Department of Speech Language Pathology
All India Institute of Speech & Hearing
Manasagangothri, Mysore 570 006

GENOTYPING AND A GENETIC ASSOCIATION STUDY IN AUTISM

Principal Investigator: Dr. Shyamala.K.C

Co-Investigator: Dr. A.Chandrashekar

Research Officer:Dhanesh Mandamkulathil

GENOTYPING AND A GENETIC ASSOCIATION STUDY IN AUTISM

Project under AIISH Research Fund (ARF)

(2009-2012)

(Ref: SH/CDN/ARF/3.69/KCS/2009-10)

Total grants: Rs 6,77,000/-

Total duration of the project: 3 years

Principal Investigator

Dr. Shyamala K.C

Professor & Head OC-(DHLS)

Department of Speech Language Pathology

Co-Investigator

Dr.A.Chandrashekar

Research Associate (Physical)

Anthropological Survey of India

Mysore

Research Officer

Dhanesh Mandamkulathil

Department of Speech Language Pathology

Acknowledgements

The authors would like to thank late Dr. Vijayalakshmi Basavaraj, the former Director, AIISH and the present Director, Dr. S.R. Savithri, All India Institute of Speech and Hearing, Mysore, for funding the project and providing infrastructure to carry out the project work. We also thanking to Dr. V.R Rao former Director of Anthropological Survey of India and Dr. K.K. Mishra the present Director for the support for successful collaboration with AIISH. We thank all the participants and their care-givers for their cooperation. We thank Mr. Srinivas Naidu Kovali (JRF-AnSI), Arjun Rao (SRF-AnSI), Dr. Venugopal P.N (Research Associate) and Ms. M.S. Vasanthalakshmi, Lecturer in Biostatistics, AIISH, Mysore, for all the Lab support and help with statistical analysis and their valuable insights. We extend our thanks to Ms. Shakuntala, Staff Nurse, Prof. H. Sundara Raju, Dr. G. Rajeshwari (Head) Department of ENT, AIISH for their assistance during blood sample collection for the study. Our sincere gratitude to the Mrs. Geetha (housing keeping staff, AIISH) Mr. Hari (Teacher - BKNM UP School, Arikkulam, Calicut, Kerala) for the support for control sample collection conducted in Mysore and Kerala.

Principal Investigator

Dr. Shyamala K.C

Professor & Head OC-(DHLS)

Dept of Speech Language Pathology.

Co-Investigator

Dr.A. Chandrashekar.

Research Associate(Physical)

Anthropological Survey of India

LIST OF TABLES

LIST OF FIGURES

LIST OF CONTENTS

CHAPTER - I

INTRODUCTION

Autism Spectrum Disorders (ASDs) are neuro-developmental disorders and the term "spectrum" refers to the broad range of symptoms, skills, and levels of impairment, or disability. DSM-IV-TR (Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition - Text Revision) is the standard diagnosing manual for ASDs and the recently updated DSM-V for ASDs will be implemented in May 2013. Severity of the symptoms may vary with children with ASDs.DSM-IV manual currently defines five disorders, i.e. pervasive developmental disorders (PDD), Autistic disorder (classic autism), Asperger's syndrome, Pervasive developmental disorder not otherwise specified (PDD-NOS), Rett's syndrome and Childhood disintegrative disorder (CDD).Autism Spectrum Disorders once considered a rarity ,have in the recent past shown a dramatic rise in prevalence. Given its complexities and debilitating nature, lack of clear understanding of its cause and consequently thereof of the treatment, it has become the focus of much public debate and research.

Autism is a neuro-developmental disorder shows character features of primary impairments in social interactions, communication, and repetitive and stereotyped behaviours (American Psychiatric Association, 2000). Approximately 20% of individuals with Autism function within the normal range on IQ tests (American Psychiatric Association, 1994; Cohen & Volkmar, 1997). Core symptoms which define autism consist of deficits of varying degrees in three areas of development; defective social or personal relatedness behaviours, language or, more broadly, communication difficulties, play or preferred activities which are characterized by preoccupations having a repetitive or stereotyped quality. The severity of autism

varies widely. Researchers agree that the cause of autism has neurobiological origins, and is not the result of poor parenting. Dysfunction of specific nerve networks appears to be responsible for autism.

This dysfunction may be secondary to a variety of disorders that affect the brain. Unfortunately, the precise cause of autism in the majority of children is not yet well understood. Many individuals with more severe autism have identifiable underlying medical conditions. These may include a variety of congenital, chromosomal, metabolic, and occasionally acquired conditions.

Kanner (1943) first identified a group of children with "autistic disturbances of affective contact" in the United States, and since then classic autism and the broader group Autism Spectrum Disorders (ASD) have been formally adopted by various professional diagnostic systems (e.g., Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR); International Classification of Diseases). A "child showing schizophrenic behaviour" was first mentioned in 1959 and autism-specific research has substantially increased since then (Daley, 2004). Sustained professional focus on autism began in the late 1980s and 1990s (Krishnamurthy, 2008), and autism is now widely recognized (Daley, 2004; Daley & Sigman, 2002).

Autism is a lifelong, often severely impairing neuro-developmental syndrome involving deficits in social relatedness, language, and behaviour. Increased prevalence, the lack of controlled studies on treatment efficacy, and the immense expenditure of education and care for these patients all combine to make autism a public health crisis. However, despite a large body of careful pharmacological, pathological, electrophysiological, and imaging investigations, the etiology of autism remains unclear. Known medical conditions account for only a fraction (5-20%) of

autism. Co-morbidity with other neurological disorders, demonstrated widespread neuropathlogical changes, and an emerging literature of structural and functional neuro-imaging differences in autism all herald an underlying central nervous system (CNS) abnormality.

Rett syndrome is a complex neurobiological disorder of development in which an infant seems to grow and develop normally at first, but then stops developing and even loses skills and abilities. In 1999, NICHD-supported scientists discovered that most girls with Rett syndrome have a change in the pattern of a single gene the Methyl-CpG-binding Protein 2 (MECP2) gene on the X chromosome. Between 90 percent and 95 percent of girls with Rett syndrome have a mutation in this gene. This gene makes methyl-CpG-binding protein 2 (MeCP2), which is necessary for the development of the nervous system-especially the brain. The mutation causes the gene to make less than the needed amounts of the protein, or to make a damaged protein that the body can't use. As a result, there may not be enough usable amount of the protein for the brain to develop normally. Beginning between 3 months and 3 years of age, most children with Rett syndrome start to show loss of purposeful hand movements, loss of speech ,balance and coordination problems, stereotypic movements and breathing problems

Childhood disintegrative disorder (CDD) is a clinical syndrome characterized by disintegration of mental functions and regression of acquired language and intellectual functions after a period of normal development up to 3-4 years in children. Occurrence of this condition, though with a different name, was described as long back as in 1930 (Heller, 1930) even before Infantile Autism, a related disorder; first described by Kanner in 1943. However, the research and development in CDD

subsequently has not kept pace with that in autism, one of the most extensively studied disorder in child psychiatry today. CDD is currently classified under the broad rubric of Pervasive Developmental Disorders (PDD) which includes autism and autistic like conditions in the major diagnostic systems operational nowadays, namely, ICD-10 (World Health Organization [WHO], 1992) and DSM-IV (American Psychiatric Association [APA], 1994). CDD is considered to be an autistic-like condition; which although recognized clinically, remains in need of better understanding (Rutter & Schopler, 1992).

A second more broader term ,Pervasive Developmental Disorder or PDD was introduced in 1980 to refer to others who shared some of the features of autism but not all of those listed by Kanner.Pervasive Developmental Disorder, Not Otherwise Specified (PDD-NOS) is a 'subthreshold' condition in which some - but not all - features of autism or another explicitly identified Pervasive Developmental Disorder are identified. PDD-NOS is often incorrectly referred to as simply "PDD." The term PDD refers to the class of conditions to which autism belongs. PDD is NOT itself a diagnosis, while PDD-NOS IS a diagnosis. The term Pervasive Developmental Disorder - Not Otherwise Specified (PDD-NOS; also referred to as "atypical personality development," "atypical PDD," or "atypical autism") is included in DSM-IV to encompass cases where there is marked impairment of social interaction, communication, and/or stereotyped behavior patterns or interest, but when full features for autism or another explicitly defined PDD are not met.

It should be emphasized that this "subthreshold" category is thus defined implicitly, that is, no specific guidelines for diagnosis are provided. While deficits in peer relations and unusual sensitivities are typically noted, social skills are less

impaired than in classical autism. The lack of definition(s) for this relatively heterogeneous group of children presents problems for research on this condition. The limited available evidence suggests that children with PDD-NOS probably come to professional attention rather later than is the case with autistic children, and that intellectual deficits are less common.

Epidemiological surveys of autism started in the mid-1960s in England (Lotter, 1966, 1967) and have since then been conducted in many countries. All epidemiological surveys have focused on a psychiatric-diagnostic approach to autism that has relied over time on different sets of criteria. Epidemiological surveys of autism have been extremely useful in establishing a baseline on the prevalence of this severe developmental disorder and in providing a relatively unbiased picture of its main correlates.

On a review of twenty-three epidemiological surveys of autism published in the English language between 1966 and 1998, in which over 4 million subjects were surveyed and 1533 subjects with autism were identified. On average, medical conditions of potential causal significance were found in 6% of subjects with autism, with tuberous sclerosis having a consistently strong association with autism. Social class and status in the society did not appear to be associated with autism. In eight surveys, occurrence for other forms of pervasive developmental disorders was two to three times higher than the rate for autism. Based on recent surveys, atleast 18.7/10000 for all types of pervasive developmental disorders was derived, which highlights the needs in special services for a large group of children (Fambonne, 1999).

Knowledge about human communication is central to theory and clinical practice in the field of autism. Milestones in language and communication play major role at almost every point in development in understanding autism. Most parents of children with autism, first begin to be concerned that something is not quite right in their child's development because of early delays or regressions in the development of speech (Short & Schopler, 1988).

The communication problems of autism differs, and it depend upon the intellectual and social development of the individual. Some may be unable to speak, while others may exhibit rich vocabularies and are able to talk about topics of interest in great depth. Despite this variation, the majority of autistic individuals have little or no problem with articulation. Most have difficulty effectively using language. Many also have problems with word and sentence meaning, intonation, and rhythm. Children who have been diagnosed with autism seldom engage in effective communication. Approximately 50 percent of children with autism never develop speech, while others develop early forms of communication and social interaction.

Often parents recognize the absence of early communication in their young children with autism sometime during the second year, when the most of children of the same age begin to have established vocabularies of a lot of words (Short & Schopler, 1988). However, normal infants show communicative behaviours even from the first weeks and months of life, including recognizing their mothers' voice, synchronizing their patterns of eye gaze, movements, facial expressions of affect, as well as vocal turn taking (Fernald, 1992). Infants typically exhibit a variety of communicative behaviours by the end of their first year that, to a known observer, is not usually seen in autism. These nonverbal communication patterns have been found

to express the same intentions for which words will be used in the coming months, such as requesting objects, rejecting offered actions, calling attention to objects or events, and commenting on their appearance (Bates, 1976; Carpenter, Nagell, & Tomasello, 1998). These intents are expressed first with simple gestures, such as reaching to indicate a request or pushing away to indicate rejection, then by more complex gestures, such as pointing to request or shaking the head to mean "no," and then gradually accompanied by and, in some cases, replaced by vocalization and speech .

As children develop, they begin to attend to items and others for longer periods of time and pay more attention in detail. Through exploring their environment they begin to understand that their actions cause other events to occur. This is also the time that children begin to respond to simple commands, gestures and expressions. Most children with autism tend to be responders to communication, rather than initiators. The majority of their communicative behaviours are limited to requests or rejections for food/drink, toys or help. They rarely communicate about their own feelings or respond empathetically to the needs of others. At times of choice making, children with autism will communicate their negative desire by a head nod from side to side, but will seldom give gestures indicating affirmative choices.

There is a lot of variation in the timing as well as the patterns of acquisition of language among children with autism. A minority of children, usually diagnosed with autism do not show any significant delays in the onset of language milestones. In contrast, most children with autism begin to speak late and develop speech at a significantly slower rate than others (Le Couteur, Bailey, Rutter, & Gottesman, 1989). Because autism is not usually diagnosed until age 3 or 4, there is relatively little

information about language in very young children with autism. Various retrospective studies using parent report and videotapes collected during infancy and the toddler years suggest that after the second year of life, the communication of most children with autism vary from other children (Dahlgren & Gillberg, 1989).

Shifting diagnostic criteria are certainly an important factor contributing to the increase in diagnoses of autism, on both 'ends' of the spectrum. Many children who have mental retardation in addition to autism may in the past have only received a diagnosis of mental retardation. For many years in India, physicians may not have seen an advantage in providing a differential diagnosis when the treatment would be the same, and when awareness of autism in India was so low. On the other end of the spectrum, many children with good communication skills may have been diagnosed with schizophrenia, ADHD, or completely missed altogether and simply considered to be 'unusual' children. Now, more and more of these children receive autism as a primary diagnosis. Environmental factors have been cited as a cause of the rise in autism. While environmental factors may not 'cause' autism, they could be acting as a trigger, and just as there are increasing numbers of other conditions with an environmental component, so might it be with autism.

There is no doubt that greater awareness of autism in India among medical professionals has played a role in the increase of children diagnosed as autistic. Following an awareness campaign among paediatricians conducted by Action for Autism and the Rajiv Gandhi Foundation between 1999-2001, there appeared to be a sharp increase in the Autistic Spectrum Disorders, number of diagnosis that followed and a drop in the mean age for diagnosis. In addition, the increased attention to the disorder by the popular media has reached both parents and professionals, Even with this increased awareness, the number of individuals receiving a diagnosis in India is

still far behind that in higher income countries. And, within India, there are fewer children being diagnosed in rural areas as compared to those in urban areas of India.

The causes of regression in autism are not yet clear. Biological correlates of regression might include epilepsy (Hrdlicka, 2004) and epileptiform EEGs (Tuchman, Rapin 1997); however, other studies have countered the findings of increased seizure activity associated with regression in autism (Baird, Robinson, Boyd, Canitano 2005; Hansen, Luyster 2005). Rogers (2004) also reported no differences between regression and non-regression groups on measures of gross motor development, which are identified as reasonable proxies for general CNS dysfunction. Possible associations between autistic regression and immunizations have not been supported within the literature. Stagnation in skill acquisition, i.e., developmental plateau or stasis, is another developmental trajectory observed in a minority of children with autism (Rogers 2004).

Developmental plateau is similar to regression in that early development is marked by typical milestone achievement; however, instead of losing skills as with regression, individuals demonstrating plateau show a pause or stagnation in development (Ozonoff, 2004). Apparently less prevalent than regression, stagnation was reported to occur in 9.3% of a clinical sample of children with autism (Siperstein and Volkmar 2005). In contrast to regression, few findings exist with respect to correlates of developmental plateau. Apart from its consistent report in the literature, language regression is of particular interest as regressive autism may represent a distinct phenotype. The same case may be seen regarding the presence of language plateau as well. If early language history were to signal phenotypic expression within the ASD spectrum, this would be clinically valuable by assisting in deciding and

planning intervention, better predicting outcomes, and anticipating co-morbid problems. Research documenting clinical characteristics, developmental outcomes, and co-morbid problems associated with regression, however, has revealed mixed findings.

Aim of the current project is to find out the association of genetic markers with ASDs in an Indian Population. The genes selected for present study is Engrailed2 and CNTANAP2.Intronic SNPs are selected for family based and case control association study. The studies with selected genes with ASDs are conducting first time in South Indian population.

Table1. Shows the selected markers for the genetic analysis with ASD.

Gene	SNPs
EN2	rs1861972
	rs1861973
	rs73163344
CNTNAP2	rs1861972
	rs1861973
	rs73163344
	rs2710108
	rs2710105
	rs2710102
	rs2538994
	rs7806512
	rs7806519
	rs2710098

CHAPTER - II

REVIEW OF LITERATURE

Autism Spectrum Disorders (ASDs) are polygenic disorder effects children at the age of 1-2, Characterized by stereotypic behaviour and interests, lack of communication and socialization. They are highly complex and show multi factorial inheritance. Folstein and Rutter (1977) noted that a person with autism rarely marry and give birth and 2% of sibs are affected .Twenty one sex matched twin pairs, eleven monozygotic (MZ) and ten dizygotic (DZ), are selected for a study in which at least one should have infantile autism, research group found that thirty six percentage similarity among the Monozygotic twins and there were no similarity in Dizygotic twins. Researchers identified for monozygotic pairs and ten percentages for dizygotic twins. In a study of Ritvo et al. (1985) found that 40 pairs of twins are concordance of 23.5 % in the group of dizygotic twins but concordance of 95.7% in monozygotic twins. Kinship for autistic children is studied by Jorde et al (1990) and found that average kinship coefficient with ASD children and normals showed strong affinity for autism to group families and the quick fall off in risk to relatives and sib risk of 4.5% with the help of Ulah Geneological database. In a familial study of Bolton et al (1994) found 2.9% familial risk and 2.9% of PDD in sibs it is seventy times higher than the risk in general population.

Twinning process is one of the factor in the development of Autism is suggested by Hallmayer et al. (2002). In a study of Pickles et al. (1995) suggested that autism has many locus and mode of inheritance involving three loci. The study is carried out in family with autism and twins. Linkage analysis done by International Molecular Genetic Study of Autism Consortium (1998) found that chromosome 2q,

7q, 6p, and 19p are overlapping with Phillippe et al. (1999) research study. Later Smalley. (1997) updated linkage study in autism. Review on chromosomal aberration, linkage study and candidate gene study was done by Lamb et al (2000). A study on ASD in 1,181 families with two effected, done with linkage and CNV revealed that 11p13-p12 and neurexins are important candidate loci. This study done by using the instrument Affymetrix 10K SNP array (Autism Genome Project Consortium, 2007).

Very complex genetic structure of ASD requires very large number of sample size to overcome the heterogeneity. In a an another genome wide study Wang et al. (2009) found that six SNPs on chromosome 5p14.1 and showed cadherine 9 (CDH9) and CDH10 are strong candidate genes in ASD with 780 families and cohort 1,204 subjects and controls from the ancestry of Europe population.Genes, CDH 9 and CDH 10 on chromosome 5p14.1 are coding neuronal cell adhesion molecules and SNP rs4307059 showed strong association with ASD. Ma et al. (2009) replicated and found chromosome 5p14.1 linkage to ASDs and rs10038113 is the significant marker, implicates strong association of cell adhesion molecule in the pathogenesis of ASDs. The significant loci of the ASDs are shown in the Table 2.

Table.2 Important loci with ASD are shown in the OMIM database (NCBI).

Susceptibility loci	Mapped to Chromosome
AUTS3	13q14
AUTS4	15q11
AUTS5	2q
AUTS6	17q11
AUTS7	17q21
AUTS8	3q25-q27
AUTS9	7q31-7q36
AUTS11	1q41
AUTS12	21p13-q11
AUTS13	12q14
AUTS14	16p11.2
AUTS15	7q35-q36
AUTS16	3q24
AUTS17	11q13

Roohi et al (2009) found CNV with gene CNTN4-chromosome 3 in three of 92 patient with ASDs used by CGH analysis. Another CNV study by Glesser et al (2009) showed susceptibility to ASDs and genotyped 550000 markers on group of 859 ASD cases and 1,409 controls. Addition of genes, NLGN1 and ASTN2 also showed CNVs in ASDs and concluded that neuronal cell adhesion molecule shows that two genes networks expressed in CNS may cause to genetic association with ASDs. Liu et al (2001) suggests that chromosome 5,X,and 19 are related to ASD in a linkage study with 335 marker(microsatellite) in 110 families with autism .Further reported that chromosome 5 ,8 and 19 also shows significant linkage with autism.

GABRA4 gene reported as a candidate gene and shows significant association with autism on chromosome 4p12.This study conducted by Vincent et al.(2006) but the study of Maestrini et al.(1999) reported that GABRA4 doesn't shows any association or linkage to ASDs. They have studied in 94 families with 174 autism subjects. SHANK3 gene on chromosome 22q13 and deletion ranges 277kb to 4.36Mb in the gene found in ASDs with three of 400 unrelated patients.NRXN1 shows chromosomal abnormality (2q16.3) in two autism spectrum disorders un related two subjects studied by Kim et al (2008).

Earlier candidate gene analysis using family based approach advocated deletions and polymorphisms in the gene SHANK3 (alternative name ProSAP2) as one of the potential causes of ASD's in the studied subjects. The SHANK3 protein plays a role in the functioning of synapses, which are the connections between nerve cells (neurons) where cell-to-cell communication occurs. Within synapses, the SHANK3 protein acts as a scaffold that connects neurons, ensuring that the signals sent by one neuron are received by another. The SHANK3 protein is also involved in

the formation and maturation of dendritic spines, are essential for the transmission of nerve impulses. This gene belongs to the SHANK gene family and is present on the chromosome 22 and its cytogenetic location is chr.22q13.3. Christelle M. Durand et al investigation suggests that abnormal gene dosage of SHANK3 is associated with severe cognitive deficits, including language and speech disorder and ASD's

Marshall et al. (2008) found that CNVs in of 427 families with ASDs. Translocation, deletion, duplication and insertion are seen in subjects but not in controls. Chromosome location of ASD was also common to mental retardation loci. Spanish patients with ASD CNV's present in12 of 96 patients and functional alteration in the genes involved in the neuronal networks was found by Cusco et al.(2009).Researchers concluded that de novo CNV is highly associated with autism

Candidate genomic regions were validated by high resolution CGH, paternity testing, cytogenetic, fluoresce in situ hybridization, and microsatellite genotyping. Confirmed de novo CNVs were significantly associated with autism (p Value=0,0005).They concluded findings established that de novo germ line mutation as a more significant risk factor for ASDs than previously recognized.

Bottini et al. (2001) worked on the gene reduced adenosine deaminase (ADA) activity in Italian population and found that low activity of the allele ADA in the total samples of autism subjects and controls .Authors concluded that the activity of ADA may cause for the development of Autism. Cheslack Postava et al.(2007) analysed the beta2- adrenergic receptor and two variants of the gene coding for ADRB2 receptor shows increase activity in AGRE population. In 331 independent Autism subjects and parents, genotyped the marker rs 1042713 and rs1042714 ,they are activity related polymorphisms. Authors found that Glu27 and ADRB2 alleles shows significant

association with Autism. To find out the association of AHI1 gene with autism, Alvarez et al. (2008) performed three stages of analysis and re-sequenced gene AHI1in AGRE families. In s third stage of analysis, they found that the AHI1 gene is associated with ASDs; previous study the gene was associated with schizophrenia.

MCC and APC genes in the chromosome 5q22.1 done with cytogenetic mapping Barber et al (1994) found that the genes are closely associated with autism. Androgen receptor gene (AR) susceptibility for autism is proved in a study of 267 subjects with ASDs and 617 controls by Henningsson et al. (2009). They analysed polymorphisms in exon 1 of AR gene specifically the CAG repeats, GGN repeats and the SNP rs6152, additionally case control study with 118 and 32, CAG, rs6152 markers are significant association with ASDs. Chakrabarti et al. (2009) reported on candidate genes with three types of functional group ,neural connectivity ,sex steroid synthesis and social emotional responsibility. Study with high functioning sub group, genes ESR2 and CYP11B1 are associated with sex steroid group with ASD, gene HOXA1, NTRK1, and NLGN4X are shown significant association with neural connectivity group. Finally the group shown social emotional characteristics are associated with the genes MAOB, AVPR1B, and WFS1.another group study with asperger syndrome , fourteen genes are associated in this CYP17A1 and CYP19A1 are related to sex steroid group and OXT gene reported as association with socio emotional behaviour.

Structural variation of chromosome includes deletion duplication translocation and insertion Marshall et al. (2008) found 277CNV's 40% of ASD families but not in control population. Karyotying analysis shown some more chromosomal abnormalities .Most of the variants are inherited.de novo changes in the chromosomes

in 27 cases,7% de novo copy number variation in the population and 13 CNV overlapping loci also found in this study. Authors identified the genes SHANK3, NLGN4, NRXN1 and new loci at DPP6, DPP10, PCDH9, ANKRD11, DPYD, and PTCHD1.

Yirmiya et al.(2006) analysed two promoter and one intronic microsatellite markers (AVPR1a-RS1,RS2 and AVR) with Autism spectrum disorders and found association with these markers.AVPR1a has important role in shaping of socialization skills in the humans and other lower vertebrates .So this gene is very important in the development of the autism features. The gene CACNA1G is found the novel candidate gene by storm et al (2010).the gene is located in 17q211-q21,shown significant linkage in previous studies. Kilpinen et al (2008) suggests the gene DISC1 on the 1q42 loci are associated autism and Asperger syndrome. They analysed Intragenic single nucleotide polymorphism D1S2709 (SNP) with autism and rs1322784 with asperger syndrome in the gene DISC1 for neuronal development suggests important role in the development in ASD. Neurexin superfamily member CNTNAP2 found as a strong candidate gene, Arking et al( )done two stages of study one is genome wide linkage and another is family based association study and found the common genetic polymorphisms which are associated with autism.FOXP2 is the gene located in 7q region is a transcription factor key regulator for embryogenesis. Gong et al (2004) found in a family study that 3 SNPs are associated with 181 chinese Han trios using the TDT analysis with autism.

Cook EH et al (1998) investigated the role of 15q11-13 region and the role of GABRB3 and suggested that further investigation is needed to confirm the association. Autism brain study conducted by Junaid et al (2004) found that increase

in polarity of enzyme Glo1 (glyoxalase) express in the brain and they have sequenced the gene and found a novel SNP (C419A).The SNP effects on protein sequence(Ala111Glu).Authors conducted a study on 71 autistic subjects and 49 controls SNP A419,revealed significantly higher frequency .Biochemical studies shows Glyoxalase enzyme activity is decreased by 38%.Gene HOXA1 and HOXB1have critical role in brain development. A study conducted by Ingram et al (2000) and found that the Gene HOXA1 plays role in autism but no association with gene HOXB1.

The SNP rs1150220 (HTR3A) effects on autism is revealed by Anderson et al (2009) .The gene is located on chromosome 11 and involved in the serotonin pathway. They have screened 403 autism families and genotyped 45 SNPs in 10 Serotonin pathways. Weiss et al (2006) suggests the role of ITGB3 gene with blood serotonin level and autism. Serotonin pathway is the important pathway because the drugs acting on serotonin pathway is very effective treatment in the case of autistic behaviour. Laumonnier et al (2006) mapped 9q23/10q22 by using FISH and cloned by BAC. Defect of calcium channel is revealed and suggested that BK (Ca) shows very keen role in pathology in the development of autism and other mental disabilities. Authors found that KCNMA1 gene, a synaptic regulator of neuronal excitability, is physically disrupted.

Rett's syndrome effects only in females and it is a neuro-devolopmental disorder comes under Autism Spectrum Disorders. Amir et al (1999) found that mutation in MECP2 gene is the main cause of this disorder. Three de novo missense mutations found in 5 in 21 patients. The region encoding the highly conserved methyl-binding domain (MBD). The mutations disrupt the TRD (Transcription

Regression Domain). Ozgen et al (2009) identified copy number variation is with microcephalin1 (MCPH1) gene at chromosome 8p23 has keen role in the development of ASD. Marui et al. (2004) investigated NF1 gene with three microsatellite polymorphisms in Japanese population with autism and normals. Six repeated alleles are not affected in cases and controls .The authors suggested for further investigation for the role NF1 in ASD. Glessner et al (2009) done a whole genome CNVs in 859 cases and 1409 controls NRXN1 and CNTN4, NLGN1 and ASTN2, were enriched with CNVs in ASD cases except in control samples. CNVs observed with ubiquitin pathways (UBE3A, PARK2, RFWD2 and FBXO40) but not in controls.

Jacob et al. (2005) have done genome wide linkage analysis with AGRE families and found that 19p13 and 17q11.2 are the susceptible loci for Autism spectrum disorders. Anne Philippi et al. (2007) studied on paired like homeodomain transcription factor (PITX1) gene on chromosome 5q31shows association with rs11959298 and rs6596189 with Autism. Fanfan et al. (2011) genotyped 7 tag-SNPs in 367 trios and done with FBAT revealed that rs4366301, rs11585959 and rs6668845 are significantly associated with autism. Authors concluded that the DISC1 gene has keen role in the development in ASDs.

A study on GABA receptors and associated markers, Dale et al (2012) revealed the novel deletions in candidate genes with ASDs. Deletions observed in the genes JAKMIP1, NRXN1, Neuroligin4Y, OXTR, and ABAT. Based on previous reported studies on OXTR marker with ASDs Daniel et al (2010) worked in 1238 pedigree with 2333 ASDs were included. The authors reported the association with autism and suggested that functional SNPs may influence in the development of ASD

in sub set of families. Katherine et al.(2011) have done a study on 42 genes related to oxidative stress total 308 SNPs in glutathione, anti oxidant in the brain .The result was 3 SNPs joint effect may contribute and comparatively less association with CTH, Gamma glutamylcysteine synthetase, alcohol dehydrogenase 5, and GLRX3. Authors suggest replication of this study is needed to confirm the association of these markers with autism.

Angelica et al (2010) reported a study on social and non social autistic children and 1st stage of the study screened autosomal SNPs from low scoring boys from general population In 2^{nd} stage of analysis rs11894053 is related with autism in social autistic children. Richard Anney et al. (2010) reported in a genome wide association analysis that the marker rs4141463, MACROD2 shows good signal for association. Research team concluded study of association with different genes KIAA0564, PLD5, POU6F2, ST8SIA2 and TAF1C. Ana I et al. (2008) found the significant association with the gene AH1 to the autism. This gene is the common variant in Joubert syndrome, recessive inherited disorder. Three stages of analysis were carried out with AGRE families.AH1 Gene has important role in cognition and behaviour. Jacob et al (2007) supported their study on OXTR gene with autism spectrum disorders. They have worked on SNPs, rs2254298, rs53579 in 57 trios in Caucasian population.

Table.3 Description of important Candidate Genes, function and chromosome location with Autism Spectrum Disorders. (NCBI)

Strong Candidate Genes	Function	Chromosome Location
ITGA4	Cell adhesion molecules	2q31.3
SLC25A12	The encoded protein has solute:cation symporter activity.	2q24
CNTN4	Neuronal cell adhesion molecule	3p26-p25
OXTR	The encoded protien has oxytocin receptor activity.	3p25
GABRA4	The encoded protein is a subunit of GABA-A receptor. Neurotransmission is predo minantly mediated by a gated chloride channel activity intrinsic to the receptor	4p12
GABRB1	The encoded protein is a subunit of GABA-A receptor. Neurotransmission is predo minantly mediated by a gated chloride channel activity intrinsic to the receptor	4p12
ADRB2	This gene encodes beta-2-adrenergic receptor which is a member of the G protein- coupled receptor superfamily.	5q31-q32
C4B	The encoded protein is part of the classical pathway of the complement system	6p21.3
GLO1	The encoded protein mediates catalysis and formation of S-lactoyl-glutathione fr om methylglyoxal and reduced glutatione.	6p21.3-p21.1
CNTNAP2	This gene encodes a member of the neurexin family which functions in the vertebr ate nervous system as cell adhesion molecules and receptors. This protein is loc alized at the juxtaparanodes of myelinated axons and associated with potassium c hannels.	7q35-q36
EN2	The encoded protein is a transcription factor.	7q36
FOXP2	The encoded protein is a Zn binding transcription factor.	7q31
MET	The encoded protein is the hepatocyte growth factor receptor with tyrosine-kinas e activity.	7q31
NRCAM	The encoded protein is a cell-cell adhesion molecule that participates in neuro nal outgrowth and guidance processes.	7q31.1-q31.2
RELN	The encoded protein is a large extracellular matrix protein that plays a pivotal role in the neuronal cell migration and in the development of neural connection	7q22

HRAS	The encoded protein binds GTP and has GTPase activity.	11p15.5
AVPR1A	The encoded protein is a receptor for arginine vasopressin. It belongs to the G -protein coupled receptor family.	12q14-q15
GABRB3	The encoded protein is a subunit of GABA-A receptor. Neurotransmission is predo minantly mediated by a gated chloride channel activity intrinsic to the receptor .	15q11.2-q12
A2BP1	sequence-specific RNA recognition, splicing regulators	16p13.3
PRKCB1	The encoded protein is one of the PKC family members.	16p11.2
SLC6A4	The encoded protein has serotonin transporter activity, serotonin:sodium symport er activity.	17q11.1-q12
ADA	The encoded protein catalyzes the hydrolysis of adenosine to inosine.	20q12-q13.11
SHANK3	The gene encodes a synaptic adaptor protein.	22q13.3
MECP2	The encoded protein has methylation-dependent transcriptional repressor activity . It is also involved in regulation of RNA splicing	Xq28
NLGN3	Neuroligins are cell-adhesion molecules at the postsynaptic side of the synapse .	Xq13.1
NLGN4X	Neuroligins are cell-adhesion molecules at the postsynaptic side of the synapse .	Xp22.32-p22.31

In the first study of the genetics of autism in India, scientists in Calcutta have found that some fathers may transfer a version of a gene that makes their children susceptible to autism. The researchers at the Manovikas Biomedical Research and Diagnostic Centre and other city institutions have found what they describe as a "possible paternal effect" that may underlie susceptibility to autism - a brain disorder marked by unusual behaviour and lack of communication abilities. Dutta et al (2008) analyzed sequences of a gene called reelin in a group of 73 autistic children and 80 children with no neurological disorders, as well as their parents who volunteered for the study.

Guhathakurta et al (2006) conducted a population-based study on 5-HTTLPR with 358 individuals, which included 79 autistic probands, 136 parents, and 143 controls from two subpopulations of east and northeast regions of India. This study shows no association or linkage of 5-HTTLPR with autism in the Indian population.

Dutta et al (2007) performed genetic analysis of three markers of GluR6 (SNP1: rs2227281, SNP2: rs2227283, SNP3: rs2235076) for possible association with autism through population, and family-based (TDT and HHRR) approaches. DSM-IV criteria and CARS/ADI-R have been utilized for diagnosis. Genotyping analysis for the SNPs has been carried out in 101 probands with autism spectrum disorder, 180 parents and 152 controls from different regions of India. Since the minor allele frequency of SNP3 was too low, the association studies have been carried out only for SNP1 and SNP2. Their results suggest that these markers of GluR6 are unlikely to be associated with autism in the Indian population

Gene Engrailed 2

The gene homeobox transcription factor (EN2) is located in the chromosome 7q36.3 is a important linked loci with ASDs. Poole et al. (1989) in Drosophilla, the engrailed proteins have very significant role in the development in segmentation. The protein sequence of EN2 mouse and human is 100% similar. Logan et al (1992) and Benayed et al (2005) explained the structure of EN2 has 8.1 kb bases and has 2 exons and 3.3 kb intronic sequences.

Schematic representation of human EN2 gene shown in *figure 1*. Important intronic SNPs are represented in order. Those SNPs are selected for present study.

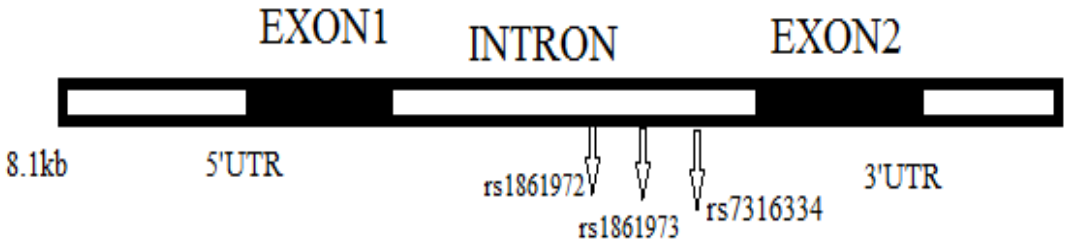


Figure 1. Schematic representation of human EN2 gene

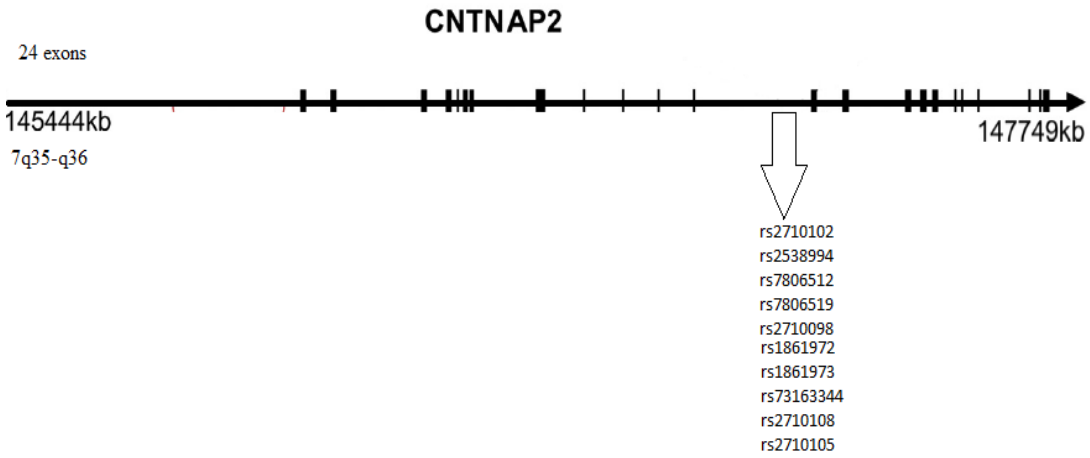
Gultknecht et al. (2001) suggested that in a genome wide linkage study chromosome 7 plays important role in the etiology of ASDs. Engrailed 2 is one of the gene included in that region. It effects the neurological development in humans. Mutant mice show similar characteristics of cerebellum with Autistic patients. Benayed et al (2005) have done a work on EN2 as a candidate gene for ASD. Subjects selected from AGRE (167families,IInd data set-222 families) and NIMH .Two effected siblings with autism, Asperger or PDD-NOS and their parents were selected, genotyped the SNPs rs1861972 and rs1861973 with samples from AGREII and NIMH. Method was pyrosequencing assay.14 SNPs were genotyped with AGREI with RFLP method. Researchers found highly significant association with the markers rs1861972 and rs1861973 with AGREII and NIMH families. This supports strongly the markers are associated with Autism phenotypes. Gharani et al. (2004) also support the gene EN2, intronic markers shows evidence for autism development. In 518 families were studied and 40% of the ASD samples showed association. Sen B et al (2010) revealed the family based association of EN2 with Indian population. Total 128 families with ASD children in that 105 trios and 23 duos were selected for the analysis. Moreover 98 families includes 83 males and 15 females from west Bengal

and 30 families with ASDs from different places from India male female ratio identified as 4:77:1.DNA extracted in slandered phenol chloroform method and RFLP method used to genotype the markers. Analysis done with the markers rs34808376 (promoter), rs6150410 (promoter),rs3735653 (exon1),rs1861972 (Intron) rs1861972 (Intron).TDT analysis observed transmission of C allele of rs1861973 from parents to the effected child. Gender based also the transmission of C alleles were observed in females in the study group, and these alleles maternally over transmitted. Del-T from rs34808376-rs1861973 shows non transmission has role in protective effect on ASD risk. These results show the genetic association of EN2 with Autism Spectrum Disorders. Kuemerle et al (2007) reported that some molecular mechanism was affected on EN1 knock out mouse and also found that anterior change in the loci of the amygdale in the central cortex. Gender difference also associated with ASDs.

Gene CNTNAP2

The gene CNTNAP2 is located in the chromosome 7q35-q36 encodes a protein calls Caspr2 (Contactin associated protein2) neurexin superfamily expressed in human brain(Frontal and Temporal lobs).This area of the brain functions for cortical function includes language processing. This protein is important for localization of voltage-activated potassium channels in the juxtaparanodal region of axons which may function to stabilize conduction and help to maintain the intermodal resting potential.

Figure 2. shows the schematic representation of the gene CNTNAP2 and SNPs selected for the study.



Several studies say that CNTNAP2 gene is more susceptible to ASDs. Vernes et al (2008) studied that FOXP2 binds in the intronic regulatory sequence and gene is directly regulate the expression of CNTNAP2 .Murrani et al (2012) mentioned that there is common frontal lobal connectivity with both Autistic traits and patients with specific language impairment. In a clinical study conducted by Murrani et al (2012)found that 450kb deletions in the gene CNTNAP2, maternally inherited with patients with delay in speech .Alarcon et al(2008) suggested the linkage association of CNTNAP2 gene with autism. Samples were recruited from AGRE. In the first stage 172 trios selected by age at first word (WORD) and in second stage independent 304 parent child with WORD. Stage 1 samples genotyped with PERLEGEN and in second stage the genotyping was done by using ILLUMINA platform. The result was in two stages of analysis genetic variation in CNTNAP2 supports the language onset male with ASDs and rs2710102 variation was associated with age at first word patients. Stein et al (2011) published that SNPs rs2710102 and rs6944808 with

mutism (Selective Mutism) and social anxiety disorder (SAD) related CNTNAP2 markers were genotyped in 99 families and 106 children and found a good significance.

Ma et al. (2009) identified in a genome wide association study with 1390 individuals with autism and validation in an additional cohort of 2390 samples from 457 families did not a significant association between autism and in Alarcon et al. (2008) study rs 270102 which was the tagging SNP . Arking et al. (2008) in a family study identified a common variant in the CNTANAP2 gene SNP rs7794745, that was associated with increased risk for autism. This SNP is located in intron2 of the gene. In an another study of Bakkaloglu et al (2008) four children with autism from three unrelated families identified an ile 869 to thr substitution in the CNTNAP2 protein. The mutation at a conserved residue occurred in the third laminine G domain and was predicted to be deleterious. Inheritance of variants from apparently affected parents in each family .negative result found among 4010 controls.

CHAPTER III

METHOD

Sampling:

A total 250 samples (ASD cases 85,62 father,85 mother,18 siblings)were collected from ASD unit, All India Institute of Speech and Hearing,Mysore.91 healthy controls samples were collected from Mysore and Kerala same not affected by Autism Spectrum Disorder. Informed consent was obtained from all participants in the study that had been approved by the Research Ethics Committee of All India Institutes of Speech and Hearing. A structured questionnaire was administered to collect socio-demographic information like ethnicity, age, sex, house status, TV exposure, economic status, migration, pre natal information etc.

Table.4 Distribution of overall samples and collected samples shown below.

ASD Subjects	Father	Mother	Siblings	Healthy Controls	Total
87	62	85	18	91	343

Differential Diagnosis Checklist for ASD

Selected Autism Spectrum Disorders (ASDs) cases were diagnosed on the basis of DDC-ASD developed by Dr. Shyamala KC (2007), All India Institute of Speech and Hearing. This classification based on DSM IV in Indian context. CARS and M-chat also used for the diagnosis of ASD.

Table.5 Total ASD Children based on classification (DDC-ASD)

ASD-Types	Autism	PDD-NOS	Rett's Syndrome	Autism with MR or CP/other complications
No. of Subjects	65	19	1	2

Data collection

From each family with ASD detail like name, address, gender, date of birth, cast/community, religion, nationality, occupation, economic status, TV exposure, House Status, Consanguinity, Family migration and prenatal history were collected with the use of a questionnaire. The Informed Consent forms were filled and signed by the parents and explained about the project before selecting for sample collection. AIISH Ethical Committee was approved the project before data collection.

Table.6.Total Distribution of Samples was selected based on Area.

		Group		
		Cases	Controls	Total
	Karnataka	30	29	59
	Kerala	49	51	100
	T.N	3	2	5
	A.P	5	1	6
	North India	8	9	17
Total		95	92	187

Blood sample collection

Approximately 3-6 ml of peripheral blood sample was collected in BD K_2 vacutainer® (BD, NJ, USA) containing EDTA as anticoagµlant. Blood samples were stored at 4^0C to avoid haemolysis and cellular damage. Samples were carried to the lab within 3 hours of collection to ensure good results. Blood samples were transferred to

labelled, sterile polypropylene centrifuge tubes. DNA was extracted from the whole blood as follows.

DNA Extraction

In the present study the DNA was extracted by Phenol chloroform method

Principle:

The cells are lysed in an alkaline lysis buffer. Incubation at 56°C in the presence of proteinase K and SDS is used to partially digest cellular proteins and loosen the association between proteins and chromosomal DNA and to degrade cellular RNA. The cell lysates is then treated with buffer-saturated phenol chloroform. The DNA remains in the aqueous phase while the cellular proteins are extracted into the organic phase. The aqueous phase containing DNA is then treated with isopropanol or ethanol to precipitate the DNA. The DNA is washed with ethanol and then dissolved in water of TE buffer (10mM Tris-HCl, 1mM EDTA, pH 8.0). DNA extracted this way is generally high molecular weight and is double stranded and is therefore suitable for PCR amplification.

Details of the buffers and reagents used for DNA extraction are given below.

1. RBC Lysis Buffer

Sucrose	109.54g
1M Magnesium Chloride	5.0ml
1M Tris-Hcl	10.0ml

Made up to 900ml with double distilled water, autoclave and add 10ml Triton X and made up to 1000ml. pH 8.0

2. Digestion Buffer

1M Tris-Hcl (pH 8.0)	40.0ml
1M Sodium Chloride	15.0ml
0.5M EDTA (Na salt)	10.0ml

3. Tris-EDTA Buffer

1M Tris-HCl (pH 8.0)	2.0ml
0.5M EDTA	4.0ml

Made upto 200ml

Extraction process:

1. To the blood sample, added 2 volume of reagent A in a polypropylene tube. Mix gently till the solution becomes clear.

2. Centrifuged at 2500 for 10 minutes to obtain a pellet, free from RBCs. The supernatant containing lysed RBCs is discarded carefully. Steps 1 and 2 are repeated 3-4 times, till the pellet is free from RBCs.

3. Disturbed the pellet thoroughly. Then add half the volume (as that of blood sample) of reagent B, proteinase K to a concentration of 100µg/ml and SDS (Sodium Dodecyl Sulphate) to make 1% concentration in the final volume. Mixed thoroughly and gently by inverting for 3-4 minutes till the solution becomes viscous. Incubated it for 3-4 hours at 56^{o}C or over night at 37^{o}C.

4. After complete digestion the solution had become clear then added reagent C ($1/4^{th}$ volume of reagent B) and mixed gently for 3-4 minutes.

5. Added equal volumes (as that of reagent B+C) of phenol (Tris saturated) + chloroform + isoamyl alcohol (25:24:1). Mixed well and centrifuged at 4000 rpm for 15mins to separate into 3 layers viz. aqueous layer, protein layer and solvent layer.

6. Transferred the aqueous layer carefully into another centrifuge

7. Added equal volumes of chloroform+ isoamyl alcohol (24:1) to the supernatant and mixed gently for 3-4 minutes and centrifuged at 4000 rpm for 15 minutes.

8. Transferred the aqueous phase to a fresh tube.

9. Added 2 volumes of chilled absolute alcohol and mixed gently to precipitate the DNA.

10. Spooled out the DNA lump in a fresh 1.5 ml tube and decanted alcohol.

11. Washed the DNA in a fresh 1.5 ml tube and decanted alcohol.

12. Dried the pellet properly and ensured that whole alcohol is dried off.

13. Dissolved the pellet properly and ensureed that whole alcohol is dried off.

14. Store the DNA samples at -20^{o}C to -80 ^{o}C as per the period of storage.

After DNA was completely dissolved in the TE buffer, its quantity and quality was checked by both spectrophotomery and gel electrophoresis.

Determination of DNA Concentration by Spectrophotometry

Prior to any analysis, DNA should be quantified and checked for purity and integrity. Based on its structure, DNA absorbs light in the ultraviolet range, specifically at a wavelength of 260nm. A value of 1 at OD_{260} is equal to 50ng/µl double stranded DNA, therefore to calculate the concentration of DNA.

The following formula was used:

Concentration DNA = $260nm_{abs}$ x 50ng/µl

Procedure

2µl DNA sample was diluted to 200 µl with Double Distilled water (Dilution 1:100). Spectrophotometer was set to auto zero with the Double Distilled water. Optical Density (OD) of the diluted DNA aliquot was measured at 260 nm and 280nm using quartz crystal cuvette.

Quality Assessment

A ratio of OD values at 260nm and 280nm indicates the purity of the extracted DNA sample. If the ratio is within range of 1.6 to 2.0 then DNA sample is considered as clear and free from contaminants like residual protein and mRNA. An OD ratio less than 1.6 indicate the residual proteins or phenol contamination, whereas ratio of more than 2.0 indicates residual RNA contamination.

Quantity Assessment

DNA quantity was estimated as the OD value at 260nm of extracted sample is 1.00 then the concentration of the DNA is 50ug/ml.

Therefore DNA concentration = OD at 260nm x 50 x Dilution factor

DNA quantitation and electrophoresis

Electrophoretic analysis of DNA using agarose gels can confirm DNA integrity. Typically, intact genomic DNA will be up to 40KB in size, depending upon the species. 1% agarose gel was prepared by adding required quantity of agarose to 1X Tris-Acetate-EDTA (TAE) buffer and mixed will. The mixture was heated in microwave oven until it became clear and care was taken to avoid over boiling and evaporation. The mixture was cooled to ~50^0 C and ehetdium bromide is added to make a final concentration of 0.001ug/ml. The entire mixture was poured in to a tray in which combs are fixed to make wells in the gel. It was cooled to form uniform gel. After gel formation, the tray was placed in buffer tank containing 1X TAE buffer for submerged gel electrophoresis and combs were removed with care to avoid rupture of wells. 1µl of each DNA sample was mixed with 1µl of loading dye and the mixture was loaded into the wells. Gel was subjected to electrophoresis at 90V for 30 minutes and visualised using gel documentation system (Syngene, UK).

Chemicals and buffers used for gel electrophoresis

1. Tris-Acetate-EDTA (TAE) buffer 20X

Tris	48.4g
Acetic Acid	11.402ml
0.5M EDTA	20ml

In 1lt of distilled water. pH 8.0, 50ml of 20X buffer was made up to 1000ml to obtain 1X TAE buffer.

2. Loading Dye (Stock)

Bromo phenom blue	25mg
Xylene cynole	25mg

Dissolved in 10 ml of distilled water.

5ml of 40% sucrose solution was added to 1ml of Lodong dye stock to make working dilution.

Preparation of DNA working Dilutions

100µl of DNA working dilutions were prepared at a concentration of 50ng/µl by dissolving required amount of stock DNA sample in milliq water (Millipore life sciences, Billerica, MA, USA). After preparation of working dilutions the uniformity of the samples were checked by performing electrophoresis on a 1% agarose gel. Samples were stored at 4°C.

Selection of sequence

As early mentioned the present study is checking the polymorphism in the EN2 and CNTNAP2. All SNPs (Single Nucleotide Polymorphisms) are in intronic region. The regions were downloaded from Ensemble genome browser. For primer designing the sequence was selected in such way the sequence covers both the side splicing site of intron regions. The sequence covers minimum of 400 bp to maximum of 900 bp regions.

Table.7. Number of Genes and List of SNPs genotyped in selected population.

Gene	SNPs	ASD -sub	Father	Mother	Sib	Controls
EN2	rs1861972	85	62	85	18	91
	rs1861973	85	62	85	18	91
	rs73163344	85	62	85	18	91
CNTNAP2	rs1861972	85	60	87	17	92
	rs1861973	85	60	87	17	92
	rs73163344	85	60	87	17	92
	rs2710108	85	60	87	17	92
	rs2710105	85	60	87	17	92
	rs2710102	85	60	87	17	92
	rs2538994	85	60	87	17	92
	rs7806512	85	60	87	17	92
	rs7806519	85	60	87	17	92
	rs2710098	85	60	87	17	92

Primer designing

The main limitation of PCR technique is one has to provide short pieces of single-stranded DNA (primers) that are complementary to a part of target sequence. With the use of human genome sequence available we can now design primers to any region of interest in the human genome. The most critical step in PCR experiment is designing oligonucleotide primers. As poor primers could result in little or even no PCR product, alternatively they could amplify unwanted DNA fragments. Either will affect the downstream analysis.

Assessing specificity of the primers

The specificity of primer is very important in view of amplifying a specific gene region. If the primer is not specific for the region of interest, problems in obtaining reliable sequence data arises. To avoid this all the primers will be checked for their specificity by performing similarity search with the entire genome of the organism for which the primers are designed. The general approach involves the use of a set of algorithms such as the BLAST programs to compare a query sequence to all the sequences in a specified database. Comparisons are made in a pair wise fashion. Each comparison is given a score reflecting the degree of similarity between the query and the sequence being compared. The higher score the greater degree of similarity. The similarity is measured and shown by aligning two sequences. Alignments can be global or local (algorithm specific). A global alignment is an optimal alignment that includes all characters from each sequence, whereas a local alignment is an optimal alignment that includes only the most similar local region or regions. Discriminating between real and artifactual matches is done using an estimate of probability that the

match might occur by chance. Primers which fit the specified criteria were checked for their specificity using NCBI BLAST.

BLAST

The BLAST programs (Basic Local Alignment Search Tools) are a set of sequence that is used to search sequence databases for optimal local alignments to a query (Stephen 1997). The BLAST programs improved the overall speed of searches while retaining good sensitivity (important as databases continue to grow) by breaking the query and database sequences into fragments, and initially seeking matches between fragments. The fragment which is having high similarity score will be extended in both the directions in an attempt to generate an alignment with a score exceeding the threshold. All the designed primers which fit the specified criteria were tested for their similarity by performing BLAST against the Homo sapiens genome build 36 version 3 (36.3). Primers with multiple hits, max score less than 40 and E value less than 0.015 were not considered. The designed oligoprimers with good score and e-value were synthesised (Sigma-Aldrich, Bangalore, India) and standardised to obtain good amplicons.

Table 8. Provides the overview of the primers designed and used in this study.

Gene	SNPs	Forward Primer	Reverse Primer
EN2	rs1861972		
	rs1861973	CCTAGTTGACTCCTGACCTG	CATGTACTTCCCTCCTTGGTG
	rs73163344	CCACAGATCCTCTTTCTCCAG	CTCTACAAGGGTGTTGACTGC
CNTNAP2	rs1861972	CAGGGAGTGAAGATTTGTTGTG	TGAGCTTATCTAGGAGGAGGAA
	rs1861973	CAGGGAGTGAAGATTTGTTGTG	TGAGCTTATCTAGGAGGAGGAA
	rs73163344	CAGGGAGTGAAGATTTGTTGTG	TGAGCTTATCTAGGAGGAGGAA
	rs2710108	CAGGGAGTGAAGATTTGTTGTG	TGAGCTTATCTAGGAGGAGGAA
	rs2710105	CAGGGAGTGAAGATTTGTTGTG	TGAGCTTATCTAGGAGGAGGAA
	rs2710102	CAGGGAGTGAAGATTTGTTGTG	TGAGCTTATCTAGGAGGAGGAA
	rs2538994	CAGGGAGTGAAGATTTGTTGTG	TGAGCTTATCTAGGAGGAGGAA
	rs7806512	CAGGGAGTGAAGATTTGTTGTG	TGAGCTTATCTAGGAGGAGGAA
	rs7806519	CAGGGAGTGAAGATTTGTTGTG	TGAGCTTATCTAGGAGGAGGAA
	rs2710098	CAGGGAGTGAAGATTTGTTGTG	TGAGCTTATCTAGGAGGAGGAA

Polymerase Chain Reaction

PCR (Polymerase Chain Reaction) is a revolutionary method developed by Kary Mullis in 1980s. It is an essential and ubiquitous tool in genetics and molecular biology. With the use of this technique we can clone DNA *invitro*.

PCR is based on using the ability of DNA polymerase to synthesize new strand of DNA complementary to the offered template strand. Because DNA polymerase can add a nucleotide only onto a pre existing 3'-OH group, it needs a primer to which it can add the first nucleotide. This requirement makes it possible to delineate a specific region of template sequence that the researcher wants to amplify. At the end of the PCR reaction, the specific sequence will be accumulated in billions of copies (amplicons). For DNA sequencing by dye termination technique requires multiple

copies of DNA due to this reason we performed PCR to generate numerous copies of DNA fragments of interest which were further used for sequencing.

The genomic DNA of each subject was amplified on a ABI 9600 thermal cycler (Applied Biosystems, USA) with an initial de-naturation at 96^0C for 3 minutes and later on for 35 cycles at 95^0C for 60 seconds, at estimated annealing temperature of the primer for 45 seconds, extension at 72^0 C for 2.30 minutes and a final extension at the end of 35^{th} cycle at 72^0C for 7 minutes in a final volume of 10 µl containing 50mM KCl, 10mm Tris (pH 8.3), 1.5mM MgCl2, 75 ng of each primer, 100 µM deoxy-NTP, and 1 U *Taq* polymerase.

Amplicon Check by Electrophoresis

After the completion of 35 cycles of polymerase chain reaction 1µl of the amplicon was electrophorased on a 2% agarose gel containing ethedium bromide. DNA bands were visualised using gel documentation system (Syngene, UK). The samples which were amplified successfully were used for the sequencing after post PCR cleanup.

DNA Sequencing

There are various methods available for DNA sequencing like chemical degradation, chain termination method, sequencing by ligation and micro fluidic Sanger sequencing Etc,. Advances in automation have opened gates to new fast and reliable automated DNA sequencing technologies. Owing to its greater expediency and speed, dye-terminator sequencing is now the mainstay in automated sequencing. Dye-terminator sequencing is a slight modification of the Sanger's chain termination method it utilizes labelling of the chain terminator ddNTPs, which permits sequencing in a single reaction. In dye-terminator sequencing, each of the four dideoxynucleotide

chain terminators is labelled with fluorescent dyes, each of which with different wavelengths of fluorescence and emission. The dye labelled DNA fragments will be capillary electrophoresed and a ditection system will identify the labelled bases when they pass through a laser that activated the dye.

Cycle sequencing

To read the sequence of the amplified DNA termination with florescent dye labelled ddNTPs. To label each base PCR amplicon was subjected to cycle reaction with one primer was done using ABI Prism® BigDye^{TM} terminator v3.1 cycle sequencing ready reaction kits (Applied Biosystems, USA) following the manufacturer's guidelines.

Sequencing Cleanup

To remove the remnants of the above reaction to the each sample 2µl of 3M sodium acetate, 50µl of 100% ethyl alcohol were added and incubated at room temperature for 15 minutes to precipitate the DNA. The samples were centrifuged at 4000rpm for 25 minutes at 4^0C. The supernatant was discarded and the reaction plate was centrifuged inversely at 300 rpm for 20 seconds. 100µl of 75% alcohol was added to each sample and centrifuged at 4000rpm for 15 minutes at 25^0C. The supernatant was discarded and plate was inversely centrifuged at 300 rpm for 20 seconds to remove alcohol completely. The plate was dried at room temperature until left out alcohol dripped off.

Sequencing run

10µl of Hi-Di formamide was added to each well of the sample plate. The samples were heated to 96^0C and immediately cooled to 4^0C to denature the DNA. Sample information sheets which contain analysis protocol along with the sample

details were prepared and imported into the data collection software. Prepared samples were analysed on ABI 3730 genetic analyzer (Applied Biosystems, USA) to generate DNA sequences.

Sequence Alignment

The generated sequences were aligned to their respective reference sequences with the use of SeqScape v2.5 software (Applied Biosystems, USA). SeqScape is one of the suits of Applied Biosystems designed for automated sequence data analysis. It performs sequence comparisons for variant identifications, SNP discovery and validation. It allows analysis of the re-sequenced data, comparing the consensus sequences to a known reference sequence. The reference sequences for the gene studied were obtained from NCBI Genbank data base.

To set clear range of the sequence, a method that considers quality values of the bases was used which removes bases from the ends of sequences until fewer than 4 bases out of 20 have QVs <20. Filter setting values to filter the inappropriate sequences were set as maximum mixed bases to 20 and minimum sample score to 25. Depending on the sequence quality and the criteria specified for filtering the samples with low quality were not assembled by the program. These unassembled samples were re-sequenced until it satisfied the quality.

Sequence Editing and Mutation Scoring

Aligned sequences were manually edited to remove inappropriate mixed base calls by checking electropherograms of sequences. Heterozygotes were confirmed by performing re-sequencing of the samples. 5% of the samples where the proportion of heterozygotes was high were re-sequenced for confirmation. All the deviations from

the reference sequences were manually checked by examining electropherograms thoroughly. Genotypes were exported from the software for further analysis.

Statistical Tools

To assess the individual trait, crosstab analysis was used. All the above analysis was performed with the use of SPSS v 12 software (SPSS for windows, USA).

Statistical analysis of genetic data

Allele frequencies were calculated for all the SNPs and were tested for Hardy-Weinberg equilibrium and allelic association with the disease (exact tests). Allelic and genotype association with the phenotypes was tested under different genetic models for both quantitative and qualitative traits by regression analysis using PLINK (Purcell et al, 2007).
(http://pngu.mgh.harvard.edu/~purcell/plink/).

Investigation of Linkage Disequilibrium (LD)

Linkage disequilibrium (LD) between all the SNPs and also for associated SNP's was estimated using Haploview
(http://www.broadinstitute.org/haploview/haploview).

The pair-wise LD statistics D' and r^2 was calculated for all markers and also for associated SNP's. Haplotype blocks defined as regions in which all SNP pairs had an LD correlation coefficient r^2 value. Further haplotype frequencies were calculated individually for affected subjects.

CHAPTER - IV

RESULT AND DISCUSSION

In the previous chapter, the methodology adopted for the present investigation was presented. This chapter is on the analysis of data and interpretation of the results. The data collected have been analyzed according to the objectives formulated earlier. For the sake of clarity, the entire chapter is divided into 2 parts:

I. Demographic data analysis of the samples collected for the present investigation.

II. Analysis of Genetic markers from the blood samples collected and details of association analysis.

Part I: Demographic data analysis of the samples collected for the present investigation

The data has been collected on various demographic parameters along with blood samples. The sample questionnaire is presented at the end of the chapter.

Age of Identification

	Frequency	Percent
1-2yrs	40	42.1
2-3yrs	49	51.6
3-4yrs	4	4.2
4-5yrs	2	2.1
Total	95	100

Table 9. Frequency distribution of age of identification of Autistic children from the present study.

TYPES OF ASD

	Frequency	Valid Percent
Autism	72	75.8
PDD-NOS	19	20
Rett's Syndrome	1	1.1
Autism+MR/CP	3	3.2
Total	95	100

Table 10: represent the types of ASD distribution in the samples collected from AIISH for the present study.

			Group		
			Case	Control	Total
TV EXPOSURE	No TV	Count	28	13	41
		%	29.50%	14.10%	21.90%
	>1hrs	Count	40	75	115
		%	42.10%	81.50%	61.50%
	1-2 hrs	Count	8	4	12
		%	8.40%	4.30%	6.40%
	2-3 hrs	Count	5	0	5
		%	5.30%	0.00%	2.70%
	3 hrs<	Count	14	0	14
		%	14.70%	0.00%	7.50%
Total		Count	95	92	187
		%	100.00%	100.00%	100.00%

Symmetric Measures			
		Value	Approx. Sig.
Nominal by Nominal	Contingency Coefficient	0.404	0
N of Valid Cases		187	

Table 11. Details of exposure to television of ASD children and normal children in a day.

			Group		Total
			Case	Control	
HOUSE STATUS	Own	Count	56	92	148
		%	58.90%	100.00%	79.10%
	Rented	Count	10	0	10
		%	10.50%	0.00%	5.30%
	Flat	Count	25	0	25
		%	26.30%	0.00%	13.40%
	Quarters	Count	4	0	4
		%	4.20%	0.00%	2.10%
Total		Count	95	92	187
		%	100.00%	100.00%	100.00%

Table 12. Details of the house status of the parents of ASD children collected from the AIISH and normal children from the present study.

			Group		Total
			Case	Control	
NEAR FACTORY	Factory	Count	6	0	6
		%	6.30%	0.00%	3.20%
	No Factory	Count	89	92	181
		%	93.70%	100.00%	96.80%
Total		Count	95	92	187
		%	100.00%	100.00%	100.00%

Table 13. Details of the families of the ASD children and normal children staying near to the industries.

			Group		Total
			Case	Control	
NEAR PLANTATIONS	Chance of Pesticide Effect	Count	4	0	4
		%	4.20%	0.00%	2.10%
	No Chance of Pesticide Effect	Count	91	92	183
		%	95.80%	100.00%	97.90%
Total		Count	95	92	187
		%	100.00%	100.00%	100.00%

Table 14. Details of the families of the ASD children and normal children staying near to the plantations.

			Group		Total
			Case	Control	
CONSANGUINITY	Consanguinity Marriage	Count	8	0	8
		%	8.40%	0.00%	4.30%
	Not Consanguinity	Count	87	92	179
		%	91.60%	100.00%	95.70%
Total		Count	95	92	187
		%	100.00%	100.00%	100.00%

Table 15. Presents the detail of consanguineous marriages of parents of ASD children and normal children from the present study.

			Group		Total
			Case	Control	
ECNOMIC STATUS	Poor	Count	4	19	23
		%	4.20%	20.70%	12.30%
	Below average	Count	35	72	107
		%	36.80%	78.30%	57.20%
	Average	Count	53	1	54
		%	55.80%	1.10%	28.90%
	Rich	Count	3	0	3
		%	3.20%	0.00%	1.60%
Total		Count	95	92	187
		%	100.00%	100.00%	100.00%

Table 16 : Economic status of families with ASD children and normal children

			Group		Total
			Case	Control	
STAYED IN FOREIGN COUNTRIES	Stayed	Count	29	0	29
		%	30.50%	0.00%	15.50%
	Not stayed	Count	66	92	158
		%	69.50%	100.00%	84.50%
Total	Count	95	92	187	
	%	100.00%	100.00%	100.00%	

Symmetric Measures			
		Value	Approx. Sig.
Nominal by Nominal	Contingency Coefficient	0.388	0
N of Valid Cases		187	

Table 17: Families with ASD children and normal children stayed outside India.

			Group		Total
			Case	Control	
MIGRATION DUE TO JOB	Migrated	Count	66	0	66
		%	69.50%	0.00%	35.30%
	No Migration	Count	29	92	121
		%	30.50%	100.00%	64.70%
Total		Count	95	92	187
		%	100.00%	100.00%	100.00%

Table 18. Presents the details of migration history of the parent of ASD children and normal children included in the present study.

Pre natal History

Information of Maternal Infection ,Gestational diabetes, Thyroid problems, Excessive hygiene ,Psychological Stress during pregnancy was noted from the questionnaire and presented in the below tables.

Table 19. MATERNAL INFECTION

	Frequency	Valid Percent
Effected	7	7.4
Not effected	88	92.6
Total	95	100

Table 20. GESTATIONAL DIABETES

	Frequency	Valid Percent
Effected	1	1.1
Not effected	94	98.9
Total	95	100

Table 21. THYROID PROBLEM

	Frequency	Valid Percent
Effected	3	3.2
Not Effected	92	96.8
Total	95	100

Table 22. EXCESSIVE HYGIENE

	Frequency	Valid Percent
Effected	1	1.1
Not effected	94	98.9
Total	95	100

Table 23. STRESS

	Frequency	Valid Percent
Effected	18	18.9
Not effected	77	81.1
Total	95	100

NPar Tests (*Table 24*)

	MATERNAL INFECTION	GESTATIONAL DIABETES	THYROID PROBLEM	EXCESSIVE HYGIENE	STRESS
Chi-Square	69.063	91.042	83.379	91.042	36.642
Asymp. Sig.	0	0	0	0	0

The result presented in the table above explains statistically significant correlation with the disease. TV exposure, House Status of the family living, family migration from one place to another within India and outside the country may contribute to the environmental factors for the development of ASDs in Indian population or especially South Indian families with ASD children.

Consanguinity and Pre natal history collected from the family reveal the association with ASDs. Parameters like meternal infection, gestational diabetes, thyroid problem, excessive hygiene and psychological stress are the prenatal contributing factors for the development of ASDs in Indian population.

II Analysis of Genetic markers from the blood samples collected and details of association analysis.

Table 25. Summary of the10 SNP and its status in Genome data base

Sl. No	Gene	SNP'S	Chromosome Loci.	Region	Base pair position	Allele	Association with ASDs.
1	EN2	rs1861972	7q36	Intron	362	G/A	Yes
2	EN2	rs1861973	7q36	Intron	514	T/C	Yes
3	EN2	rs73163344	7q36	Intron	611	C/T	Yes
4	CNTNAP2	rs2710108	7q35-q36	Intron	133	T/C	No
5	CNTNAP2	rs2710105	7q35-q36	Intron	195	G/T	No
6	CNTNAP2	rs2710102	7q35-q36	Intron	349	A/G	No
7	CNTNAP2	rs2538994	7q35-q36	Intron	468	C/T	No
8	CNTNAP2	rs7806512	7q35-q36	Intron	514	C/T	No
9	CNTNAP2	rs7806519	7q35-q36	Intron	533	C/T	No
10	CNTNAP2	rs2710098	7q35-q36	Intron	631	G/T	No

A total of 10 variations (10 SNPs which are listed in the dbSNP) are identified by re-sequencing in the present study. A total of 3 variations are found to be associated with Autism Spectrum Disorders. All 3 variations, **(**rs1861972**,** rs1861973**,** rs73163344,) are present in the Intronic region of EN2gene.The variations in CNTNAP2 are not associated with ASDs.

Table 26. Presents the case control association details of the genetic markers genotyped from the ASD children and normal children. CHISQ test was performed to see the differences of allele frequencies between cases and controls. This basic association test did not warranted any significant results.

CHR	SNP	BP	A1	F_A	F_U	A2	CHISQ	P	OR	SE	L95	U95
7	rs2710108	1.48E+08	C	0.494	0.4891	T	0.008496	0.9266	1.02	0.2135	0.6712	1.55
7	rs2710105	1.48E+08	G	0.4759	0.5163	T	0.5698	0.4503	0.8507	0.2143	0.559	1.295
7	rs2710102	1.48E+08	G	0.5118	0.4725	A	0.5415	0.4618	1.17	0.2135	0.77	1.778
7	rs2538994	1.48E+08	C	0.2941	0.3315	T	0.5744	0.4485	0.8402	0.2299	0.5354	1.318
7	rs7806512	1.48E+08	T	0.4941	0.4728	C	0.1604	0.6888	1.089	0.2129	0.7175	1.653
7	rs7806519	1.48E+08	T	0.4824	0.4565	C	0.2368	0.6266	1.109	0.2132	0.7304	1.685
7	rs2710098	1.48E+08	T	0.5118	0.462	G	0.8775	0.3489	1.221	0.2131	0.804	1.854
7	rs1861972	1.55E+08	G	0.3529	0.3187	A	0.463	0.4962	1.166	0.226	0.7489	1.816
7	rs1861973	1.55E+08	T	0.375	0.3242	C	0.9938	0.3188	1.251	0.2247	0.8053	1.943
7	rs73163344	1.55E+08	T	0.2059	0.1319	C	3.451	0.0632	1.707	0.2898	0.9672	3.012

Table 27. Details of Logistic regression of the genotypes in ASD children and normal children. Though initial CHISQ was not found to be significant, however logistic regression using a Dominant genetic model showed significance. The genotypes of the SNPs rs1861972, rs1861973 and rs73163344 are found to be statistically significant with ASD. The significant p-values are presented in bold.

CHR	SNP	BP	A1	TEST	NMISS	OR	SE	L95	U95	STAT	P
7	rs2710108	1.48E+08	C	DOM	176	0.8333	0.3569	0.414	1.677	-0.5108	0.6095
7	rs2710105	1.48E+08	G	DOM	175	0.7662	0.3673	0.373	1.574	-0.7249	0.4685
7	rs2710102	1.48E+08	G	DOM	176	1.117	0.3638	0.5473	2.278	0.3033	0.7617
7	rs2538994	1.48E+08	C	DOM	177	0.7146	0.302	0.3953	1.292	-1.113	0.2659
7	rs7806512	1.48E+08	T	DOM	177	1.021	0.3537	0.5106	2.043	0.05994	0.9522
7	rs7806519	1.48E+08	T	DOM	177	0.9578	0.3507	0.4817	1.905	-0.1229	0.9022
7	rs2710098	1.48E+08	T	DOM	177	1.241	0.3584	0.6146	2.505	0.6019	0.5473
7	rs1861972	1.55E+08	G	DOM	176	1.922	0.3065	1.054	3.504	2.131	**0.03306**
7	rs1861973	1.55E+08	T	DOM	175	2.117	0.3111	1.151	3.895	2.411	**0.0159**
7	rs73163344	1.55E+08	T	DOM	176	2.253	0.3371	1.164	4.362	2.41	**0.01597**

Table 28. Analysis of the SNPs that are targeted and or identified from the re-sequencing data subjected to the TDT. The rs1861972 is found to the statistically significant (p = 0.039) implying that there may be a effect of this SNP on ASD in South Indian population (families) selected for the present study. However, the other SNPs are not found to be statistically significant. The other two SNPs which are in agreement with the ASD in case-control analysis using dominant model in logistic regression didn't showed any significance in TDT.

CHR	SNP	BP	A1	A2	T	U	OR	L95	U95	CHISQ	P
7	rs2710108	1.48E+08	T	C	16	26	0.6154	0.3301	1.147	2.381	0.1228
7	rs2710105	1.48E+08	G	T	18	28	0.6429	0.3556	1.162	2.174	0.1404
7	rs2710102	1.48E+08	A	G	17	28	0.6071	0.3323	1.109	2.689	0.1011
7	rs2538994	1.48E+08	C	T	21	17	1.235	0.6517	2.341	0.4211	0.5164
7	rs7806512	1.48E+08	C	T	19	28	0.6786	0.379	1.215	1.723	0.1893
7	rs7806519	1.48E+08	C	T	20	30	0.6667	0.3786	1.174	2	0.1573
7	rs2710098	1.48E+08	G	T	19	30	0.6333	0.3565	1.125	2.469	0.1161
7	rs1861972	1.55E+08	G	A	30	16	1.875	1.022	3.44	4.261	**0.039**
7	rs1861973	1.55E+08	T	C	26	21	1.238	0.6967	2.2	0.5319	0.4658
7	rs73163344	1.55E+08	T	C	26	25	1.04	0.6006	1.801	0.01961	0.8886

CHR	SNP	TEST	A1	A2	GENO	O(HET)	E(HET)	P
7	rs2710108	ALL	C	T	38/97/41	0.5511	0.4999	0.2272
7	rs2710108	AFF	C	T	20/43/21	0.5119	0.4999	1
7	rs2710108	UNAFF	C	T	18/54/20	0.587	0.4998	0.1432
7	rs2710105	ALL	G	T	37/100/38	0.5714	0.5	0.07072
7	rs2710105	AFF	G	T	16/47/20	0.5663	0.4988	0.2747
7	rs2710105	UNAFF	G	T	21/53/18	0.5761	0.4995	0.2094
7	rs2710102	ALL	G	A	36/101/39	0.5739	0.4999	0.06982
7	rs2710102	AFF	G	A	20/47/18	0.5529	0.4997	0.3899
7	rs2710102	UNAFF	G	A	16/54/21	0.5934	0.4985	0.09377
7	rs2538994	ALL	C	T	20/71/86	0.4011	0.4305	0.3833
7	rs2538994	AFF	C	T	10/30/45	0.3529	0.4152	0.1907
7	rs2538994	UNAFF	C	T	10/41/41	0.4457	0.4432	1
7	rs7806512	ALL	T	C	36/99/42	0.5593	0.4994	0.1334
7	rs7806512	AFF	T	C	19/46/20	0.5412	0.4999	0.5196
7	rs7806512	UNAFF	T	C	17/53/22	0.5761	0.4985	0.2082
7	rs7806519	ALL	T	C	32/102/43	0.5763	0.4981	0.04957
7	rs7806519	AFF	T	C	18/46/21	0.5412	0.4994	0.5185
7	rs7806519	UNAFF	T	C	14/56/22	0.6087	0.4962	0.03782
7	rs2710098	ALL	T	G	36/100/41	0.565	0.4996	0.09895
7	rs2710098	AFF	T	G	20/47/18	0.5529	0.4997	0.3899
7	rs2710098	UNAFF	T	G	16/53/23	0.5761	0.4971	0.1491
7	rs1861972	ALL	G	A	25/68/83	0.3864	0.4457	0.08999
7	rs1861972	AFF	G	A	8/44/33	0.5176	0.4567	0.3404
7	rs1861972	UNAFF	G	A	17/24/50	0.2637	0.4342	0.000237
7	rs1861973	ALL	T	C	24/74/77	0.4229	0.4541	0.4047
7	rs1861973	AFF	T	C	8/47/29	0.5595	0.4688	0.1046
7	rs1861973	UNAFF	T	C	16/27/48	0.2967	0.4382	0.003392
7	rs73163344	ALL	T	C	6/47/123	0.267	0.279	0.5875
7	rs73163344	AFF	T	C	2/31/52	0.3647	0.327	0.5044
7	rs73163344	UNAFF	T	C	4/16/71	0.1758	0.229	0.04272

Table. 29 depicts the test results of Hardy Weinberg Equilibrium. All the SNP obeyed Hardy Weinberg equilibrium except rs1861972 ($p = 0.000237$), rs1861973 ($p = 0.003392$), rs73163344 ($p = 0.04272$) . Interestingly these above SNPs are in equilibrium in cases but significantly deviated in controls. Probably low sample size may be contributing to the above trend. Interestingly all these SNPs are in HWE in equilibrium in the populations from 1000 genomes. Through two of the above three SNPs are earlier said to the associated with ASD, their deviation from HWE in controls and observed allelic richness in cases may be due to low sample size.

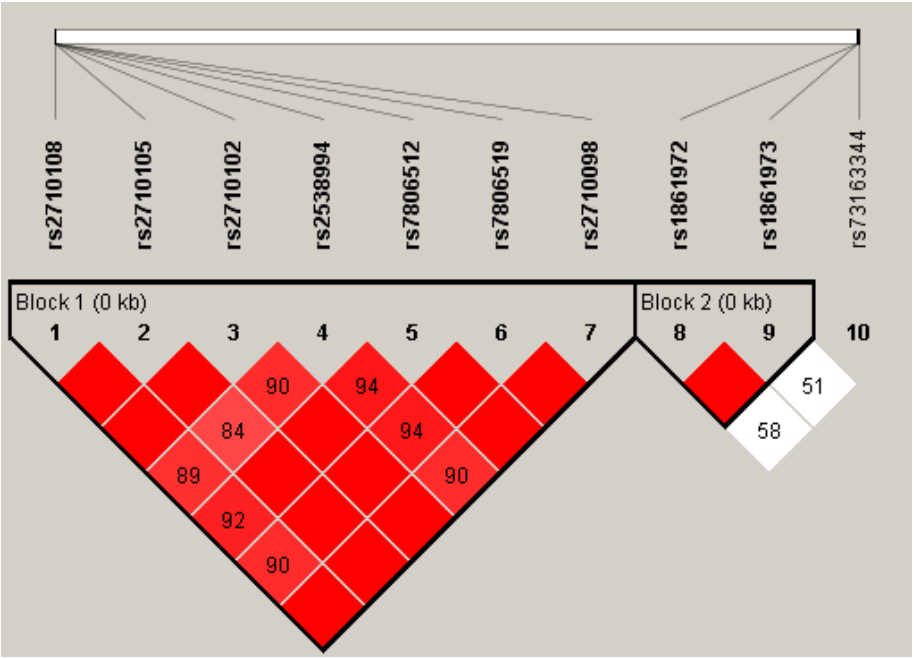


Figure 3. Linkage Disequilibrium plot in ASD cases from SNPs analyzed in the present study. The SNPs rs1861972, rs1861973 are in tight LD which were seen in block 2. (D'=1, r^2= 0.949). However, the SNP rs73163344 which was also found to be associated with the ASD is not in LD with any of the SNPs from block 2. Block 1 explains the LD between the SNPs, their D' and r^2 values are tabulated in the table 30 None of the SNPs from the block 1 are found to the in association with the ASD.

Table 30. Block 1 explains the LD between the SNPs, their D' and r^2 values are tabulated

Allele	test	r^2
rs2710108	rs2710105	0.885
rs2710105	rs2710105	1
rs2710102	rs2710105	0.976
rs2538994	rs2538994	1
rs7806512	rs2710105	0.908
rs7806519	rs2710105	0.865
rs2710098	rs2710105	0.976
rs1861972	rs1861973	0.949
rs1861973	rs1861973	1
rs73163344	rs73163344	1

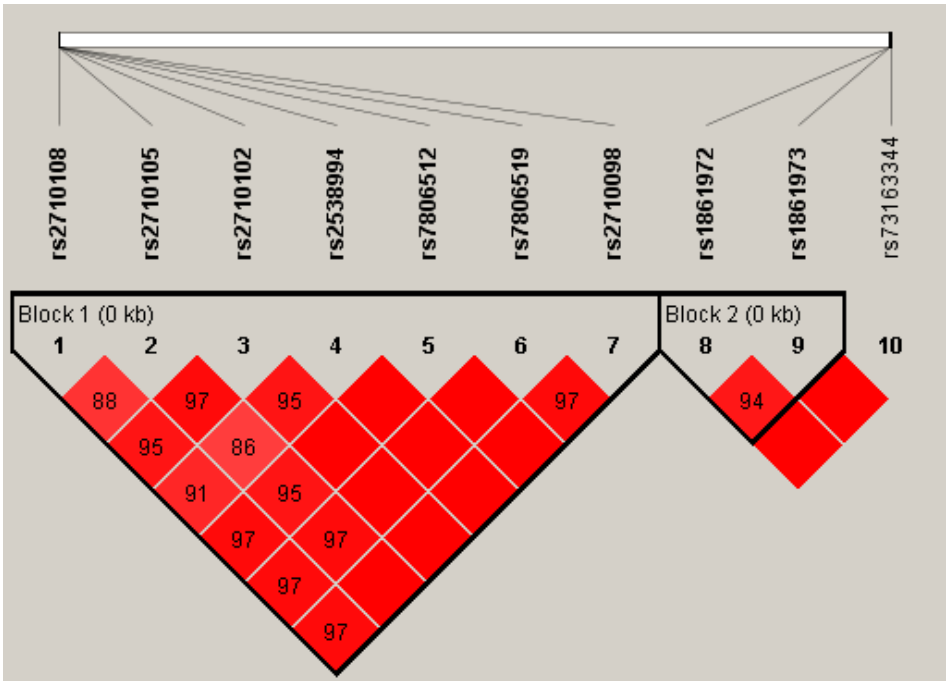


Figure 4.Linkage Disequilibrium plot in controls from SNPs analyzed in the present study. The SNPs rs1861972, rs1861973 are in tight LD which were seen in block 2. (D'= 0.948, r^2= 0.877). The SNP rs73163344 which was also found to be associated with the ASD is in weak LD with SNPs from block 2. Block 1 and block 2 explains the LD between the SNPs, their D' and r^2 values are tabulated in the table 31.

Table 31. Block 1 and block 2 explains the LD between the SNPs, their D' and r^2 values are tabulated

Allele	Test	r^2
rs2710108	rs2710102	0.87
rs2710105	rs2710102	0.893
rs2710102	rs2710102	1
rs2538994	rs2538994	1
rs7806512	rs2710102	0.978
rs7806519	rs2710102	0.936
rs2710098	rs2710102	0.978
rs1861972	rs1861973	0.877
rs1861973	rs1861973	1
rs73163344	rs73163344	1

Electrophorogram representation of SNPs which are used for the present study.

SNP-rs1861972

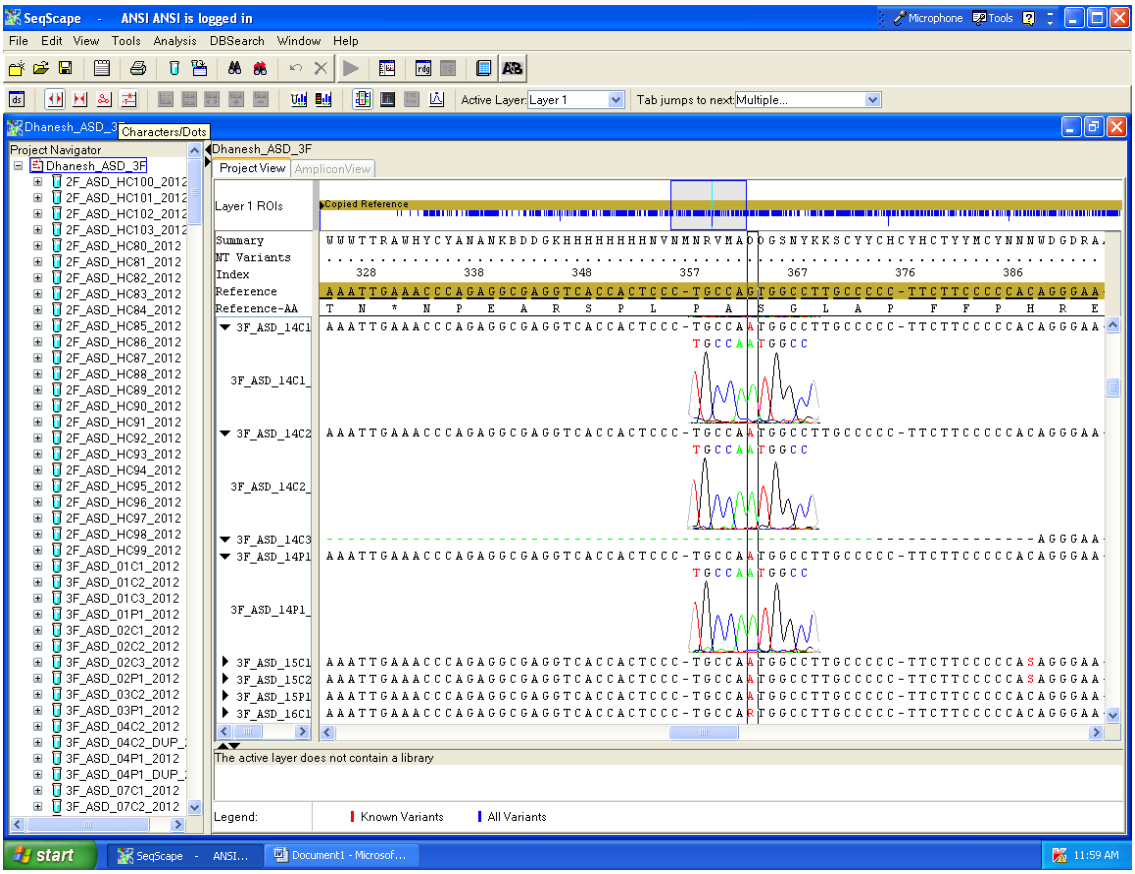


Figure 5. Electrophorograms of the sequence analyzed in the present study. Three electrophorograms are opened in the above project view, first graph show a A>G homozygous mutation. Whereas graph 2 and 3 shows a herterozygous mutation i.e., of the two homologous chromosome one contains G and the other contains A at the same position.

SNP-rs1861973

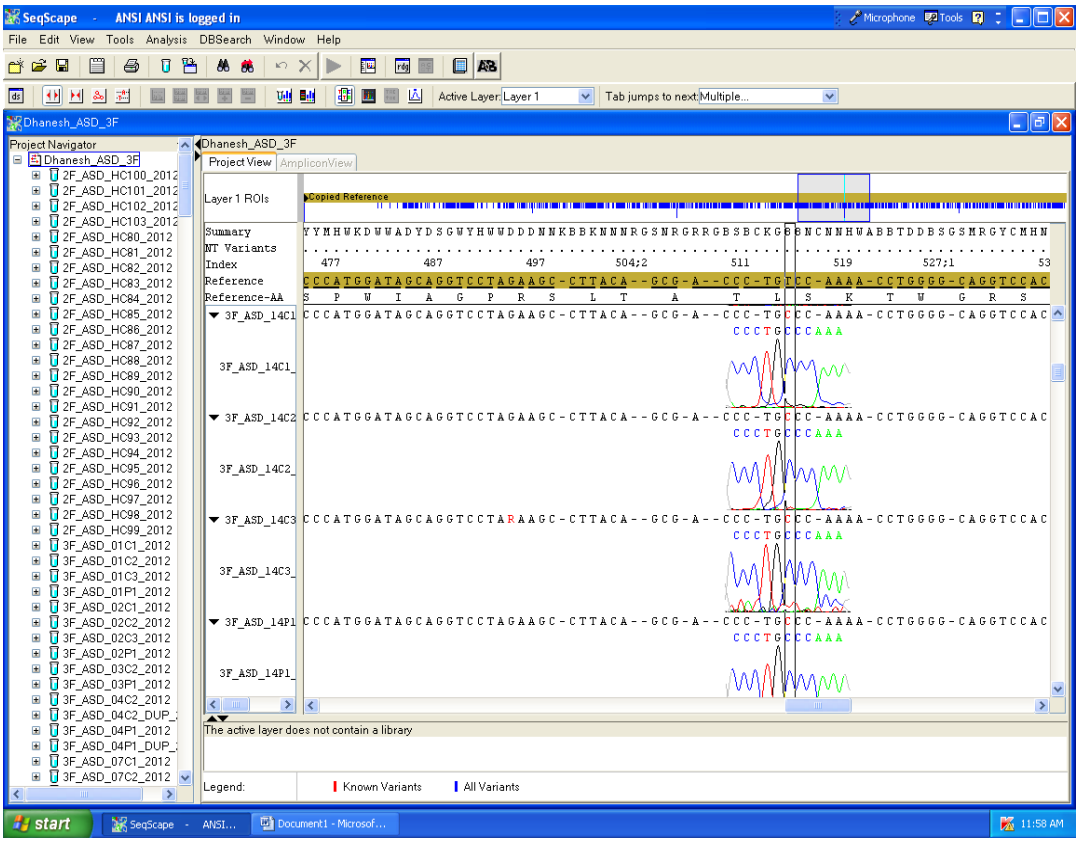


Figure 6. Electrophorograms of the sequences analyzed in the present study. Four electrophorograms are opened in the above project view for the representation purpose, all the four graphs show a mutation in from T>C which indicates that all the four samples shown are homozygous for the minor allele rs1861973.

SNP-rs73163344

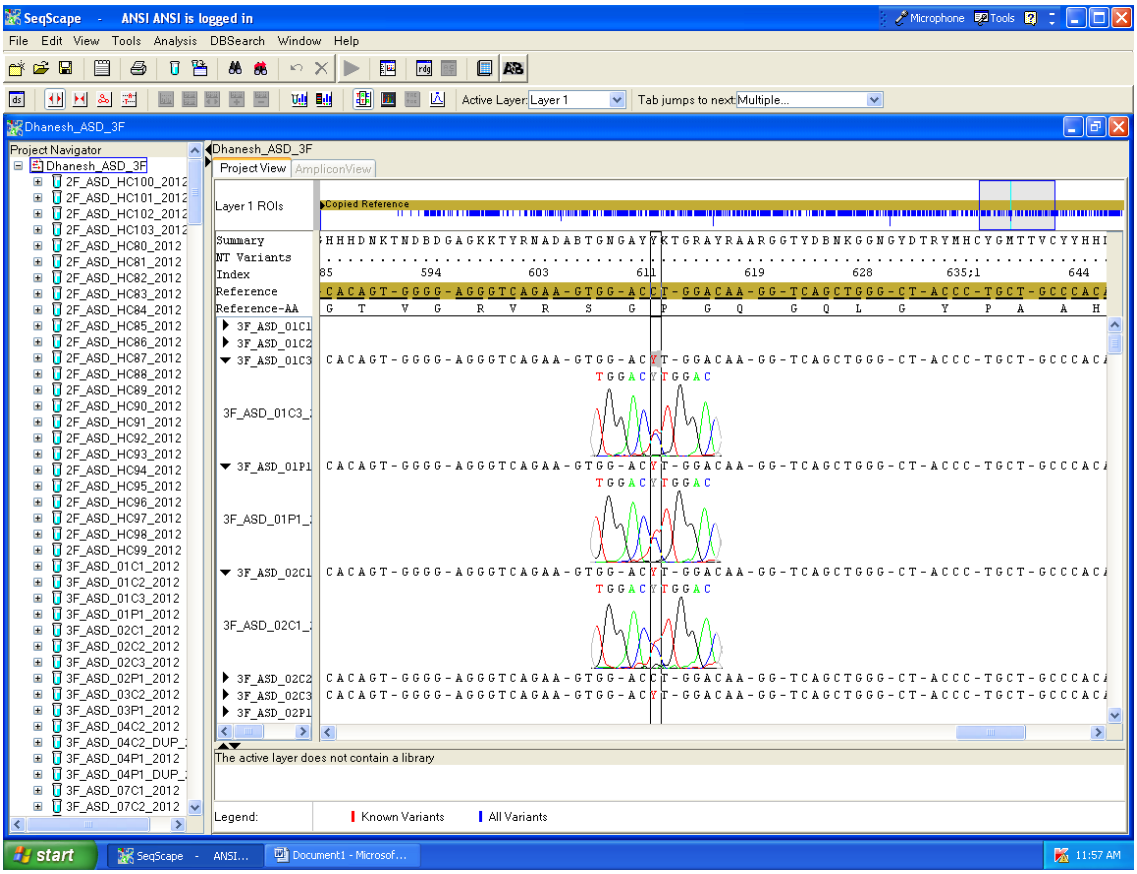


Figure 7. Electrophorograms of the sequences analyzed in the present study. The above figure shows three electrophorograms where all the 3 regions showing a heterozygous mutation i.e., out of two homologous chromosomes one contains a C allele and the other contains mutant T allele for the SNP rs73163344.

SNP-rs2710102

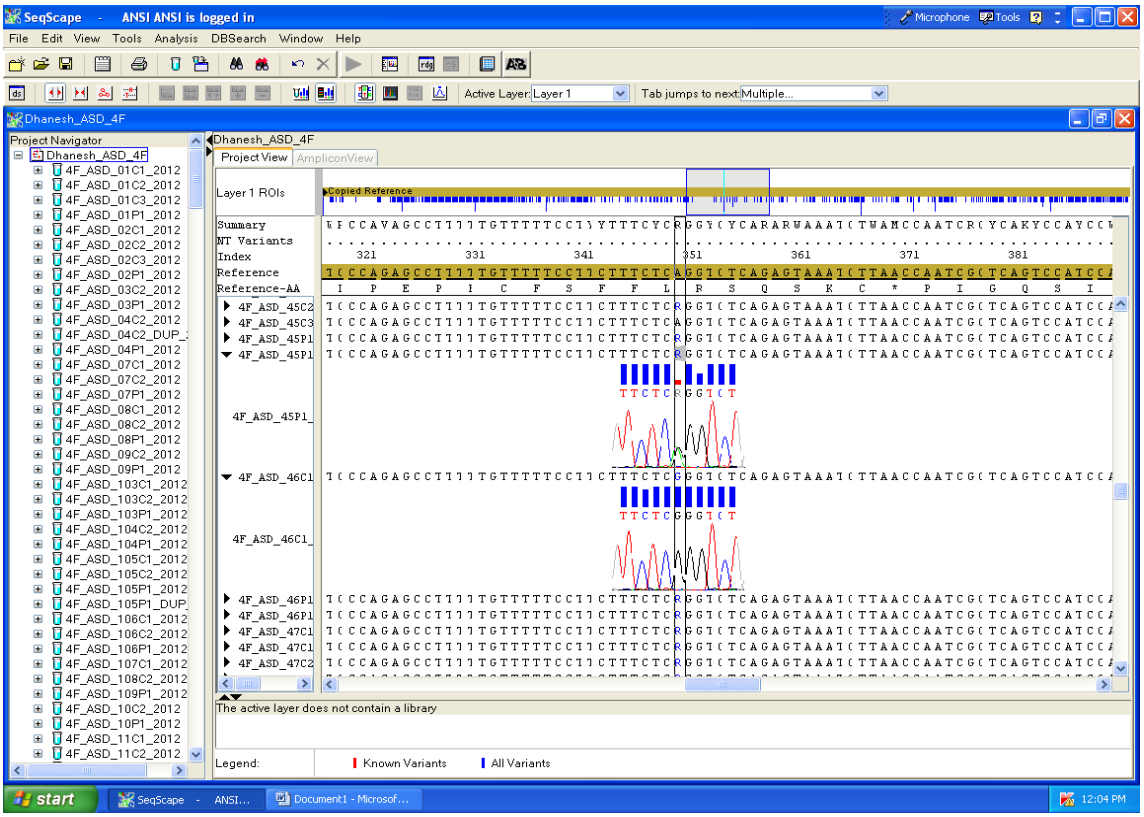


Figure 8. Electrophorograms of the sequences analyzed in the present study. First electrophorogram shows a heterozygous mutation, and the second electrophorogram shows a homozygous mutations A>G.

DISCUSSION

Autistic spectrum disorders (ASDs) are the continuum of cognitive and neurobehavioral disorders including autism. The prevalence of autism varies significantly with case ascertainment, ranging from 0.7 to 21.1 per 10 000 children (median 5.2 per 10 000) while the prevalence of autistic spectrum disorder is projected to be 1 to 6 per 1000 (World Health Organization, 2002). Autism is strongly genetically determined, as demonstrated by its increased prevalence in siblings. A review from American Medical Association suggests that 20 Monozygotic twins show 60% to 90% concordance, that means 60% to 90% of cases in which one has autism, the other twin does also. Concordance in dizygotic twins and siblings are 5%-10%. The genetic control of autism is an extremely active area of research. Combined with rapidly advancing technology, an enormous amount of genetic data attempting to explain autism has emerged. Several different genetic abnormalities have been found in autistic individuals, and affected individuals in the same family tend to carry the same genetic abnormality.

The results of the demography and lifestyle data suggests that the factors like exposure to the Television is more in ASD cases when compared to the normal children. House status, economic status, family migration from one place to another within India and foreign countries, consanguiness marriages and pre natal histories (Maternal Infection, gestational diabetes, excessive hygiene, thyroid problems, psychological stress) may be some of the cause of environmental factors for the development of autism in south indian population. We suggest proper genetic counselling to avoid consanguiness marriage and proper awareness of ASDs and the correlating environmental factors to the public and also we suggests more studies has to

be conducted with more samples and different ethnic population to reveal the causation and epigenetic effect of Autism in Indian scenario.

Several loci have been identified so far , some or all of which may contribute to the phenotype. Among these 7q regions is one of the candidate regions successfully replicated in many case control studies in developed World. Richler E, Reichert JG, Buxbaum JD, McInnes LA (2006) suggests that chromosome location 7q contains a bulk of genes that playes important role in in transcription and development. This 7q region is also for ultra conserved non coding sequences, defined as human–rodent sequences that are 100% aligned over 200 base pairs, which have a high probability of being functional.The 7q35-q36 harbours NRCAM, EN2, CNTNAP2, FOXP2, MET genes which are highly susceptible candidate genes for ASD(OMIM). Gharani et al (2004) reported the gene EN2 encodes engrailed-2 protein which acts as a transcription factor for homeo box genes. Homeobox genes are key determinants during the embryonic development. Brielmaier et al (2012), mice studies indicate that abnormal function of this gene causes neurodevelopmental abnormalities in cerebellar circuits and cell numbers. Cerebellar abnormalities are one of the most common histopathological findings in humans affected with ASDs. In continuation with the animal studies, and human histopathological findings, several studies reported association of intronic mutations from this EN2 gene with Autism in humans.

Logan et al (1992) isolated chicken and human genomic clones of the Engrailed (EN1) and Engrailed2 (EN2) genes and confirmed that as in mouse and chicken, the predicted cloning region of the human EN2 gene is interrupted by a single intron. Benayed et al (2005) concluded that the human EN2 gene span 8.1 kb genomic DNA and consist of two exons separated by a 3.3kb intron. During development expressed

mouse EN2 ectopically in cortical precursors. Results of LD studies suggested that EN2 may play a role in susceptibility to ASDs. Ladd-Acosta et al (2007) found that EN2 expression is decreased in brain tissue with less methylation. Poole et al (1989) mapped the EN2 gene to human chromosome 7 regional mapping by in situ hybridization localized in to 7q36.

SNPs rs1861972 and rs1861973 are the intronic markers associated with ASDs in many populations. Gharanni et al (2004) suggests the same markers have important role in ASDs. According to the research carried out in chinees population, Yang et al (2010) suggests haplotype A-C of rs1861972 and rs1861973 is the core element of the observed haplotype association study, which plays a role as a protective factor against autism; in addition, the haplotype G-A-C is less frequent in male cases compared to controls (38.64 vs. 52.51%), which plausibly modulate disease vulnerability to autism. Sen et al (2010) worked on the SNPs associated with ASDs in Indian population family-based analysis of these markers using TDT showed that the two intronic markers, which are in absolute LD, are associated with autism/ASD with a significant higher transmission of C allele of rs1861973 (hence A allele of rs1861972) to affected offspring, Present TDT family study supports the association of two intronic SNPs (rs1861972 and rs1861973) and additional SNP(rs73163344) in Indian population.

The gene Contactin Associated Protein-Like 2(CNTNAP2) is very important gene, encodes a member of the neurexin family which functions in the vertebrate nervous system as cell adhesion molecules and receptors. This gene encompasses almost 1.5% of chromosome 7 and is one of the largest genes in the human genome. It is directly bound and regulated by forkhead box protein P2 (FOXP2), a transcription factor related to speech and language development. This gene has been implicated in

multiple neurodevelopmental disorders, including Gilles de la Tourette syndrome, schizophrenia, epilepsy, autism, ADHD and mental retardation.

CNTNAP2 variants affect early language development in the general population and it alters brain function during linguistic processing in healthy individuals. rs2710102, a common SNP in the CNTNAP2 gene, was found to be significantly associated ($p<0.028$) with a delayed onset of speech, as measured by the age at which a child speaks their first words, in children with autism. This effect is primarily seen in males, perhaps correlated with the 4-5x over representation of males with autism compared with females. Speech development rs4431523, rs17236239 and significant associations (with P values from 0.01 to 5.0 x10 – 5) between nonsense-word repetition and nine intronic SNPs (rs851715, rs10246256, rs2710102, rs759178, rs1922892, rs2538991, rs17236239, rs2538976, and rs2710117).

Verkerk et al (2003) hypothesised in a family study that, decreased expression of CNTNAP2 could lead to a disturbed distribution of potassium channels in the nervous system, there by influencing conduction or repolarisation of action potential, causing unwanted action or movements in GTS(Gilles de la Torette Syndrome). Alarcon et al (2008), Arking et al (2008), Bakkaloglu et al.(2008) identified SNP in the CNTNAP2 gene that were associated with significant association with autism. In a two stage analysis of 10 MB quantitative triat locus for autism related traits on 7q35-q36 using parent child trios, Alarcon et al (2008) revealed an association between variation at rs2710102 in the CNTNAP2 gene and age at first word in autism spectrum disorders sample from male-only families. The authors concluded that SNP association did not imply that variation at rs2710102 is causally related to ASD, but rather that variation here likely to to be in LD with an untested functional variant.

Ma et al. (2009) identified in a genome wide association study with 1390 individuals with autism and validation in an additional cohort of 2390 samples from 457 families did not a significant association between autism and in Alarcon et al (2008) study rs 270102 which was the tagging SNP . Arking et al. (2008)in a family study identified a common variant in the CNTANAP2 gene SNP rs7794745,that was associated with increased risk for autism. This SNP is located in intron2 of the gene. In a study of Bakkaloglu et al. (2008) four children with autism from three unrelated families identified an i869 to their substitution in the CNTNAP2 protein. The mutation at a conserved residue occurred in the third laminine G domain and was predicted to be deleterious. Inheritance of variants from apparently affected parents in each family .negative result found among 4010 control chromosome.

In the present study SNP rs2710102 was targeted for re-sequencing. Our re-sequencing identified rs2710108, rs2710105, rs2538994, rs7806512, rs7806519, rs2710098 which are in the close vicinity of the targeted variant rs2710102. However, none of these SNPs are found to the associated with the ASD in both case-control and family TDT analysis. The results are presented in the tables (26,27,28).

The present study with the intronic markers of the gene CNTNAP2 could not found any association with the ASD in both case-control and family TDT analysis. Besides differences in the ethnicity low sample size may be one of the reason that can be attributed for this reason. The SNP rs1861972 and rs1861973 from the gene EN2 which was already told to be associated with the ASD in the Indian populations and as well as in the different populations in the world wide were also found to be significantly associated with ASD in the present study in case control analysis (P-Values,rs1861972=0.03306; rs1861972=0.0159; rs73163344=0.01597). In contrast only

SNP rs1861972 was found to be associated with ASD in family TDT analysis (p=0.039). In addition, the present study identified a new association with the ASD in case-control analysis with the marker rs73163344 (p=0.0159) from EN2 gene which is present in the vicinity of rs1861973. The present study reporting a marginal to moderate significance of the above discussed variants with ASD. The study is one of its kinds to be reported from Karnataka region of South India. The present study warrants for further studies to resequence the entire genes with elevated number of samples from different ethinic populations by which more functional variants may be found for the better explanation of the genes EN2 and CNTNAP2 role in disease progression of ASD.

In the present study genotypes of SNPs rs1861972, rs1861973, rs73163344 were associated with the autism in logistic regression dominant model. All these 3 SNP's i.e., SNPs rs1861972 (p = 0.000237), rs1861973 (p = 0.003392), rs73163344 (p = 0.04272) were deviated from the Hardy Weinberg equilibrium in HW test in controls. Interestingly these 3 SNPs obeyed HWE in cases rs1861972 (p = 0.3404), rs1861973 (p = 0.1046), rs73163344 (p = 0.5044), further when pooled the case and controls data and test for HWE the same trend was observed as in the case of the cases. There are many factors contributing to the deviation of HWE in general, in general populations. Also, it was a well established fact that no population exist in HWE. As the selection of our either cases or controls is strictly random, sampling bias can be overruled in the present study. Further we cross checked our data for possible genotyping error to overcome from this situation. Samples were randomly re-sequenced and the concordance rate was 99.99%, where even genotyping error can also

be overruled. Hence it may be concluded that low sample size contributed to the deviation of HWE in controls.

CHAPTER - V

SUMMARY AND CONCLUSION

Autism disorder, Asperger's disorder, childhood Disintegrative Disorder (CDD), Rett's disorder and PDD-NOS (Pervasive Development Disorder Not Otherwise Specified) are the disorders under Autism Spectrum. An ASD (American Psychiatric Association, 2000) appears usually around 3 years of age characterized by the triad of limited or absent verbal communication, a lack of reciprocal social interaction or responsiveness, and restricted, repetitive, and ritualized patterns of interests and behavior (Baily et al., 1996; Rish et al., 1999). Folstein and Rosen-Sheidley (2001) found male to female ratio of idiopathic autism estimated at 4-10:1, and with an increase in this ratio as the intelligence of the affected individual increases. Prevalence of Autism Spectrum Disorder is increasing because of genetic and environmental factors including the air we breathe, our food, our water, and even the social community where are parented and raised. Autism has a strong genetic basis, although the genetics of autism are complex and it is unclear whether ASD is explained more by a rare mutation with major effects, or by rare multi gene interactions of common genetic variants.

There is no breakthrough or single gene that is a major new cause of autism. But the role of genetics becomes even more evident when these single base changes are considered i.e., ASD is not a monogenic disorder which can be attributed to a single gene or chromosomal location. For instance, an individual with autism is nearly 6-fold more likely to have a functional variant in genes expressed in the brain. Sanders et al. estimate as many as 14 percent of affected individuals have such a risk variant. This 14 percent is in addition to the 10–20 percent with a large copy number variant or

identified genetic syndrome. O'Roak et al (2002). find that 39 percent of these variants are related to a specific biochemical pathway, important for brain signalling and Neale et al.(2012), while cautioning that the net effect of all of these changes still leave much of the risk for autism unexplained, pointing the roles of a few specific genes as genuine candidates.

Primary aim of the present study was to genotype established markers from the candidate genes which have a role in disease progression of autism and to evaluate them by genetic association study in Indian scenario. Markers (SNPs) are identified from the genes EN2 and CNTNAP2. Selected genes are strong candidate genes with ASD and they play a very important role in the development of the parts of brain and language processing. Second aim was to compare the genotype and allele frequency of EN2 and CNTNAP2 genes with ASD subjects and controls. Third aim was to find out the role of these genes in the family and the final aim was to delineate environmental factors underlying ASD by analysis of socio demographic factors and other factors if any.

The selection criteria of ASD was based on DSM IV criteria, developed check list in the ASD unit (DDC-ASD based on DSM IV criteria) of All India Institute of Speech and Hearing, Mysore. Autistic children and parents are recruited to the ENT department and staff nurse would withdraw 5-6ml blood samples from the child, parents, and sibling .The completed questionnaire and the informed consent were taken from each of the parents. A structured questionnaire was administered to collect socio-demographic information like ethnicity, age, sex, house status, TV exposure, economic status, migration, pre natal information etc. The blood samples were stored in the storage facility in AnSI (Anthropological Survey of India) and DNA was isolated,

quantified. Quality check was done by gel electrophoresis as well as by using spectrophotometer (Hitachi). DNA was extracted using standard phenol–chloroform method described earlier. Stock DNA was prepared basing on the readings of spectrophotometer using TE buffer. 200 µl of stock DNA was prepared for each sample and 100 µl of working DNA solution using Milliq water.

Total three primers were designed for each SNPs selected for the study and polymerase chain reaction (PCR) done for total 250 samples (ASD cases 85, 62 fathers, 85 mothers, 18 siblings) were collected from ASD unit, All India Institute of Speech and Hearing,Mysore.91 healthy controls samples were collected from Mysore and Kerala (Not affected by Autism Spectrum Disorder). After getting good PCR product from the reaction the samples were subjected to cycle sequencing Big-Dye terminator V.

After cycle sequencing the samples were processed to remove remained Dye and other potential contaminants by post cycle sequencing clean up method as described elsewhere.

The sequences are up loaded to the SeqScape v2.5 (Applied Biosystems) software for mutation checking. Sequences are aligned to the reference sequence from NCBI or Ensemble and there after edited where ever required. Mutations are confirmed and noted in a spread sheet. Genetic analysis software PLINK is used for the analysis of genetic data and SPSS used for demographic data.

Results of the study are summarized as follows:

- ✓ TV exposure, House Status (live in own house/apartments/flat/rental or any) of the family living, family migration from one place to another within India and outside

the country may contribute to the environmental factors for the development of ASDs in Indian population especially South Indian families with ASD children.

- ✓ Consanguinity and Pre natal history collected from the family reveal the association with ASDs, parameters like maternal infection, gestational diabetes, thyroid problem, excessive hygiene and psychological stress may be the prenatal contributing factors for the development of ASDs in Indian population.

- ✓ A case control association detail of the genetic markers was genotyped from the ASD children and normal children. CHISQ test was performed to see the differences of allele frequencies between cases and controls individually. This basic association test did not warrant any significant results.

- ✓ Logistic regression of the genotypes in ASD children and normal children was seen. Though initial CHISQ was not found to be significant, logistic regression using a dominant genetic model showed significance. The genotypes of the SNPs rs1861972, rs1861973 and rs73163344 were found to be statistically significant with ASD in Indian population.

- ✓ Novel genetic association was found with the intronic marker-rs73163344 (EN2 gene) with ASDs in Indian population. This SNP was not reported to be a candidate marker for ASD elsewhere. This is the first study which found association of this marker with ASD with statistically significant value.

- ✓ Analysis of the SNPs that are targeted and or identified from the re-sequencing data was subjected to the TDT. The rs1861972 was found to the statistically significant ($p = 0.039$) implying that there may be effect of this SNP on ASD in South Indian population (families) selected for the present study. However, the

other SNPs are not found to be statistically significant. The other two SNPs which are in agreement with the ASD in case-control analysis using dominant model in logistic regression didn't show any significance in TDT.

Implications of the present study:

- The present study evidences effect of genetic markers of EN2 and CNTNAP2 with ASDs in Indian population especially with south Indian ethnicity.
- The research helped to identify the mutation associated with ASD in selected ASD children and their parents.
- The demographic data revealed the correlating factors which may contribute for the development of ASD in south Indian population.
- This study help to develop database of genetic markers which effects in ASDs in Indian population and it helps in the genetic diagnosis of ASDs and genetic counselling in future.
- This study may helps further more functional proteomic study and which help to the discovery of drug targets for the development of molecular based medicine for ASDs.

Future Directions

- Increase in the sample size and number of markers with special reference to ASDs might help to evaluate the role of genetic basis of ASD in Indian scenario.

- Suggests more proteomic studies in ASD with functional markers to improve the development of drug targets.

- Recommending to develop genetic based diagnosis criteria for ASDs in Indian population including cytogenetic study. This may help to differentiate the syndromic and non syndromic ASD children.

- The study urges the need to awareness programs about ASDs consanguineous marriages and other factors which affect for the developmental of the delay in speech language, socialization other complications with ASDs.

- Research may be conducted in the area of pre natal Diagnosis of ASDs to avoid incidence of ASD.

- Suggests the need for demographic and genetic research in Autism with different ethnic populations, special focus on environmental factors which may elevate the genetic affects with the individuals carrying the susceptible alleles.

REFERENCES

Abrahams, B. S.; Tentler, D.; Perederiy, J. V.; Oldham, M. C.; Coppola, G.; Geschwind, D. H.(2007).Genome-wide analyses of human perisylvian cerebral cortical patterning. *Proc. Nat. Acad. Sci*. 104: 17849-17854.

Alarcon, M.; Abrahams, B. S.; Stone, J. L.; Duvall, J. A.; Perederiy, J. V.; Bomar, J. M.; Sebat, J.; Wigler, M.; Martin, C. L.; Ledbetter, D. H.; Nelson, S. F.; Cantor, R. M.; Geschwind, D. H.(2008) Linkage, association, and gene-expression analyses identify CNTNAP2 as an autism-susceptibility gene. *American Journal of Human Genetics.* 82: 150-159.

Alvarez Retuerto, AI., Cantor ,RM., Gleeson, JG., Ustaszewska,A., Schackwitz ,WS., Pennacchio, LA., Geschwind, DH.(2008) Association of common variants in the Joubert syndrome gene (AHI1) with autism. *Human Molecular Genetics*. 17(24):3887-96.

Amir, RE., Van den Veyver, IB., Wan, M., Tran, CQ., Francke, U., Zoghbi, HY.(1999) Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nature Genetics*. 185-8.

Anderson, BM., Schnetz-Boutaud, NC., Bartlett, J., Wotawa, AM., Wright, HH., Abramson ,RK., Cuccaro, ML., Gilbert, JR., Pericak-Vance, MA., Haines, JL.(2009) Examination of association of genes in the serotonin system to autism. *Neurogenetics*. 209-16. doi: 10.1007/s10048-009-0171-7.

Anney, R., Klei, L., Pinto, D., Regan, R., Conroy, J., Magalhaes, TR., Correia, C., Abrahams ,BS., Sykes, N., Pagnamenta, AT., Almeida, J., Bacchelli, E., Bailey,

AJ., Baird, G.,Battaglia ,A., Berney, T., Bolshakova, N., Bölte, S., Bolton, PF., Bourgeron, T., Brennan, S., Brian, J., Carson, AR., Casallo, G., Casey, J., Chu, SH., Cochrane, L., Corsello, C.,Crawford ,EL., Crossett, A., Dawson, G., de Jonge, M., Delorme, R., Drmic, I., Duketis, E., Duque ,F., Estes, A., Farrar, P., Fernandez, BA., Folstein ,SE., Fombonne, E., Freitag, CM., Gilbert, J, Gillberg C, Glessner JT, Goldberg J, Green J, Guter SJ, Hakonarson H, Heron EA, Hill M., Holt, R., Howe, JL., Hughes, G., Hus, V., Igliozzi, R., Kim, C.,Klauck, SM., Kolevzon ,A., Korvatska, O., Kustanovich, V., Lajonchere, CM., Lamb, JA., Laskawiec, M., Leboyer, M., Le Couteur, A., Leventhal, BL., Lionel, AC., Liu, XQ., Lord, C.,Lotspeich ,L., Lund ,SC., Maestrini, E., Mahoney, W., Mantoulan, C., Marshall ,CR., McConachie ,H., McDougle ,CJ., McGrath, J., McMahon, WM., Melhem, NM., Merikangas, A.,Migita, O, Minshew NJ, Mirza GK, Munson J, Nelson SF, Noakes C, Noor A, Nygren G, Oliveira G., Papanikolaou, K., Parr, JR., Parrini, B., Paton, T., Pickles, A., Piven, J.,Posey, DJ., Poustka, A., Poustka, F., Prasad, A., Ragoussis, J., Renshaw, K., Rickaby, J., Roberts ,W., Roeder ,K., Roge, B., Rutter, ML., Bierut ,LJ., Rice, JP., Salt, J., Sansom, K., Sato ,D., Segurado, R., Senman, L., Shah, N., Sheffield ,VC., Soorya, L., Sousa, I., Stoppioni, V., Strawbridge, C., Tancredi, R., Tansey, K., Thiruvahindrapduram, B.,Thompson ,AP., Thomson, S., Tryfon, A., Tsiantis, J., Van Engeland, H., Vincent, JB., Volkmar ,F., Wallace, S., Wang, K., Wang, Z., Wassink, TH., Wing, K., Wittemeyer ,K.,Wood ,S., Yaspan, BL., Zurawiecki, D., Zwaigenbaum, L., Betancur ,C., Buxbaum, JD., Cantor ,RM., Cook ,EH., Coon, H., Cuccaro, ML., Gallagher, L., Geschwind, DH., Gill, M.,Haines, JL., Miller, J., Monaco ,AP., Nurnberger ,JI Jr., Paterson ,AD., Pericak-Vance ,MA., Schellenberg, GD., Scherer, SW., Sutcliffe, JS., Szatmari, P., Vicente, AM.,

Vieland ,VJ., Wijsman, EM., Devlin, B., Ennis ,S., Hallmayer, J.(2010) A genome-wide scan for common alleles affecting risk for autism. *Human Molecular Genetics*. 19 (20): 4072-82.

Arking DE, Cutler DJ, Brune CW, Teslovich TM, West K, Ikeda M, Rea A, Guy M, Lin S, Cook EH, Chakravarti A.(A common genetic variant in the neurexin superfamily member CNTNAP2 increases familial risk of autism. *American Journal of Human genetics*. 160-4.

Barber, JC., Ellis, KH., Bowles, LV., Delhanty, JD., Ede, RF., Male, BM., Eccles, DM.(1994) Adenomatous polyposis coli and a cytogenetic deletion of chromosome 5 resulting from a maternal intrachromosomal insertion. *Journal of Medical Genetics* 312-6.

Bates, E. (1976). Language in context. New York: Academic Press.

Belloso, J. M.; Bache, I.; Guitart, M.; Caballin, M. R.; Halgren, C.; Kirchhoff, M.; Ropers, H.-H.; Tommerup, N.; Tumer, Z.(2007).Disruption of the CNTNAP2 gene in a t(7;15) translocation family without symptoms of Gilles de la Tourette syndrome. *Europ. J. Hum. Genet.* 15: 711-713.

Bottini,N.,DeLuca,N.,Saccucci,P.,Fiumara,A.,Elia,M.,Porfirio,MC.,Lucarelli,P.Curatolo,P.(2001) Autism: evidence of association with adenosine deaminase genetic polymorphism. *Neurogenetics*,3 (2):111-3.

Bowers, K., Li, Q., Bressler, J., Avramopoulos, D., Newschaffer ,C., Fallin, MD.(2011) Glutathione pathway gene variation and risk of autism spectrum disorders.

Journal of Neurodevelopmental Disorder. 3(2):132-43. doi: 10.1007/s11689-011-9077-4.

Campbell, DB., Datta, D., Jones, ST., Batey Lee, E., Sutcliffe, JS., Hammock, EA., Levitt, P.(2011). Association of oxytocin receptor (OXTR) gene variants with multiple phenotype domains of autism spectrumdisorder. *Journal Neurodevelopmental Disorder*. 3(2):101-12. doi: 10.1007/s11689-010-9071-2.

Chakrabarti, B., Dudbridge, F., Kent, L., Wheelwright, S., Hill-Cawthorne, G., Allison, C., Banerjee-Basu, S., Baron-Cohen, S. (2009) Genes related to sex steroids, neural growth, and social-emotional behavior are associated with autistic traits, empathy, and Asperger syndrome. *Autism Research* .157-77.

Cheslack Postava, K., Fallin, MD., Avramopoulos, D., Connors, SL., Zimmerman, AW., Eberhart, CG., Newschaffer, CJ.(2007). beta2-Adrenergic receptor gene variants and risk for autism in the AGRE cohort. *Mol Psychiatry* .283-91.

Cook EH, Jr., Courchesne, RY., Cox, NJ., Lord ,C., Gonen, D., Guter, SJ., Lincoln, A., Nix ,K., Haas, R., Leventhal, BL., Courchesne, E.(1998). Linkage-disequilibrium mapping of autistic disorder, with 15q11-13 markers. *American Journal of Human Genetics*. 1077-83

Dahlgren, S. O., & Gillberg, C. (1989). Symptoms in the first two years of life: A preliminary population study of infantile autism. *European Archives of Psychiatric and Neurological Science*, 283, 169–174.

Daley, T. C., & Sigman, M. D. (2002). Diagnostic conceptualization of autism among Indian psychiatrists, psychologists, and pediatricians. *Journal of Autism and Developmental Disorders*, 32, 12–23.

Folstein, S., Rutter, M.(1997) Genetic influences and infantile autism. *Nature* 265: 726-728.

Glessner, J. T., Wang, K., Cai, G., Korvatska, O., Kim, C. E., Wood, S., Zhang, H., Estes, A., Brune, C. W., Bradfield, J. P., Imielinski, M., Frackelton, E. C., and 47 others.(2009) Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. *Nature* 459: 569-573.

Gong, .X, Jia M, Ruan, Y., Shuang, M., Liu, J., Wu, S., Guo, Y., Yang, J., Ling, Y., Yang, X., Zhang, D.(2004) Association between the FOXP2 gene and autistic disorder in Chinese population. *American Journal of Medical Genetics B Neuropsychiatric Genetics*.113-6.

Hallmayer, J., Glasson, E. J., Bower, C., Petterson, B., Croen, L., Grether, J., Risch, N. On the twin risk in autism.(2002) *American Journal of Human Genetics*. 71: 941-946.

Hallmayer, J., Hebert, J. M., Spiker, D., Lotspeich, L., McMahon, W. M., Petersen, P. B., Nicholas, P., Pingree, C., Lin, A. A., Cavalli-Sforza, L. L., Risch, N., Ciaranello, R. D. (1996)Autism and the X chromosome: multipoint sib-pair analysis. *Arch. Gen. Psychiat.* 53: 985-989.

Hedges, DJ., Hamilton-Nelson, KL., Sacharow, SJ., Nations, L., Beecham, GW., Kozhekbaeva, ZM., Butler, BL., Cukier, HN., Whitehead, PL., Ma, D., Jaworski,

JM.,Nathanson, L., Lee, JM., Hauser, SL., Oksenberg, JR., Cuccaro, ML., Haines, JL., Gilbert, JR., Pericak-Vance, MA.(2012). Evidence of novel fine-scale structural variation at autism spectrum disorder candidate loci. *Molecular Autism*.

Henningsson, S., Jonsson, L., Ljunggren, E., Westberg, L., Gillberg, C., Råstam, M., Anckarsäter, H., Nygren, G., Landén, M., Thuresson ,K., Betancur, C., Leboyer, M.,Gillberg, C., Eriksson, E., Melke, J.(2009) Possible association between the androgen receptor gene and autism spectrum disorder. *Psychoneuroendocrinology* .752-61

Hrdlicka, M., Komarek, V., Propper, L., Kulisek, R., Zumrova, A., & Faladova, L. (2004). Not EEG abnormalities but epilepsy is associated with autistic regression and mental functioning in childhood autism. *European Child and Adolescent Psychiatry*, 13, 209–213.

Ingram, JL., Stodgell, CJ., Hyman, SL., Figlewicz, DA., Weitkamp, LR., Rodier, PM.(2000) Discovery of allelic variants of HOXA1 and HOXB1: genetic susceptibility to autism spectrum disorders. *Teratology*. 393-405.

Junaid, MA., Kowal, D., Barua, M., Pullarkat, PS., Sklower Brooks, S., Pullarkat, RK.(2004) Proteomic studies identified a single nucleotide polymorphism in glyoxalase I as autism susceptibility factor. *American Journal of Medical Genetics* A.11-7.

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child*, 2, 217–253.

Kilpinen, H., Ylisaukko-Oja, T., Hennah, W., Palo, OM., Varilo, T., Vanhala, R., Nieminen-von Wendt, T., von Wendt,. L, Paunio, T., Peltonen, L.(2008) Association of DISC1 with autism and Asperger syndrome. *Molecular Psychiatry* .187-96.

Kim, H.-G., Kishikawa, S., Higgins, A. W., Seong, I.-S., Donovan, D. J., Shen, Y., Lally, E., Weiss, L. A., Najm, J., Kutsche, K., Descartes, M., Holt, L., and 14 others.(2008) Disruption of neurexin 1 associated with autism spectrum disorder. Am. J. Hum. Genet. 82: 199-207.

Kuemerle, B., Gulden, F., Cherosky, N., Williams, E., Herrup, K.(2007) The mouse Engrailed genes: a window into autism. Behav Brain Res. 176(1):121-32.

Lamb, J. A., Moore, J., Bailey, A., Monaco, A. P. Autism: recent molecular genetic advances. *Human Molecular Genetics*. 9: 861-868, 2000.

Laumonnier, F., Roger, S., Guérin, P., Molinari, F., M'rad, R., Cahard, D., Belhadj, A., Halayem, M., Persico, AM., Elia., M, Romano, V., Holbert, S., Andres, C., Chaabouni, H., Colleaux, L., Constant, J., Le Guennec, JY., Briault, S. (2006) Association of a functional deficit of the BKCa channel, a synaptic regulator of neuronal excitability, with autism and mental retrdation. *American Journal of Psychiatry*. 1622-9.

Le Couteur, A., Bailey, A., Rutter, M., & Gottesman, I. (1989),. Epidemiologically based twin study of autism. Paper presented at the First World Congress on Psychiatric Genetics, Churchill College, Cambridge, England.

Liu, J., Nyholt, D. R., Magnussen, P., Parano, E., Pavone, P., Geschwind, D., Lord, C., Iversen, P., Hoh, J., Autism Genetic Resource Exchange Consortium, Ott, J., Gilliam, T. C.(2001) A genomewide screen for autism susceptibility loci. Am. J. Hum. Genet. 69: 327-340, 2001. Note: Erratum: *American Journal of Human Genetics*. 69: 672 only.

Lotter, V. (1966). Epidemiology of autistic conditions in young children: I. Prevalence. *Social Psychiatry* 1, 124±137.

Lotter, V. (1967). Epidemiology of autistic conditions in young children: II. Some characteristics of the parents and children. *Social Psychiatry* , 4, 163-173.

Ma, D., Salyakina, D., Jaworski, J. M., Konidari, I., Whitehead, P. L., Andersen, A. N., Hoffman, J. D., Slifer, S. H., Hedges, D. J., Cukier, H. N., Griswold, A. J., McCauley, J. L., and 9 others.(2009) A genome-wide association study of autism reveals a common novel risk locus at 5p14.1. *American Journal of Human Genetics*. 73: 263-273.

Ma, D.; Salyakina, D.; Jaworski, J. M.; Konidari, I.; Whitehead, P. L.; Andersen, A. N.; Hoffman, J. D.; Slifer, S. H.; Hedges, D. J.; Cukier, H. N.; Griswold, A. J.; McCauley, J. L.; and 9 others.(2009).A genome-wide association study of autism reveals a common novel risk locus at 5p14.1. *American Journal of Human Genetics*. 73: 263-273.

Maestrini, E., Lai, C., Marlow, A., Matthews, N., Wallace, S., Bailey, A., Cook, E. H., Weeks, D. E., Monaco, A. P.(1999) International Molecular Genetic Study of Autism (IMGSA) Consortium. Serotonin transporter (5-HTT) and gamma-aminobutyric acid receptor subunit beta-3 (GABRB3) gene polymorphisms are

not associated with autism in the IMGSA families. Am. J. Med. Genet. (Neuropsychiat. Genet.) 88B: 492-496.

Marshall, CR., Noor, A., Vincent, JB., Lionel, AC., Feuk, L., Skaug, J., Shago, M., Moessner ,R., Pinto,D., Ren,Y.,Thiruvahindrapduram,B., Fiebig,A., Schreiber,S., Friedman,J., Ketelaars,CE., Vos,YJ., Ficicioglu,C., Kirkpatrick,S., Nicolson,R., Sloman,L., Summers,A., Gibbons, CA., Teebi,A., Chitayat, D., Weksberg, R., Thompson, A., Vardy, C.,Crosbie, V., Luscombe, S., Baatjes,R., Zwaigenbaum,L., Roberts,W., Fernandez,B., Szatmari, P., Scherer, SW.(2008) Structural variation of chromosomes in autism spectrum disorder. *American Journal of Human Genetics*. 477-88.

Marui, T., Hashimoto, O., Nanba, E., Kato, C., Tochigi, M., Umekage, T., Ishijima, M., Kohda, K., Kato, N., Sasaki, T.(2004) Association between the neurofibromatosis-1 (NF1) locus and autism in the Japanese population. *American Journal of Medical Genetics B Neuropsychiatric Genetics*. 43-7.

McCauley, JL., Li ,C., Jiang, L., Olson, LM., Crockett, G., Gainer, K., Folstein., SE, Haines ,JL., Sutcliffe, JS. (2005).Genome-wide and Ordered Subset linkage analyses provide support for autism loci on 17q and 19p with evidenceof phenotypic and interlocus genetic correlates. *BMC Med Genetics*.6:1.

Ozgen , HM., van Daalen, E., Bolton, PF., Maloney, VK., Huang, S., Cresswell, L., van den Boogaard, MJ., Eleveld, MJ., van 't Slot, R., Hochstenbach, R., Beemer FA, Barrow M, Barber JC, Poot M.(2009) Copy number changes of the microcephalin 1 gene (MCPH1) in patients with autism spectrum disorders. *Clinical Genetics*. 348-56

Ozonoff, S., Williams, B. J., & Landa, R. (2005). Parental report of the early development of children with regressive autism. *Autism*, 9, 461–486

Philippi, A., Tores, F., Carayol, J., Rousseau, F., Letexier, M., Roschmann, E., Lindenbaum ,P., Benajjou, A., Fontaine, K., Vazart, C., Gesnouin, P., Brooks. P., Hager, J.(2007) Association of autism with polymorphisms inthe paired like homeodomain transcription factor 1 (PITX1)on chromosome 5q31: a candidate gene analysis. *BMC Med Genetics*. 8:74.

Pickles, A., Bolton, P., Macdonald, H., Bailey, A., Le Couteur, A., Sim, C.-H., Rutter, M. (1995)Latent-class analysis of recurrence risks for complex phenotypes with selection and measurement error: a twin and family history study of autism. *American Journal of Human Genetics*. 57: 717-726.

Rogers, S. J. (2004). Developmental regression in autism spectrum disorders. *Mental Retardation & Developmental Disabilities Research Reviews*, 10, 139–143.

Ronald, A., Butcher, LM., Docherty, S., Davis, OS., Schalkwyk, LC., Craig, IW., Plomin R.(2010) A genome-wide association study of social and non-social autistic-like traits in the general population usingpooled DNA, 500 K SNP microarrays and both community and diagnosed autism replication samples. Behaviour Genetics. 40(1):31-45. doi: 10.1007/s10519-009-9308-6.

Short, C. B., & Schopler, E. (1988). Factors relating to age of onset in autism. Journal

Smalley, S. L. Genetic influences in childhood-onset psychiatric disorders: autism and attention-deficit/hyperactivity disorder.(1997) *American Journal of Human Genetics*. 60: 1276-1282.

Strom, SP., Stone, JL., Ten Bosch, JR., Merriman, B., Cantor, RM., Geschwind, DH., Nelson, SF.(2010) High-density SNP association study of the 17q21 chromosomal region linked to autism identifies CACNA1G as a novel candidate gene. *Molecular Psychiatry*. 996-1005.

Vernes, S. C.; Newbury, D. F.; Abrahams, B. S.; Winchester, L.; Nicod, J.; Groszer, M.; Alarcon, M.; Oliver, P. L.; Davies, K. E.; Geschwind, D. H.; Monaco, A. P.; Fisher, S. E.(2008).A functional genetic link between distinct developmental language disorders. *New English Journal of Medicine*. 359: 2337-2345.

Vincent, J. B., Horike, S. I., Choufani, S., Paterson, A. D., Roberts, W., Szatmari, P., Weksberg, R., Fernandez, B., Scherer, S. W.(2006) An inversion inv(4)(p12-p15.3) in autistic siblings implicates the 4p GABA receptor gene cluster. J. Med. Genet. 43: 429-434.

Wang, K., Zhang, H., Ma, D., Bucan, M., Glessner, J. T., Abrahams, B. S., Salyakina, D., Imielinski, M., Bradfield, J. P., Sleiman, P. M. A., Kim, C. E., Hou, C., and 44 others. (2009) Common genetic variants on 5p14.1 associate with autism spectrum disorders. *Nature*. 459: 528-533.

Weiss, LA., Kosova, G., Delahanty, RJ., Jiang ,L., Cook, EH, Ober, C., Sutcliffe, JS.(2006) Variation in ITGB3 is associated with whole-blood serotonin level and autism susceptibility. European Journal of Human Genetics .923-31.

Yirmiya, N., Rosenberg, C., Levi, S., Salomon, S., Shulman, C., Nemanov, L., Dina, C., Ebstein, RP. (2006) Association between the arginine vasopressin 1a receptor (AVPR1a) gene and autism in a family-based study: mediation by socialization skills. *Molecular Psychiatry* .488-94.

Zheng, F., Wang, L., Jia, M., Yue, W., Ruan, Y., Lu, T., Liu, J., Li, J., Zhang, D.(2011). Evidence for association between Disrupted-in-Schizophrenia 1 (DISC1) gene polymorphisms and autism in Chinese Han population: a family-based association study. *Behaviour Brain Function*. 7:14. doi: 10.1186/1744-9081-7-14.

Dutta, S., Sinha, S., Ghosh, S., Chatterjee, A., Ahmed, S., Usha, R.(2008) Genetic analysis of reelin gene (RELN) SNPs: no association with autism spectrum disorder in the Indian population. *Neuroscience Letter*. 441(1):56-60.

Dutta, S., Das, S., Guhathakurta,S., Sen, B., Sinha, S., Chatterjee, A., Ghosh, S., Ahmed, S., Ghosh, S., Usha, R.(2007) Glutamate receptor 6 gene (GluR6 or GRIK2) polymorphisms in the Indian population: a genetic association study on autism spectrum disorder. *Cell Mol Neurobiology*. 1035-47.

Guhathakurta, S.,Ghosh, S., Sinha, S., Chatterjee, A., Ahmed, S., Chowdhury, SR., Gangopadhyay ,PK., Ghosh, S., Singh, M., Usha, R.(2006) Serotonin transporter promoter variants: Analysis in Indian autistic and control population. *Brain Research*. 28-35.

Biological Data Base –

- NCBI (http://www.ncbi.nlm.nih.gov/)
- Ensembl (http://asia.ensembl.org/index.html)
- Autdb (http://autism.mindspec.org/autdb/Welcome.do)

APPENDIX

Questionnaire

Case No:

Name:

ASD Type:

Sex: M ( ) F ( )

Age: Age of Identification:

Date of Birth: Place of Birth:

Height of child: Weight of the child:

Present Address:	**Native Address:**
Contact No:	

Religion:

Caste: Sub caste:

How many hours the child watch TV daily:

i. Residential Information:

Own house/flat/rental etc:

Nearby Industry/factory etc:

(Mention the type of industry)

Nearby any plantations:

(Chance of Pesticides/if you know the type, explain?)

ii. Family History:

Joint: ( ) Nuclear: ( )

Consanguinity (If yes explain the relationship): Yes ( ) No ( )

iii. Parents Information:

	Father	Mother
Name		
Age		
Occupation & Salary per Month		
Place of Occupation, Since How long?		
Any Psychiatric illnesses?		
Any Other Health problems like Diabetes,CVD,Hypertension,Blood Pressure etc?		
Do you Smoke/Alcoholic? If yes Since how long?		
Usage of Medicine per day? Reason?		
Time you spend with child daily?		
Any Other?		

Sibling Information:

No of Siblings:

Age/Sex:

Any Health problems:

iv. Family Size:

Sl. No	***Name of the member***	***Relation with Child***	***M/F***	***Age***	***Marital status***	***Education***	***Occupation***	***Health Status***
1								
2								
3								
4								
5								
6								
7								
8								
9								
10								

v. Vaccination History:

Type of vaccine: At which Place:

Date/Year:

vi. Diet(In family):

Vegetarian ( ) Mixed ( )

	Normal Days	During Pregnancy
Menu		
Any Fast Food (Specify)		

vii. Pre-natal (during pregnancy) History:

Maternal infection: Yes ( ) No ( )
If Yes, Type of Infection:

Gestational diabetes: Yes ( ) No ( )

Thyroid problems: Yes ( ) No ( )

Excessive hygiene: Yes ( ) No ( )

Stress: Yes ( ) No ( )

Any other Remarkable Incident Yes ( ) No ( )

If Yes; explain,

Did you take any medication (drugs) daily? Yes ( ) No ( )
If Yes, State daily dosage:

Type of medication: Homoeopathy ( ) allopathic ( ) Traditional (Ayurveda) ( )

viii. Breast Feeding:

No Breast Feeding ( )	Breast feeding <2 months ( )
Breastfeeding 2–6 months ( )	Breast feeding >6 months ( )

Have you sent hair, blood or any other sample of your child to a Laboratory for testing? If so, please give details:

Remarks:

Blood sample Ref. No:

Signature of the investigator with Date: