

## **Background and Significance**

Autistic spectrum disorders (ASDs) represent a neurobiologically diverse group of conditions with heterogeneous clinical representation. Characteristic signs of patients with ASD include (1) social impairment, (2) language or communication deficits, and (3) patterns of restricted and/or repetitive behavior. Beyond these core symptoms, autism can present itself with highly variable pathological symptoms from one individual to the next, representing a continual spectrum rather than discrete diagnostic categories<sup>6</sup>.

There is a clear genetic predisposition to idiopathic autism, beginning with a disproportion of cases in males compared to females (4:1), suggesting X-linked inheritance and/or genetic imprinting. Furthermore, studies of monozygotic and dizygotic twins show that autism is highly heritable with a monozygotic concordance rate between 70-90% and a dizygotic rate around 7%. A sibling recurrence risk of about 5-6% is also suggestive of genetic predisposition. The data are consistent with a multifactorial inheritance pattern, where multiple genetic loci, genetic heterogeneity, epistasis and gene-environment interactions contribute to the pathology of the diseases. In recent years, epidemiological studies indicate that the incidence of autism and associated disorders has increased approximately ten-fold to a rate of 20-60 in 10,000<sup>6</sup>. The increase is thought to be due to a combination of the changes in diagnostic and reporting practices as well as environmental and epigenetic factors acting on a genetically predisposed population.

There are a number of characterized syndromes whose pathophysiology lie within the autism spectrum include Asperger, Fragile X, Angelman, Rett, Williams, Prader-Willi, Hyperlexia, and Landau-Kleffner syndrome. However, only about 15% of diagnosed cases of autism have an underlying cause identified<sup>7</sup>. There is a strong belief, and some recent evidence suggesting an incidence of concurrent symptoms that may indicate clusters of conditions that stem from the same root cause<sup>10</sup>. Genetic studies so far have identified a handful of single gene mutations in a small percentage of the ASD population, generally thought to represent about 1% of all cases (Figure 1)

The extensive heterogeneity observed among patients with nonsyndromic ASD presents major difficulties from the clinical standpoint of providing a diagnosis and effective treatment, as well as from the research standpoint of understanding the biological causes and genetic contributions. Here we propose that through a combination of extensive and consistent phenotypic clinical evaluation and the use of high-throughput comparative genomic hybridization microarrays, we can classify subgroups within the ASDs that will allow us to identify multigenetic factors leading to a specific group of symptoms.

## **Experimental design**

The first stage of the research project will require a comprehensive patient workup by qualified physicians specializing in the area of neurological disorders. It is important that the same standards of evaluation be set for all patients. Patients will be recruited through the UCSD Autism Research Program and pediatric neurology department at Rady Children's Hospital in San Diego. The subjects must be evaluated based on the standardized criterion outlined here. This will include a medical evaluation, behavioral and psychological evaluation, metabolic profiling, immunological evaluation, imaging of brain structure, electrical brain activity, and extensive genetic workup.

- Basic medical evaluation – physicians will obtain standard physical measurements, such as height, weight, blood pressure, and head circumference. Close attention will be paid to any signs of dysmorphology, as well as visual and auditory abnormalities. In addition, physicians will document the medical history of the patient including medications, vaccination records, and physical illnesses, such as gastrointestinal problems and food intolerance, which have shown some prevalence in patients with ASD<sup>10</sup>.
- Psychological evaluation – This part will include an assessment of characteristic behavioral patterns in the patient, including specific social skills, progression of language development, and patterns of aggression and mood disorders. We will also look at trends in IQ as a measure of mental function.
- Metabolic profiling – Some evidence suggests that patients with ASD exhibit abnormal patterns of oxidative stress and methylation capacity<sup>5</sup>. A blood sample will be drawn to measure metabolites in the

methionine transmethylation and transsulfuration pathways (i.e. methionine, S-adenosylmethionine, adenosine, etc.), as well as the ratio of oxidized to reduced glutathione.

- Immunological evaluation – A possible connection between the immune system and central nervous system may influence the development of ASD as early on as embryogenesis<sup>1</sup>. Patients will be monitored for aberrant immune activity, including abnormal T helper cell type I cytokine profiles, decreased lymphocyte numbers, decreased T cell mitogen response and imbalance of serum immunoglobulin levels.
- Neuroanatomical imaging – Eric Courchesne, PhD., at UCSD has been a pioneer in studying neuroanatomical features of children with autism. His work has identified various anatomical abnormalities associated with the disease using structural and functional MRI techniques. All patients will be evaluated in the Courchesne lab under the current protocol.
- Electrical brain activity – Autism has a relatively high incidence of comorbid epilepsy (prevalence 0-45%) and epileptiform EEG abnormalities (prevalence 10-70%)<sup>4</sup>. It has already been suggested that this represents a subgroup of autistic children. We will conduct EEG studies to measure brain activity and monitor for signs of epilepsy and seizure.
- Genomic profiling – Here is where we will be assaying for the unknown. Patients DNA will be analyzed by comparative genome hybridization (CGH) microarray techniques. There are three types of arrays to be considered. Initially we will focus on the bacterial artificial chromosome (BAC) arrays and oligonucleotide arrays, which can assay for microdeletions, micro duplications, or copy number variants (CNVs). Alternatively or in addition we may also look for single nucleotide polymorphisms using the SNP arrays, but this will ultimately require much more in depth data processing techniques.

Data will be collected from a cohort of autistic patients, along with a control set from the general population. We aim to collect 100+ affected samples. When possible we will also collect full datasets from the parents of affected patients. Based on the phenotype data we will identify subgroups within the affected individuals that present the same symptoms. Multiple sub-groupings will be analyzed. Some concurrent phenotypes have already been suggested in previous studies<sup>4,10</sup>, which will be considered here. To identify candidate genes from the CGH microarray data we will perform whole-genome linkage analysis using standard techniques. In addition to considering multiple sub-groupings of the affected individuals, if a parent of parents exhibit a phenotype characterized within the subgroup of autistic patients, we will compare the results when including the parental data in the analysis.

In addition, or as an alternative approach, we can use a machine learning algorithm for classification of subgroups and linkage analysis. Powerful new statistical data mining approaches allow for rapid data processing, as well as the ability to pull out, otherwise obscure, information from large data sets. Many data clustering techniques have been developed and applied for the purpose of predictive proteomics<sup>2</sup>. The use of these techniques on large microarray datasets is proving efficacious. In collaboration with the super computer center and systems biology department at UCSD we will use supervised and/or unsupervised machine-learning algorithms to analyze the genomic microarray data for predictive markers and to further locate interesting subgroups within the data. Support vector machines and Random Forest™ decision trees are two algorithms that have been successfully used in the classification of microarray data, however other algorithms may be applicable<sup>8</sup>. Software for setting up the data analysis is generally available to the public or through special request. The use of powerful machine-learning programs will be necessary when moving into SNP data analysis, as the datasets for these microarrays generally provide millions of results. The ability to pull out the important hits will require the computational power and speed of these algorithms.

## **Conclusion**

As the prevalence of ASDs increases, so does our need to understand the biological basis for the disease(s), as well as the potential for therapeutics. The wide diversity in phenotypic outcomes of the disease is suggestive of multiple etiologies that may result from variable multigenomic underpinnings. Further classification of the autistic disorders into distinct subgroups will allow us a foundation for future research on the pathological basis of the disease, as well as the potential to develop better treatment plans and novel therapeutics.

<b>Candidate Gene(s)</b>	<b>Reference</b>	<b>Mutations/deletions</b>	<b>Occurrence</b>	<b>Clinical Phenotype</b>
<b>Neurologin 3 &amp; 4</b>	Jamain <i>et al.</i>	NLGN4 D396X, NLGN3 R451C	2/158	Autism or Asperger syndrome
	Blasi <i>et al.</i>	-	0/124	-
	Vincent <i>et al.</i>	-	0/196	-
	Laumonnier <i>et al.</i>	NLGN4 D429X	13/13 (same family)	Mental retardation, autism, PDD-NOS
	Ylisaukko-oja <i>et al.</i>	-	0/30	-
	Yan <i>et al.</i>	NLGN4 G99s, K378R 403M r704C	3/148	mild-severe autism, PDD-NOS
	Gauthier <i>et al.</i>	-	0/96	-
	Wermter <i>et al.</i>	-	0/107	-
	Lawson-Yuen <i>et al.</i>	NLGN4 del exons 4,5,6	1/1	Autism w/ motor tics
<b>SHANK3</b>	Durand <i>et al.</i>	142 kb and 800 kb del at 22q13, E409X	3/227	Autism with severe language and social deficits
	Moessner <i>et al.</i>	227 kb, 3.2 Mb, & 4.36 Mb del at 22q13, Q321R	4/400	Autism with non-verbal communication and social deficits
<b>Neurexin1</b>	Autism Genome Project Consortium	300 kb del at 2p16	2/196	Autism with language deficits
	Feng <i>et al.</i>	-	0/192	Autism with seizures and facial dysmorphism
	Kim <i>et al.</i>	-	0/57	-
	Yan <i>et al.</i>	-	0/116	-
<b>MECP2</b>	Lam <i>et al.</i>	IVS2+ 2delTAAG	1/21	Autism and MR with no regression, epilepsy or microcephaly
	Vourc'h <i>et al.</i>	-	0/59	-
	Beyer <i>et al.</i>	-	0/202	-
	Carney <i>et al.</i>	1157del41, R294X	2/69	Autism, MR and history of regression
	Zappella <i>et al.</i>	R133C R453X	2/19	Variant of Rett syndrome
	Shibayama <i>et al.</i>	-	0/24	-
	Lobo-Menendez <i>et al.</i>	-	0/99	-
	Li <i>et al.</i>	-	0/65	-
	Xi <i>et al.</i>	-	0/31	-
	Harvey <i>et al.</i>	-	0/401	-
Coutinho <i>et al.</i>	-	0/172	-	
<b>PTEN</b>	Butler <i>et al.</i>	H93R, D252G, F241S	3/18	Extreme macrocephaly and macrosomy
	Herman <i>et al.</i>	530 insT; R130X	2/71	Macrocephaly, autism, developmental delay
	Buxbaum <i>et al.</i>	D326N	1/88	Macrocephaly, polydactyly, autism, MR, language delay
<b>TOTAL</b>			<b>39/3400</b> <b>(1.1%)</b>	

Figure 1: Summary of genetic studies on autistic phenotypes adapted from Lintas *et al.*<sup>6</sup>

## References

- <sup>1</sup> Ashwood P, Wills S, Van de Water J. The immune response in autism: a new frontier for autism research. *J Leukoc Biol.* 2006 Jul;80(1):1-15. Epub 2006 May 12.
- <sup>2</sup> Barla A, Jurman G, Riccadonna S, Merler S, Chierici M, Furlanello C. Machine learning methods for predictive proteomics. *Brief Bioinform.* 2008 Mar;9(2):119-28. Epub 2008 Feb 29.
- <sup>3</sup> Caronna EB, Milunsky JM, Tager-Flusberg H. Autism spectrum disorders: clinical and research frontiers. *Arch Dis Child.* 2008 Jun;93(6):518-23. Epub 2008 Feb 27.
- <sup>4</sup> Hrdlicka M. EEG abnormalities, epilepsy and regression in autism: a review. *Neuro Endocrinol Lett.* 2008 Aug;29(4):405-9.
- <sup>5</sup> James SJ, Cutler P, Melnyk S, Jernigan S, Janak L, Gaylor DW, Neubrandner JA. Metabolic biomarkers of increased oxidative stress and impaired methylation capacity in children with autism. *Am J Clin Nutr.* 2004 Dec;80(6):1611-7.
- <sup>6</sup> Lintas C, Persico AM. Autistic phenotypes and genetic testing: state-of-the-art for the clinical geneticist. *J Med Genet.* 2009 Jan;46(1):1-8. Epub 2008 Aug 26.
- <sup>7</sup> Mendelsohn NJ, Schaefer GB. Genetic evaluation of autism. *Semin Pediatr Neurol.* 2008 Mar;15(1):27-31.
- <sup>8</sup> Resson HW, Varghese RS, Zhang Z, Xuan J, Clarke R. Classification algorithms for phenotype prediction in genomics and proteomics. *Front Biosci.* 2008 Jan 1;13:691-708.
- <sup>9</sup> Schaefer GB, Mendelsohn NJ. Genetics evaluation for the etiologic diagnosis of autism spectrum disorders. *Genet Med.* 2008 Jan;10(1):4-12.
- <sup>10</sup> Xue Ming, Brimacombe M, Chaaban J, Zimmerman-Bier B, Wagner GC. Autism spectrum disorders: concurrent clinical disorders. *J Child Neurol.* 2008 Jan;23(1):6-13.