ABSTRACT

Title of Document:	GENETIC STUDIES OF STUTTERING IN A LARGE WEST AFRICAN FAMILY
	Bailey Levis, MS, 2006
Directed By:	Associate Professor, Sarah Tishkoff, Department of Biology

A large family of West African origin displaying persistent developmental stuttering was ascertained. This family contains 106 individuals, 45 of whom display the phenotype. A genome wide scan was conducted using 366 microsatellite markers, with additional markers in regions of interest. Parametric linkage analyses were conducted under several different models of inheritance patterns. The highest two point LOD scores were obtained under a fully penetrant model on chromosome 1p31 at marker D1S1588 (LOD = 2.97, $\theta = 0.062$), and chromosome 20 at marker D20S873 (LOD = 2.97, $\theta = 0.15$). Multipoint LOD scores provided significant evidence for linkage on chromosome 1 (LOD=4.39) with markers D1S2774 and D1S3471 (114 cM and 117 cM) respectively and suggestive evidence of linkage on chromosome 20 (LOD=2.22) with markers D20S482 and D20S851 (12 cM and 25 cM, respectively). These data provide suggestive evidence for an association of PDS to multiple chromosomal locations.

GENETIC STUDIES OF STUTTERING IN A LARGE WEST AFRICAN FAMILY

By

Bailey Levis

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Master of Science
2006

Advisory Committee:
Associate Professor Sarah Tishkoff, Chair
Professor Nan Bernstein Ratner
Adjunct Professor, Dennis T. Drayna

Dedication

This work is dedicated to my Grandpa Ben Levis and in memory of Raif Simon

Acknowledgements

I would like to thank my parents, Marilee and Lenny for always encouraging me to follow my heart and for always being supportive of the decisions I have made in life and through all of the struggles (and rewards!) that those decisions have brought on. I would like to thank my brother Jason for an incredible sense of reality that often grounds me, and my sister Sophia for the will to survive.

I would especially like to thank my advisory committee, Dr. Nan Bernstein Ratner, Dr. Dennis Drayna, and Dr. Sarah Tishkoff for imparting their knowledge during my course of studies, providing invaluable guidance and especially for their insightful comments on earlier drafts of this thesis. I would like to extend special thanks to Dr. Dennis Drayna who has provided me the opportunity to pursue my passion for stuttering research. Dr. Drayna has taught me what it means to be a good scientist and has offered a different perspective on life as well as unique insights into human nature. I would also like to extend a special thanks to Dr. Nan Bernstein Ratner, one of the most knowledgeable people in the field of stuttering research. Dr. Bernstein Ratner opened my eyes to the wide world of everything stuttering.

I would also like to thank the NIDCD / NIH for continued support throughout my studies. The research presented in this thesis was carried out exclusively in Dr. Drayna's lab at the NIDCD.

Table of Contents

Chapter 1: Introduction

Characterization

Stuttering is a speech disorder in which affected individuals know what they want to say but are unable to say it. The World Health Organization ICD-10 defines stuttering as "Speech that is characterized by frequent repetition or prolongation of sounds or syllables or words, or by frequent hesitations or pauses that disrupt the rhythmic flow of speech" (WHO, 2004). According to the American Psychological Association's Diagnostic and Statistical Manual (DSM-IV), the diagnostic criteria for stuttering are:

A. Disturbance in the normal fluency and time patterning of speech (inappropriate for the individual's age), characterized by frequent occurrences of one or more of the following:
    (1) Sound and syllable repetitions
    (2) Sound prolongations
    (3) Interjections
    (4) Broken words (e.g., pauses within a word)
    (5) Audible or silent blocking (filled or unfilled pauses in speech)
    (6) Circumlocutions (word substitutions to avoid problematic words)
    (7) Words produced with an excess of physical tension
    (8) Monosyllabic whole-word repetitions (e.g., "I-I-I-I see him")
B. The disturbance in fluency interferes with academic or occupational achievement or with social communication.

C. If a speech-motor or sensory deficit is present, the speech difficulties are in excess of those usually associated with these problems. (DSM-IV, 1994, p. 65)

Charles Van Riper defines stuttering as "A deviation in the ongoing fluency of speech, an inability to maintain the connected rhythms of speech." (Van Riper, 1982)

The incidence of stuttering is about 5% (Bloodstein, 1995). About 80% of those cases spontaneously recover and about 1% of the population persists into adulthood. Stuttering exerts a deep emotional and psychological effect on those affected with the disorder. People who stutter (PWS) are prone to economic hardships as well (Klein & Hood, 2004). Many PWS believe that their stuttering decreases their chances of getting hired or promoted. Some PWS may also turn down job promotions because more public speaking would be involved (Klein & Hood, 2004). Because of the difficulties and challenges that PWS face in their lives there is a strong need to better understand the disorder and its causes.

Etiological Theories of Stuttering

The nature vs. nurture debate has a long history in stuttering, with theories formulated from both perspectives. Neither of these perspectives has been able to completely and satisfactorily explain both the development and persistence of stuttering, and many do not take into account potential and likely combinations of factors. The majority of theories of the etiology of stuttering have been based on potential environmental, psychological or physical causes.

Role of the Environment

Two popular theories that fall into this category include, Johnson's (1942) diagnosogenic theory and anticipatory struggle theories. Both of these will be discussed in turn.

Diagnosogenic Theory

The diagnosogenic theory (Johnson, 1942), put simply, suggests that stuttering develops due to the mis-diagnosis of a child's normal speech disfluency by his or her parents. It was believed that the parents made the child so self-conscious about his/her speech that the normal developmental disfluencies would persist as stuttering. Evidence in support of the theory has been sparse but has widely influenced doctors', parents' and even speech-language pathologists' reluctance to work directly with children at onset of stuttering. The diagnosogenic theory was tested by one of Johnson's students in an unpublished master's thesis, discussed many years later in a paper by Silverman (1988). Tudor (1939) intentionally identified six children at an orphanage as displaying stuttering-like behavior when in fact they were normally fluent, in an effort to see if mis-labeling would have an effect on speech fluency. The caretakers at the orphanage were told to monitor the children and label normal disfluencies as problematic and undesirable. Ambrose & Yairi, (2002) performed a reanalysis of Tudor's data. Upon doing so it became clear that identifying normally fluent children as showing stuttering like behavior did not lead to the development of stuttering. In a newspaper article published in the San Jose Mercury News, Dryer (2001) claims that at least one individual (participant #11)

went on to develop major speech disorders and an extreme shyness regarding her speech. Ambrose and Yairi (2002) point out that there was no scientific evidence confirming that her speech disorder was a result of this study. At the beginning of this study she was asked if she knew anyone who stuttered and said that she did. However it was not clear if she was referring to someone else or perhaps herself, as her percent disfluency was observed to be ~5%, a range that is considered moderate. Participant 11 showed an increase in prolongations towards the end of the study, and a decrease in repetitions. Taken together it is difficult to determine whether her overall fluency improved or worsened. Furthermore, because of her disfluency at the beginning of the study, it is impossible to postulate with any certainty that her disfluencies were due to the experiment. The overall results of this study do not support Johnson's Diagnosogenic theory of stuttering. Additional evidence that the environment does not play a primary role in either the origins or persistence of developmental stuttering comes from a recent highly successful therapeutic approach, the Lidcombe program (Harris et al., 2002). This approach provides feedback to children's early stuttering behavior by their parents. When the child speaks smoothly and without any stuttering like behaviors the parent(s) provide positive feedback. When a child unambiguously displays stuttering like behaviors the parent(s) will make a comment along the lines of, "That sounded like it was a little bumpy for you. Would you try saying it again with no bumps?" (Harris et al., 2002). According to the diagnosogenic theory, by making the child aware of their stuttering, their fluency should worsen. This in fact is not the case.

Anticipatory Struggle

Anticipatory struggle theories can be boiled down to the simple hypothesis that PWS have a belief that speech is difficult. According to the theory, PWS will anticipate stuttering, start to fear that moment and tense up. This tension is then manifested as stuttering. The triggers in anticipatory struggle theories can come from within the child and also from external sources. Some suggest that parents put a high value on correct and proper speech. The child may begin to be very conscious of his/her speech, and hypersensitive to any disfluencies. These disfluencies can lead to a greater fear of speaking, thus leading back to increased stuttering. Van Riper developed the preparatory set hypothesis (Van Riper, 1937, 1954). His theory consists of a set of physical and psychological states leading to the moment of stuttering. The three aspects of the preparatory set include first, an awareness of an upcoming word or sound that is difficult; second, a series of rehearsals by the speech muscles; and finally, the arrival of difficult moment (Van Riper, 1937). The anticipation of stuttering then leads to the stuttering behavior.

Psychological Theories of Stuttering

There are two ways in which psychological theories of stuttering can be addressed. One has to do with the moment of stuttering, while the other has to do with the development and persistence of stuttering. The repressed needs hypothesis states that the moment of stuttering represents unconscious desires that need to be satisfied (see summary in Bloodstein and Bernstein Ratner, 2007). This hypothesis has been built by

several psychoanalysts during the early to mid part of the 20th century. For example, Fenichel, (1945) believed that "the anal-sadistic universe of wishes [of people who stutter]" (Fenichel, 1945 p.312) was the basis for stuttering. He compared the "expulsion and retention of speech" of stuttering to the "expulsion and retention of feces." (p.312). Other psychoanalysts also believed that stuttering was in some way related to early sexual needs (Coriat, 1927). Additionally, Fenichel (1945) believed that stuttering was due to an unconscious hostility towards listeners. Research into the psychological profiles of PWS, and their families, together with the lack of demonstrated therapeutic success of psychological treatments for stuttering, strongly suggests that these theories are not valid. Studies of PWS do not find that they are more likely to be emotionally disturbed than the general population. Furthermore, psychotherapy has little success in improving fluency in PWS (for review, see Bloodstein & Bernstein Ratner, 2007).

Additional psychological theories of stuttering include theories that invoke classical conditioning and operant conditioning. The most traditional example of classical conditioning is that of Pavlov and his dog, which was trained to salivate at the ring of a bell. An application of this to developmental stuttering would be the child's feeling that s/he is receiving negative input from a parent following a speech disfluency. This could create tension or fear. This tension might manifest itself as stuttering only when speaking to the disapproving parent. Over time, however, this tension and fear can make its way into speaking situations where the child had previously felt comfortable. As this continues the child develops an aversion to all speaking situations.

Operant conditioning makes use of positive and negative reinforcements. In stuttering this type of conditioning has been associated with secondary behaviors such as eye-blinks, facial tics, or clenched fists. A child would recognize s/he is stuttering (and struggling) and realizes that by clenching a fist, s/he can say the word s/he is struggling with. This becomes a form of negative reinforcement. The fist makes the bad situation better. Over time, the effect of the clenched fist might wear off, but the behavior has been habituated through negative reinforcement. This type of learning theory explains secondary behaviors associated with stuttering much better than the initial symptoms of the disorder. This theory is therefore not incompatible with the developed symptoms, but does not do a good job of explaining the origins of the primary speech symptoms in childhood.

Physical / Constitutional Theories of Stuttering

Physical or constitutional theories of stuttering postulate that stuttering can be due to atypical cerebral dominance, poor motor control, or disordered timing of neural signals. Such problems are readily compatible with a genetic etiology for the disorder. For the most part, these theories postulate dysfunctions within the brain that result in stuttering behaviors.

Atypical Cerebral Dominance

Orton and Travis first introduced the theory of cerebral dominance in the 1930's (Travis, 1931). This theory reflected the long-held belief that stuttering often followed forced shifts to right-handedness, and the observation that many PWS were either left-

handed or ambidextrous. It was thought that this resulted in a conflict between the two hemispheres of the brain. Specifically, the hypothesis speculated that if one hemisphere were not dominant over the other, then the two hemispheres might be poorly synchronized. This would lead to poorly synchronized signals being sent to the speech musculature. However, more recent research has emerged suggesting that PWS are not more likely to be left-handed or ambidextrous than the general population. Additionally, it has been reported that many children who have been forced to right-handedness do not stutter (Bloodstein & Bernstein Ratner, 2007).

However, the theory has more recent support from other types of research. More recent studies of stuttering have investigated hemispheric lateralization of speech using structural and functional imaging techniques (e.g., Foundas, Bollich, Corey, Hurley, & Heilman, 2001; Salmelin, Schnitzler, Schmitz, & Freund, 2000; Wu et al., 1997). Findings from these studies have shown differences in neural activation patterns in PWS compared to people who do not stutter (PWNS). These differences occur in many areas of the brain, in both right and left hemispheres. One of the consistent findings is that language function in PWS seems to be more bilateral than in PWNS. However, it is not entirely clear if the differences cause stuttering or if they are in response to the stuttering (i.e. compensatory or emotional).

Motor and Pre-motor Theories of Stuttering

There is a large body of literature suggesting deficits in neural processing in PWS. These deficits have been identified not only in language, but also in areas of motor and pre-motor function and will be discussed below.

It is well understood that the basal ganglia are involved in motor movements (Gazzaniga, 2002). Research has suggested that the basal ganglia play a role in stuttering. Evidence of this comes not only from studies of developmental stuttering (Salmelin, Schnitzler, Schmitz, & Freund, 2000; Wu et al., 1997) but also from lesion studies of acquired stuttering (Ludlow & Loucks, 2003). When one considers the fact that the basal ganglia and thalamus, with their connections to the motor cortex, have been repeatedly implicated in stuttering, it is reasonable to hypothesize that there must be at least some degree of motor dysfunction involved in the underlying etiology of stuttering events. The hypothesis of a more general motor disorder is further supported by results obtained by Jones et al. (2002) who observed deficits in PWS during various movement and tracking tasks, in which PWS performed slower or less accurately than control subjects. One possible explanation comes from Brotchie, et al. (1991). They argue that the basal ganglia have a mechanism for switching off sustained pre-movement activity and switching on the preparatory phase for the next movement. This provides the ability to prime the next motor movement, and to shift the planned movement from a planning phase to an execution phase. This transition must occur rapidly and thus requires proper neural synchronization and processing. If this process is disrupted it could lead to the observed differences in the movement and tracking tasks reported by Jones, et al. (2002). Research in the past few years has lead to the elucidation of stages of motor control and planning for speech, (Ludlow & Loucks, 2003 for review). Salmelin et al. (2000) performed studies of single word reading in PWS and controls and measured brain activation using Magnetoencephalography (MEG). Neural responses to visually presented words were analyzed over the 400 ms time period post stimulus onset. As the

task in this study involved a spoken response, this time period is thought to be involved in the planning of speech. They found that both groups show activity in similar regions, but the PWS activate the motor execution regions of the left central middle inferior region prior to the left inferior region associated with articulation planning (Salmelin et al., 2000).

Evidence for motor dysfunction in PWS comes from research that has measured orofacial movements. Zimmermann, (1980) compared articulatory movements in PWS during fluent and stuttered speech to normally fluent individuals. He found that articulator movements in PWS showed patterns that were unlike those in the normally fluent subjects. He also observed that normally fluent individuals showed more synchronous articulator movement than PWS during both fluent and non-fluent speech. Additional evidence for disorders of timing comes from Guitar et al. (1988). Using electromyography (EMG), they measured the movement of two lip muscles during a single - word speaking task. They found that the sequence of muscle movement was reversed in PWS compared to normally fluent individuals. Additional evidence exists as well that supports the theory of speech motor dysfunction in PWS (Alfonso, 1991; Hulstijn & Van Lieshout, 1998; Van Lieshout, 1995). A confounding factor in these studies is that not all PWS appear to show the same motor dysfunction (Denny & Smith, 1992; McClean, Tasko, & Runyan, 2004). This suggests that motor dysfunction in PWS is not the same in all individuals.

Dopamine hypothesis of stuttering

Evidence exists suggesting dopamine (DA), a neurotransmitter, plays a role in stuttering. Evidence of this comes from studies of anti-psychotics that have had a positive effect on improving fluency. In particular, Haloperidol has been quite successful in improving fluency in PWS, and works specifically on dopamine type 2 receptors (D2Rs) (Brady, 1991). Even though it has improved the fluency in many individuals, the side effects are usually so severe that PWS will not continue to remain on the drug (Brady, 1991). Wu et al. (1997) have suggested that PWS show an increase in DA activity in the basal ganglia. Furthermore Alm (2004) suggests that a correlation may exist between levels of type 1 and type 2 DA receptors and child development near the age of onset of stuttering. Increases in disfluencies in patients with Parkinson's disease after taking Levodopa, a precursor to DA have also been reported (Anderson, Hughes, Rothi, Crucian, & Heilman, 1999; Lois, Winfield, Fahn, & Ford, 2001). In another study of stuttering following Levodopa treatment, Goberman & Blomgren (2003) did not observe an increase in stuttering. This study provided some evidence against the DA hypothesis of stuttering. Even though conflicting evidence has been reported, the majority studies support a dopamine hypothesis of stuttering.

Genetic Effects

Research has also demonstrated a genetic component in stuttering. Historically, there have been observations that stuttering "runs in families" (West et al., 1939). Such findings also weaken support for environmental theories of stuttering onset or persistence, especially when children who stutter have not spent a significant amount of

time with relatives who stutter. Thus, imitation or common environmental demands cannot account for the multiple cases of stuttering. In addition, stuttering at onset does not resemble the symptoms of adult stuttering (Manning, 2001), further weakening the potential evidence that stuttering develops as a response to environmental models or demands.

Support for the hypothesis that stuttering is heritable comes from twin studies (Andrews, et al., 1991; Felsenfeld et al., 2000; Howie, 1981), segregation analyses (Ambrose et al., 1993, 1997; Cox et al., 1984; Viswanath et al., 2004), and linkage analyses of stuttering (Riaz et al., 2005; Shugart et al., 2004; Suresh et al., 2006). Although the results of these studies are not always consistent with each other, they do provide convincing evidence of a genetic effect.

Few of the previous theories include molecular mechanisms that contribute to stuttering. Even the dopamine hypothesis of stuttering requires a way to account for the differences in dopamine activity in PWS compared to PWNS. Genetic approaches to understanding the etiology of persistent developmental stuttering have the advantage of addressing, at a fundamental level, potential molecular mechanisms that lead to persistent developmental stuttering.

Genetic Methods for Studying Disease

Linkage Studies

A genetic linkage study attempts to identify a specific region on a chromosome that is responsible for a disease of interest; that is, what region of the genome is linked to or co-inherited with a disease. It is the first step in identifying a causative gene when no

prior knowledge exists of what the gene or causative mechanism may be. Linkage studies make use of genetic markers that assay natural genetic variation that occurs throughout the genome, and are conducted in families in which a particular trait segregates. These variations are known as polymorphisms and can be highly variable within a population.

Linkage analyses can be performed to identify a significant association of a particular allele of one of these polymorphic markers to the disease. These analyses are based on Mendel's second law, the law of independent assortment, stating basically that two genetic loci will be transmitted from parent to offspring independently of each other. However, this is not always the case. When two loci are close enough to each other they can be transmitted together, lacking independence from each other, they are said to be "linked". This can also occur between a genetic marker and a disease gene, thus violating Mendel's law of independent assortment. These polymorphic markers are passed along to offspring and segregate in families, and have been well characterized with their precise locations in the genome known. The variation at these polymorphic sites is assayed when conducting a linkage study. Ideally, one of the polymorphic sites will lie close enough to the gene responsible for the disease that the two will always co-segregate, and thus will be transmitted together from parent to offspring. It is this violation of Mendel's second law that provides the theoretical framework for linkage analysis. A linkage analysis produces LOD scores, which is a measure of significance of linkage. A LOD score is a $\log_{10}$ of an odds ratio. This is a ratio of the likelihood that two loci (e.g. marker and disease, or two markers) are linked to the likelihood that they are not. A LOD score of 3 is considered significant evidence for linkage and means that the

data are 1000 times more likely to support linkange than not. A LOD score of -2 is considered significant evidence against linkage.

Mapping Methods

There are two types of polymorphic markers commonly used in linkage studies. One type of polymorphism is known as short tandem repeat polymorphisms (STRPs) or microsatellite markers. STRPs consist of varying numbers of short repeating units. For example a single site may contain the tri-nucleotide repeat ATA, notated as $(ATA)_n$ and typically have 10 – 20 repeats at a single locus (Weber & May, 1989)

. These repeats are assayed by the use of the polymerase chain reaction (PCR) to amplify a DNA segment containing the repeat. The number of repeats determines the size of the PCR product, which can be determined using electrophoretic methods. One benefit of this type of marker is that they typically display many alleles in the population and thus they are highly informative in family studies. Although there are on the order of hundreds of thousands of STRPs in the human genome (Toth, Gaspari, & Jurka, 2000; Zhao, Heil, & Weber, 1999), a particularly useful subset of these are typically used for genotyping. This subset is based on a combination of characteristics, including, but not limited to, high heterozygosity, PCR amplification properties, and precisely known locations in the genome (Ghebranious et al., 2003). A standard screening set of STRP markers used in a genome wide scan for linkage is on the order of 400 markers, with average spacing of ~10 cM (equivalent to roughly 10 million bases) (Ghebranious et al., 2003). This is often sufficient for the initial identification of a genetic locus responsible for a disease, especially for Mendelian traits. However, additional markers will often

need to be genotyped to obtain a more refined genetic location. Furthermore, there is the possibility of not observing any significant evidence for linkage, even if a locus might actually be linked to the trait. This possibility becomes more likely in complex genetic traits where a single gene may have only a small effect on the phenotype. Larger STRP screening sets containing closer to 800 markers with an average spacing of ~ 5 cM have been developed and may be useful for complex traits (Ghebranious et al., 2003). It has been shown that more densely spaced microsatellite screening sets have a higher information content than sets with markers every ~10 cM (Evans & Cardon, 2004). Evans and Cardon (2004) also suggest that even microsatellite screening sets with markers every ~ 3 cM do not extract the full information content possible out of a pedigree. In the case of complex traits where the genetic effect is less clear than in a strict Mendelian trait, it is beneficial to extract as much information out of a pedigree as possible.

Another option is the use of single nucleotide polymorphism (SNP) markers. SNPs differ from microsattelite markers in that they consist of variation at the level of single nucleotides, whereas STRPs differ in the number of repeats present. SNPs affect only the DNA sequence. One advantage of SNPs is that there is estimated to be on the order of 3 million SNPs in the human genome, far more than the number of STRPs. SNP genotyping sets can contain as many as 500,000 dimorphic markers that have only two variants occurring in a population. In a recent study of alcoholism Zhang et al. (2005) obtained greater evidence for linkage using an Affymatrix SNP genotyping set consisting of 11,050 SNPs compared to 328 STRPs. Using SNPs for conducting genome wide scans has been growing in popularity as the technology and cost have made it more

feasible to assay large numbers of markers in large numbers of individuals. Furthermore, as Evans and Cardon (2004) have shown, the inheritance information possible with a 15,000 SNP screening set is about twice that of a 300-400 microsatellite set.

The large number of SNPs in the human genome allows for a much more densely spaced genome wide scan. The disadvantage is that there is much less information content in SNPs than in microsatellites, as SNPs are typically only dimorphic. This loss of information is made up for by the density of markers. SNPs can still be used quite successfully to map a disease gene to a locus, especially when haplotypes containing several SNPs are analyzed (Tishkoff & Verrelli, 2003). The study conducted by Zhang et al. (2005) compared LOD scores, (measures of evidence for linkage) obtained from a series of genome wide scans using several different SNP screening sets and STRPs. They found loci for which the LOD score obtained with microsatellite markers was significantly lower than that observed with SNPs. In some cases the microsatellite data provided no support for linkage, when SNP markers at the same locus provided suggestive evidence for linkage. This is likely due to the increase in information that can be obtained from a densely packed SNP screening set, as mentioned in the previous section.

Linkage Analysis Methods

Linkage studies are carried out in families and there are several ways that these studies can be designed. One method is to use a number of unrelated families segregating the same disease. In such studies multiple offspring, both affected and unaffected, are studied to measure co-inheritance of marker alleles, or a lack thereof, thus determining an

overall LOD score. An overall LOD score is a combination of LOD scores from all families used in the analysis. LOD scores for the individual families are also calculated and it is possible to determine the extent to which each family contributes to the overall LOD score.

However, a problem can arise when using multiple families in a linkage study due to the presence of genetic heterogeneity (i.e. when there is more than one gene causing the observed phenotype in different families). There are many examples of this, including hereditary hearing loss, where hundreds of mutations have been identified in scores of different genes. Such heterogeneity reduces the ability to determine a significant effect at a single locus when using multiple families in a linkage study. The effects of heterogeneity can be addressed by performing an HLOD (Heterogeneity LOD) analysis, a method of performing linkage analysis under the assumption that genetic heterogeneity exists. Another method for addressing the problem of genetic heterogeneity is by the use of a single large family, as has been employed in the current study in West Africa. It is likely that the observed phenotype in such families is due to the same genetic effect. A third method of overcoming genetic heterogeneity is to use specialized populations, such as consanguineous families. These last two approaches aim to reduce the amount of genetic heterogeneity in the study population, thereby reducing the effects of genetic heterogeneity on a linkage study (Wright, Carothers, & Pirastu, 1999).

There are two main methods for performing genetic linkage analyses, parametric and non-parametric. In a parametric analysis, it is necessary to have a good hypothesis about the mode of inheritance, disease frequency and other factors that might have an effect on

the phenotype, such as age of onset. These factors can usually be determined based on visual inspection of the pedigree and are often straightforward for traits that follow clear Mendelian inheritance. It may be possible to assume a particular mode of inheritance, but it's not always clear-cut. When performing parametric analyses it is important to specify the correct parameters (i.e. mode of inheritance, and disease allele frequency). Mis-specifying parameters can have drastic effects on the significance of the results obtained as demonstrated in Terwilliger and Ott (1994). In complex traits, those that do not follow simple Mendelian inheritance patterns, such as stuttering, it is often difficult to state the mode of inheritance with any certainty.

Several methods exist to combat the challenge of correctly specifying the model in parametric linkage analyses of complex traits. One such method is to conduct segregation analyses prior to performing linkage analyses. Segregation analyses provide an estimate of the most likely mode of inheritance of the phenotype in a particular family or population. The results of such analyses can be incorporated into a parametric linkage analysis. Another method is the use of non-parametric linkage analysis in which the genetic model does not need to be specified. Complex traits in which the mode of inheritance does not follow strict Mendelian patterns can benefit from non-parametric analyses. However, there is a loss of power as these methods generally rely only on affected individuals by comparing marker alleles shared only in affected individuals (Strauch, Fimmers, Baur, & Wienker, 2003).

Chapter 2: Literature Review

Familial Aggregation of Stuttering

There has been a longstanding question regarding the relative influence of nature versus nurture in the origin of stuttering. It was commonplace to think that if stuttering clustered in families, it was perhaps due to imitation rather than a biological component. As early as the 1930's, researchers began to document familial clustering of stuttering. West, Nelson, & Berry (1939) present some of the first lines of evidence demonstrating familial aggregation of persistent developmental stuttering (PDS). Several additional reports of familial aggregation of stuttering also surfaced around this same time (Gray, 1940; Wepman, 1939). Up until 1939 there had been very little, if any, published research addressing what may be the cause of this clustering. Was it transmitted biologically or was it due to social transmission?

West et al. (1939) reported the results of a large study of 204 individuals with a family history of stuttering that were initially identified by Nelson (1939). Family members of these individuals were also ascertained. The sample population for the control group consisted of a total of 6600 individuals. West et al. (1939) also ascertained 204 control individuals who do not stutter, and their family members for a sample size of 6266 individuals. The numbers of individuals who stutter in each of these groups were compared. Of the relatives of the 204 affected probands, 210 stuttered while among the

relatives of the control individuals, only 37 stuttered. This represents a 6-fold increase in the number of affected individuals in the families of the probands. In an attempt to control for the effects of shared environment, they compared the number affected individuals who were close to and spent much time with other affected family members (n=75, or 37% of the stuttering probands) to the number of control individuals who had a close association with other PWS, such as childhood friends (n=60, or 29% of the non-stuttering controls). By taking the difference in percentages of these two groups, they suggest that 8% of stuttering can be "explained by social transmission" (West et al., 1939, p. 25), that is, by non-genetic factors. In addition, they argued that because 63% of affected probands had no or little contact with other PWS, there is likely to be "some more potent factor of transmission than mere association to explain the greater number of stutterers in the family lines of the stutterers than in the family lines of non-stutterers." (West et al., 1939, p. 25). West et al. suggested a tendency towards familial clustering based on sex, that is, female probands tended to have more female relatives who stuttered and male probands tended to have more male relatives who stuttered. Based on these observations the authors acknowledged that it is quite possible that heredity could play a factor in stuttering. They went on to pose a question that was remarkably prescient in view of the current research efforts on genetic factors in stuttering;

> "Would it not be reasonable, in view of what we have seen thus far, to say that apparently the thing that is transmitted from generation to generation is not stuttering, but the tendency to stutter? … Thus not all who have inherited the tendency to stutter actually stutter. Not all who can transmit this tendency to their offspring stutter themselves."p. 27-8

In 1939 Wepman identified 250 “stammerers” and compared the incidence of stuttering in their families to that in an equal number of families of “non-stammerers”. They found an increased incidence of stuttering in the relatives of the stammering subjects of about 4:1, providing further evidence for familial aggregation in this disorder. However he did not discuss whether or not the trait is heritable, or if some other factor is involved.

In 1940, Gray reported on a large family of five generations, known as Family X, consisting of two branches, one residing in Iowa, the other in Kansas. In the Iowa branch, 12/37 (32%) of the individuals were affected, and although less information was gathered on the Kansas branch, the incidence of stuttering was apparently somewhat lower in this branch, with 3/26 (11.5%) of the individuals affected. Despite the difference in the incidence of stuttering in the two branches of this family, their ancestors were traced back to a single patriarch who stuttered, and whose wife was an unaffected daughter of an affected mother. Gray attributed the difference in the incidence of stuttering in the two branches as largely due to differences in behavior and attitude towards stuttering among the two branches, in referring to the Iowa branch, she stated that,

> “This family might with some reason be regarded as “stuttering conscious,” quick to diagnose as stuttering the hesitant, repetitious first speech attempts of children learning to talk, and prone to develop in the child the familial anxieties and generally negative evaluations regarding hesitant or repetitious speech.” p. 346

Gray had a difficult time accounting for the observed differences between the Iowa and Kansas branches of the family by "an hereditarian interpretation" (p. 347), and preferred to invoke, what she calls the semantogenic interpretation, quoted above.

Although it was not entirely clear to any of these investigators whether familial clustering of stuttering in families was biological or due to social factors, it became clear that stuttering has a tendency to run in families. These observations inspired further investigations into the heritability of stuttering as a genetic trait.

One of the first studies that focused specifically on genetic factors in stuttering was conducted by Kidd, Heimbuch, & Records (1981). Using more advanced statistical measures, they demonstrated an increase in the risk for developing stuttering in relatives of individuals who stutter, especially if their parents also stutter, leading them to the conclusion that stuttering is transmitted vertically in families. Although this study did not specify a mode of transmission for the disorder, the authors point out that the observation of vertical transmission is a necessary pre-requisite for further studies to determine modes of inheritance.

In another study with a strong genetic focus, MacFarlane, Hanson, Walton, & Mellon (1991) reported a large family of European origin residing in Utah and Idaho containing over 1200 individuals in 5 generations. Two branches of this family totaling approximately 400 individuals contain no reported stuttering. However the incidence of stuttering in some of the other branches of this family was as much as 15 times higher than that of the general population. Five branches, containing near 500 individuals, have a prevalence of stuttering that is 5 – 10 times that of the general population, while another branch of this family displays persistent developmental stuttering at a rate that is 10-15

times that of the general population. The segregation analysis of this single branch suggested a monogenic, autosomal dominant inheritance pattern with reduced penetrance.

Twin Studies

One of the most powerful methods for separating inherited factors from environmental factors in a familial trait is a twin study (Sneider & MacGregor, 2003). In such studies, concordance for a trait is measured in identical (monozygotic, MZ) twins and compared to that in fraternal (dizygotic, DZ) twins. Typically both MZ and DZ twins are reared together, providing a common environment for both types of twin pairs. The fact that monozygotic twins share 100% of their genetic makeup and dizygotic twins share 50% of their genetic makeup, no more than any other full siblings, can be exploited to determine the heritability of a trait. If a trait were genetic in nature then one would expect to observe an increased concordance in MZ compared to DZ twins. Heritability is defined as the fraction of phenotypic variation that is due to genetic effects. With some simplifying assumptions, heritability can be simply estimated using a ratio of the correlation of phenotype to the expected correlation of the genotype. However, in the case of twins, heritability estimates tend to rely on more sophisticated methods that can involve analysis of variance (ANOVA) or computing intraclass correlations (Vitzthum, 2003).

The two common methods of determining concordance rates are pairwise and case-(or proband) wise. "Pairwise concordance measures the proportion of pairs in which at least one of the co-twins expresses the disease. Casewise concordance measures

the proportion of co-twins of affected twin indivduals ('cases') that express the disease or trait themselves" (Snieder & MacGregor, 2003).

Howie (1981) reported on 29 twin pairs in which at least one was either reported to stutter or had recovered from stuttering. The proband-wise concordance rates were 0.77 in MZ twins and 0.32 in DZ twins. In another study, Andrews et al. (1991) reported on a much larger subject population that was gathered through an Australian twin study, in which 3810 twin pairs responded to a questionnaire, including one item regarding stuttering. This study included a total of 135 twin pairs (50 MZ, 85 DZ) in which at least one of the twins was affected. The inclusion criteria in Andrews et al. (1991) were either a positive self-report or the observance of stuttering during a telephone interview. The concordance rates observed in this study were 0.20 for MZ twins and 0.035 of DZ twins. Although these values are considerably smaller than reported by Howie (1981), the trends are the same, in that MZ twins are far more concordant for stuttering than are DZ twins. A third twin study was conducted in a similar manner to Andrews et al. (1991), in that the participants were ascertained through a questionnaire given to a different cohort of Australian twins (Felsenfeld et al., 2000). A total of 457 telephone-based interviews were carried out from a total of 1567 twin pairs and 634 singletons from the Australian Twin Registry who reported themselves as a PWS. A total of 91 complete twin pairs were interviewed where at least one twin was considered affected. The results of this study followed the same trends; MZ twins had a higher concordance rate than DZ twins, with concordance rates in the interviewed sample of 0.45 in MZ and 0.15 in DZ twins.

The results of these three studies are summarized in Table 2.1.

	MZ Concordance	DZ Concordance	Computing Method
Howie, 1981	0.77	0.32	Proband - Wise
	0.63	0.19	Pairwise
Andrews et al., 1991	0.33	0.07	Proband - Wise*
	0.20	0.035	Pairwise
Felsenfeld et al., 2000	0.45	0.15	Proband - Wise
	0.62	0.26	Pairwise

Table 2.1 Twin concordance rates in monozygotic (MZ) and dizygotic (DZ) twins provide an estimate of the genetic effect of diseases. * As calculated by Felsenfeld (2000)

The variation in reported concordance rates between the three studies could be due to differences in sample size, ascertainment biases or diagnoses/inclusion criteria. For example, in Howie's study, subjects were recruited from speech clinics and newspaper ads. Inclusion criteria included a diagnosis of stuttering by a trained clinician that was based on two 500-word recorded speech samples from each participant. It has been observed in other studies (Felsenfeld et al., 2000) that self-report leads to an overrepresentation of the number of affected individuals. Andrews et al. (1991) recruited subjects through a questionnaire mailed out to over 5900 twin pairs registered with the Australian Twin Registry. Inclusion criteria in this report were "either heard to stutter when [a clinician] telephoned them or, if recovered, reported having sought professional advice or treatment or reported that the stutter used to interfere significantly with their lives and activities" (Andrews, 1991 p. 1034). They claim that this method of ascertainment is preferable to the methods used by Howie (1981). Andrews et al. (1991) suggest that an ascertainment bias is introduced when twins are recruited from clinics or advertisements, as was done by Howie (1981). Andrews et al. (1991) state that Howie's method leads to an "increased pairwise concordance in MZ … vs DZ twins" (Andrews et al., 1991 p. 1034). They suggest that since the subjects in Howie's (1981) study were

from a selected population this bias might distort concordance values. Andrews et al. (1991) provide no support for the claim that this method of ascertainment introduces a bias in concordance estimates. However, their claim is used to justify the difference in concordance rates observed in these two studies. The subjects in Felsenfeld et al.'s study were also gathered through the Australian Twin Registry, however it was a more recent cohort than that used by Andrews et al. (1991). Their inclusion criteria were perhaps the most thorough with the least amount of bias so far. Twin pairs were only included in the concordance estimates if both individuals in a pair were interviewed by a clinician, thereby reducing the number of diagnostic errors.

Although these three studies vary with respect to the absolute concordance values, MZ twins consistently show a higher concordance for stuttering than do DZ twins, strongly supporting genetic factors in this disorder. Given the inconsistency in concordance rates, however, it is difficult to determine the precise magnitude of the genetic contribution to PDS. Heritability estimates, which must be viewed with caution due to the diversity of findings, range from 0.36 to 0.71 (Felsenfeld et al., 2000; Andrews et al., 1991).

Segregation Analyses

Segregation analyses, which test various models of transmission including combinations of genetic and non-genetic effects, have been used to investigate genetic contribution and mode of inheritance of PDS. Although these studies have produced inconsistent results in terms of specific modes of inheritance, they all support a genetic component to the disorder.

Kidd (1977) conducted one of the first segregation analyses. He tested two models that were commonly used in psychiatric genetics at the time, primarily a multifactorial-polygenic model and a single-major-locus model. Neither of these models could be ruled out but a major genetic effect was supported in both models. The main conclusion of this study was that "genetic models that incorporate environmental factors and sex differences in susceptibility can quite adequately explain the familial concentration in stuttering". Segregation analyses of 386 probands and their first degree relatives, conducted by Cox, Kramer, & Kidd (1984) suggested potential genetic heterogeneity. This indicates that there are likely to be different genes involved in PDS in different families and/or populations. They observed no support for a single major gene effect or for a polygenic effect in these subjects.

The only segregation analysis performed to date on a single large family was conducted on the single branch of the Utah / Idaho family with the highest prevalence of PDS (Mellon, Hanson, Hasstedt, Leppert, & White, 1991). Their results suggest a monogenic, autosomal dominant trait with reduced penetrance. Another analysis by Ambrose, Cox, & Yairi, (1997) suggested a single major gene effect with underlying polygenic contributions. They suggest that persistent and non-persistent stuttering may share the same genetic component (i.e. the same gene), but additional genetic effects may control persistence or recovery of stuttering. More recently Viswanath, Lee, & Chakraborty (2004) performed segregation analyses in a group of 56 families. They found a single major gene effect that is influenced by the sex of the individual and the parents' affection status best explains the mode of transmission of PDS. This means that there are additional factors (gender and parental affection status) that have an influence

on a single gene in PDS. Taken together these results are consistent with the results of the twin studies, and confirm that there is a genetic component to PDS. However, they also suggest that the genetic effect on stuttering may be different in different families. A number of the genetic characteristics of PDS are still somewhat uncertain, and it is clear that PDS is a complex genetic trait, and does not seem to follow strict Mendelian transmission patterns. One way to address this complexity and to aid in the search for a causative or pre-disposing gene (or genes) is to use large individual families. Multifamily studies may be complicated by the possibility that different families have different genetic contributions to stuttering. If this is the case it becomes more difficult to find significant linkage. Using a single large family may reduce this genetic heterogeneity, making it easier to identify a causative gene.

Linkage Analyses

To date there have been three published linkage analyses of stuttering (Shugart et al., 2004; Riaz et al., 2005; Suresh et al., 2006). Shugart et al. (2004) performed non-parametric linkage analysis on 68 families of North American and European decent. Results from this study provided suggestive evidence for a linked region on chromosome 18, with a total non-parametric linkage (NPL) score of 5.35 and a single family contributing an NPL score of 4.72. It is therefore possible that there is a predisposing gene on this chromosome in one of the families and perhaps loci on other chromosomes contribute to PDS in other families.

One method of reducing the genetic heterogeneity in such studies is to use specialized populations. In 2005 Riaz et al. published a linkage analysis of PDS in an

inbred Pakistani population, where cousin – cousin marriages are common (Hussain & Bittles, 1998). Matings between individuals that share a common ancestor reduce the genetic diversity, which can increase genetic homogeneity. These populations have minimal diversifying effects from population admixture. The increase in genetic homogeneity can help to identify individual genetic loci. Significant evidence for linkage was obtained on chromosome 12 in this population, which was especially strong in a subset of families. Chromosomes 1, 5 and 7 also provided suggestive evidence of linkage with LOD scores over 2. However, the evidence for linkage on chromosomes 5 and 7 was obtained with markers at the end of the chromosome. Because there were no markers between these and the end of the chromosome, multipoint analysis was only performed on one side of these markers. Multipoint linkage analysis uses information from multiple markers by placing the disease locus at different intervals between two markers. This is not done with the last marker on a chromosome and therefore the results obtained should be interpreted with caution (Riaz et al., 2005). This was the first study to produce genome-wide significant evidence of linkage for stuttering. So, while the results of this study provided additional support for the view that stuttering can be strongly genetic, they also supported the view that PDS is a genetically heterogeneous trait. It is therefore likely that different genes, even within a single population of similar geographic origins, can be involved in the predisposition of this disorder.

The unequal sex ratio, or sex-bias, of stuttering has long been recognized. Gray (1940) reported a ratio of affected males (n=8) to females (n=5) in Family X, of 1.6:1. She acknowledges that this ratio is in disagreement with previously published reports (Nelson, 1939) of male to female ratios ranging from 3:1 to 5:1. Mellon et al. (1991)

report a male to female ratio of PDS in a single large family that is also 1.6:1. Drayna et al. (1999) report a similar ratio (1.57:1) of affected males to females in familial cases of PDS, but contrasts this to a ratio near 4:1 in sporadic cases, which have no family history of PDS. Male to female ratios of sporadic cases ranging from 2.5:1 to 5:1 have been observed in additional studies (Bloodstein, 1995; Janssen, Kloth, Kraaimaat, & Brutten, 1996; Kidd, Heimbuch, & Records, 1981) as well. All of these results support a difference in the ratio of affected males to females between cases with a family history of PDS and sporadic cases with no reported family history.

To address the issue of sex-bias in PDS Suresh et al. (2006) conducted linkage analyses that took these differences into account. They ascertained 100 families of European decent from the US, Sweden and Israel. There were a total of 252 individuals who exhibited PDS, 45 that were recovered and 19 that were too young to classify. The 19 that were too young to classify may not have developed speech yet, or they may be too young to classify as persistent. Children generally develop stuttering between the ages of 2 and 5, and many will recover spontaneously, usually by puberty, and many within the first 12 months of onset (Guitar, 1998). Depending on how the phenotype is defined (i.e. ever stuttered or persistent stuttering) for subject enrollment and / or linkage analysis, it is important to correctly diagnose subjects. This can only be done if children are old enough.

Linkage analyses were initially carried out with a broad phenotype of "ever stuttered" and a narrower phenotype of "persistent stuttering". Suggestive evidence of linkage was found on chromosome 9 in the broad phenotype (LOD = 2.3), and chromosome 15 for the narrow phenotype, (LOD = 1.95). The disproportionate male to

female ratio (sex-bias) of 2:1 to ~ 4:1 observed in stuttering led these investigators to perform a sex- specific analysis. This was done by setting the affection status to unknown in females for the male only data set and males set to unknown in the female only data set. They found an increased LOD score (increasing from 0.04 to 4.5) on chromosome 21 in the female data set, and on chromosomes 7 and 20 (LOD = 2.99 and 2.18 respectively) in the male data set. Based on previous suggestions that persistent developmental stuttering might be a polygenic trait, they performed conditional linkage analyses, in order to "assess potential interactions between linkage signals of interest and the rest of the genome." (Suresh et al., 2006 p. 557). Based on these conditional analyses, interactions between two loci could have positive or negative interactions. An interaction is positive if the conditional analyses produced a higher LOD score than previously obtained at that marker. Conversely an interaction is negative if the LOD score decreases. A positive interaction was observed between chromosome 13 and chromosome 15 linkage signals. An increase in the LOD score (LOD = 1.78 to LOD = 2.8) on chromosome 13 was observed when conditioned on the evidence for linkage on chromosome 15. A strong negative interaction with chromosome 20 was present, when conditioned on chromosome 15. When analysis was conditional on chromosome 9, a significant positive interaction was observed with a locus on chromosome 2. The same locus on chromosome 2, however, showed a negative association when conditional analysis was performed based on the evidence for linkage at chromosome 7.

The results of these three published studies support the view that stuttering is a complex genetic trait. However, progress has been made over the past 50 years towards a

better understanding of the genetic causes of stuttering. Despite this progress, genes responsible for persistence or recovery in PDS have not yet been identified.

Specialized populations can be used to help identify specific genes in the face of such complexities. Riaz et al. (2005) employed this strategy by using highly consanguineous families from Pakistan. Another specialized study population consists of large families in which the trait behaves as a Mendelian disorder, presumably under control of a single gene of large effect. The current study employs a single large family with multiple generations displaying a high prevalence of PDS. The availability of such a large family provides additional opportunities for understanding the genetic contributions to stuttering, and is especially powerful for genetic linkage studies, because it allows application of the well-developed methods for finding Mendelian disease genes, thereby circumventing many of the problems inherent in linkage studies of complex traits.

Chapter 3: Subjects

Enrollment

A large 5-generation polygamous family from the Republic of Cameroon was ascertained by Dennis Drayna. This family, designated CAMST01, is of Bantu lineage from Northwest Province, an Anglophonic region of the country (Levis et al., In preparation). The family was initially ascertained through contact made during the annual International Online Conference on Stuttering sponsored by the Stuttering Homepage (http://www.mnsu.edu/comdis/kuster/stutter.html) in 2002. Subjects were enrolled under NIH IRB approved protocol 97-DC-0057, additionally reviewed and approved by the IRB of the Institute of Tropical Medicine, Kumba, Southwest Province, Cameroon. Eligibility criteria were as follows; a) subject is a blood relative in a family with 2 or more people who stutter; b) subject is over the age of 7; c) subject has stuttered beyond the age of 7; d) other family members have stuttered beyond the age of 7; e) subject has stuttered for a period of 6 months or more and f) subject is not cognitively impaired. Age is an important inclusion criterion because of the many instances of spontaneous recovery. Although estimates of the percentage of children who spontaneously recover vary, many children will have done so by the age of 7 (Bloodstein & Bernstein Ratner, 2007, for review). A subject who has stuttered beyond the age of 7 and for more than 6 months will, therefore, likely remain persistent in their stuttering.

Stuttering Diagnosis

Diagnosis was performed using audio recordings of speech samples obtained using standard diagnostic texts (Stuttering Severity Instrument - 3, Plate VI, 160 syllables or Plate III, Shipping Scene), (Riley, 1980). In order to avoid diagnostic errors, all subjects in this study are from an Anglophonic region of Cameroon and all speak English. A single clinician at Hollins Communications Research Institute in Roanoke, VA, USA, performed speech diagnoses. Stuttering disfluencies were counted and scored as both the percentage of disfluent syllables and words. Syllables were scored as disfluent if the speaker displayed audible struggle behavior concurrent with speech initiation, silent stops, repetitions of sounds, prolongations, omissions or substitutions (Webster, 1978). Individuals were classified as affected if they demonstrated a reading or free speech percent disfluency of 4% or greater, which corresponds to the upper limit of the mild stuttering range. The proband (individual 135, Fig. 3.1) displayed moderate stuttering and had a percent disfluency of 30% on words and 23% on syllables in a speaking task. In a reading task the percent disfluency was 9.4% on words and 7.4% on syllables.

The pedigree of CAMST01 is shown in Figure 3.1. A total of 38 individuals in this family were diagnosed with stuttering (shaded symbols, Fig. 3.1). The founding parents (individuals 101 and 102) produced a total of 4 offspring in generation II, one of which is reported to have stuttered (individual 106). However, all 4 offspring in this generation were deceased at the time of ascertainment, and it is impossible to know with certainty the affection status of these four individuals. Stuttering exists in the descendents of all four siblings of the second generation. The most striking branch of

this family descends from a single individual in the second generation, individual 112, who has 19 children, 16 of whom stutter.

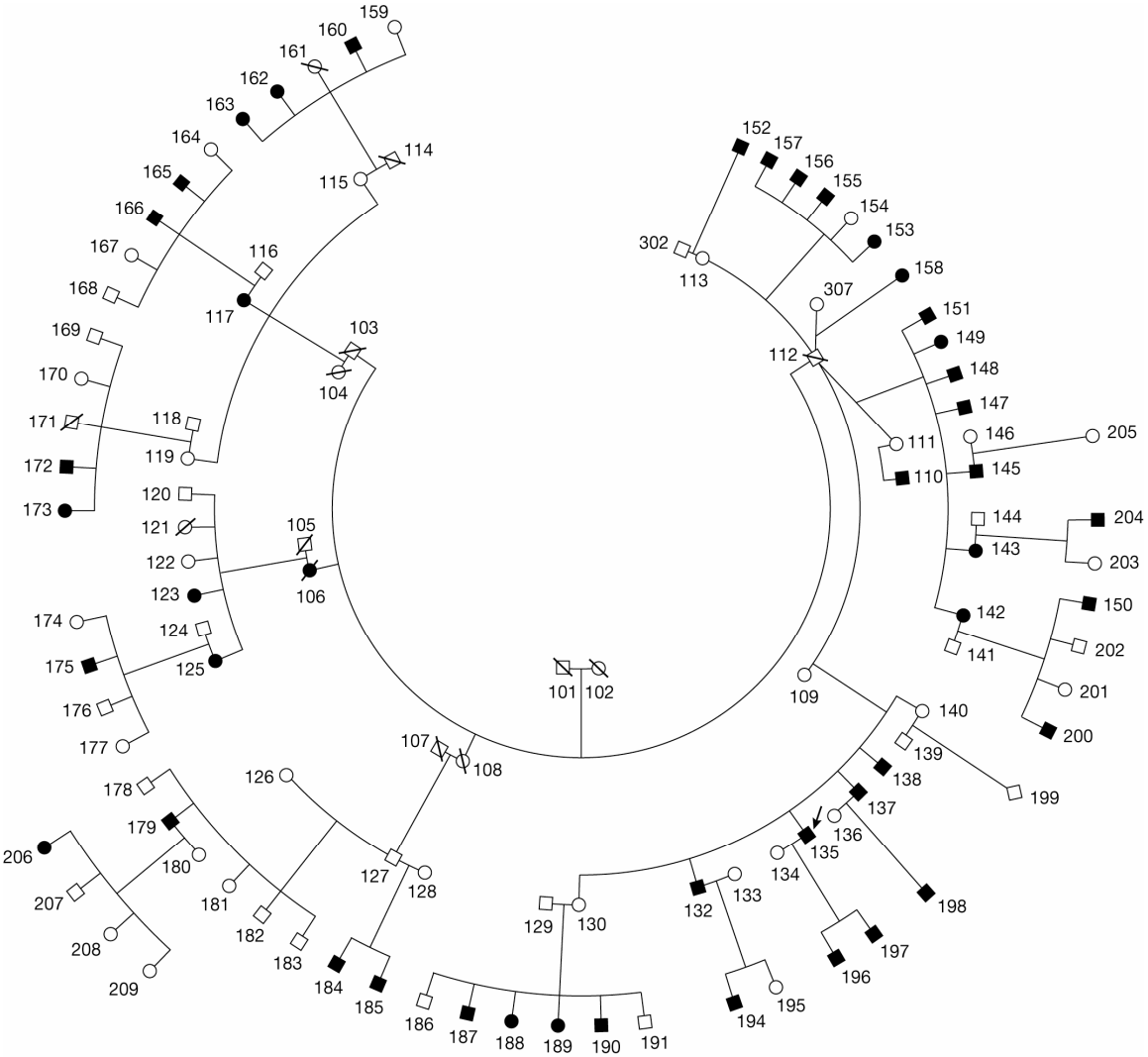


Figure 3.1 Pedigree of CAMST01. Circles represent females; squares represent males. Filled symbols are affected, unfilled are either unknown or unaffected individuals. Diagonal hash marks indicate the individual is deceased.

An important issue in such a family is to determine whether other symptoms occur, indicative of a syndromic disorder. Although it was not feasible to obtain detailed clinical examinations of family members beyond the proband, no gross intellectual,

motor, or other neurological disorders were observable in the 50 family members enrolled in this study. An audiological, neurological and general medical exam was performed on the proband at the NIH Clinical Center in Bethesda, MD.

A complete audiological assessment was conducted with clinical audiometers in sound attenuated test suites meeting American National Standards Institute (ANSI) approved conditions for audiological testing (ANSI, 1989, 1991). Pure-tone thresholds for air and bone conduction were normal for the entire test frequency range of 250-8000 Hz and ranged from 0 to 20 dB HL with no air-bone gaps in excess of 10 dB HL. Word recognition ability was normal in each ear. Tympanometry indicated normal middle ear pressure and peak compensated static compliance. The middle ear muscle reflex to acoustic stimuli occurred at normal levels and without time-based decay.

The history and results of the general medical and neurological examinations of the proband (individual 135, Figure 3.1) were as follows:

> This 38-year-old right-handed male, a native of Cameroon, was admitted for evaluation of persistent developmental stuttering. Developmental stuttering is familial and persists. He reported no known allergies and was on no medications at the time of examination. Review of systems was noncontributory. Neurological examination: speech was characterized by frequent repetitions and blocks characteristic of developmental stuttering. Receptive and expressive language abilities were intact, as was short-term memory. Cranial nerves 2-12 were intact and there were no motor or sensory deficits noted. The patient was diffusely hyporeflexic, but reflexes were equal, 1+, with reinforcement throughout. There were no pathological reflexes and toes were downgoing bilaterally. Cerebellar exam and gait were within normal limits.

Overall, the observations are consistent with persistent developmental stuttering as the sole neurologic, auditory, or speech disorder in this family.

Ascertainment Bias

CAMST01 was not randomly ascertained for this study, introducing an ascertainment bias. This family was ascertained through an affected individual who presented a large family with many affected individuals. In fact it was precisely this reason that CAMST01 was of particular interest. It has been suggested that this bias should not have a large effect on the linkage analysis and that the benefits of a pedigree, such as CAMST01, outweigh the minimal effect of ascertainment bias (Slager & Vieland, 1997).

Chapter 4: Methods

Genome Wide Scan

A genome wide scan was performed using 63 individuals, including 39 family members. 18 random Cameroonian (RC) and 6 African American (AA) unaffected individuals were used to determine marker allele frequency. Genotyping was carried out using the Marshfield-Weber panel 9 screening set, containing 387 markers with an average spacing of ~10 cM. Only autosomal markers were genotyped for a total of 366 markers. As there is no evidence that stuttering is a sex-linked trait, the 21 markers on the X and Y chromosomes were excluded. PCR was performed in duplex reactions. In cases where one or both markers failed to amplify in duplex reactions, singleplex reactions were performed. PCR was carried out on MJ Research Tetrad thermal cyclers, and Perkin Elmer / Applied Biosystems 9700 thermal cyclers. The standard PCR reaction contained 50 ng template DNA, 1.5 mM $MgCl_2$, 0.28 μM each primer, 0.5 mM each dNTP, 0.03 units either Taq polymerase or Tsp Polymerase and 1X Tris buffer. Some markers failed under these conditions and required PCR optimization. Amplification of these markers was usually successful with 2 mM or 3 mM $MgCl_2$. Genotyping was carried out on Applied Biosystems, ABI Prism 377 DNA sequencers using Genescan v. 3.1 and Genotyper v. 2.0 genotype allele calling software. Additional markers were genotyped in regions of interest using Applied Biosystems Capillary Electrophoresis DNA sequencers, 3730 and 3130 XL. An additional 8 individuals were

later sampled and genotyped using microsatellite markers on 3730 and 3130 XL capillary electrophoresis DNA sequencers.

Pedigree Structure Verification

Marker genotypes were arranged for each chromosome by genetic map position. The software program PREST (Pedigree Relationship Statistical Test) (McPeek & Sun, 2000) was used to verify the family structure as recorded when CAMST01 was initially ascertained. This was run with the assistance of Geoffrey Hayes and Nancy Cox at the University of Chicago.

Verification of Mendelian Inheritance

PedCheck (O'Connell & Weeks, 1997) was used to check for genotyping errors and inconsistencies in Mendelian inheritance. This was run on Helix (helix.nih.gov), a shared memory symmetrical multiprocessor computer. Three levels of PedCheck were run. Level 1 checks for inconsistencies among parents and offspring, checking only at the nuclear family level. Level 2 runs the Lange – Goradia algorithm (Lange & Goradia, 1987) to perform genotype elimination. This determines which genotypes do not need to be included in likelihood calculations, thus speeding up the pedigree checking process. This level is highly effective in identifying Mendelian inconsistencies. Level 3 determines which of the inconsistent individuals can have their genotypes set to unknown, allowing for a consistent pedigree.

Linkage Analysis

Maximum Estimated LOD Score

FASTSLINK version 2.51 (Ott, 1989; Weeks & Ott, 1990) was used to estimate the maximum LOD score possible in CAMST01 using the initially ascertained individuals and again including the individuals ascertained in the second enrollment phase. These analyses were performed on the Helix computer at the National Institutes of Health.

Parametric Linkage Analysis

Two point and multipoint linkage analyses were performed using MLINK and LINKMAP, respectively, from the FASTLINK software package v. 4.1 (Cottingham, Idury, & Schaffer, 1993; Schaffer, Gupta, Shriram, & Cottingham, 1994). Although the prevalence of stuttering in Cameroon is believed to be similar to that in other populations (Lukong, 2002) 2-point analyses were run using a disease allele frequency of 0.10 and 0.01. It is clear that stuttering does not follow a perfectly Mendelian dominant inheritance pattern in this family (note individuals 115, 118, 127, Figure 3.1). This suggests that the causative allele in this family may not be completely penetrant. Therefore the disease penetrance was set to 0.40 for individuals heterozygous at the disease locus, 0.80 for homozygous individuals, and 0.0 for unaffected individuals. These values were chosen because ~ 40% of the individuals in CAMST01 are affected. This value is in agreement with families from a previous linkage study of PDS (personal communication, Drayna, 2006; Shugart et al., 2004). In an autosomal dominant trait it is expected that 50% of the individuals would be affected. This provides support for using

a reduced penetrance in our analyses. Allele frequencies were estimated from 18 RCs and 6 AAs. Linkage analyses were also performed using other models of inheritance as seen in table 4.1. A total of six different models were tested, two co-dominant models, two that are more fully penetrant and two that more closely resemble a recessive trait.

	Penetrance		
Model	**dd**	**Dd**	**DD**
Co-dominant with medium penetrance	0.80	0.40	0.00
Co-dominant with lower penetrance	0.50	0.20	0.00
Fully Penetrant 1	0.99	0.99	0.00
Fully Penetrant 2	0.90	0.90	0.00
Recessive 1	0.90	0.30	0.05
Recessive 2	0.90	0.10	0.05

Table 4.1: Parametric linkage analysis models. **dd** = homozygous for disease allele, **Dd** = heterozygous for disease allele and normal allele, **DD** = homozygous for normal allele.

Parametric multipoint linkage analysis was also performed with SIMWALK2 (Sobel & Lange, 1996; Sobel, Papp, & Lange, 2002; Sobel, Sengul, & Weeks, 2001) using the same parameters as FASTLINK. Non-parametric analyses with SIMWALK2 were unsuccessful due to the large size of the pedigree. SIMWALK2 would only accept a maximum of 20 affected individuals in this analysis. CAMST01 contains 45 affected individuals. It is not entirely clear what the effects of breaking up the pedigree would be and how NPL (non-parametric linkage) scores would be combined.

Refining Linkage Regions

In an attempt to refine regions of interest based on results from linkage analyses, 8 additional individuals (CAMST01: 110, 141, 158, 163, 193, 194, 200, 205) were sampled and ~20 additional microsatellite markers were genotyped in the region of interest on chromosome 1p22.1-31.1 at a resolution of 1-2 cM. In some cases multiple

markers at the same genetic map position were genotyped. This was done to insure that all of the possible genotypic information was being obtained, as some markers may not be informative, or they may not amplify well. Additional linkage analyses were performed with the new markers in all sampled members of CAMST01 (n=47).

Haplotype Analysis

Haplotypes were created by inspection for the region of interest on chromosome 1 spanning approximately 40 cM and were recreated for individual 112 using genotypes from the offspring. In a further attempt to refine the linkage region, these haplotypes were inspected for any obvious shared regions spanning at least 3 markers throughout the extended pedigree.

Chapter 5: Results

Pedigree Relationship Statistical Test

Results from PREST suggested that the pedigree as initially drawn was inconsistent with the genotype data. In particular there were two individuals (150 and 192) who required their relationships be respecified. Table 1 shows the PREST output for these individuals. Relationships as specified were rejected at $p < 0.01$ for all three statistics calculated. This included individual 150 who was initially identified as the offspring of 111 & 112, but based on allele sharing statistics, it is clear that 150 is neither the offspring of 111, nor a full sibling of individuals 143-151, the other offspring of 111 &112.

Ind. 1	Ind. 2	Relationship	# Markers	EIBD - Obs	P0, Est.	P1, Est.	P2, Est.	p-value, EIBD	p-value, AIBS	p-value, IBS
150	109	6	350	NA	0.9999	0.0001	0.0000	NA	NA	0.74072
150	111	10	346	NA	0.6615	0.3385	0.0000	NA	0.00000	0.00000
150	113	6	357	NA	1.0000	0.0000	0.0000	NA	NA	0.37590
150	118	5	351	0.2336	0.8164	0.1836	0.0000	0.23896	0.27118	0.06555
150	123	5	358	0.2341	0.8086	0.1785	0.0129	0.25199	0.57805	0.05505
150	125	5	358	0.2425	0.7826	0.2086	0.0088	0.58762	0.66905	0.10456
150	129	6	356	NA	0.9893	0.0000	0.0106	NA	NA	0.93584
150	130	2	354	0.4392	0.6786	0.2996	0.0218	0.00503	0.03346	0.01861
150	132	2	354	0.4353	0.6766	0.3166	0.0068	0.00287	0.01969	0.00224
150	133	6	349	NA	1.0000	0.0000	0.0000	NA	NA	0.12953
150	135	2	356	0.4082	0.7490	0.2510	0.0000	0.00002	0.00013	0.00003
150	137	2	357	0.4574	0.6171	0.3724	0.0106	0.04807	0.15632	0.09824
150	142	1	357	0.9580	0.0123	0.9695	0.0182	0.37444	0.60249	0.29226
150	143	1	356	0.6449	0.6072	0.3871	0.0057	0.00000	0.00000	0.00000
150	145	1	357	0.7192	0.4724	0.5115	0.0161	0.00000	0.00000	0.00000
150	147	1	357	0.7548	0.4148	0.5852	0.0000	0.00000	0.00000	0.00000
150	148	1	352	0.6673	0.5347	0.4653	0.0000	0.00000	0.00000	0.00000
150	149	1	357	0.7483	0.4178	0.5729	0.0093	0.00000	0.00000	0.00000
150	151	1	355	0.6436	0.6136	0.3858	0.0007	0.00000	0.00000	0.00000
150	152	2	338	0.3309	0.9942	0.0000	0.0057	0.00000	0.00000	0.00000
150	153	2	358	0.4624	0.6034	0.3964	0.0002	0.08142	0.20110	0.04850
150	154	2	356	0.4073	0.7557	0.2442	0.0000	0.00002	0.00008	0.00001
150	155	2	357	0.4676	0.5968	0.3890	0.0142	0.13451	0.16342	0.17179
150	156	2	355	0.4444	0.6552	0.3448	0.0000	0.01030	0.01791	0.00367
150	157	2	355	0.4062	0.7548	0.2452	0.0000	0.00001	0.00011	0.00001
150	187	7	355	0.2203	0.8761	0.0976	0.0264	0.04574	0.14285	0.04471
150	188	7	356	0.2304	0.8237	0.1763	0.0000	0.18688	0.16285	0.04478
150	192	7	354	0.2023	0.9267	0.0733	0.0000	0.00137	0.00338	0.00014
192	109	3	351	0.3085	1.0000	0.0000	0.0000	0.00000	0.00000	0.00000
192	111	6	346	NA	0.9987	0.0000	0.0013	NA	NA	0.41987
192	113	6	357	NA	1.0000	0.0000	0.0000	NA	NA	0.04085
192	129	6	355	NA	0.9869	0.0000	0.0131	NA	NA	0.61616
192	130	4	354	0.3740	0.8404	0.1596	0.0000	0.00000	0.00000	0.00000
192	132	10	354	NA	0.8558	0.1442	0.0000	NA	0.00000	0.00000
192	133	6	347	NA	1.0000	0.0000	0.0000	NA	NA	0.75511
192	135	4	356	0.3447	0.9233	0.0767	0.0000	0.00000	0.00000	0.00000
192	137	4	357	0.3828	0.8113	0.1887	0.0000	0.00000	0.00000	0.00000
192	142	7	357	0.2223	0.8534	0.1415	0.0051	0.06157	0.10349	0.01856
192	143	7	356	0.2152	0.8810	0.1190	0.0000	0.01907	0.01597	0.00262
192	145	7	357	0.1996	0.9160	0.0840	0.0000	0.00068	0.00380	0.00007
192	147	7	357	0.2332	0.8139	0.1861	0.0000	0.25753	0.39797	0.09591
192	148	7	352	0.2431	0.7848	0.1914	0.0238	0.64260	0.99526	0.95698
192	149	7	357	0.2213	0.8587	0.1413	0.0000	0.05278	0.08910	0.01830
192	150	7	354	0.2023	0.9267	0.0733	0.0000	0.00137	0.00338	0.00014
192	151	7	355	0.2343	0.8115	0.1885	0.0000	0.29096	0.34130	0.12015
192	152	7	339	0.1784	1.0000	0.0000	0.0000	0.00000	0.00000	0.00000
192	153	7	358	0.2177	0.8663	0.1337	0.0000	0.02937	0.07551	0.00117
192	154	7	355	0.2220	0.8558	0.1432	0.0011	0.05996	0.09585	0.00811
192	155	7	358	0.2182	0.8713	0.1287	0.0000	0.03189	0.03251	0.00717
192	156	7	355	0.2124	0.8755	0.1245	0.0000	0.01140	0.02039	0.00006
192	157	7	355	0.2137	0.8792	0.1208	0.0000	0.01471	0.01310	0.00019
192	187	5	355	0.2198	0.8686	0.1173	0.0141	0.02916	0.08793	0.02092
192	188	5	356	0.2082	0.9254	0.0715	0.0030	0.00253	0.00380	0.00373

Table 5.1: Output from prest_out2. Highlighted in gray are the most statistically significant cases in which the specified relationships are inconsistent with that expected from the genotype data. These were the pivotal results in determining that individuals 150 and 192 were mis-specified. Key to relationships, and expected P0, P1, P2: 1 = fullsib (0.25, 0.5, 0.25); 2 = half-sib (0.5, 0.5, 0); 3 = grandparent – child (0.5, 0.5, 0); 4 = avuncular (0.5, 0.5, 0); 5 = first cousin (0.75, 0.25, 0); 6 = unrelated (1, 0, 0); 7 = half-avuncular (0.75, 0.25, 0); 8 = half first cousin (0.875, 0.125, 0); 9 = half sib plus first cousin (0.375, 0.5, 0.125); 10 = parent – offspring (0, 1, 0)

The most likely relationship for 150 was as the offspring of individual 142 (table 5.2).

DNA from individual 141 had not been obtained at the time this analysis was performed.

Thus it was not possible to empirically determine the relationship of 150 to 141, and

therefore could only assume that 141 was the biological parent of 150. Additional evidence (discussed shortly) did however emerge suggesting that 141 was not the biological parent of 150. Individual 192 was also re-specified as evidence suggested this individual was not the offspring of individual 132. Alleles in 192 were consistent with 131 but not consistent with individual 132. Individual 192 was assigned a ghost father and removed from further analyses. The output from prest_out3, one of the PREST output files, is shown in Table 5.2 and contains relative pairs in which parent – offspring relationships may be mis-specified based on the estimated probability of sharing 1 allele (p1). Individuals are included in prest_out3 under two circumstances, 1) when the estimated p1 < 0.9 in individuals that are specified to be parent – offspring relative pairs, but may not actually be; or 2) p1 > 0.75 in individuals that are specified as other than parent – offspring relative pairs but who appear to be parent - offspring. The expected value for p1 in parent – offspring pairs is 1 since a child will always get one allele from each parent, therefore sharing one of the alleles from each parent 100% of the time.

				Alleles Shared - Expected			Alleles Shared - Estimated		
Ind. 1	Ind. 2	Relationship	# Markers	p0	p1	p2	p0	p1	p2
111	150	10	346	0	1	0	0.6615	0.3385	0.0000
132	192	10	354	0	1	0	0.8558	0.1442	0.0000
142	150	1	357	0.25	0.5	0.25	0.0123	0.9695	0.0182

Table 5.2: Output from prest_out3. Parent-offspring pairs that were either specified (relationship = 10) with a probability of sharing a single allele of < 0.9, or specified other than parent-offspring, but have a probability of sharing a single allele of > 0.75. Col: 1, 2 – individual IDs; 3 – relationship specified in prest according to pedigree (10 = parent–offspring, 1 = full sibling); 4-6 – expected probability of sharing 0,1 or 2 alleles based on specified relationship; 7-9 – estimated probability of sharing 0,1 or 2 alleles, as calculated by PREST.

Additional PREST analyses were performed with the 8 new individuals and 25 markers in the region of interest on chromosome 1. However due to the small number of markers,

none of the results were statistically significant, and therefore no additional information was available for determining actual relationships of the new individuals.

Due to the small number of markers genotyped in the newly ascertained individuals PEDCHECK was run to identify Mendelian inconsistencies in the genotypes among these individuals. Visual haplotype analysis of these individuals was also performed. This led to the discovery of several other inconsistencies. First, individual 141 did not appear to be the parent of individuals 150 and 200. These offspring were often consistent with each parent (141 and 142) separately, but not together. Since individual 142 showed proper inheritance from her parents, all genotypes for 141 were assigned as unknown for future analyses. As individual 141 married into the family and is reported as being unaffected, assigning the genotypes as unknown should not have a large effect on linkage analyses. The genotypes of individual 152 appeared inconsistent with those of his reported father (112), but consistent with the mother (113). Therefore a ghost father (302) was assigned to individual 152. The final inconsistency observed was individual 158, whose genotypes were consistent with those of the father (112) but not the mother (113) as originally thought. This type of inconsistency is relatively uncommon; nonetheless evidence does suggest that this is the case with individuals 158 and 113. A ghost mother (307) was therefore assigned to individual 158 for future analyses (figure 5.1).

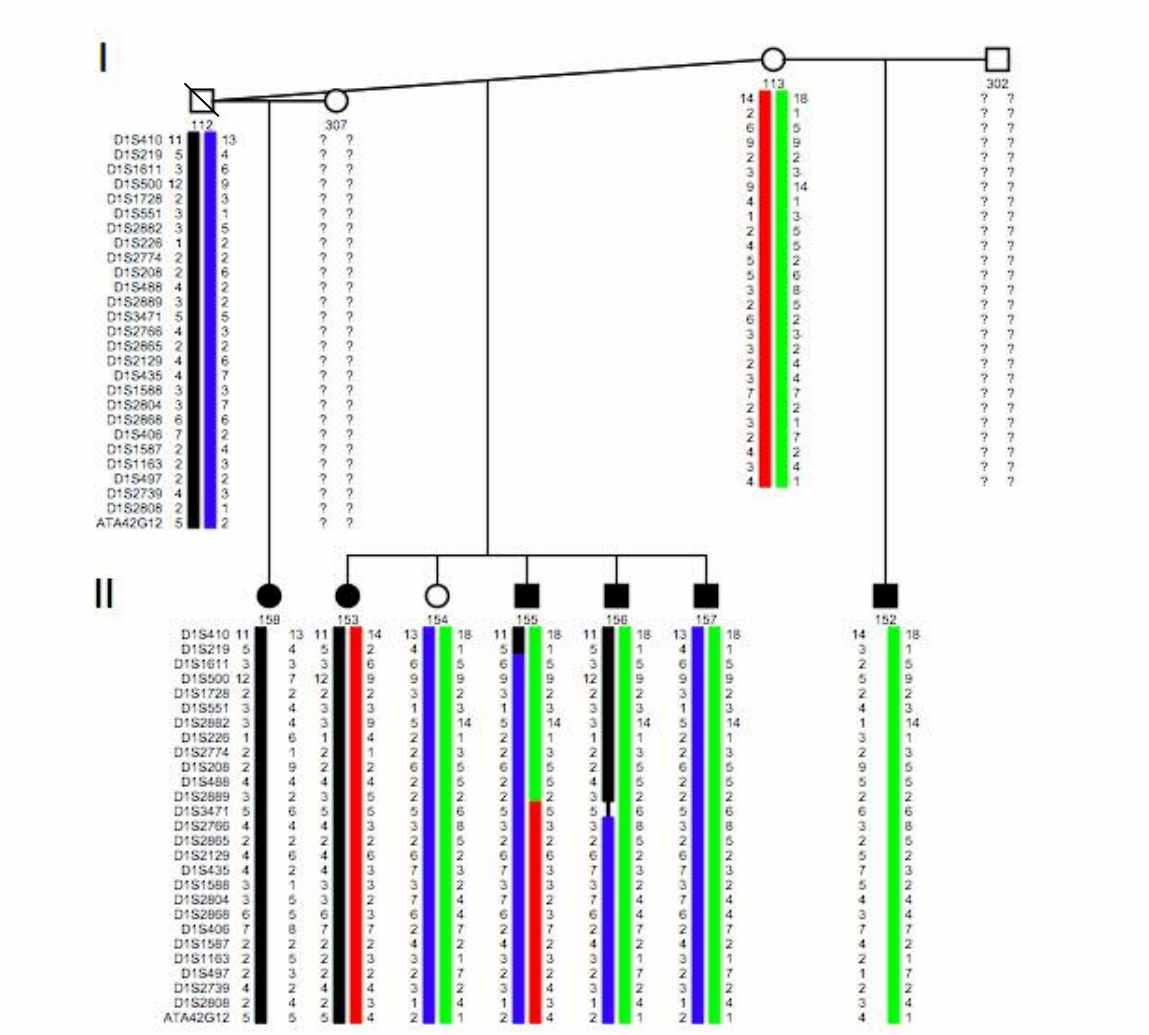


Figure 5.1: The black and red haplotypes are from the father (112), while the red and green haplotypes are from the mother who was reported to have born all offspring pictured in generation II. All individuals in generation II were reported to be offspring of 112 and 113. It is clear that there is mis-parentage in individuals 152 and 158.

PEDCHECK Results

For the initial marker data from the genome wide scan, Mendelian inconsistencies identified by PEDCHECK were checked against original genotype plots in case any genotypes had been mis-called. In the cases where the genotype plots could not resolve the inconsistencies, genotypes were coded as unknown. Markers that showed gross

inconsistencies were not included in the genome wide linkage analysis. This resulted in the removal of 6 markers from initial linkage analyses, ending up with genotype data from ~98% of the markers in the initial autosomal screening set. Of the remaining markers (n=360), the majority were genotyped with a success rate of ≥ 95% and no markers had a success rate below 90%. As additional markers in the region of interest were genotyped, attempts were made to fill in as many missing genotypes possible, with more than 99% of the genotypes obtained in a total of 27 markers in the region of interest on chromosome 1. Based on results from PEDCHECK, several individuals showed numerous Mendelian inconsistencies and were removed altogether from linkage analyses, including 141, 158, 193.

FASTSLINK Results

Analyses were initially performed to estimate the power to detect linkage in CAMST01. The MSIM component of FASTSLINK was initially run with the original set of DNAs obtained. MSIM provided a maximum estimated LOD score of 6.14 in this family. As a LOD score of 3 is considered significant evidence of linkage, it was evident that this family could support linkage. This analysis was re-run after sampling eight additional individuals. A maximum estimated LOD score of over 14 (Table 5.3) was obtained.

	Max Estimated LOD Score
From Initial Sampling	6.13
w/ additional Individuals	14.85

Table 5.3. Maximum estimated LOD scores from FASTSLINK

Linkage Analysis Results

A genome-wide linkage analysis for stuttering was performed. Results of the linkage analysis with the FASTLINK software package produced evidence for linkage on chromosomes 1 and 20 (figure 5.2). Initial two-point results produced a LOD score of 2.27 at D1S551 (114 centimorgans, (cM), band 1p31.1) and 1.58 at D20S482 (12 cM, band 20p13).

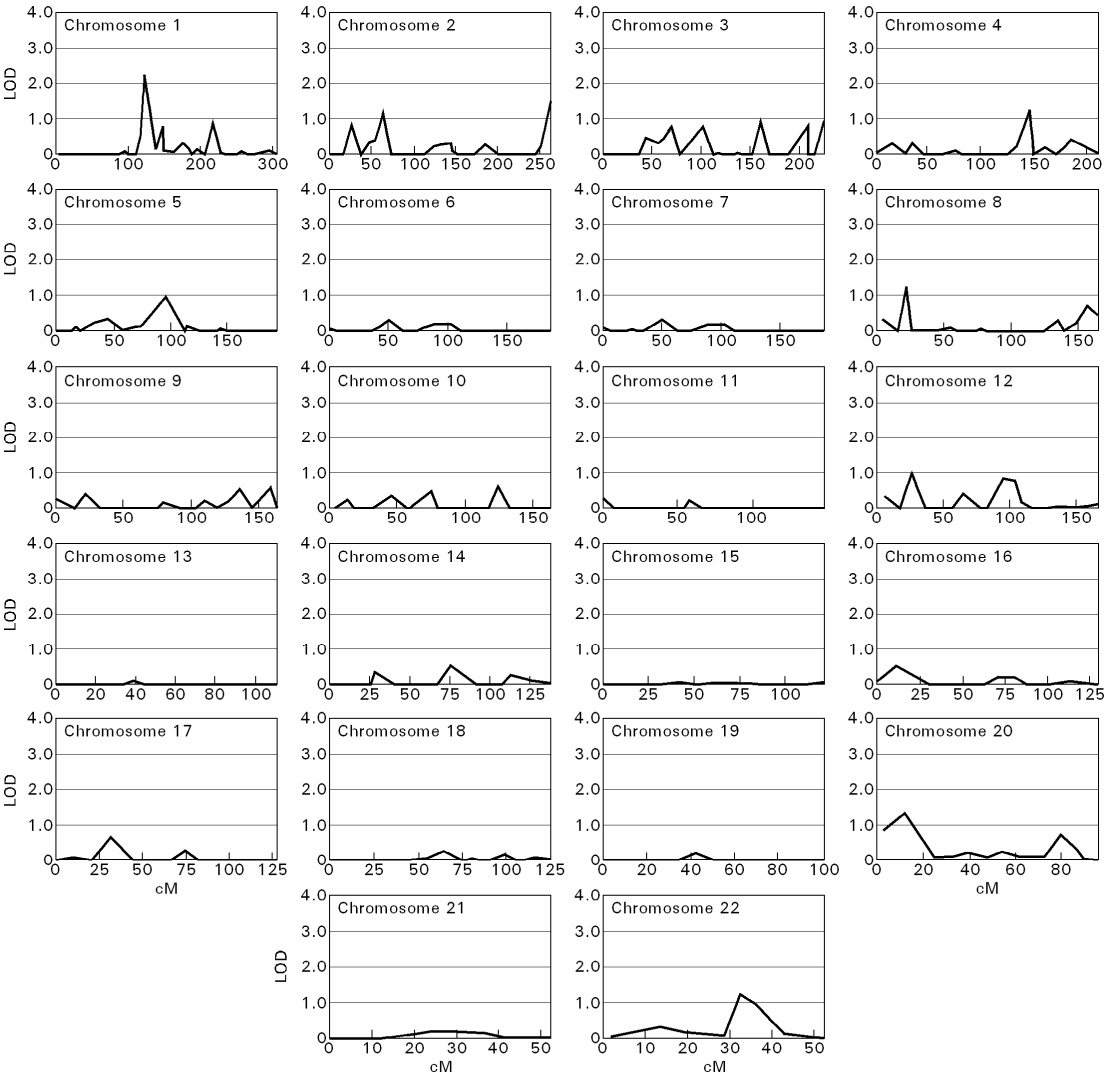


Figure 5.2: Two point LOD scores from genome wide scan using Marshfield – Weber microsatellite screening set 9. Penetrance, 0.8, 0.4, 0.0, Disease allele frequency = 0.01

Because the results of chromosome 1 seemed most promising, 20 additional markers were genotyped in a 40 cM region surrounding D1S551. Additional two-point analysis resulted in a maximum LOD score of 2.16 at D1S1588 (125 cM). Multipoint analysis in this region on chromosome 1 yielded a maximum LOD score of 4.39 with the disease locus between markers D1S2774 and D1S3471 (114 cM and 117 cM) respectively, providing evidence for linkage of this region to persistent developmental stuttering in this family.

Additional markers on chromosome 20 surrounding D20S482 were also genotyped. Two-point analyses with these markers did not produce any higher LOD scores. However, multipoint analysis produced a LOD score of 2.22 with the disease locus between markers D20S482 at 12 cM and D20S851 at 25 cM, providing suggestive evidence of linkage of this region to PDS in this family.

To provide additional information, parametric analysis with SIMWALK2 was also run. SIMWALK2 uses a different algorithm in determining the statistical significance of chromosomal associations with a given phenotype. Whereas FASTLINK uses an exact calculation method, SIMWALK2 uses an estimation method. There are many possible configurations of the genotypic data that need to be tested when performing a linkage analysis. Many calculations need to be performed to test all possible combinations of genotype data when calculating LOD scores. The Markov chain Monte carlo (MCMC) estimation method recognizes that any given configuration is theoretically possible, but excludes those that are highly unlikely. This is beneficial for linkage analyses of extended pedigrees because it reduces the number of calculations that need to be done, thereby reducing computational time. This becomes less of a problem

as computing power has greatly increased. Even though there is a savings in computing time, it is still preferable to use the exact methods when possible.

If linkage truly exists at a locus, it should be detectable regardless of the method used. Results from parametric SIMWALK2 analysis produced a location score, which is directly comparable to a LOD score, of 2.56 on chromosome 1 with markers D1S208 and D1S1588 (114 cM and 125 cM, respectively). Although this does not reach the level of significance found with FASTLINK, it does provide support for linkage of this 11 cM region on chromosome 1 to PDS in this family. The results on chromosome 20 produced a location score of 2.29 at D20S482, also supporting the previous results. No other support for linkage was found using SIMWALK2.

In an attempt to further refine the region of interest speech and blood samples from 8 additional individuals (CAMST01: 110, 141, 158, 163, 193, 194, 200, 205) were obtained. These individuals were previously unsampled due to age or unavailability, or because they yielded little or poor quality DNA from the initial sampling. The new individuals were genotyped with markers in the regions of interest. Additional microsatellite markers surrounding D1S1588 (D1S406, D1S435, D1S2868, D1S2804) and chromosome 20 (D20S906, D20S179, D20S97, D20S849, D20S873, D20S916, D20S901) were also genotyped in the entire family. 2-point linkage analyses were repeated on the larger data set, under multiple hypotheses (Table 5.4).

It is important to note that testing multiple models, as has been done in this study, can lead to an inflated maximum LOD score. To account for this it has been suggested that the LOD score required for significant evidence of linkage be increased by $\log_{10}(n)$, where n is the number of models tested. Therefore a LOD score of $3 + \log_{10}(6) = 3.78$

should be considered significant evidence of linkage in this study (Terwilliger & Ott, 1994). Likewise, when only a single model has been tested a LOD score of 2 is considered suggestive evidence of linkage (Lander & Kruglyak, 1995). Therefore a LOD score of 2.78 should be considered suggestive evidence of linkage in the current study.

The markers immediately surrounding D1S1588 did not produce any evidence for linkage, as would be expected if this locus were contributing to the phenotype. This prompted further inspection of the genotype data and allele frequencies. It became apparent that two of the allele frequencies had been transposed. Alleles 3 and 4 had incorrectly been assigned frequencies of 0.01 and 0.2292, respectively. The correct frequencies were 0.2292 and 0.01 for alleles 3 and 4, respectively. It was clear that allele 3 appeared linked to PDS and this allele was actually far more common than had initially been entered into the linkage analysis. Upon correcting this error, two point linkage analyses were re-run and evidence for linkage at this marker D1S1588 went down. Initially a LOD score of >5 had been achieved at this marker, after the corrections were made to the input data, the maximum LOD score at this marker was 2.97 under the "Fully Penetrant 1" model of nearly complete penetrance (Table 5.4) as mentioned above. Under the initial co-dominant model, a LOD score of 1.54 was obtained.

Model	Penetrance			Chromosome 1		Chromosome 20	
	dd	Dd	DD	Marker	LOD Score	Marker	LOD Score
Co-dominant with medium penetrance	0.80	0.40	0.00	D1S500 D1S551 D1S1588 D1S1163	1.37 <1 1.52 1.27	D20S873	1.68
Co-dominant with lower penetrance	0.50	0.20	0.00	D1S500 D1S551 D1S1588 D1S1163	1.13 <1 1.13 0.98	D20S873	0.83
Fully Penetrant 1	0.99	0.99	0.00	D1S500 D1S551 D1S1588	-12.78 -13.25 2.97	D20S873 D20S901	2.97 2.54
Fully Penetrant 2	0.90	0.90	0.00	D1S500 D1S551 D1S1588	1.03 <1 2.48	D20S873 D20S901	2.16 1.5
Recessive 1	0.90	0.30	0.05	D1S500 D1S551 D1S1588	1.81 <1 1.79	D20S873 D20S901	2.12 226
Recessive 2	0.90	0.10	0.05		Nothing over 1	D20S873 D20S901	0.42 1.76

Table 5.4 Pairwise linkage analysis with complete dataset, all sampled individuals, and corrected allele frequencies. **dd** = homozygous disease allele, **Dd** = heterozygous **DD** = homozygous normal allele.

With the penetrance set at 0.9, 0.9, 0 for 2, 1 and 0 copies of the disease allele, respectively, a LOD score of 2.48, $\theta = 0.10$ at D1S1588 was obtained (using allele frequencies obtained from 44 random Cameroonians) and 2.16, $\theta = 0.10$ at D20S873. Under a model in which the trait is almost completely penetrant (penetrance = 0.99, 0.99, 0) a 2-point LOD score of 2.97, $\theta = 0.062$ was achieved at D1S1588 and 2.97, $\theta = 0.15$ at D20S873. Under this same model, using allele frequencies from 44 random Cameroonians, table 5.5, strong evidence against linkage (which would be a LOD score $\leq$ -3.78 in this analysis) was found at D1S551 (LOD = -13.25, $\theta = 0.0$). This marker previously provided suggestive evidence of linkage from the initial genome wide scan. However, this did not include the 8 individuals from the second phase of sampling. Additionally, the initial results were assuming a model of co-dominant inheritance, with reduced penetrance. Furthermore, the allele frequencies used in the initial linkage analysis were obtained using a sample consisting of 18 random Cameroonians (RCs) and 6 African Americans (AAs). At the time of the initial sampling, blood samples from only

18 RCs were collected. The 6 African Americans were added to increase the sample size in order to obtain a better estimate of the allele frequencies in the Cameroonian population. It was assumed that the African Americans would be genetically similar to the Cameroon population and would not significantly skew the allelic distribution. Due to the possibility that this assumption was not valid, a larger and more representative sample of 44 RCs was genotyped. These individuals were ascertained at a later date and were not available for the initial genome wide scan. The allele frequencies obtained from these two groups were not significantly different (Table 5.5) and should not have a large effect on the LOD score. The change in the LOD score for D1S551 was most likely due to the additional individuals ascertained.

Allele Size (bp)	Allele #	Frequencies[a] ± std. err.	Frequencies[b] ± std. err.
164	1	0.0208 ± 0.021	0.0000
168	2	0.0833 ± 0.040	0.0795 ± 0.029
172	3	0.3125 ± 0.067	0.2159 ± 0.044
176	4	0.3125 ± 0.067	0.3523 ± 0.051
180	5	0.2083 ± 0.059	0.3182 ± 0.050
184	6	0.0417 ± 0.029	0.0341 ± 0.019
188	7	0.0208 ± 0.021	0.0000

Table 5.5: **Allele frequencies D1S551** a. Allele frequencies determined from 18 random Cameroonians (RCs) and 6 African Americans. b. Allele frequencies from 44 Random Cameroonians. Allele's 1 and 7 found in group a, were from the African Americans and were not present in any of the Cameroonian individuals, including the family, CAMST01.

Haplotype Analysis Results

Haplotypes for genotyped individuals in CAMST01 were compared to individual 112, the patriarch who is hypothesized to carry the mutation (Figure 5.3). The goal of performing this analysis was to see if there were any obvious regions of shared haplotypes in affected individuals. Documenting shared haplotypes is important for two reasons. First,it is possible that an associated allele can come from the founder of this family and be passed along, in which case the allele would be identical by decent (IBD). The associated allele can also come from an individual who marries in to the family. If this allele were passed on to future generations, it would be identical by state (IBS). If an associated allele is IBS in some individuals it is possible to get an overestimate of the LOD score, especially if that allele is not associated with stuttering outside the ancestry of CAMST01. Therefore it was important to identify markers that might be IBD as opposed to IBS. The second purpose of the haplotype analysis was to narrow down the critical interval to one that contains a smaller number of candidate genes for further study.

As can be seen in some of the more distantly related individuals (e.g. 163, 175) there is substantial sharing of 112's haplotypes. However, this sharing occurs in many small blocks. This analysis did not provide information that allowed unambiguous narrowing of the disease gene region. One reason for this would be if there were multiple loci involved in persistent developmental stuttering in this family. An epistatic effect may exist in which one locus modifies a second locus thereby affecting the phenotype. It may be the case that this interaction is not required for persistent developmental stuttering, but might increase one's predisposition to developing PDS. Another

explanation could be that there are two independent loci on chromosome 1 in this region that are involved in PDS. The fact that evidence, albeit only modest evidence, for two linkage peaks 10 cM apart (D1S500, LOD = 1.81, 114 cM and D1S1588, LOD = 2.97, 125 cM) have been observed supports the possibility of two separate loci. The possibility also exists that the evidence for linkage on chromosome 1 is a false positive. This is unlikely as several markers in this region have provided some degree of evidence for linkage, with the most significant evidence being multi-point LOD score of > 4 with markers D1S2774 (114 cM) and D1S3471 (117 cM).

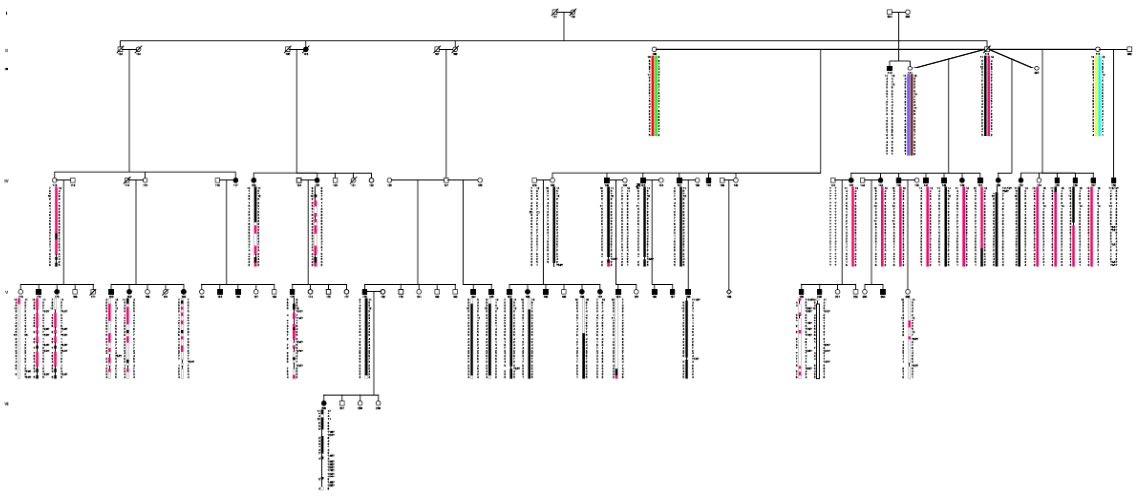


Figure 5.3: Chromosome 1 haplotypes. Black and pink haplotypes are from the individual 112. White haplotypes are not shared with 112, solid line represents indeterminate haplotypes. **? =** untyped markers; **?1/2?** = alleles whose phase is unknown.

Chapter 6: Discussion

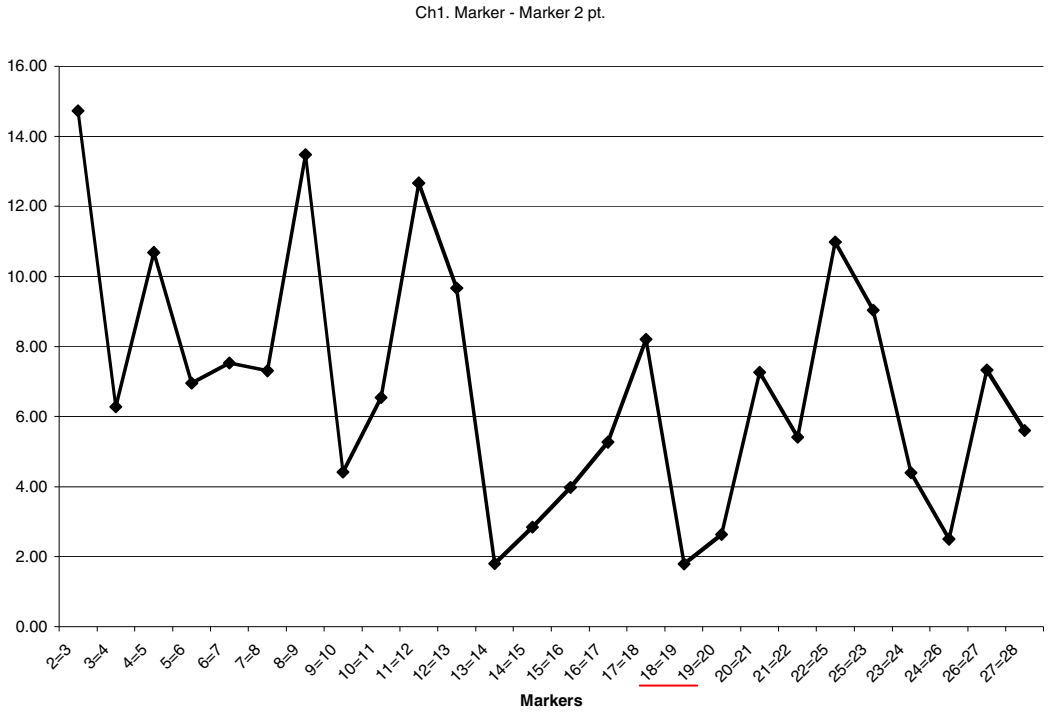
Although there is substantial evidence supporting a genetic contribution to stuttering, the precise effect and mode of transmission of these factors is not entirely clear. Stuttering in CAMST01 does not follow a pure Mendelian mode of transmission. For example, autosomal dominant traits result in a pedigree in which half of all individuals are affected, affected individuals are observed in every generation, and no affected individuals come from two unaffected parents. Although family CAMST01 contains sibships in which about half of the offspring are affected (e.g. individuals 114 & 115 and their offspring and 118 & 119 and their offspring (Figure 3.1)), there are instances where PDS appears to skip at lest one generation. On the other hand there are also occurrences in which affected as well as unaffected parents produce offspring where 1/4 or 1/5 are affected (offspring of 124 & 125 and 126 & 127, Figure 3). Perhaps the most striking observation is the fact that 16 of the 19 offspring from the polygamous individual 112 are affected. Although individual 112 was deceased at the time this family was ascertained, and thus his phenotype could not be confirmed, it is not unreasonable to hypothesize that he was affected based on the number of affected offspring.

These observations support the conclusion that PDS in this family does not follow a simple inheritance pattern. Additionally, the excess of affected males seen in CAMST01 (M:F = 2.14:1), and previous familial studies of stuttering (Drayna et al. 1999; Mellon et al., 1991) often leads to the question of stuttering being an X-linked trait. This is argued against by the several instances of apparent male to male transmission of stuttering in CAMST01 (see individuals 132 → 194, 135 → 196 & 197).

The transmission pattern of stuttering in this family was initially hypothesized to be autosomal co-dominant with reduced penetrance. Under this model, individuals with one copy of the gene would possess a certain susceptibility to developing stuttering, while individuals with two copies would have a higher risk of developing stuttering. In the current study of CAMST01 initial linkage analyses were performed under this model with the disease penetrance set at 0.4 for individuals that have a single copy of the disease gene and 0.8 for individuals possessing two copies of the disease gene. Initial linkage results provided strongest evidence for linkage on chromosome 1. Based on these results additional genotyping in this region was performed. Secondary analyses with the new markers lead to a second linkage peak about 10 cM away from the initial linkage peak at 114 cM. These two linkage peaks could have been two separate linkage signals, not necessarily related to each other. Additional markers were genotyped within 1 megabase of D1S1588 to obtain more information on the linkage pattern observed. Two-point linkage analysis with the additional neighboring markers and the disease did not lead to any significant or even suggestive evidence for linkage. This could have been due to lack of informativity with these markers or the linkage peak we observed at D1S1588 may not have been a true linkage signal. A marker is considered informative if recombination information is able to be unambiguously determined. This is important because linkage analysis works using principles of recombination.

In order to check marker-to-marker informativity, marker-to-marker two-point linkage analysis was performed. Due to the proximity of D1S1588 and neighboring markers (assuming they are informative) marker-to-marker LOD scores should be near the maximum estimated LOD score of ~14 in CAMST01. The observed LOD scores in

this analysis were far below the maximum possible LOD score (red outlines, Figure 6.1). This suggests that D1S1588 is not a highly informative marker in this family. Another explanation for the observed LOD scores is that these markers may not actually be linked. In order to confirm the possibility that D1S1588 is uninformative in CAMST01, genotypes for individual 112 were reconstructed. Upon visual inspection of the genotype data it became apparent that individual 112 was homozygous at D1S1588. All marker informativity was therefore lost in the offspring of individual 112. This accounts for the relatively low observed LOD scores in both marker-to-marker linkage and marker to disease linkage analyses.



Locus	Marker	Map Position (cM)
2	D1S410	100.39
3	D1S219	101.48
4	D1S1611	104.79
5	D1S500	107.56
6	D1S1728	109.04
7	D1S551	113.69
8	D1S2882	113.69
9	D1S226	113.69
10	D1S2774	113.69
11	D1S208	113.83
12	D1S488	114.24
13	D1S2889	116.72
14	D1S3471	116.72
15	D1S2766	118.14
16	D1S2865	120.28
17	D1S2129	~122.00
18	D1S435	125.51
19	D1S1588	125.51
20	D1S2804	126.16
21	D1S2868	126.16
23	D1S406	126.16
23	D1S1587	128.73
24	D1S1163	129.37
25	D1S497	129.37
26	D1S2739	130.73
27	D1S2808	131.87
28	ATA42G12	139.02

Figure 6.1 Marker-to-marker 2 point LOD scores. The markers tested are listed above. Markers with LOD scores near the estimated maximum LOD score of 14 suggest that these markers are more highly informative. Markers with low LOD scores, as seen with marker 19, D1S1588 and its neighbors, suggest low informativity. The map position for D1S2129 (locus 17) was estimated based on the physical position, as it is not mapped on the Marshfield – Weber genetic map used for the other markers. For the purpose of linkage analysis, it is necessary to assign locus 1 as the disease locus. Because this analysis considers only marker loci, the disease locus is not included.

A similar exercise was carried out for marker D1S551. The genotype data were displayed on a pedigree and were visually inspected for any obvious inconsistencies or recombination events. The genotype data and the allele frequencies appeared to be correct. However, because these allele frequencies were based on a much smaller sample size (18RCs, 6 AAs) all of the available RCs (n=44) were genotyped. The allele frequencies changed slightly, but were not significantly different. This may or may not have had an impact on the results of the linkage analyses. Depending on which allele of D1S551 is providing the evidence of linkage, the LOD score could either go up or down. In fact, the LOD score did go down (2.27 to < 1) when 2-point linkage analysis was performed with the 8 additional individuals and the allele frequencies obtained from 44 RCs. It is unclear how much of this change is due to the new individuals and how much might be due to the new allele frequencies.

Under the nearly complete penetrance model, strong evidence against linkage was found for D1S551, with a LOD score of -13.25. Under all other models of inheritance, D1S551 failed to produce a LOD score over 1. This is likely due to the additional individuals more so than the allele frequencies for two reasons. One reason is that the difference between allele frequencies as determined by the smaller sample size (18 RCs, 6 AAs) and the larger sample (44 RCs) was found to be non-significant. Secondly, and perhaps with a greater effect, many of the additional individuals were more distantly related to individual 112. These individuals can provide stronger evidence in favor of or against linkage because of the many number of meioses between them and individual 112, providing more power to detect linkage.

The other marker of interest on chromosome 1 is D1S500. This marker produced very little evidence for linkage to PDS, and approached suggestive evidence under a recessive model (LOD=1.81, θ=0.01, dd = 0.9, Dd = 0.3, DD = 0.05). However, under a dominant model, strong evidence against linkage was observed (LOD = -12.78, θ = 0, dd = 0.99, Dd = 0.99, DD = 0). Interestingly though, under a closely related model in which the penetrance was lowered to 0.9, 0.9, 0 a LOD score of 1.03 was observed. Although this is hardly evidence for linkage, it is also not the highly negative LOD score from the more completely penetrant model.

D1S1163 also produced LOD scores worth noting, under the co-dominant model, a LOD score of 1.27 was observed. Although this is only weak evidence for linkage, it is worth mentioning due to the fact that it is among other markers in this region that have produced similar LOD scores suggesting that it might be worthwhile to focus future attention in this region of the genome.

Genotyping was not performed with markers on the X chromosome nor were linkage analyses conducted under an X-linked recessive model due to the observations of male-to-male transmission mentioned previously.

The linkage results obtained to date are indicative of a complex trait, which is not particularly surprising considering results from previous linkage analyses (Riaz et al., 2005; Shugart et al., 2004; Suresh et al., 2006) and segregation analyses (Ambrose, Cox, & Yairi, 1997; Cox, Kramer, & Kidd, 1984; Mellon, Hanson, Hasstedt, Leppert, & White, 1991; Kidd, 1977; Viswanath, Lee, & Chakraborty, 2004). It is also clear that further analyses will need to be conducted, perhaps using conditional analyses as was done by Suresh et al. and/or non-parametric analyses as done by Riaz et al. (2005) and

Shugart et al. (2004). SIMWALK2 is a popular linkage analysis software package that has an option for non-parametric analysis. However, attempts at performing this type of analysis were unsuccessful due to the size of the pedigree. While the program would run, not all of the necessary statistics were calculated. This did not seem to be due to inadequate memory or problems with the input files. In particular the statistic NPL_ALL, which measures overrepresentation of a small number of founder alleles in affected individuals in the pedigree was not calculated. This statistic is particularly useful for additive traits, as has been hypothesized in CAMST01. The maximum number of affected individuals this program will handle is 20, and CAMST01 contains more than that in a single branch. Breaking up that branch would result in an undesirable loss of information. Under the assumption that individual 112 is homozygous for the disease – causing allele, the power in this family comes from affected individuals that are more distantly related to individual 112. There have been many more meioses in these distantly related individuals. If they share a disease causing allele with individual 112 the possibility of observing strong linkage increases. As mentioned previously, parametric linkage analyses using SIMWALK2 corroborated the findings on chromosomes 1 and 20, providing additional support loci on chromosomes 1 and 20 showing linkage with PDS in this family.

The high male:female ratio among PWS may have several explanations. An inherited X-linked recessive allele could be one such cause, but in family CAMST01, this is unlikely due to the many instances of male to male transmission. As discussed earlier, recent evidence for sex-specific genetic factors of stuttering comes from Suresh et al. (2006) who have performed linkage analyses that were conditional on sex. Although the

actual underlying genetic mechanisms are yet to be identified, it is possible that males and females differ in predisposing causes of PDS. This is consistent with Ambrose et al.'s (1997) hypothesis that males and females respond differently to the same underlying genetic variation. Although not done in this study, it is possible to incorporate a sex-specific model, as Suresh et al. (2006) have done. This should be considered for future analyses.

Chapter 7: Conclusion

In conclusion, a large multi-generational, polygamous family from the Northwest Province of Cameroon has been identified. It is clear that persistent developmental stuttering segregates in this family. A genome wide scan and linkage analyses were performed on this family. Suggestive evidence for linkage to chromosomes 1 and 20 was found with maximum LOD scores of 2.97 on each chromosome. Questions still remain about the exact mode of inheritance and why a LOD score close to the maximum expected LOD score was not obtained. Considering the fact that stuttering is a complex trait that is likely comprised of genetic and non-genetic factors, it is not surprising that only suggestive evidence for linkage was obtained. These results, taken into consideration with previous linkage studies of stuttering follow patterns similar to other complex traits (e.g. Schizophrenia) (Riley & Kendler, 2006). In particular, multiple loci have been identified and evidence for linkage with varying degrees of significance has been found. As genetic studies of stuttering are still in their infancy there have not been any other independent studies attempting to replicate published findings. It is likely that as more genetic linkage studies of persistent developmental stuttering are conducted, even among populations that have been previously studied, additional loci will continue to be identified. Furthermore, it will be necessary to continue employing specialized populations and large pedigrees segregating PDS.

Chapter 8: Future Studies

It is apparent that the estimated maximum LOD score in this family has not been obtained. It is clear that further analyses need to be conducted. One such analysis would be non-parametric linkage analysis, and although it might not help to narrow down the regions of interest, it may be possible to obtain more significant LOD scores. Attempts at these methods were unsuccessful, but additional development in the analysis programs may provide ways to accomplish non-parametric analyses. Another possibility would be to conduct sex-specific analyses and also conditional analyses in which linkage analysis is performed conditional on loci that have shown previous evidence for or against linkage, as was done by Suresh et al. (2006). The linked regions on chromosomes 1 and 20 could not be narrowed to any less than ~10 Mb. There are ~60 genes in the region on chromosome 1 between 114 cM and 125 cM. Fifty of these are known and 10 are predicted. Once stronger evidence for linkage is obtained and the region on chromosome 1 is narrowed down to a manageable size, a candidate gene study will be conducted. One method to narrow these regions would be through ascertainment of additional families from the same region of Cameroon, assuming that they are genetically very similar. If this is so it is possible that stuttering in these additional families is due to a mutation at the same genetic locus. Based on evidence from population genetics studies in Africa (personal communication, Tishkoff, 2006), the assumption that other Bantu speaking families are genetically very similar is reasonable. However, the possibility that stuttering is genetically heterogeneous in this population needs to be considered. This

does not rule out the possibility that additional families may still be of use in narrowing down regions of interest on chromosomes 1 and 20. Another approach would be to perform a case / control association study using one of the large SNP genotyping sets available. These SNP genotyping sets currently contain 100,000 to 500,000 SNPs and as demonstrated by Zhang et al. (2005), might be able to identify and narrow down linked regions better than microsatellite markers. However due to initial evidence of genetic heterogeneity in this, as well as other, populations, it may be difficult to observe a significant association.

Bibliography

Alfonso, P. J. (1991). Implications of the concepts underlying task-dynamic modeling on kinematic studies of stuttering. In H. F. M. Peters, W. Hulstijn & C. W. Starkweather (Eds.), Speech motor control and stuttering. Amsterdam: Elseviers Publishers.

Alm, P. A. (2004). Stuttering and the basal ganglia circuits: a critical review of possible relations. Journal of Communication Disorders, 37(4), 325-369.

Ambrose, N. G., Yairi, E., & Cox, N. (1993). Genetic aspects of early childhood stuttering. J Speech Hear Res, 36(4), 701-706.

Ambrose, N. G., Cox, N. J., & Yairi, E. (1997). The genetic basis of persistence and recovery in stuttering. J Speech Lang Hear Res, 40(3), 567-580.

Ambrose, N., & Yairi, E. (2002). The Tudor Study: Data and Ethics. American Journal of Speech-Language Pathology, 11, 190-203.

Anderson, J. M., Hughes, J. D., Rothi, L. J. G., Crucian, G. P., & Heilman, K. M. (1999). Developmental stuttering and Parkinson's disease: the effects of levodopa treatment. J Neurol Neurosurg Psychiatry, 66(6), 776-778.

Andrews, G., Morris-Yates, A., Howie, P., & Martin, N. G. (1991). Genetic factors in stuttering confirmed. Arch Gen Psychiatry, 48(11), 1034-1035.

ANSI. (1989). American National Standard Specification for Audiometers. New York: American National Standards Institute.

ANSI. (1991). American National Standard Maximum Permissible Ambient Noise Levels for Audiometers. New York: American National Standards Institute.

Bloodstein, O. (1995). A handbook on stuttering. 5th ed. San Diego: Singular Publishing Group.

Bloodstein, O., & Bernstein Ratner, N. (2007). A handbook on stuttering (6 ed.). Clifton Park, NY: Thomson.

Brady, J. P. (1991). The pharmacology of stuttering: a critical review. Am J Psychiatry, 148(10), 1309-1316.

Brotchie, P., Iansek, R., & Horne, M. K. (1991). Motor function of the monkey globus pallidus. 2. Cognitive aspects of movement and phasic neuronal activity. Brain, 114 (Pt 4), 1685-1702.

Coriat, I. H. (1927). Stammering: A Psychoanalytic Interpretation. Nervous and Mental Disease Monographs, 47, 1-68.

Cottingham, R. W., Jr., Idury, R. M., & Schaffer, A. A. (1993). Faster sequential genetic linkage computations. Am J Hum Genet, 53(1), 252-263.

Cox, N. J., Kramer, P. L., & Kidd, K. K. (1984). Segregation analyses of stuttering. Genet Epidemiol, 1(3), 245-253.

Denny, M., & Smith, A. (1992). Gradations in a Pattern of Neuromuscular Activity Associated With Stuttering. J Speech Hear Res, 35(6), 1216-1229.

Drayna, D., Kilshaw, J., & Kelly, J. (1999). The sex ratio in familial persistent stuttering. Am J Hum Genet, 65(5), 1473-1475.

Drayna, D. (2006). Personal Communication.

Dryer, J. (2001, June 11). An experiment leaves a lifetime of anguish the study's young victims were left in ignorance, to cope alone. Experts debate whether the benefits justified the harm. San Jose Mercury News, p. 1A.

DSM-IV. (1994). Diagnostic and Statistical Manual of Mental Disorders: American Psychiatric Association.

Evans, D. M., & Cardon, L. R. (2004). Guidelines for genotyping in genomewide linkage studies: single-nucleotide-polymorphism maps versus microsatellite maps. Am J Hum Genet, 75, 687-692.

Felsenfeld, S., Kirk, K. M., Zhu, G., Statham, D. J., Neale, M. C., & Martin, N. G. (2000). A study of the genetic and environmental etiology of stuttering in a selected twin sample. Behav Genet, 30(5), 359-366.

Fenichel, O. (1972 (1945)). The psychoanalytic theory of neurosis. New York: W. W. Norton and Company, Inc.

Foundas, A. L., Bollich, A. M., Corey, D. M., Hurley, M., & Heilman, K. M. (2001). Anomalous anatomy of speech-language areas in adults with persistent developmental stuttering. Neurology, 57(2), 207-215.

Gazzaniga, M. S., Ivry, R. B., & Mangun, G. R. (2002). Cognitive Neuroscience: the biology of the mind (2 ed.). New York: W.W. Norton & Company, Inc.

Ghebranious, N., Vaske, D., Yu, A., Zhao, C., Marth, G., & Weber, J. L. (2003). STRP screening sets for the human genome at 5 cM density. BMC Genomics, 4(1), 6.

Goberman, A., & Blomgren, M. (2003). Parkinsonian speech disfluencies: effects of L-Dopa-related fluctuations. Journal of Fluency Disorders, 28(1), 55-70.

Gray, M. (1940). The X Family: A clinical and laboratory study of a "stuttering" family. Journal of Speech Disorders(5), 343-348.

Guitar, B. (1998). Stuttering: An integrated approach to its nature and treatment (2 ed.). Baltimore: Williams & Wilkins.

Guitar, B., Guitar, C., Neilson, P., O'Dwyer, N., & Andrews, G. (1988). Onset Sequencing of Selected Lip Muscles in Stutterers and Nonstutterers. J Speech Hear Res, 31(1), 28-35.

Harris, V., Onslow, M., Packman, A., Harrison, E., & Menzies, R. (2002). An experimental investigation of the impact of the Lidcombe Program on early stuttering. Journal of Fluency Disorders, 27(3), 203-214.

Howie, P. M. (1981). Concordance for stuttering in monozygotic and dizygotic twin pairs. J Speech Hear Res, 24(3), 317-321.

Hulstijn, W., & Van Lieshout, P. H. H. M. (1998). A Motor skill approach to stuttering. In W. Ziegler & K. Deger (Eds.), Clinical Phonetics and Linguistics. London: Whurr Publishers.

Hussain, R., & Bittles, A. H. (1998). The Prevalence and Demographic Characteristics of Consanguineous Marriages in Pakistan. Journal of Biosocial Science, 30, 261-275.

Janssen, P., Kloth, S., Kraaimaat, F., & Brutten, G. J. (1996). Genetic factors in stuttering: a replication of Ambrose, Yairi and Cox's (1993) study with adult probands. J Fluency Disord, 21, 105-108.

Johnson, W. (1942). A study of the onset and development of stuttering. Journal of Speech Disorders, 7, 251-257.

Jones, R. D., White, A. J., Lawson, K. H. C., & Anderson, T. J. (2002). Visuoperceptual and visuomotor deficits in developmental stutterers: An exploratory study. Human Movement Science, 21, 603-619.

Kidd, K. K. (1977). A genetic perspective on stuttering. Journal of Fluency Disorders, 2(4), 259-269.

Kidd, K. K., Heimbuch, R. C., & Records, M. A. (1981). Vertical transmission of susceptibility to stuttering with sex-modified expression. Proc Natl Acad Sci U S A, 78(1), 606-610.

Klein, J. F., & Hood, S. B. (2004). The impact of stuttering on employment opportunities and job performance. Journal of Fluency Disorders, 29(4), 255-273.

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

Lange, K., & Goradia, T. M. (1987). An algorithm for automatic genotype elimination. Am J Hum Genet, 40(3), 250-256.

Levis, B., Mundorff, J., Lukong, J., Kuster, J., Braun, A. R., Brewer, C., et al. (In preparation).

Lois, E. D., Winfield, L., Fahn, S., & Ford, B. (2001). Speech dysfluency exacerbated by levodopa in Parkinson's disease. Movement Disorders, 16(3), 562-565.

Ludlow, C. L., & Loucks, T. (2003). Stuttering: a dynamic motor control disorder. J Fluency Disord, 28(4), 273-295; quiz 295.

Lukong, J. (2002). Speak Clear Association of Cameroon.

MacFarlane, W. B., Hanson, M., Walton, W., & Mellon, C. D. (1991). Stuttering in five generations of a single family : A preliminary report including evidence supporting a sex-modified mode of transmission. Journal of Fluency Disorders, 16(2-3), 117-123.

Manning, W. H. (2001). Clinical Decision Making in Fluency Disorders (2 ed.). Vancouver: Singular/Thomson Learning.

McClean, M. D., Tasko, S. M., & Runyan, C. M. (2004). Orofacial Movements Associated With Fluent Speech in Persons Who Stutter. J Speech Lang Hear Res, 47(2), 294-303.

McPeek, M. S., & Sun, L. (2000). Statistical tests for detection of misspecified relationships by use of genome-screen data. Am J Hum Genet, 66(3), 1076-1094.

Mellon, C., Hanson, M. L., Hasstedt, S., Leppert, M., & White, R. (1991). Early Findings from the Study of a Large Pedigree Segregating a Stuttering Phenotype. Am J Hum Genet, 49 (supl)(4), 351.

Nelson, S. (1939). The Role of Heredity in Stuttering. The Journal of Pediatrics, 14, 642-654.

O'Connell, J., & Weeks, D. E. (1997). PedCheck: A program for identifying genotype incompatibilities in linkage analysis. Journal of Human Genetics, 63, 259-266.

Ott, J. (1989). Computer-simulation methods in human linkage analysis. Proc Natl Acad Sci U S A, 86(11), 4175-4178.

Riaz, N., Steinberg, S., Ahmad, J., Pluzhnikov, A., Riazuddin, S., Cox, N. J., et al. (2005). Genomewide significant linkage to stuttering on chromosome 12. Am J Hum Genet, 76(4), 647-651.

Riley, B., & Kendler, K. S. (2006). Molecular genetic studies of schizophrenia. Eur J Hum Genet, 14(6), 669-680.

Riley, G. (1980). Stuttering Severity Instrument for Children and Adults (3 ed.). Los Angeles: Western Psychological Services.

Salmelin, R., Schnitzler, A., Schmitz, F., & Freund, H. J. (2000). Single word reading in developmental stutterers and fluent speakers. Brain, 123 (Pt 6), 1184-1202.

Schaffer, A. A., Gupta, S. K., Shriram, K., & Cottingham, R. W., Jr. (1994). Avoiding recomputation in linkage analysis. Hum Hered, 44(4), 225-237.

Shugart, Y. Y., Mundorff, J., Kilshaw, J., Doheny, K., Doan, B., Wanyee, J., et al. (2004). Results of a genome-wide linkage scan for stuttering. Am J Med Genet A, 124(2), 133-135.

Silverman, F. H. (1988). The "monster" study. Journal of Fluency Disorders, 13, 225-231.

Slager, S. L., & Vieland, V. J. (1997). Investigating the numerical effects of ascertainment bias in linkage analysis: Development of methods and preliminary results. Genetic Epidemiology, 14(6), 1119-1124.

Snieder, H., & MacGregor, A. J. (2003). Twin Methodology. In D. N. Cooper (Ed.), Encyclopedia of the Human Genome. London; New York: Nature Publishing Group.

Sobel, E., & Lange, K. (1996). Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker-sharing statistics. Am J Hum Genet, 58(6), 1323-1337.

Sobel, E., Papp, J. C., & Lange, K. (2002). Detection and integration of genotyping errors in statistical genetics. Am J Hum Genet, 70(2), 496-508.

Sobel, E., Sengul, H., & Weeks, D. E. (2001). Multipoint estimation of identity-by-descent probabilities at arbitrary positions among marker loci on general pedigrees. Hum Hered, 52(3), 121-131.

Strauch, K., Fimmers, R., Baur, M. P., & Wienker, T. F. (2003). How to Model a Complex Trait. Human Heredity, 55(4), 202-210.

Suresh, R., Ambrose, N., Roe, C., Pluzhnikov, A., Wittke-Thompson, J. K., Ng, M. C., et al. (2006). New complexities in the genetics of stuttering: significant sex-specific linkage signals. Am J Hum Genet, 78(4), 554-563.

Terwilliger, J. D., & Ott, J. (1994). Handbook of Human Genetic Linkage. Baltimore: Johns Hopkins University Press.

Tishkoff, S. A. (2006), personal Communication.

Tishkoff, S. A., & Verrelli, B. C. (2003). Patterns of human genetic diversity: implications for human evolutionary history and disease. Annu Rev Genomics Hum Genet, 4, 293-340.

Toth, G., Gaspari, Z., & Jurka, J. (2000). Microsatellites in Different Eukaryotic Genomes: Survey and Analysis. Genome Res., 10(7), 967-981.

Travis, L. (1931). Speech Pathology. New York: Appleton-Century.

Tudor, M. (1939). An experimental study of the effect of evaluative labeling on speech fluency. University of Iowa, Iowa City.

Van Lieshout, P. H. H. M. (1995). Motor planning and articulation in fluent speech of stutterers and nonstutterers. Nijmegen: University of Nijmegen.

Van Riper, C. (1937). The Preparatory Set in Stuttering. Journal of Speech Disorders, 2, 149-154.

Van Riper, C. (1954). Speech Correction: Principles and Methods (3 ed.). New York: Prentice-Hall, Inc.

Van Riper, C. (1982). The Nature of Stuttering (2 ed.). Englewood Cliffs, N.J.: Prentice-Hall.

Viswanath, N., Lee, H. S., & Chakraborty, R. (2004). Evidence for a major gene influence on persistent developmental stuttering. Hum Biol, 76(3), 401-412.

Vitzthum, V. J. (2003). A Number No Greater than the Sum of Its Parts: The Use and Abuse of Heritability. Human Biology, 75(4), 539-558.

Weber, J. L., & May, P. E. (1989). Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Am J Hum Genet, 44(3), 388-396.

Webster, M. (1978). Stuttering. Can Med Assoc J, 119(4), 305.

Weeks, D. E., & Ott, J. (1990). SLINK: a general simulation program for linkage analysis. Am J Hum Genet, 47(3), A204 (supplement).

Wepman, J. (1939). Familial Incidence of Stammering. Journal of Heredity, 30, 207-210.

West, R., Nelson, S., & Berry, M. (1939). The Heredity of Stuttering. The Quarterly Journal of Speech(25), 23-30.

WHO. (2004). International Statistical Classification of Diseases and Health Related Problems, ICD-10: World Health Organization.

Wright, A. F., Carothers, A. D., & Pirastu, M. (1999). Population choice in mapping genes for complex diseases. Nat Genet, 23(4), 397-404.

Wu, J. C., Maguire, G., Riley, G., Lee, A., Keator, D., Tang, C., et al. (1997). Increased dopamine activity associated with stuttering. Neuroreport, 8(3), 767-770.

Zhang, C., Cawley, S., Liu, G., Cao, M., Gorrell, H., & Kennedy, G. C. (2005). A genome-wide linkage analysis of alcoholism on microsatellite and single-nucleotide polymorphism data, using alcohol dependence phenotypes and electroencephalogram measures. BMC Genet, 6 Suppl 1, S17.

Zhao, C., Heil, J., & Weber, J. L. (1999). A genome-wide portrait of short tandem repeats. American Journal of Human Genetics (65(supplement)), A102.

Zimmermann, G. (1980). Articulatory Behaviors Associated with Stuttering: A Cinefluorographic Analysis. J Speech Hear Res, 23(1), 108-121.