidden Markov model. The SNP threshold was then gradually lowered to a single SNP using both algorithms.

5.3.4 Fine Mapping with Microsatellites

One or two polymorphic microsatellite markers that reside in each region of shared SNP homozygosity were selected to confirm each of the five regions. Once a homozygous locus was verified with the first pass of microsatellite genotyping, additional microsatellite markers were selected to narrow the minimum interval of shared homozygosity. Microsatellite markers were selected based on the highest heterozygosity (defined by Centre d'Etude du Polymorphisme Humain (CEPH) Genotype database, http://www.cephb.fr/cephdb/php/eng/search.php), spacing within the interval and the number of repeats. The seventeen DNA samples from these families were organized into a list with a standardized order of samples to ensure the person performing the experiment was blinded to gender and family composition. Quality control (QC) standards were included at specified location and randomly selected duplicated individuals were used as internal QC's. DNA from patients' blood samples was amplified with the microsatellite primers. Genotyping was performed with 10 ng of DNA per reaction, utilizing a genotyping method on acrylamide gels previously described [248].

5.3.5 Sequencing of Candidate Genes

UCSC's "Known Genes" track was used to identify genes in the 12q region of interest. Primers were designed using Primer3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) with genomic sequence from UCSC's bioinformatics website. We sequenced a subset of the DNA samples using the five case samples from 2118 and 2281 because we expected all affected individuals would exhibit the same genetic mutation for NAM. PCR product was purified using Sephadex G-50 fine (Amersham Biosciences, Piscataway, NJ) and QuickStep2 SOPE Resin (Edge BioSystems, Gaithersburg, MD). The samples were then sequenced in both forward and reverse directions using the BigDye Terminator v3.1 Cycle Sequencing Ready Reaction kit (Applied Biosystems, Foster City, CA) and purified using Sephadex G-50 fine. Sequence analysis was performed using ABI 3100 Data Collection Software Version 1.01 and ABI Sequencing Analysis 3.7. PCR products covered all the exons and respective intronic boundaries. The primers were designed to flank the exons and covered 25-50 base pairs of sequence flanking the exons to include potential splicing regions. All sequences for candidate genes *ITGA7*, *PIP5K2C*, *PDE1B*, and *MLC1sa* were visually inspected with this software to identify noteworthy variants. The sequence was compared to unaffected parents to ascertain if any genetic variant detected was transmitted or a de novo mutation.

5.4 Results

5.4.1 Genetic Mapping

The SNP screen identified five regions of homozygosity across 3 or more contiguous SNPs in the five NAM affected individuals (Table 5.1). The longest homozygous haplotype length and largest number of contiguous homozygous SNPs was observed at 12q13.13-14.1.

Coincidentally, a balanced paracentric inversion of chromosome 1 (q43-44) was observed in the proband of family 2118 (C. Powell, personal communication). Affected individuals without this inversion, however, do not share homozygous SNPs around this region of chromosome 1q, suggesting that this inversion is unlikely to contain the NAM gene.

Ten microsatellite markers were selected to cover the five regions identified by the SNP screen. These ten markers are described in Table 5.2. The microsatellite repeat marker data narrowed the region of shared homozygosity to one region on chromosome 12. The other four regions in chromosomes 2, 3, 5 and 6 were eliminated as potential regions of interest by heterozygous calls observed for the polymorphic microsatellite markers genotyped internal to the regions in the five patients with NAM. The remaining region of shared homozygosity maps to 12q13.13-14.1 spanning a total homozygous segment of 5.6 Mb flanked by D12S398 and RS3842936 markers, a 0.3 Mb decrease from the SNP homozygosity mapping (Table 5.1). The 12q haplotypes for the five NAM families are shown in Figure 5.1. All three sampled unaffected siblings from families DUK 2118, 3008, and 3013 did not share any region of the 12q homozygous haplotype, and only one unaffected individual (100 from family 3008) was heterozygous for the

same alleles on the 12q haplotype as the affected proband (Figure 5.1). The overall results confirm homozygosity in region 12q13.13 to 12q14.2 in the five individuals with NAM.

No significant copy number variants were detected. Whole genome array-CGH, providing approximately 50-100K resolution, was performed on DNA from individual 1 in family 2281 and no deletions in the five SNP segments described in Table 5.1 were noted. Loss of heterozygosity (LOH) estimates from CNAT4.0 identified a long stretch of LOH in SNPs at 12q supporting our findings from the homozygosity mapping.

5.4.2 Candidate Genes

The 12q13.13-14.1 chromosomal region is gene dense and contains approximately 135 known genes (UCSC Genome Browser http://genome.ucsc.edu/; NCBI Build 35). These genes were evaluated by a literature search using Online Mendelian Inheritance of Man (OMIM) and PubMed. Four candidate genes were selected for sequence analysis based on their biological plausibility. *Integrin α-7* (*ITGA7*) mutations have been shown to cause a human congenital myopathy [249]. *MLC1sa* is a myosin with a hexameric ATPase cellular motor protein that is expressed in both slow-twitch skeletal muscle and non-muscle tissue. Two genes involved in calcium regulation were of interest because they might explain the molecular mechanisms for both the myopathy and malignant hyperthermia. These were *PIP5K2C*, a protein involved in actin cytoskeleton and phosphatidylinositol signaling, and *PDE1B*, a calcium/ calmodulin-stimulated cyclic nucleotide phosphodiesterase. All four genes were sequenced (*ITGA7, MLC1sa, PIP5K2C,* and *PDE1B*) but no coding mutations were

detected. A known polymorphism for *MLC1sa* was detected in the mother (individual 1001) for DUK family 2281. She is heterozygous for the known deletion/insertion polymorphism rs3841624. The other four individuals genotyped did not exhibit this polymorphism.

5.5 Discussion

In this genetic mapping study of NAM, whole genome SNP genotyping identified five shared homozygous segments in affected cases, with chromosome 12 having the largest homozygous segment. Fine mapping with microsatellite repeat markers eliminated four of the regions of interest and confirmed and narrowed our focus to a 5.6 MB region spanning D12S398 to RS3842936 at 12q13.13 to 12q14.1. The 12q microsatellite data provided an elegant follow-up to the SNP data, as most of the parental polymorphic microsatellite data were heterozygous, allowing informative tracing of the allelic transmission patterns that contributed to the homozygous haplotype in the affected individuals. This shared homozygous haplotype supports our hypothesis of autozygosity for NAM.

While these NAM mapping results provide evidence for a NAM disease gene on chromosome 12q, it is possible that this region represents chance homozygosity. For instance, a homozygous genomic segment may not be autozygous because two copies of the genomic segment may be identical by chance mutation and recombination rates [250]. Moreover, estimates of chance IBD sharing within the Lumbee population are not possible at this time, because the expected allele frequencies for the Lumbee population

are currently unavailable, and it is challenging to correctly specify the degree of relatedness or kinship coefficient for the NAM pedigrees. Despite having historical pedigrees for these five families, most of them only have pedigree structure extending to the grandparental generation, and it is probable that a significant founder effect within this population and unaccounted familial relatedness would lead to an underestimate of the extent of inbreeding. Moreover, homozygous segments have recently been identified in apparently outbred populations that are also thought to be from ancestral haplotypes [144,188-190]. This further limits our ability to calculate chance homozygosity because a certain degree of IBD sharing may be present within the general population that would be difficult to model at this time.

Several homozygosity mapping studies have successfully identified a relevant disease gene [192,193,247,251], while others have observed allelic heterogeneity within consanguineous kindreds and therefore failed to map the underlying disease mutation [194-196]. Therefore, we cannot entirely exclude the possibility that the homozygosity observed at 12q13.13-14.1 is unrelated to the NAM phenotype.

Once the narrowed homozygous 12q locus was characterized, we began identifying and testing candidate genes. Mutational screening was performed on these four biologically relevant candidate genes: *ITGA7, PIP5K2C, PDE1B,* and *MLC1SA*. Since NAM is considered a Mendelian disease, we expected to observe an obvious mutation leading to a functional change such as a premature termination codon, a frameshift mutation(s) or a missense mutation(s), and in the genes evaluated such functional mutations were not observed.

Individuals with NAM have a higher susceptibility to MH, a process appearing to result from abnormal calcium ion balance in skeletal muscle. This suggests that the NAM gene may affect calcium regulation. Predisposition to malignant hyperthermia is typically inherited as an autosomal dominant trait and has extensive genetic heterogeneity with six different malignant hyperthermia susceptibility (MHS) loci, designated as MHS1-6 [29-37]. Two MHS genes have been identified, including *RYR1* and *CACNA1S*. MHS1 maps to 19q13.1 and is caused by mutations in the *RYR1* gene [167,168]. MHS5 maps to the 1q32 locus [175], and a putative causative mutation in the *CACNA1S* gene was identified in a single large French family [176]. MHS2 maps to 17q11.2 - q24 [170], and three families have mapped to the *SCN4A* gene but no mutation has been identified [165,171]. The other MHS loci map to the following locations: MHS3 to 7q21-22 [172,173], MHS4 to 3q13.1 [174], and MHS6 to 5p [175]. None of these MHS loci are in the homozygous regions identified in our SNP screen, suggesting a novel locus for the gene responsible for NAM. Because malignant hyperthermia is part of the phenotypic profile of NAM, we searched the 12q region for genes homologous to the MHS-associated genes *RYR1, CACNL2A*, and *SCN4A*. We also searched for homology to the protein tyrosine phosphatase, non-receptor type II *(PTPN11*) gene that maps to 12q24.1, a gene associated with Noonan syndrome [MIM 163950] [252]. Noonan syndrome is an autosomal dominant disorder, and affected individuals have short stature and dysmorphic features that resemble NAM. However, patients with Noonan syndrome frequently have structural cardiac defects not seen in NAM and typically do not have congenital myopathy, cleft palate or congenital joint contractures as seen in NAM. No homology was found in the region of interest on chromosome 12 for any of these genes.

The 12q region is very gene dense, with well over 100 annotated genes in the 5.6 Mb segment. Because of this, identifying the NAM gene may be a challenging undertaking. Future efforts to identify the NAM gene might include ascertaining additional families to identify recombinant individuals, sequencing across the entire segment, or using stored muscle biopsies from NAM patients to identify candidate genes based on histological findings or RNA expression profiles.

In addition to homozygosity mapping, the SNP data were used to screen for CNVs and LOH. Although no deletions or amplifications were identified, it is still possible a smaller deletion exists on the 12q homozygous background that would be detected with denser arrays. Importantly, the LOH predictions for the SNP chip data from the five individuals with NAM also identified 12q as having loss of heterozygosity. Therefore, this LOH approach could be used in other studies as a quick screening method to predict regions of shared homozygosity that then could be validated using more polymorphic markers such as microsatellites.

In summary, we utilized whole genome SNP-array technology to identify regions of shared homozygosity between five Lumbee individuals with NAM. The SNP results were followed up by genotyping microsatellites to confirm and narrow the homozygosity to one critical region. Identification of the underlying genetic basis of NAM may lead to better treatments and presymptomatic identification of at-risk individuals for malignant hyperthermia. Additionally, identification of the NAM gene may suggest new genetic etiologies for phenotypic components of other conditions such as cleft palate and malignant hyperthermia. This study demonstrates how SNP screens can be used in

homozygosity mapping for rare Mendelian recessive diseases even when only one affected individual is available per family.

5.6 Acknowledgements

The authors would like to thank the patients and their family members for their generous contributions. This work would not be possible without their participation. The authors thank Ms. Bei Zhao and Ms. Jianzhen Xie for technical assistance. We also would like to thank Jessica Connelly, Ph.D. and Simon Gregory, Ph.D. for their technical expertise with the whole genome array-CGH. The authors thank Karen Mohlke, Ph.D. and Kari North, Ph.D. for reading the manuscript and providing helpful comments. We gratefully acknowledge support from the Muscular Dystrophy Association.

Table 5.1 Shared homozygous segments observed in affected individuals identified by SNP chip.

Region	SNPs flanking region	Size (Mb)	# Contiguous homozygous SNPs	Physical Distance (Mb)
2p11.2	rs725046 – rs1074883	2.8	6	86.3 -89.1
3q21.3	rs977683 – rs725132	1.8	4	128.2-130.0
5q13.2-q13.3	rs277939 – rs1363577	3.6	5	70.9-74.5
6q22.1	rs3897507 – rs339346	1.1	3	116.1-117.3
12q13.13-q14.1	rs373374 – rs3842936	5.9	18	51.2-57.1

Table 5.2 Information on microsatellites used to evaluate regions of shared SNP homozygosity.

Marker ID	Chr	Location (bp)	Forward Primer	Reverse Primer	HET$^{\alpha}$
D2S417	2	86899572	AGATGCTTGTGGTTTAGATGG	ATAAAAGCATCTTTTACTGGCTG	72.34%
D2S2216	2	88248932	GGGGTAAATTATTTCCAGCC	TCCCTGACAAAGTGGGC	60.71%
D3S3607	3	128682908	ACAGGCTCAGGCCCAC	CTAAGAATAGCTTCCCAGAAACGG	64.29%
D5S1982	5	73322387	AGACAGAATCTGGGGGCA	TCTCTGGGCTGCTGGA	69.23%
AFM268VH5	6	117775687	GCTCTGATAAATTTTTGAAAAGTG	TGAGTATTAATTTCACTACATCCCT	N/A
D12S398	12	51483266	GGTGGATGGATGGATGAATA	AGTGTTTGGACTGTCATTTGC	70.37%
D12S1586	12	52433022	CTGCACCAAGGCAGGAAC	AATTCCAGGCCCACACTTT	92.86%
D12S1724	12	53156125	CTCTGGAGGCTGAGGTGG	ATCCGTGCTGGTTCTATCTGTGTA	82.14%
D12S1632	12	54701605	GCCTAATCAAGATGTCACCA	GCTAGGGAGCCAATTCA	61.54%
D12S1700	12	58302944	AGATGCTAATGCCATGACA	CCCTGTGGATGACCAGTT	82.14%

Table key:
Chr = chromosome
α= Heterozygosities from CEPH Genotype Database (http://www.cephb.fr/cephdb/php/eng/search.php)
N/A= Heterozygosity information not available
Location based on NCBI Build 35 (http://www.ncbi.nlm.nih.gov/)

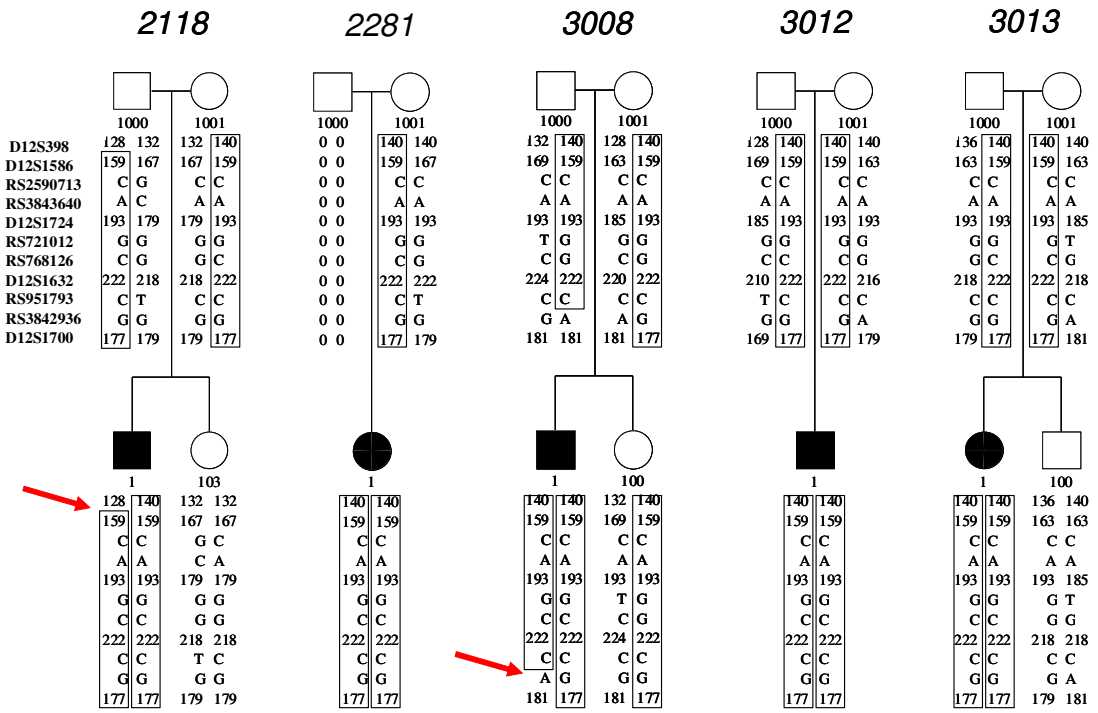


Figure 5.1 The 12q13.13 to 12q14.1 haplotype for the five NAM families.

The disease-associated haplotype is outlined in black boxes and is consistent with an autosomal recessive inheritance pattern. This diagram represents only a subset of the markers; several additional SNP markers demonstrated homozygosity in this region. The red arrows indicate key recombinant haplotypes.

CHAPTER 6

Discussion of Future Directions and Potential Limitations of Projects

The discussion which follows summarizes the research findings and comments on the potential limitations and future directions for the two projects investigating:

1) Genetic susceptibility to neural tube defects in a large multiplex family 8776;
2) Phenotypic features of NAM and homozygosity mapping outcomes for NAM.

6.1.1 Genetic Susceptibility to Neural Tube Defects in Family 8776

The first project involved genetic analysis of neural tube defects in one large family (family 8776) with multiple affected individuals from the NTD Collaborative Study. Our original linkage findings from this family suggested two broad regions of interest for NTD susceptibility genes on chromosomes 2 and 7. To follow up these linkage peaks, we ascertained the 1005/1006/102 branch and performed fine mapping that led to substantially narrowing the disease-associated haplotypes to 2q33.1-q35 and 7p21.1-21.3 (Chapter 3). Unfortunately, however, we have been unable to map the underlying disease causing mutations. As with any study evaluating the genetics of complex disease, several factors may have affected our ability to successfully identify the

actual disease loci including: 1) genetic heterogeneity; 2) misspecification of the model of inheritance; 3) different environmental exposures; and 4) chance linkage findings.

We considered several of these complicating factors in our efforts to map these NTD genes, for example the evidence for phenotypic and genetic heterogeneity was assessed by excluding individual 100, an individual with an unique NTD phenotype, a congenital dermoid cyst. All other affected individuals in this family displayed the lumbosacral myelomenginocele phenotype. However, the strongest evidence for linkage was obtained in analyses that included individual 100, suggesting a similar genetic susceptibility to NTDs in family 8776, irrespective of the exact NTD phenotype.

Initial investigation of family 8776 did not indicate any obvious inheritance pattern and to avoid issues related to misspecification of the disease model both parametric and nonparametric or genetic model-free (allele sharing) methods were implemented. The strongest support for linkage was obtained when implementing nonparametric methods.

The results from the haplotype analysis were suggestive of an autosomal dominant segregating disease in both the 2q and 7p regions of interest. Future genetic analysis of family 8776 would benefit from additional ascertainment of the last living affected individual 1075, who is shown in Figure 6.1

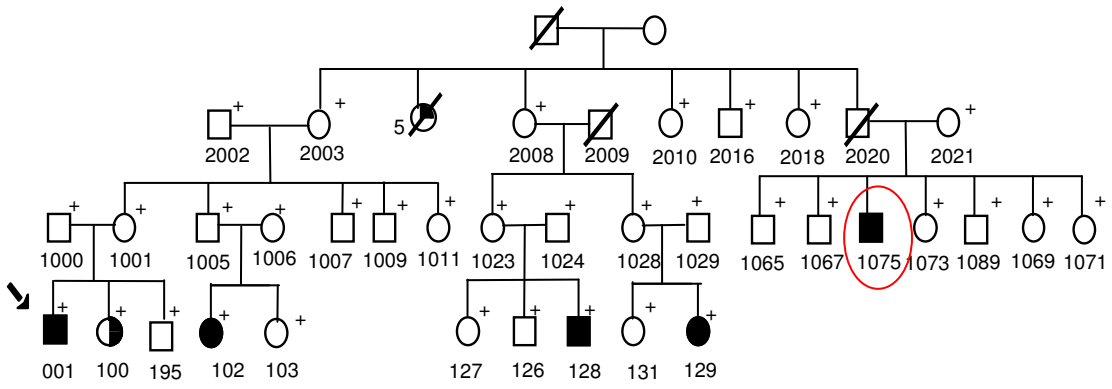


Figure 6.1 Pedigree plot of family 8776.
The plus (+) symbol means that DNA is available for that individual. The red ellipse represents that individual 1075 is the only affected individual within this family who is not currently sampled.

For example, if individual 1075 also carried the 2q and 7p haplotypes, both allele sharing and model-based methods would provide increased support for linkage. Moreover, if an additional recombination event were observed, a further narrowing of the 2q and 7p interval could be achieved.

Notably, some of the presumed unaffected individuals, with presumably unaffected offspring, also carried significant portions of either the 2q and/or 7p haplotype. For example, the 2q and 7p haplotypes are carried by the two unaffected offspring 103 and 126. Individuals 2010 and 2016 in the grandparental generation also carry the 7p haplotype. According to our records, individual 2010, who has cerebral palsy, did not have any children. Individuals 2016 and 2017 had 13 pregnancies, of which 6 resulted in live children with no documented NTD, while the remaining 7 (54%) pregnancies resulted in miscarriages—a much higher miscarriage rate than the expected 15-20% chance of miscarriage observed in the general population. This increased

frequency of spontaneous abortions could be due to the 7p haplotype or any number of etiologies including an unrelated chromosomal anomaly, a teratogenic exposure, and/or a congenital malformation (e.g. anencephaly) in the fetus.

Individuals 2010 and 2018, in the grandparental generation, and 5 offspring (1065, 1067, 1071, 1073, and 1089) of affected individual 1075, carry the 2q haplotype. Interestingly, individual 2018, who carries the 2q haplotype but doesn't carry the 7p haplotype, had 6 normal children without any reported miscarriages. These data on individuals 2016 and 2018 are suggestive that NTD risk may be modulated by epistatic interactions between the genes residing on chromosomes 2 and 7. An alternative explanation may be that a gene on 2q may confer protection against spontaneous abortion, perhaps by leading to a less severe form of a NTD. To explore the role of the 2q and 7p haplotypes in disease susceptibility further, an avenue of future research may involve collecting DNA samples on the twelve surviving children of these branches of the family to evaluate whether they carry the disease-associated haplotypes and to determine their carrier frequency rate. For example, if the live offspring from 2016 (who had 7 spontaneous abortions) carry the 7p haplotype at a lower frequency than expected by chance, it may provide further support that the 7p haplotype increase the risk of NTDs.

Since chromosomal aberrations are observed in 2-16% of nonsyndromic NTDs [22], we evaluated family 8776 for chromosomal abnormalities using array-CGH, copy number variant detection on the 10K SNP chip data, and subtelomeric *FISH*. Despite all three methods showing normal findings, the possibility remains that smaller cytogenetic abnormalities are present and were undetected in our analyses. Table 6.1 shows known

cytogenetic abnormalities in 2q and 7p regions associated with NTDs. Interestingly, all of these aberrations were only associated with spina bifida and encephalocele and not with anencephaly or craniorachischisis [22]. Though encephalocele may or may not be associated with genetic susceptibility to open spina bifida, these abnormalities suggest that genes mapping to those segments may be involved with neural tube closure. Additionally, denser SNP chips are more frequently used for clinical diagnostic purposes, facilitating detection of more subtle chromosomal abnormalities. Such diagnostic tests may lead to an expansion of the known NTD associated chromosomal aberrations, and perhaps will define smaller aberrations within the 2q and 7p locations.

Table 6.1 Cytogenetic abnormalities in chromosomal 2q and 7p regions associated with NTDs.

Chromosomal Abnormality	Spina Bifida	Encephalocele
Trisomy 2	+	+
Trisomy 7		+
Duplication 7p		+
Deletion 2q24.3		+
Deletion 2q35-36	+	

Table adapted from information from Lynch 2005[22].

The chromosome 7 haplotype maps to an 8.2 Mb interval at 7p21.1-21.3, a region with approximately 9 annotated genes. Three of these genes (*DGKB*, *ETV1*, and *MEOX2*) have already been sequenced; the remaining 6 genes (*NXPH1*, *ICA1 NDUFA4*, *PHF14*, *SCIN*, *ARL4A*) should be evaluated in future sequencing efforts of family 8776. Particularly *ICA1* (islet cell autoantigen 1), as there is an associated risk of maternal diabetes with having a child affected with a NTD, the *ICA1* gene modulates expression of neuronal cells, it was down regulated in case lymphoblast cell lines (Chapter 3) and it is

expressed in the fetus during Carnegie Stages 12 and 13 around the time of neural tube closure [241].

The mapped 2q region has approximately 89 genes, of which 15 have been screened for obvious coding functional alterations. Also based on preliminary SAGE work coupled with biological plausibility additional genes are currently being sequenced in the 2q region which include: *BMPR2*, *XRCC5* and *ICA1L*. The sequencing efforts for family 8776 focused on the coding regions since obvious alterations in amino acid structure could be readily argued to have an effect on protein function because it is difficult to understand the functional consequence of non-coding variants when analyzing only one family. However, intronic, promoter, and 3' UTR variants may be functionally relevant to gene regulation in family 8776 and other NTD families.

Another project currently underway at Duke University is investigating the genetic basis of Chiari II malformations using family based gene mapping approaches. Since Chiari II malformation co-occurs with open spina bifida, any genetic mutations associated with Chiari II could also be tested in family 8776 and in our large collection of families with open spina bifida. Alternatively, once an interesting genetic variant was detected in family 8776, the next steps would be to carry out association studies of this gene in the larger NTD data set as well as in the Chiari II population to determine the relevance of this gene in other families.

6.1.2 Environmental Exposures and Other Potential Risk Factors for NTD

After the identification of main genetic effects in Family 8776 and using the larger NTD dataset, it will be important to consider gene-environment interactions using

the data collected from the environmental risk factor questionnaire. Specifically folic acid status could be evaluated based on information on maternal perinatal intake of folic acid and daily multivitamins. Other potential risk factors to consider include maternal weight, maternal diabetes, smoking and medication use during pregnancy. The two critical regions with similar evidence in favor of linkage for family 8776 suggest that there may be polygenic effects; therefore, potential gene-gene and gene-environment interactions could be studied in several NTD families.

Epigenetic mechanisms may also play a role in normal and dysregulated neural tube closure. In family 8776, epigenetic mechanisms may be involved in disrupted neural tube closure. For instance, the sex ratios for transmitting members of the family to affected children has a 5:2 female to male ratio suggesting a stronger maternal transmission pattern. Although this observation is not statistically significant and may reflect ascertainment biases toward maternal interviews, it illustrates the potential for maternal nutritional status in family 8776 affecting imprinting mechanisms and/or general methylation reactions that may increase the risk of having a child affected with an NTD. At this time, no imprinted genes in the 2q and 7p regions of interest were noted [253]. If future studies reveal imprinted genes in these regions they should be evaluated.

To estimate the degree of folate derivates consumed by women around the time of conception through the first month of pregnancy, a risk factor questionnaire with specific questions pertaining to folic acid, prenatal vitamins and multivitamins use was completed by NTD study participants. Although these types of questionnaires are subject to recall bias, a rough measure of maternal folic acid status at the time of neural tube closure was obtained. The folic acid questionnaire outcomes in family 8776 of the five mothers who

had a child affected with an NTD are illustrated in Table 6.2. All five mothers reported that they took folic acid, prenatal vitamins, or multivitamins first two months of pregnancy. Individuals 1006, 1023, and 2021 reported taking prenatal or multivitamins for an NTD-affected pregnancy before 1990 when there was less public awareness of the role of folic acid in preventing NTDs. Thus, the amount of folic acid present in the multivitamin formulations at that time may have been inadequate for full protection against NTDs. It has been additionally difficult to determine if the mothers commenced taking folic acid prior to conception or more specifically, prior to neural tube closure. To address this issue, the questionnaire has recently been changed to ask: "How far along before you knew you were pregnant?" This question is a way to better estimate the mother's actual folic acid intake at the time of neural tube closure, and the mothers in family 8776 have not been asked this question yet. Another possible explanation is that the mothers from family 8776 had adequate folic acid consumption according to current recommendations, and this family may alternatively have genetic variant(s) conferring NTD risk that either lead to the intrauterine environment being folic acid resistant or increase the metabolism/excretion of folic acid or are completely unrelated to folic acid.

Table 6.2 Folic acid questionnaire results for family 8776

Mother	Child	Year Child Born	FA 3m before preg?	FA in first 2m of preg?	PV 3m before preg?	PV in first 2m of preg?	MV 3m before preg?	MV in first 2m of preg?
1001	1	1998	Yes	yes	no	yes	no	No
1001	100	1999	Yes	yes	no	yes	no	No
1006	102	1987	No	no	no	yes	yes	No
1023	128	1991	No	no	no	yes	no	No
1028	129	1983	No	no	no	yes	no	No
2021	1075	1957	No	no	no	no	yes	Yes

FA, folic acid; PV, prenatal vitamins; MV, multivitamins; m, months; preg, pregnancy

Recently, autoantibodies against folate receptors to placental membranes in women who had NTD-affected pregnancies were identified that may be blocking the cellular uptake of folate [89]. It would be interesting to examine if the mothers with affected children in family 8776 also express these autoantibodies, and more generally to investigate if the autoantibodies are correlated to all NTDs or only specific form of NTDs in a larger NTD dataset. The presence of other predisposing factors for NTD development in the presence of these autoantibodies to folate receptors is supported by the observation that two women with normal pregnancies also had autoantibodies against folate receptors. If these mothers from family 8776 demonstrate autoantibodies to folate receptors, it is possible that the chromosome 2 and 7 regions of interest may contain modifier genes that increase NTD risk in the developing fetus. Therefore, it would be additionally informative to collect serum on family members who also have children that are not affected by an NTD. Although the titer patterns of these autoantibodies is currently unknown, such that levels may change during different stages of pregnancy or alternatively, there may be variability from one pregnancy to the next. At this time, it is unclear whether these autoantibodies are environmental triggered or genetically determined or a combination of the two.

6.1.3 Commentary on Future Genetic Studies of Human NTDs

Since linkage often maps to large ~10-20 Megabase segments resulting in multiple interesting candidate genes, a staged approach has been commonly applied to

genetic analysis of complex human disorders. These staged approaches can involve fine mapping under the linkage peaks that is then followed by association analysis to provide more precise localization of candidate genes. At this time, only limited linkage data is available for human NTDs and of the 43 other multiplex families studied in the previous screen, none had strong evidence of linkage to the 2q and 7p regions, preventing association analysis in those regions. However, future whole genome linkage or association analysis within the ~330 multiplex and ~860 simplex American Caucasian families enrolled in the Duke NTD study may reveal promising genes mapping to the 2q and 7p region of interest. In the field of human genetics, whole genome association (WGA) studies are becoming a more popular form of genetic analysis than traditional linkage studies, and it is therefore worthwhile to consider the rationale behind performing WGA studies in NTDs. Given the number of tests in a dense WGA, many regions of follow-up may be identified, and many of which will be false positive findings. In order to reduce the number of false positives identified, WGA studies need to be adequately powered with large sample sizes, but replication samples are also needed. With the NTD study, even though enough sampled individuals exist to conduct WGA at this scale, it would be a very costly research endeavor and would require replication in another population or dataset to understand the broader significance of an association of polymorphism(s) with NTD susceptibility.

Moreover, the SNP arrays currently available for WGA may or may not be the most informative set for studying the genetic basis of NTDs. Many platforms preferentially select coding or exonic SNPs; however, this preference may lead to missing key promoter or intronic SNPs that may be functionally important for neural tube

closure. In addition, SNP selection for these platforms is based on the assumption that common diseases are caused by common variants with minor allele frequencies >5%, and this hypothesis may not hold true for more rare complex diseases, such as NTDs. I personally believe a combination of both common and rare SNPs predispose to NTDs and the constituent combination will vary in different populations and in different phenotypes. From a more optimistic perspective, common SNPs may be in LD with some of the rare high risk SNPs such that WGA may still be a powerful tool for gene mapping studies in NTDs. Moreover, the density of the arrays continue to increase with 500K and 1M chips now available, and these denser chips may adequately capture more of the non-coding variation in the human genome.

Overall, successful identification of an interesting NTD variant using linkage, association and/or sequencing approaches would provide valuable insight into the molecular mechanisms underlying NTDs, and may allow further investigation of these mechanisms in appropriate model organisms to elucidate therapeutic options to prevent and/or ameliorate disrupted neural tube closure.

6.2 Native American Myopathy Study

The second project studied Native American myopathy, a Mendelian congenital myopathy associated with high mortality and morbidity. For the purposes of this dissertation, we conducted a clinical evaluation (Chapter 4) and the first ever gene mapping study of NAM (Chapter 5).

6.2.1 NAM: Clinical Evaluation and Natural History of Disease

Chapter 4 provides the most extensive clinical description of NAM to date, among fourteen individuals of Lumbee Indian descent. In terms of the natural history of NAM, we were able to document for the first time the mortality rate for this condition, a rate of 36% by age 18. Also, we were able to document that malignant hyperthermia, a potentially dangerous condition, occurred in 29% of patients. Moreover, pedigree inspection of the fourteen supported an autosomal recessive inheritance pattern for NAM suggesting that homozygosity mapping may be a productive approach for gene identification studies. Overall, this information will further our understanding of the clinical presentation and natural history of NAM, and ideally lead to improved diagnosis and counseling of patients and their families in the Lumbee population.

Although this clinical study of Lumbee Indian family members is the most extensive evaluation of NAM to date, future work to characterize the muscle pathology of patients with NAM is warranted. Often the classification of muscular disorders is done through histopathological examination of skeletal muscle tissue from muscle biopsies. Although the muscle biopsy procedure is somewhat invasive, the information could not only contribute to the clinical categorization of the disease as NAM relates to other congenital myopathies, but also could be used for candidate gene selection. To avoid potential risks associated with new muscle biopsies, stored muscle biopsies on patients with NAM could alternatively be assessed.

Furthermore, given the clinical risks associated with NAM, future work to identify the NAM gene is also warranted, as the identification of a susceptibility NAM gene could aid in genetic counseling for families with Lumbee ancestry. Since the

original dataset for NAM included only two families and collecting more families would allow for a more powerful homozygosity mapping study, multiple clinical ascertainment strategies were employed for study recruitment of NAM families. These approaches included re-contacting families already involved in the study by telephone, referral from clinics, sending letters to physicians in Robeson County, and the posting of an advertisement on Lumbee affiliated web browsers. Unfortunately, we had very limited success in our ascertainment strategies with only 3 new families being recruited, two of which were families previously seen by Cynthia Powell, a medical geneticist at the University of North Carolina at Chapel Hill. Altogether five simplex families were available to us for homozygosity mapping studies. The challenges we encountered with attempts to recruit additional NAM families may reflect inherent culture fears within the Lumbee population towards genetic studies and/or a lack of diagnosis of NAM by physicians. Future ascertainment could involve closer collaboration with the Lumbee Tribal government.

6.2.2 Homozygosity Mapping in NAM

Homozygosity mapping of the five affected NAM individuals using a genome-wide 10K SNP screen identified five regions (2p, 3q, 5q, 6q, and 12q) of shared homozygosity, with 3 or more contiguous SNPs (Chapter 5). Additional genotyping in these segments narrowed the interval to one region that also was the longest region of homozygosity on chromosome 12q13.13 to 12q14.1 (5.6 Mb segment).

The observation that 12q maps to the longest homozygosity segment is a potentially promising finding based on a study by Woods and colleagues, who evaluated the degree of homozygous segments in individuals with recessive disease whose parents were first cousins. They found the longest homozygous segment was the disease-associated segment 17% of the time—a 12% increase over the expected 5% (1/20 segments) [254]. Thus, since the 12q homozygous segment was the longest segment observed in the patients with NAM and the only confirmed homozygous segment, and because NAM appears to be following a recessive inheritance pattern, we are hypothesizing that the underlying susceptibility gene is localized on chromosome 12q.

However, it is important to acknowledge the possibility that the 12q region represents chance autozygosity. At this time, estimates of chance IBD sharing within the Lumbee population are not possible because we do not know the expected allele frequencies of the markers genotyped for the Lumbee population and it is challenging to correctly specify the degree of relatedness within the NAM pedigrees to estimate the extent of expected IBD allele sharing. In the future, ascertainment of additional NAM families or Lumbee individuals enrolled for other studies could be used to estimate allele frequencies. However, misspecification of the degree of relatedness may remain a problem since the Lumbee population is likely to have experienced a significant founder effect and may have unspecified family relationships that could lead to an underestimate of the extent of IBD sharing.

6.2.3 Identification of NAM gene

The NAM homozgosity mapping work revealed a 5.6 Mb segment shared among the five patients with NAM, mapping to 12q13.13 to 12q14.1. This region appears to be very gene dense with well over 100 annotated genes. As a result, future sequencing efforts in this 12q region may be difficult. Thus far, *ITGA7, PIP5K2C, PDE1B,* and *MLC1SA* have been sequenced and no functional mutations were identified in these genes. No other obvious candidate genes were noted in this region. Nonetheless, several approaches could be used in the future to help identify the causative mutation residing in the putative NAM locus such as ascertaining additional families to identify a recombinant individual and/or ascertaining individuals with MH only to possibly narrow the interval and localize the NAM gene. Another approach to minimize sequencing efforts may be to use muscle biopsies from NAM patients to prioritize candidate genes based on histological finding or RNA expression profiles. Perhaps in the future more affordable sequencing platforms may become readily available such that determining the sequence across the entire 5.6 Mb segment would be feasible.

Even though the homozygosity mapping to 12q appears promising for identification of a novel gene for NAM, the disease gene may map to another location. Several factors may have prevented the identification of the actual NAM locus such as genetic heterogeneity or a smaller homozygous segment flanking the disease gene in another region that falls outside of our 3 or more contiguous SNP threshold. In addition, it is possible that the disease gene may be in a region with a paucity of SNP coverage on the 10K chip. Based on the possibility of chance autozygosity at 12q, the MHS associated genes *RYR1* and *CACNA1S* can still be considered potential candidate genes

for NAM. However, the 19q13.1 locus for the *RYR1* gene did not reveal shared homozygosity among the five affected individuals with NAM (even between two contiguous SNPs). Nonetheless, it is still possible that one of the individuals may be heterozygous for the mutation rather than homozygous. Moreover, more than one mutation in the NAM gene may be segregating in the Lumbee population such that an individual with NAM could be a compound heterozygote. Although NAM presents with features similar to other diseases as described in Chapter 1, we believe this disorder is clinically distinct. However, mutations in the MHS genes *RYR1* or *CACNA1S* may still be responsible for the NAM phenotype and cannot be excluded, especially since both autosomal recessive and dominant mutations in *RYR1* are associated with variable phenotypic presentations of congenital myopathies and/or malignant hyperthermia susceptibility. Because mutations in the *RYR1* cause the majority of known mutations associated with MHS and/or congenital myopathy, sequencing of the *RYR1* gene in one of the individuals with NAM, who has a history of MH is being performed by PreventionGenetics (Marshfield, WI).

Once the NAM gene is identified, future analysis could involve characterizing genotype-phenotype correlations to investigate whether an individual with one versus two mutated genes leads to differences in phenotypic presentation. Carrier testing may be particularly helpful if it is discovered that heterozygotes are at greater risk for malignant hyperthermia, a potentially life-threatening illness. Moreover, to definitively determine if heterozygotes were susceptible to MH, the IVCT test could be done on muscle biopsy specimens to clarify the phenotype because MH will not be observed unless the patient has received halogenated agents or succinylcholine during anesthesia. Overall, the

identification of a NAM gene could contribute to our understanding of the etiology of NAM and may provide an avenue for the development of a genetic test for this devastating disease.

REFERENCES

1. Campbell LR, Dayton DH, and Sohal GS (1986) Neural tube defects: a review of human and animal studies on the etiology of neural tube defects. *Teratology* 34 (2):171-187

2. *Neural Tube Defects: From Origin to Treatment*. 2006. Wyszynski DF (ed).Oxford University Press, Inc.

3. Centers for Disease Control (1989) Current trends economic burden of spina bifida -- United States, 1980-1990. *Morbidity and Mortality Weekly* 38 (15):264-267

4. Mathews, T. J. Trends in Spina Bifida and Anencephalus in the United States, 1991-2004. 2007. Center for Disease Control and Prevention's National Center for Health Statistics.

5. Canfield MA, Collins JS, Botto LD, Williams LJ, Mai CT, Kirby RS, Pearson K, Devine O, and Mulinare J (2005) Changes in the birth prevalence of selected birth defects after grain fortification wiht folic acid in the United States: Findings from a multi-state population-based study. *Birth Defects Res.A Clin.Mol.Teratol.* 73 (10):679-689

6. Suarez L, Hendricks KA, Cooper SP, Sweeney AM, Hardy RJ, and Larsen RD (2000) Neural tube defects among Mexican Americans living on the US-Mexico border: effects of folic acid and dietary folate. *Am.J.Epidemiol.* 152 (11):1017-1023

7. Missmer SA, Suarez L, Felkner M, Wang E, Merrill AH, Jr., Rothman KJ, and Hendricks KA (2006) Exposure to fumonisins and the occurrence of neural tube defects along the Texas-Mexico border. *Environ.Health Perspect.* 114 (2):237-241

8. Terwilliger, J. D. and Ott, J. Handbook of Human Genetic Linkage. 1994. Baltimore, Johns Hopkins University Press.

9. Agarwal, S. S. Folic Acid for Prevention of Neural Tube Defects. 1998. Nutrition Foundation of India.

10. Verma M, Chhatwal J, and Singh D (1991) Congenital malformations--a retrospective study of 10,000 cases. *Indian J.Pediatr.* 58 (2):245-252

11. Li Z, Ren A, Zhang L, Ye R, Li S, Zheng J, Hong S, Wang T, and Li Z (2006) Extremely high prevalence of neural tube defects in a 4-county area in Shanxi Province, China. *Birth Defects Res.A Clin.Mol.Teratol.* 76 (4):237-240

12. Holmes LB, Driscoll SG, and Atkins L (1976) Etiologic Heterogeneity of Neural-Tube Defects. *N Engl J Med* 294:365-369

13. Khoury MJ, Erickson JD, and James LM (1982) Etiologic heterogeneity of neural tube defects. II. Clues from family studies. *Am.J.Hum.Genet.* 34:980-987

14. Khoury MJ, Erickson JD, and James LM (1982) Etiologic heterogeneity of neural tube defects: clues from epidemiology. *Am.J.Epidemiol.* 115 (4):538-548

15. Hunter AG (1984) Neural tube defects in Eastern Ontario and Western Quebec: demography and family data. *Am.J.Med.Genet.* 19 (1):45-63

16. Hall JG, Friedman JM, Kenna BA, Popkin J, Jawanda M, and Arnold W (1988) Clinical, genetic, and epidemiological factors in neural tube defects. *Am.J.Hum.Genet.* 43:827-837

17. Simpson JL, Mills J, Rhoads GG, Cunningham GC, Conley MR, and Hoffman HJ (1991) Genetic heterogeneity in neural tube defects. *Ann.Genet.* 34 (3-4):279-286

18. Hume RF, Drugan A, Reichler A, Lampinen J, Martin LS, Johnson MP, and Evans MI (1996) Aneuploidy among prenatally detected neural tube defects. *Am.J.Med.Genet.* 61:171-173

19. Kennedy D, Chitayat D, Winsor EJT, Silver M, and Toi A (1998) Prenatally diagnosed neural tube defects: Ultrasound, chromosome, and autopsy or postnatal findings in 212 cases. *Am.J.Med.Genet.* 77:317-321

20. Philipp T and Kalousek DK (2002) Neural tube defects in missed abortions: embryoscopic and cytogenetic findings. *Am.J.Med.Genet.* 107 (1):52-57

21. Sepulveda W, Corral E, Ayala C, Be C, Gutierrez J, and Vasquez P (2004) Chromosomal abnormalities in fetuses with open neural tube defects: prenatal identification with ultrasound. *Ultrasound Obstet.Gynecol.* 23 (4):352-356

22. Lynch SA (2005) Non-multifactorial neural tube defects. *Am.J.Med.Genet.C.Semin.Med.Genet.* 135 (1):69-76

23. Gomez MR (1981) Possible teratogenicity of valproic acid. *J.Pediatr.* 98 (3):508-509

24. Bjerkedal T, Czeizel A, Goujard J, Kallen B, Mastroiacova P, Nevin N, Oakley G, Jr., and Robert E (1982) Valproic acid and spina bifida. *Lancet* 2 (8307):1096

25. Lindhout D and Schmidt D (1986) In-utero exposure to valproate and neural tube defects (letter). *Lancet* 1:1392-1393

26. Shaw GM, Velie EM, and Schaffer D (1996) Risk of neural tube defect-affected pregnancies among obese women. *JAMA* 275 (14):1093-1096

27. Watkins ML, Scanlon KS, Mulinare J, and Muin JK (1996) Is maternal obesity a risk factor for anencephaly and spina bifida? *Epidemiology* 7 (5):507-512

28. Marasas WF, Riley RT, Hendricks KA, Stevens VL, Sadler TW, Gelineau-van Waes J, Missmer SA, Cabrera J, Torres O, Gelderblom WC, Allegood J, Martinez C, Maddox J, Miller JD, Starr L, Sullards MC, Roman AV, Voss KA, Wang E, and Merrill AH, Jr. (2004) Fumonisins disrupt sphingolipid metabolism, folate transport, and neural tube development in embryo culture and in vivo: a potential risk factor for human neural tube defects among populations consuming fumonisin-contaminated maize. *J.Nutr.* 134 (4):711-716

29. Loeken MR (2005) Current perspectives on the causes of neural tube defects resulting from diabetic pregnancy. *American Journal of Medical Genetics Part C-Seminars in Medical Genetics* 135C (1):77-87

30. Robert E and Guibaud P (1982) Maternal valproic acid and congenital neural tube defects. *Lancet* 2:937

31. Lammer EJ, Sever LE, and Oakley GP (1987) Teratogen update: Valproic acid. *Teratology* 35:465-473

32. Rosa FW (1991) Spina bifida in infants of women treated with carbamazepine during pregnancy. *N Engl J Med* 324:674-677

33. Finnell RH (1991) Genetic differences in susceptibility to anticonvulsant drug-induced developmental defects. *Pharmacol.Toxicol.* 69 (4):223-227

34. Lindhout D, Meinardi H, Meijer JWA, and Nau H (1992) Antiepileptic drugs and teratogenesis in two consecutive cohorts: Changes in prescription policy paralleled by changes in pattern of malformation. *Neurology* 42:94-110

35. Medveczky E, Puho E, and Czeizel EA (2004) The use of drugs in mothers of offspring with neural-tube defects. *Pharmacoepidemiol.Drug Saf* 13 (7):443-455

36. Alsdorf R and Wyszynski DF (2005) Teratogenicity of sodium valproate. *Expert.Opin.Drug Saf* 4 (2):345-353

37. Ross SA, McCaffery PJ, Drager UC, and De Luca LM (2000) Retinoids in embryonal development. *Physiol Rev.* 80 (3):1021-1054

38. Lindhout D and Meinardi H (1984) Spina bifida and in-utero exposure to valproate. *Lancet* 2 (8399):396

39. Dansky LV, Rosenblatt DS, and Andermann E (1992) Mechanisms of teratogenesis: folic acid and antiepileptic therapy. *Neurology* 42 (4 Suppl 5):32-42

40. Duncan S, Mercho S, Lopes-Cendes I, Seni MH, Benjamin A, Dubeau F, Andermann F, and Andermann E (2001) Repeated neural tube defects and valproate monotherapy suggest a pharmacogenetic abnormality. *Epilepsia* 42 (6):750-753

41. Bologa M, Tang B, Klein J, Tesoro A, and Koren G (1991) Pregnancy-induced changes in drug metabolism in epileptic women. *J.Pharmacol.Exp.Ther.* 257 (2):735-740

42. (1991) Prevention of neural tube defects: results of the Medical Research Council Vitamin Study. MRC Vitamin Study Research Group. *Lancet* 338 (8760):131-137

43. Centers for Disease Control and Prevention (1992) Recommendations for the use of folic acid to reduce the number of cases of spina bifida and other neural tube defects. *Morbidity and Mortality Weekly* 41:1-7

44. Mathews TJ, Honein MA, and Erickson JD. Spina bifida and anencephaly prevalence - United States 1991-2000. 51, 8-10. 2002.

45. Williams LJ, Mai CT, and Edmonds LD (2002) Prevalence of spina bifida and anencephaly during the transition to mandatory folic acid fortification in the United States. *Teratology* 66:33-39

46. Robbins JM, Tilford JM, Bird TM, Cleves MA, Reading JA, and Hobbs CA (2006) Hospitalizations of newborns with folate-sensitive birth defects before and after fortification of foods with folic acid. *Pediatrics* 118 (3):906-915

47. Evans MI, Llurba E, Landsberger EJ, O'Brien JE, and Harrison HH (2004) Impact of folic acid fortification in the United States: markedly diminished high maternal serum alpha-fetoprotein values. *Obstet.Gynecol.* 103 (3):474-479

48. Frey L and Hauser WA (2003) Epidemiology of neural tube defects. *Epilepsia* 44 Suppl 3:4-13

49. Northrup H and Volcik KA (2000) Spina bifida and other neural tube defects. *Curr.Probl.Pediatr.* 30 (10):313-332

50. Toriello HV and Higgins JV (1985) Possible causal heterogeneity in spina bifida cystica. *Am.J.Med.Genet.* 21:13-20

51. Hall JG and Keena BA (1986) Adjusting recurrence risks for neural tube defects based on B.C. data. *Am.J.Hum.Genet.* 39:A64

52. Frecker MF, Fraser FC, and Heneghan WD (1988) Are 'upper' and 'lower' neural tube defects aetiologically different? *J Med.Genet.* 25 (7):503-504

53. Drainer E, May HM, and Tolmie JL (1991) Do familial neural tube defects breed true? *J.Med.Genet.* 28 (9):605-608

54. Garabedian BH and Fraser FC (1993) Upper and lower neural tube defects: an alternate hypothesis. *J Med.Genet.* 30 (10):849-851

55. Cowchock S, Ainbender E, Prescott G, Crandall B, Lau L, Heller R, Muir WA, Kloza E, Feigelson M, Mennuti M, and Cederquist L (1980) The recurrence risk for neural tube defects in the United States: a collaborative study. *Am.J.Med.Genet.* 5 (3):309-314

56. Elwood, J. M., Little, J., and Elwood, J. H. Epidemiology and Control of Neural Tube Defects. 1992. Oxford, Oxford University Press.

57. Li H, Zhao Y, Li S, Xu Y, Huang B, Cui M, and Zheng G (2002) [Epidemiology of twin and twin with birth defects in China]. *Zhonghua Yi.Xue.Za Zhi.* 82 (3):164-167

58. Harris MJ and Juriloff DM (2007) Mouse mutants with neural tube closure defects and their role in understanding human neural tube defects. *Birth Defects Res.A Clin.Mol.Teratol.* 79 (3):187-210

59. Fuhrmann W (1971) [Genetic counseling as a problem of pediatric practice. Introduction]. *Monatsschr.Kinderheilkd.* 119 (4):154-156

60. Partington MD and McLone DG (1995) Hereditary factors in the etiology of neural tube defects. *Pediatr Neurosurg* (23):311-316

61. Stamm DS, Rampersaud E, Slifer SH, Mehltretter L, Siegel DG, Xie J, Hu-Lince D, Craig DW, Stephan DA, George TM, Gilbert JR, and Speer MC (2006) High-density single nucleotide polymorphism screen in a large multiplex neural tube defect family refines linkage to loci at 7p21.1-pter and 2q33.1-q35. *Birth Defects Res.A Clin.Mol.Teratol.* 76 (6):499-505

62. Carter CO (1974) Clues to the aetiology of neural tube malformations. *Dev.Med.Child Neurol.* 16 (6 Suppl 32):3-15

63. Copp AJ and Bernfield M (1994) Etiology and pathogenesis of human neural tube defects: insights from mouse models

1. *Curr.Opin.Pediatr.* 6 (6):624-631

64. Fineman RM, Jorde LB, Martin RA, Hasstedt SJ, Wing SD, and Walker ML (1982) Spinal dysraphia as an autosomal dominant defect in four families. *Am.J.Med.Genet.* 12:457-464

65. Demenais F, Le Merrer M, Briard ML, and Elston RC (1982) Neural tube defects in France: segregation analysis. *Am.J.Med.Genet.* 11:287-298

66. Juriloff DM and Harris MJ (2000) Mouse models for neural tube closure defects. *Hum.Mol.Genet.* 9 (6):993-1000

67. Rampersaud E, Melvin EC, Siegel D, Mehltretter L, Dickerson ME, George TM, Enterline D, Nye JS, Speer MC, and NTD Collaborative Group (2003) Updated investigations of the role of methylenetetrahydrofolate reductase in human neural tube defects. *Clin.Genet.* 63 (3):210-214

68. Mitchell LE, Adzick NS, Melchionne J, Pasquariello PS, Sutton LN, and Whitehead AS (2004) Spina bifida. *Lancet* 364 (9448):1885-1895

69. Boyles AL, Hammock P, and Speer MC (2005) Candidate gene analysis in human neural tube defects. *Am.J.Med.Genet.C.Semin.Med.Genet.* 135 (1):9-23

70. Whitehead AS, Gallagher P, Mills JL, Kirke PN, Burke H, Molloy AM, Weir DG, Shields DC, and Scott JM (1995) A genetic defect in 5,10 methylenetetrahydrofolate reductase in neural tube defects. *Quarterly Journal of Medicine* 88:763-766

71. van der Put NM, Steegers-Theunissen RPM, Frosst P, Trijbels FJM, Eskes TKAB, van den Heuvel LP, Mariman ECM, den Heyer M, Rozen R, and Blom HJ (1995) Mutated methylenetetrahydrofolate reductase as a risk factor for spina bifida. *Lancet* 346:1070-1071

72. van der Put NM, Gabreels F, Stevens EM, Smeitink JA, Trijbels FJ, Eskes TK, van den Heuvel LP, and Blom HJ (1998) A second common mutation in the methylenetetrahydrofolate reductase gene: An additional risk factor for neural-tube defects? *Am.J.Hum.Genet.* 62:1044-1051

73. Centers for Disease Control and Prevention (CDC) (2004) Spina bifida and anencephaly before and after folic acid mandate--United States, 1995-1996 and 1999-2000. *MMWR Morb.Mortal.Wkly.Rep.* 53 (17):362-365

74. Melvin EC, George TM, Worley G, Franklin A, Mackey J, Viles K, Shah N, Drake CR, Enterline DS, McLone D, Nye J, Oakes WJ, McLaughlin C, Walker ML, Peterson P, Brei T, Buran C, Aben J, Ohm B, Bermans I, Qumsiyeh M, Vance J, Pericak-Vance MA, and Speer MC (2000) Genetic studies in neural tube defects. NTD Collaborative Group. *Pediatr.Neurosurg.* 32 (1):1-9

75. Harris MJ (2001) Why are the genes that cause risk of human neural tube defects so hard to find? *Teratology* 63 (5):165-166

76. Earley KJ, Blanco JD, Prien S, and Willis D (1991) Patient attitudes toward testing for maternal serum alpha-fetoprotein values when results are false-positive or true-negative. *South.Med.J.* 84 (4):439-442

77. O'Rahilly R and Muller F (1994) Neurulation in the normal human embryo. *Ciba Found.Symp.* 181:70-82

78. Stover PJ (2004) Physiology of folate and vitamin B12 in health and disease. *Nutr.Rev.* 62 (6 Pt 2):S3-12

79. van der Put NM and Blom HJ (2000) Neural tube defects and a disturbed folate dependent homocysteine metabolism. *Eur.J.Obstet.Gynecol.Reprod.Biol.* 92 (1):57-61

80. Steegers-Theunissen RP, Boers GH, Trijbels FJ, Finkelstein JD, Blom HJ, Thomas CM, Borm GF, Wouters MG, and Eskes TK (1994) Maternal hyperhomocysteinemia: a risk factor for neural-tube defects? *Metabolism* 43 (12):1475-1480

81. Zhao W, Mosley BS, Cleves MA, Melnyk S, James SJ, and Hobbs CA (2006) Neural tube defects and maternal biomarkers of folate, homocysteine, and glutathione metabolism. *Birth Defects Res.A Clin.Mol.Teratol.* 76 (4):230-236

82. Nijhout HF, Reed MC, Budu P, and Ulrich CM (2004) A mathematical model of the folate cycle: new insights into folate homeostasis. *J.Biol.Chem.* 279 (53):55008-55016

83. Reed MC, Nijhout HF, Sparks R, and Ulrich CM (2004) A mathematical model of the methionine cycle. *J.Theor.Biol.* 226 (1):33-43

84. Pitkin RM (2007) Folate and neural tube defects. *Am.J.Clin.Nutr.* 85 (1):285S-288S

85. Milunsky A, Jick H, Jick SS, Bruell CL, MacLaughlin DS, Rothman KJ, and Willett W (1991) Multivitamin/folic acid supplementation in early pregnancy reduces the prevalence of neural tube defects. *JAMA* 262:2847-2852

86. Piedrahita JA, Oetama B, Bennett GD, van Waes J, Kamen BA, Richardson J, Lacey SW, Anderson RG, and Finnell RH (1999) Mice lacking the folic acid-binding protein Folbp1 are defective in early embryonic development. *Nat.Genet.* 23 (2):228-232

87. Finnell RH, Junker WM, Wadman LK, and Cabrera RM (2002) Gene expression profiling within the developing neural tube. *Neurochem.Res.* 27 (10):1165-1180

88. da Costa M, Sequeira JM, Rothenberg SP, and Weedon J (2003) Antibodies to folate receptors impair embryogenesis and fetal development in the rat. *Birth Defects Res.A Clin.Mol.Teratol.* 67 (10):837-847

89. Rothenberg SP, da Costa MP, Sequeira JM, Cracco J, Roberts JL, Weedon J, and Quadros EV (2004) Autoantibodies against folate receptors in women with a pregnancy complicated by a neural-tube defect. *N.Engl.J.Med.* 350 (2):134-142

90. Reik W, Constancia M, Fowden A, Anderson N, Dean W, Ferguson-Smith A, Tycko B, and Sibley C (2003) Regulation of supply and demand for maternal nutrients in mammals by imprinted genes. *J.Physiol* 547 (Pt 1):35-44

91. Paoloni-Giacobino A and Chaillet JR (2004) Genomic imprinting and assisted reproduction. *Reprod.Health* 1 (1):6

92. Waterland RA and Jirtle RL (2004) Early nutrition, epigenetic changes at transposons and imprinted genes, and enhanced susceptibility to adult chronic diseases. *Nutrition* 20 (1):63-68

93. Jaenisch R and Bird A (2003) Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat.Genet.* 33 Suppl:245-254

94. Chatkupt S, Lucek PR, Koenigsberger MR, and Johnson WG (1992) Parental sex effect in spina bifida: a role for genomic imprinting? *Am.J.Med.Genet.* 44 (4):508-512

95. Mariman EC and Hamel BC (1992) Sex ratios of affected and transmitting members of multiple case families with neural tube defects. *J.Med.Genet.* 29 (10):695-698

96. Byrne J, Cama A, Reilly M, Vigliarolo M, Levato L, Boni L, Lavia N, and Andreussi L (1996) Multigeneration maternal transmission in Italian families with neural tube defects. *Am.J.Med.Genet.* 66 (3):303-310

97. Byrne J and Carolan S (2006) Adverse reproductive outcomes among pregnancies of aunts and (spouses of) uncles in Irish families with neural tube defects. *Am.J.Med.Genet.A* 140 (1):52-61

98. Purves D, Lichtman JW. 1985. *Principles of Neural Development.* Sunderland: Sinauer Assocs

99. Sakai Y (1989) Neurulation in the mouse: manner and timing of neural tube closure. *Anat.Rec.* 223 (2):194-203

100. Juriloff DM, Harris MJ, Tom C, and Macdonald KB (1991) Normal mouse strains differ in the site of initiation of closure of the cranial neural tube. *Teratology* 44 (2):225-233

101. Golden JA and Chernoff GF (1993) Intermittent pattern of neural tube closure in two strains of mice. *Teratology* 47:73-80

102. van Allen M, Kalousek DK, Chernoff GF, Juriloff D, Harris M, McGillivray BC, Yong SL, Langlois S, Macleod PM, Chitayat D, Freidman JM, Wilson D, McFadden D, Pantzar J, Ritchie S, and Hall JG (1993) Evidence for multi-site closure of the neural tube in humans. *Am.J.Med.Genet.* 47:723-743

103. Van Allen MI (1996) Multisite neural tube closure in humans. *Birth Defects Orig.Artic.Ser.* 30 (1):203-225

104. Sulik KK and Sadler TW (1993) Postulated mechanisms underlying the development of neural tube defects. Insights from in vitro and in vivo studies. *Ann.N.Y.Acad.Sci.* 678:8-21

105. Nakatsu T, Uwabe C, and Shiota K (2000) Neural tube closure in humans initiates at multiple sites: evidence from human embryos and implications for the pathogenesis of neural tube defects. *Anat.Embryol.(Berl)* 201 (6):455-466

106. O'Rahilly R and Muller F (2002) The two sites of fusion of the neural folds and the two neuropores in the human embryo. *Teratology* 65 (4):162-170

107. Copp AJ, Greene ND, and Murdoch JN (2003) The genetic basis of mammalian neurulation. *Nat.Rev.Genet.* 4 (10):784-793

108. Fleming A and Copp AJ (2000) A genetic risk factor for mouse neural tube defects: defining the embryonic basis. *Hum.Mol.Genet.* 9 (4):575-581

109. Wallingford JB and Harland RM (2001) Xenopus Dishevelled signaling regulates both neural and mesodermal convergent extension: parallel forces elongating the body axis. *Development* 128 (13):2581-2592

110. Wallingford JB and Harland RM (2002) Neural tube closure requires Dishevelled-dependent convergent extension of the midline. *Development* 129 (24):5815-5825

111. De Marco P, Merello E, Mascelli S, and Capra V (2006) Current perspectives on the genetic causes of neural tube defects. *Neurogenetics.* 7 (4):201-221

112. Ybot-Gonzalez P, Savery D, Gerrelli D, Signore M, Mitchell CE, Faux CH, Greene ND, and Copp AJ (2007) Convergent extension, planar-cell-polarity signalling and initiation of mouse neural tube closure. *Development* 134 (4):789-799

113. Kibar Z, Torban E, McDearmid JR, Reynolds A, Berghout J, Mathieu M, Kirillova I, De Marco P, Merello E, Hayes JM, Wallingford JB, Drapeau P, Capra V, and Gros P (2007) Mutations in VANGL1 associated with neural-tube defects. *N.Engl.J.Med.* 356 (14):1432-1437

114. Harris MJ and Juriloff DM (1999) Mini-review: toward understanding mechanisms of genetic neural tube defects in mice. *Teratology* 60 (5):292-305

115. Haldane JBS (1919) The combination of linkage values and the calculation of distances between two loci of linked factors. *Journal of Genetics* 8:299-309

116. Kosambi DD (1944) The estimation of map distances from recombination values. *Annal of Eugenics* 12:172-175

117. Clerget-Darpoux F, Bonaiti-Pellie C, and Hochez J (1986) Effects of misspecifying genetic parameters in lod score analysis. *Biometrics* 42 (June):393-399

118. Greenberg DA (1989) Inferring mode of inheritance by comparison of lod scores. *Am.J.Med.Genet.* 34:480-486

119. Greenberg DA (1990) Linkage analysis assuming a single-locus mode of inheritance for traits determined by two loci: inferring mode of inheritance and estimating penetrance. *Genet Epidemiol.* 7:467-479

120. Hodge SE and Elston RC (1994) Lods, wrods, and mods: the interpretation of lod scores calculated under different models. *Genet Epidemiol.* 11:329-342

121. Lander E and Kruglyak L (1995) Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat.Genet.* 11 (3):241-247

122. Penrose LS (1935) The detection of autosomal linkage in data which consists of pairs of brothers and sisters of unspecified parentage. *Annals of Eugenics* 6:133-138

123. Haseman JK and Elston RC (1972) The investigation of linkage between a quantitative trait and a marker locus. *Behavior Genetics* 2 (1):3-19

124. Lander ES and Green P (1987) Construction of multilocus genetic linkage maps in humans. *Proc.Natl.Acad.Sci.USA* 84:2363-2367

125. Kruglyak L, Daly MJ, Reeve-Daly MP, and Lander ES (1996) Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am.J.Hum.Genet.* 58 (6):1347-1363

126. Whittemore AS and Halpern J (1994) A class of tests for linkage using affected pedigree members. *Biometrics* 50 (1):118-127

127. Kong A and Cox NJ (1997) Allele-sharing models: LOD scores and accurate linkage tests. *Am.J.Hum.Genet.* 61 (5):1179-1188

128. Elston RC and Stewart J (1971) A general model for the genetic analysis of pedigree data. *Hum.Hered.* 21 (6):523-542

129. Ott J (1974) Estimation of the recombination fraction in human pedigrees: Efficient computation of the likelihood for human linkage studies. *Am.J.Hum.Genet.* 26:588-597

130. Lathrop GM, Lalouel JM, Julier C, and Ott J (1984) Strategies for multilocus linkage analysis in humans. *Proc.Natl.Acad.Sci.U.S.A* 81 (11):3443-3446

131. Lange K, Weeks DE, and Boehnke M (1988) Programs for pedigree analysis: MENDEL, FISHER, AND dGENE. *Genet Epidemiol.* 51:235-249

132. O'Connell JR and Weeks DE (1995) The VITESSE algorithm for rapid exact multilocus linkage analysis via genotype set-recoding and fuzzy inheritance. *Nat.Genet.* 11 (4):402-408

133. Gudbjartsson DF, Jonasson K, Frigge ML, and Kong A (2000) Allegro, a new computer program for multipoint linkage analysis. *Nat.Genet.* 25 (1):12-13

134. Rampersaud E, Bassuk AG, Enterline DS, M T, Siegel DG, Melvin EC, Aben J, Allen J, Aylsworth A, Brei T, Bodhurtha J, Buran C, Floyd LE, Hammock P, Iskandar B, Ito J, Kessler JA, Lasarsky N, Mack P, Mackey J, McLone D, Meeropol E, Mehltretter L, Mitchell LE, Oakes WJ, Nye JS, Powell C, Sawin K, Stevenson R, Walker M, West SG, Worley G, Gilbert JR, and Speer MC (2005) Whole genome-wide linkage screen for Neural Tube Defects reveals regions of interest on chromosomes 7 and 10. *J.Med.Genet.*

135. Yonan AL, Alarcon M, Cheng R, Magnusson PKE, Spence SJ, Palmer AA, Grunn A, Juo SHH, Terwilliger JD, Liu JJ, Cantor RM, Geschwind DH, and Gilliam TC (2003) A genomewide screen of 345 families for autism-susceptibility loci. *Am.J.Hum.Genet.* 73 (4):886-897

136. Hauser MA, Li YJ, Takeuchi S, Walters R, Noureddine M, Maready M, Darden T, Hulette C, Martin E, Hauser E, Xu H, Schmechel D, Stenger JE, Dietrich F, and Vance J (2003) Genomic convergence: identifying candidate genes for Parkinson's disease by combining serial analysis of gene expression and genetic linkage. *Hum.Mol.Genet.* 12 (6):671-677

137. Krauthammer M, Kaufmann CA, Gilliam TC, and Rzhetsky A (2004) Molecular triangulation: bridging linkage and molecular-network information for identifying candidate genes in Alzheimer's disease. *Proc.Natl.Acad.Sci.U.S.A* 101 (42):15148-15153

138. Marnellos G (2003) High-throughput SNP analysis for genetic association studies. *Curr.Opin.Drug Discov.Devel.* 6 (3):317-321

139. Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, Cho EK, Dallaire S, Freeman JL, Gonzalez JR, Gratacos M, Huang J, Kalaitzopoulos D, Komura D, MacDonald JR, Marshall CR, Mei R, Montgomery L, Nishimura K, Okamura K, Shen F,

Somerville MJ, Tchinda J, Valsesia A, Woodwark C, Yang F, Zhang J, Zerjal T, Zhang J, Armengol L, Conrad DF, Estivill X, Tyler-Smith C, Carter NP, Aburatani H, Lee C, Jones KW, Scherer SW, and Hurles ME (2006) Global variation in copy number in the human genome. *Nature* 444 (7118):444-454

140. Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P, Maner S, Massa H, Walker M, Chi M, Navin N, Lucito R, Healy J, Hicks J, Ye K, Reiner A, Gilliam TC, Trask B, Patterson N, Zetterberg A, and Wigler M (2004) Large-scale copy number polymorphism in the human genome. *Science* 305 (5683):525-528

141. Tuzun E, Sharp AJ, Bailey JA, Kaul R, Morrison VA, Pertz LM, Haugen E, Hayden H, Albertson D, Pinkel D, Olson MV, and Eichler EE (2005) Fine-scale structural variation of the human genome. *Nat.Genet.* 37 (7):727-732

142. Sharp AJ, Locke DP, McGrath SD, Cheng Z, Bailey JA, Vallente RU, Pertz LM, Clark RA, Schwartz S, Segraves R, Oseroff VV, Albertson DG, Pinkel D, and Eichler EE (2005) Segmental duplications and copy-number variation in the human genome. *Am.J.Hum.Genet.* 77 (1):78-88

143. McCarroll SA, Hadnott TN, Perry GH, Sabeti PC, Zody MC, Barrett JC, Dallaire S, Gabriel SB, Lee C, Daly MJ, and Altshuler DM (2006) Common deletion polymorphisms in the human genome. *Nat.Genet.* 38 (1):86-92

144. Simon-Sanchez J, Scholz S, Fung HC, Matarin M, Hernandez D, Gibbs JR, Britton A, De Vrieze FW, Peckham E, Gwinn-Hardy K, Crawley A, Keen JC, Nash J, Borgaonkar D, Hardy J, and Singleton A (2007) Genome-wide SNP assay reveals structural genomic variation, extended homozygosity and cell-line induced alterations in normal individuals. *Hum.Mol.Genet.* 16 (1):1-14

145. Wong KK, deLeeuw RJ, Dosanjh NS, Kimm LR, Cheng Z, Horsman DE, MacAulay C, Ng RT, Brown CJ, Eichler EE, and Lam WL (2007) A comprehensive analysis of common copy-number variations in the human genome. *Am.J.Hum.Genet.* 80 (1):91-104

146. Stewart CR, Kahler SG, and Gilchrist JM (1988) Congenital myopathy with cleft palate and increased susceptibility to malignant hyperthermia: King syndrome? *Ped.Neurol.* 4:371-374

147. Bryant A, Jr. and LaFromboise TD (2005) The racial identity and cultural orientation of Lumbee American Indian high school students. *Cultur.Divers.Ethnic.Minor.Psychol.* 11 (1):82-89

148. Berry B (1963) Almost White. New York: MacMillan

149. McMillan H. 1888. *Sir Walter Raleigh's Lost Colony: An Historical Sketch of the Attempts of Sir Walter Raleigh to Estabilish a Colony in Virgina, with the Traditions of an Indian Tribe in North Carolina.* Wilson, NC: Advance Press

150. Pollitzer WS (1964) Analysis of a Tri-Racial Isolate. *Hum.Biol.* 36:362-373

151. Pollitzer WS, Menegaz-Bock RM, and Herion JC (1966) Factors in the microevolution of a triracial isolate. *Am.J.Hum.Genet.* 18 (1):26-38

152. Lumbee Tribe of North Carolina. Lumbee Federal Recognition. 2007.

153. Bailey AG and Bloch EC (1987) Malignant hyperthermia in a three-month-old American Indian infant. *Anesth.Analg.* 66 (10):1043-1045

154. Meluch AM, Sibert KS, and Bloch EC (1989) Malignant hyperthermia following isoflurane anesthesia in an American Lumbee Indian. *N.C.Med.J.* 50 (9):485-487

155. Habib AS, Millar S, Deballi P, III, and Muir HA (2003) Anesthetic management of a ventilator-dependent parturient with the King-Denborough syndrome. *Can.J.Anaesth.* 50 (6):589-592

156. Abel DE and Grotegut CA (2003) King syndrome in pregnancy. *Obstet.Gynecol.* 101 (5 Pt 2):1146-1149

157. Robinson R, Carpenter D, Shaw MA, Halsall J, and Hopkins P (2006) Mutations in RYR1 in malignant hyperthermia and central core disease. *Hum.Mutat.* 27 (10):977-989

158. Froster-Iskenius UG, Waterson JR, and Hall JG (1988) A recessive form of congenital contractures and torticollis associated with malignant hyperthermia. *J.Med.Genet.* 25 (2):104-112

159. Sinkovich DD and Mitch-Resignalo AE (1991) Malignant hyperthermia. *Orthop.Nurs.* 10 (1):39-43

160. Britt BA (1972) Recent advances in malignant hyperthermia. *Anesth.Analg.* 51 (5):841-850

161. Kolb ME, Horne ML, and Martz R (1982) Dantrolene in human malignant hyperthermia. *Anesthesiology* 56 (4):254-262

162. Krause T, Gerbershagen MU, Fiege M, Weisshorn R, and Wappler F (2004) Dantrolene--a review of its pharmacology, therapeutic use and new developments. *Anaesthesia* 59 (4):364-373

163. (1984) A protocol for the investigation of malignant hyperpyrexia (MH) susceptibility. The European Malignant Hyperpyrexia Group. *Br.J.Anaesth.* 56 (11):1267-1269

164. Litman RS and Rosenberg H (2005) Malignant hyperthermia: update on susceptibility testing. *JAMA* 293 (23):2918-2924

165. Moslehi R, Langlois S, Yam I, and Friedman JM (1998) Linkage of malignant hyperthermia and hyperkalemic periodic paralysis to the adult skeletal muscle sodium channel (SCN4A) gene in a large pedigree. *Am.J.Med.Genet.* 76 (1):21-27

166. Sudbrak R, Golla A, Hogan K, Powers P, Gregg R, Du C, I, Lehmann-Horn F, and Deufel T (1993) Exclusion of malignant hyperthermia susceptibility (MHS) from a putative MHS2 locus on chromosome 17q and of the alpha 1, beta 1, and gamma subunits of the dihydropyridine receptor calcium channel as candidates for the molecular defect. *Hum.Mol.Genet.* 2 (7):857-862

167. Gillard EF, Otsu K, Fujii J, Khanna VK, de Leon S, Derdemezi J, Britt BA, Duff CL, Worton RG, and MacLennan DH (1991) A substitution of cysteine for arginine 614 in the ryanodine receptor is potentially causative of human malignant hyperthermia. *Genomics* 11 (3):751-755

168. Manning BM, Quane KA, Ording H, Urwyler A, Tegazzin V, Lehane M, O'Halloran J, Hartung E, Giblin LM, Lynch PJ, Vaughan P, Censier K, Bendixen D, Comi G, Heytens L, Monsieurs K, Fagerlund T, Wolz W, Heffron JJA, Muller CR, and McCarthy TV (1998) Identification of Novel Mutations in the Ryanodine-Receptor Gene (RYR1) in Malignant Hyperthermia: Genotype-Phenotype Correlation. *Am.J.Hum.Genet.* 62:599-609

169. Zhou H, Jungbluth H, Sewry CA, Feng L, Bertini E, Bushby K, Straub V, Roper H, Rose MR, Brockington M, Kinali M, Manzur A, Robb S, Appleton R, Messina S, D'Amico A, Quinlivan R, Swash M, Muller CR, Brown S, Treves S, and Muntoni F (2007) Molecular mechanisms and phenotypic variation in RYR1-related congenital myopathies. *Brain*

170. Levitt RC, Nouri N, Jedlicka AE, McKusick VA, Marks AR, Shutack JG, Fletcher JE, Rosenberg H, and Meyers DA (1991) Evidence for genetic heterogeneity in malignant hyperthermia susceptibility. *Genomics* 11 (3):543-547

171. Levitt RC, Olckers A, Meyers S, Fletcher JE, Rosenberg H, Isaacs H, and Meyers DA (1992) Evidence for the localization of a malignant hyperthermia susceptibility locus (MHS2) to human chromosome 17q. *Genomics* 14:562-566

172. Iles DE, Lehmann-Horn F, Scherer SW, Tsui LC, Olde WD, Suijkerbuijk RF, Heytens L, Mikala G, Schwartz A, Ellis FR, and . (1994) Localization of the gene encoding the alpha 2/delta-subunits of the L-type voltage-dependent calcium channel to chromosome 7q and analysis of the segregation of flanking markers in malignant hyperthermia susceptible families. *Hum.Mol.Genet.* 3 (6):969-975

173. Schleithoff L, Mehrke G, Reutlinger B, and Lehmann-Horn F (1999) Genomic structure and functional expression of a human alpha(2)/delta calcium channel subunit gene (CACNA2). *Genomics* 61 (2):201-209

174. Sudbrak R, Procaccio V, Klausnitzer M, Curran JL, Monsieurs K, Van Broeckhoven C, Ellis R, Heyetens L, Hartung EJ, Kozak-Ribbens G, and . (1995) Mapping of a further malignant hyperthermia susceptibility locus to chromosome 3q13.1. *Am.J.Hum.Genet.* 56 (3):684-691

175. Robinson RL, Monnier N, Wolz W, Jung M, Reis A, Nuernberg G, Curran JL, Monsieurs K, Stieglitz P, Heytens L, Fricker R, Van Broeckhoven C, Deufel T, Hopkins PM, Lunardi J, and Mueller CR (1997) A genome wide search for susceptibility loci in three European malignant hyperthermia pedigrees. *Hum.Mol.Genet.* 6 (6):953-961

176. Monnier N, Procaccio V, Stieglitz P, and Lunardi J (1997) Malignant-hyperthermia susceptibility is associated with a mutation of the alpha 1-subunit of the human dihydropyridine-sensitive L-type voltage-dependent calcium-channel receptor in skeletal muscle. *Am.J.Hum.Genet.* 60 (6):1316-1325

177. Stewart SL, Hogan K, Rosenberg H, and Fletcher JE (2001) Identification of the Arg1086His mutation in the alpha subunit of the voltage-dependent calcium channel (CACNA1S) in a North American family with malignant hyperthermia. *Clin.Genet.* 59 (3):178-184

178. Robinson R, Hopkins P, Carsana A, Gilly H, Halsall J, Heytens L, Islander G, Jurkat-Rott K, Muller C, and Shaw MA (2003) Several interacting genes influence the malignant hyperthermia phenotype. *Hum.Genet.* 112 (2):217-218

179. Rosenberg H, Davis M, James D, Pollock N, and Stowell K (2007) Malignant hyperthermia. *Orphanet.J.Rare.Dis.* 2:21

180. Fujii J, Otsu K, Zarzato F, de Leon S, Khanna VK, Weiler JE, O'Brien PJ, and MacLennan DH (1991) Identification of a mutation in porcine ryanodine receptor associated with malignant hyperthermia. *Science* 253:448-451

181. Sambuughin N, McWilliams S, de Bantel A, Sivakumar K, and Nelson TE (2001) Single-amino-acid deletion in the RYR1 gene, associated with malignant hyperthermia susceptibility and unusual contraction phenotype. *Am.J.Hum.Genet.* 69 (1):204-208

182. Fill M, Coronado R, Mickelson JR, Vilven J, Ma JJ, Jacobson BA, and Louis CF (1990) Abnormal ryanodine receptor channels in malignant hyperthermia. *Biophys.J.* 57 (3):471-475

183. Jungbluth H, Sewry CA, and Muntoni F (2003) What's new in neuromuscular disorders? The congenital myopathies. *Eur.J.Paediatr.Neurol.* 7 (1):23-30

184. Wu S, Ibarra MC, Malicdan MC, Murayama K, Ichihara Y, Kikuchi H, Nonaka I, Noguchi S, Hayashi YK, and Nishino I (2006) Central core disease is due to RYR1 mutations in more than 90% of patients. *Brain* 129 (Pt 6):1470-1480

185. Treves S, Anderson AA, Ducreux S, Divet A, Bleunven C, Grasso C, Paesante S, and Zorzato F (2005) Ryanodine receptor 1 mutations, dysregulation of calcium homeostasis and neuromuscular disorders. *Neuromuscul.Disord.* 15 (9-10):577-587

186. Lander ES and Botstein D (1987) Homozygosity Mapping: a way to map human recessive traits with the DNA of inbred children. *Science* 236:1567-1570

187. Houwen RHJ, Baharloo S, Blankenship K, Raeymaekers P, Juyn J, Sandkuijl LA, and Freimer NB (1994) Genome screening by searching for shared segments: mapping a gene for benign recurrent intrahepatic cholestasis. *Nat.Genet.* 8:380-386

188. Broman KW and Weber JL (1999) Long homozygous chromosomal segments in reference families from the centre d'Etude du polymorphisme humain. *Am.J.Hum.Genet.* 65 (6):1493-1500

189. Gibson J, Morton NE, and Collins A (2006) Extended tracts of homozygosity in outbred human populations. *Hum.Mol.Genet.* 15 (5):789-795

190. Li LH, Ho SF, Chen CH, Wei CY, Wong WC, Li LY, Hung SI, Chung WH, Pan WH, Lee MT, Tsai FJ, Chang CF, Wu JY, and Chen YT (2006) Long contiguous stretches of homozygosity in the human genome. *Hum.Mutat.* 27 (11):1115-1121

191. Puffenberger EG, Hu-Lince D, Parod JM, Craig DW, Dobrin SE, Conway AR, Donarum EA, Strauss KA, Dunckley T, Cardenas JF, Melmed KR, Wright CA, Liang W, Stafford P, Flynn CR, Morton DH, and Stephan DA (2004) Mapping of sudden infant death with dysgenesis of the testes syndrome (SIDDT) by a SNP genome scan and identification of TSPYL loss of function. *Proc.Natl.Acad.Sci.U.S.A* 101 (32):11689-11694

192. Kelsell DP, Norgett EE, Unsworth H, Teh MT, Cullup T, Mein CA, Dopping-Hepenstal PJ, Dale BA, Tadini G, Fleckman P, Stephens KG, Sybert VP, Mallory SB, North BV, Witt DR, Sprecher E, Taylor AE, Ilchyshyn A, Kennedy CT, Goodyear H, Moss C, Paige D, Harper JI, Young BD, Leigh IM, Eady RA, and O'Toole EA (2005) Mutations in ABCA12 underlie the severe congenital skin disease harlequin ichthyosis. *Am.J.Hum.Genet.* 76 (5):794-803

193. Strauss KA, Puffenberger EG, Craig DW, Panganiban CB, Lee AM, Hu-Lince D, Stephan DA, and Morton DH (2005) Genome-wide SNP arrays as a diagnostic tool: clinical description, genetic mapping, and molecular characterization of Salla disease in an Old Order Mennonite population. *Am.J.Med.Genet.A* 138 (3):262-267

194. Pannain S, Weiss RE, Jackson CE, Dian D, Beck JC, Sheffield VC, Cox N, and Refetoff S (1999) Two different mutations in the thyroid peroxidase gene of a

large inbred Amish kindred: Power and limits of homozygosity mapping. *Journal of Clinical Endocrinology and Metabolism* 84:1061-1071

195. Miano MG, Jacobson SG, Carothers A, Hanson I, Teague P, Lovell J, Cideciyan AV, Haider N, Stone EM, Sheffield VC, and Wright AF (2000) Pitfalls in homozygosity mapping. *Am.J.Hum.Genet.* 67 (5):1348-1351

196. Frishberg Y, Ben Neriah Z, Suvanto M, Rinat C, Mannikko M, Feinstein S, Becker-Cohen R, Jalanko H, Zlotogora J, and Kestila M (2007) Misleading findings of homozygosity mapping resulting from three novel mutations in NPHS1 encoding nephrin in a highly inbred community. *Genet.Med.* 9 (3):180-184

197. Lary JM and Edmonds LD (1996) Prevalence of spina bifida at birth-United States, 1983-1990: a comparison of two surveillance systems. *MMWR Morb Mortal Wkly Rep* 45:15-26

198. MRC Vitamin Study Research Group (1991) Prevention of neural tube defects: Results of the Medical Research Council Vitamin Study. *Lancet* 338:131-137

199. Khoury MJ and Wagener DK (1995) Epidemiological evaluation of the use of genetics to improve the predictive value of disease risk factors. *Am.J.Hum.Genet.* 56:835-844

200. Greene ND and Copp AJ (2005) Mouse models of neural tube defects: investigating preventive mechanisms. *Am.J.Med.Genet.C.Semin.Med.Genet.* 135 (1):31-41

201. Rampersaud E, Bassuk AG, Enterline DS, George TM, Siegel DG, Melvin EC, Aben J, Allen J, Aylsworth A, Brei T, Bodurtha J, Buran C, Floyd LE, Hammock P, Iskandar B, Ito J, Kessler JA, Lasarsky N, Mack P, Mackey J, McLone D, Meeropol E, Mehltretter L, Mitchell LE, Oakes WJ, Nye JS, Powell C, Sawin K, Stevenson R, Walker M, West SG, Worley G, Gilbert JR, and Speer MC (2005) Whole genomewide linkage screen for neural tube defects reveals regions of interest on chromosomes 7 and 10. *J.Med.Genet.* 42 (12):940-946

202. Evans DM and Cardon LR (2004) Guidelines for genotyping in genomewide linkage studies: single-nucleotide-polymorphism maps versus microsatellite maps. *Am.J.Hum.Genet.* 75 (4):687-692

203. Middleton FA, Pato MT, Gentile KL, Morley CP, Zhao X, Eisener AF, Brown A, Petryshen TL, Kirby AN, Medeiros H, Carvalho C, Macedo A, Dourado A, Coelho I, Valente J, Soares MJ, Ferreira CP, Lei M, Azevedo MH, Kennedy JL, Daly MJ, Sklar P, and Pato CN (2004) Genomewide linkage analysis of bipolar disorder by use of a high-density single-nucleotide-polymorphism (SNP) genotyping assay: a comparison with microsatellite marker assays and

finding of significant linkage to chromosome 6q22. *Am.J.Hum.Genet.* 74 (5):886-897

204. John S, Shephard N, Liu G, Zeggini E, Cao M, Chen W, Vasavda N, Mills T, Barton A, Hinks A, Eyre S, Jones KW, Ollier W, Silman A, Gibson N, Worthington J, and Kennedy GC (2004) Whole-genome scan, in a complex disease, using 11,245 single-nucleotide polymorphisms: comparison with microsatellites. *Am.J.Hum.Genet.* 75 (1):54-64

205. Schaid DJ, Guenther JC, Christensen GB, Hebbring S, Rosenow C, Hilker CA, McDonnell SK, Cunningham JM, Slager SL, Blute ML, and Thibodeau SN (2004) Comparison of microsatellites versus single-nucleotide polymorphisms in a genome linkage screen for prostate cancer-susceptibility Loci. *Am.J.Hum.Genet.* 75 (6):948-965

206. Huang Q, Shete S, and Amos CI (2004) Ignoring linkage disequilibrium among tightly linked markers induces false-positive evidence of linkage for affected sib pair analysis. *Am.J.Hum.Genet.* 75 (6):1106-1112

207. Boyles AL, Scott WK, Martin ER, Schmidt S, Li YJ, Ashley-Koch A, Bass MP, Schmidt M, Pericak-Vance MA, Speer MC, and Hauser ER (2005) Linkage disequilibrium inflates type I error rates in multipoint linkage analysis when parental genotypes are missing. *Hum.Hered.* 59 (4):220-227

208. Melvin EC, Scott WK, Speer MC, Wolpert C, and Pericak-Vance MA (1998) Development of a standardized family history questionnaire for research purposes. *Journal of Genetic Counseling* 7 (6):475

209. Boehnke M (1986) Estimating the power of a proposed linkage study: a practical computer simulation approach. *Am.J.Hum.Genet.* 39:513-527

210. Ploughman LM and Boehnke M (1989) Estimating the power of a proposed linkage study for a complex genetic trait. *Am.J.Hum.Genet.* 44 (4):543-551

211. Rampersaud E, Bassuk AG, Enterline DS, George TM, Siegel DG, Melvin EC, Aben J, Allen J, Aylsworth A, Brei T, Bodurtha J, Buran C, Floyd LE, Hammock P, Iskandar B, Ito J, Kessler JA, Lasarsky N, Mack P, Mackey J, McLone D, Meeropol E, Mehltretter L, Mitchell LE, Oakes WJ, Nye JS, Powell C, Sawin K, Stevenson R, Walker M, West SG, Worley G, Gilbert JR, and Speer MC (2005) Whole genomewide linkage screen for neural tube defects reveals regions of interest on chromosomes 7 and 10. *J.Med.Genet.* 42 (12):940-946

212. Goring HHH and Ott J (1995) Verification of sib relationship without knowledge of parental genotypes. *Am.J.Hum.Genet.* 57:A192

213. McPeek MS and Sun L (2000) Statistical tests for detection of misspecified relationships by use of genome-screen data. *Am.J.Hum.Genet.* 66 (3):1076-1094

214. Haynes C, Speer MC, Peedin M, Roses AD, Haines JL, Vance JM, and Pericak-Vance MA (1995) PEDIGENE: A comprehensive data management system to facilitate efficient and rapid disease gene mapping. *Am J Hum Genet* 57 (4):A193

215. O'Connell JR and Weeks DE (1998) PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am.J.Hum.Genet.* 63 (1):259-266

216. Carlson CS, Eberle MA, Rieder MJ, Yi Q, Kruglyak L, and Nickerson DA (2004) Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. *Am.J.Hum.Genet.* 74 (1):106-120

217. Weeks DE, Sobel E, O'Connell JR, and Lange K (1995) Computer programs for multilocus haplotyping of general pedigrees. *Am.J.Hum.Genet.* 56 (6):1506-1507

218. Xu H, Vance JM, and Stenger JE (2002) Integrating genomic laboratory data with whole genome annotation using Endembl-DAS. *Am.J.Hum.Genet.* 71:393-1301A

219. Lawrence N and Morel V (2003) Dorsal closure and convergent extension: two polarised morphogenetic movements controlled by similar mechanisms? *Mech.Dev.* 120 (11):1385-1393

220. Mankoo BS, Skuntz S, Harrigan I, Grigorieva E, Candia A, Wright CV, Arnheiter H, and Pachnis V (2003) The concerted action of Meox homeobox genes is required upstream of genetic pathways essential for the formation, patterning and differentiation of somites. *Development* 130 (19):4655-4664

221. Tanzi RE and Bertram L (2001) New frontiers in Alzheimer's disease genetics. *Neuron* 32 (2):181-184

222. Winn MP, Conlon PJ, Lynn KL, Farrington MK, Creazzo T, Hawkins AF, Daskalakis N, Kwan SY, Ebersviller S, Burchette JL, Pericak-Vance MA, Howell DN, Vance JM, and Rosenberg PB (2005) A mutation in the TRPC6 cation channel causes familial focal segmental glomerulosclerosis. *Science* 308 (5729):1801-1804

223. Wang L, Fan C, Topol SE, Topol EJ, and Wang Q (2003) Mutation of MEF2A in an inherited disorder with features of coronary artery disease. *Science* 302 (5650):1578-1581

224. Ashley-Koch AE, Shao Y, Rimmler JB, Gaskell PC, Welsh-Bohmer KA, Jackson CE, Scott WK, Haines JL, and Pericak-Vance MA (2005) An autosomal genomic screen for dementia in an extended Amish family. *Neurosci.Lett.* 379 (3):199-204

225. Puranam RS, Jain S, Kleindienst AM, Saxena S, Kim MK, Kelly CB, Padma MV, Andrews I, Elston RC, Tiwari HK, and McNamara JO (2005) A locus for generalized tonic-clonic seizure susceptibility maps to chromosome 10q25-q26. *Ann.Neurol.* 58 (3):449-458

226. Blanton SH, Liang CY, Cai MW, Pandya A, Du LL, Landa B, Mummalanni S, Li KS, Chen ZY, Qin XN, Liu YF, Balkany T, nance WE, and Liu XZ (2002) A novel locus for autosomal dominant non-syndromic deafness (DFNA41) maps to chromosome 12q24-qter. *J.Med.Genet.* 39 (8):567-570

227. Tlili A, Mannikko M, Charfedine I, Lahmar I, Benzina Z, Ben Amor M, Driss N, Ala-Kokko L, Drira M, Masmoudi S, and Ayadi H (2005) A Novel Autosomal Recessive Non-Syndromic Deafness Locus, DFNB66, Maps to Chromosome 6p21.2-22.3 in a Large Tunisian Consanguineous Family. *Hum.Hered.* 60 (3):123-128

228. Rampersaud E, Bassuk AG, Enterline DS, George TM, Siegel DG, Melvin EC, Aben J, Allen J, Aylsworth A, Brei T, Bodurtha J, Buran C, Floyd LE, Hammock P, Iskandar B, Ito J, Kessler JA, Lasarsky N, Mack P, Mackey J, McLone D, Meeropol E, Mehltretter L, Mitchell LE, Oakes WJ, Nye JS, Powell C, Sawin K, Stevenson R, Walker M, West SG, Worley G, Gilbert JR, and Speer MC (2005) Whole genomewide linkage screen for neural tube defects reveals regions of interest on chromosomes 7 and 10. *J.Med.Genet.* 42 (12):940-946

229. Nye JS, Balkin N, Lucas H, Knepper P, McLone D, and Charrow J (1998) Meningomyelocele and Waardenburg syndrome (type 3) in patients with interstitial deletion of 2q35 and the PAX3 gene: possible digenic inheritance of a neural tube defect. *Am.J.Med.Genet.* 75:401-408

230. Detrait ER, George TM, Etchevers HC, Gilbert JR, Vekemans M, and Speer MC (2005) Human neural tube defects: developmental biology, epidemiology, and genetics. *Neurotoxicol.Teratol.* 27 (3):515-524

231. (1991) Prevention of neural tube defects: results of the Medical Research Council Vitamin Study. MRC Vitamin Study Research Group. *Lancet* 338 (8760):131-137

232. O'Connell JR and Weeks DE (1998) PedCheck: a program for identification of genotype incompatibilities in linkage analysis. *Am.J.Hum.Genet.* 63 (1):259-266

233. Boehnke M and Cox NJ (1997) Accurate inference of relationships in sib-pair linkage studies. *Am.J.Hum.Genet.* 61 (2):423-429

234. Epstein MP, Duren WL, and Boehnke M (2000) Improved inference of relationship for pairs of individuals. *Am.J.Hum.Genet* 67 (5):1219-1231

235. Fiegler H, Carr P, Douglas EJ, Burford DC, Hunt S, Smith J, Vetrie D, Gorman P, Tomlinson IP, and Carter NP (2003) DNA microarrays for comparative genomic hybridization based on DOP-PCR amplification of BAC and PAC clones. *Genes Chromosomes.Cancer* 36 (4):361-374

236. Pinkel D, Segraves R, Sudar D, Clark S, Poole I, Kowbel D, Collins C, Kuo WL, Chen C, Zhai Y, Dairkee SH, Ljung BM, Gray JW, and Albertson DG (1998) High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat.Genet.* 20 (2):207-211

237. Ogino S, Gulley ML, Den Dunnen JT, and Wilson RB (2007) Standard mutation nomenclature in molecular diagnostics: practical and educational challenges. *J.Mol.Diagn.* 9 (1):1-6

238. Linardopoulou EV, Williams EM, Fan Y, Friedman C, Young JM, and Trask BJ (2005) Human subtelomeres are hot spots of interchromosomal recombination and segmental duplication. *Nature* 437 (7055):94-100

239. Edwards JH (1969) Inheritance of chronic granulomatous disease. *Lancet* 2 (7625):850-851

240. Hata K, Okano M, Lei H, and Li E (2002) Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development* 129 (8):1983-1993

241. Dellinger A., Thomas S., Xu P., Etchevers H.C., Vekemans M., Gilbert J.R., and Speer M.C. (2007) Global Analysis of Four Neural Tube Defect LongSAGE Libraries to Identify Candidate Genes for Anencephaly and Spina Bifida. *American Society for Human Genetics*

242. Basinger AA, Booker JK, Frazier DM, Koeberl DD, Sullivan JA, and Muenzer J (2006) Glutaric acidemia type 1 in patients of Lumbee heritage from North Carolina. *Mol.Genet.Metab* 88 (1):90-92

243. Minegishi Y, Lavoie A, Cunningham-Rundles C, Bedard PM, Hebert J, Cote L, Dan K, Sedlak D, Buckley RH, Fischer A, Durandy A, and Conley ME (2000) Mutations in activation-induced cytidine deaminase in patients with hyper IgM syndrome. *Clin.Immunol.* 97 (3):203-210

244. Jurkat-Rott K, McCarthy T, and Lehmann-Horn F (2000) Genetics and pathogenesis of malignant hyperthermia. *Muscle Nerve* 23 (1):4-17

245. Strazis KP and Fox AW (1993) Malignant hyperthermia: a review of published cases. *Anesth.Analg.* 77 (2):297-304

246. Robinow M and Johnson GF (1981) The Gordon syndrome: autosomal dominant cleft palate, camptodactyly, and club feet. *Am.J.Med.Genet.* 9 (2):139-146

247. Van Bogaert P, Azizieh R, Desir J, Aeby A, De Meirleir L, Laes JF, Christiaens F, and Abramowicz MJ (2007) Mutation of a potassium channel-related gene in progressive myoclonic epilepsy. *Ann.Neurol.*

248. Haines JL, Pericak-Vance MA. 1998. *Approaches to Gene Mapping in Complex Human Diseases.* Vol. 1. Haines JL, Pericak-Vance MA (eds). New York, N.Y.: Wiley & Sons, Inc.

249. Hayashi YK, Chou FL, Engvall E, Ogawa M, Matsuda C, Hirabayashi S, Yokochi K, Ziober BL, Kramer RH, Kaufman SJ, Ozawa E, Goto Y, Nonaka I, Tsukahara T, Wang JZ, Hoffman EP, and Arahata K (1998) Mutations in the integrin alpha 7 gene cause congenital myopathy. *Nat.Genet.* 19 (1):94-97

250. Clark AG (1999) The size distribution of homozygous segments in the human genome. *Am.J.Hum.Genet.* 65 (6):1489-1492

251. Crisponi L, Crisponi G, Meloni A, Toliat MR, Nurnberg G, Usala G, Uda M, Masala M, Hohne W, Becker C, Marongiu M, Chiappe F, Kleta R, Rauch A, Wollnik B, Strasser F, Reese T, Jakobs C, Kurlemann G, Cao A, Nurnberg P, and Rutsch F (2007) Crisponi syndrome is caused by mutations in the CRLF1 gene and is allelic to cold-induced sweating syndrome type 1. *Am.J.Hum.Genet.* 80 (5):971-981

252. Yoshida R, Hasegawa T, Hasegawa Y, Nagai T, Kinoshita E, Tanaka Y, Kanegane H, Ohyama K, Onishi T, Hanew K, Okuyama T, Horikawa R, Tanaka T, and Ogata T (2004) Protein-tyrosine phosphatase, nonreceptor type 11 mutation analysis and clinical assessment in 45 patients with Noonan syndrome. *J.Clin.Endocrinol.Metab* 89 (7):3359-3364

253. Murphy SK and Jirtle RL (2003) Imprinting evolution and the price of silence. *Bioessays* 25 (6):577-588

254. Woods CG, Cox J, Springell K, Hampshire DJ, Mohamed MD, McKibbin M, Stern R, Raymond FL, Sandford R, Malik SS, Karbani G, Ahmed M, Bond J, Clayton D, and Inglehearn CF (2006) Quantification of homozygosity in consanguineous individuals with autosomal recessive disease. *Am.J.Hum.Genet.* 78 (5):889-896