Children's Hospital of Philadelphia

Annual Progress Report: 2008 Nonformula Grant

Reporting Period

July 1, 2009 – June 30, 2010

Nonformula Grant Overview

The Children's Hospital of Philadelphia received $4,708,555 in nonformula funds for the grant award period June 1, 2009 through May 31, 2013. Accomplishments for the reporting period are described below.

Research Project: Project Title and Purpose

CHOP/Penn Center of Excellence for Autism Research - This grant establishes a Center of Excellence for Autism Research at Children's Hospital of Philadelphia (CHOP), in collaboration with the University of Pennsylvania, Temple University, Lincoln University and the Philadelphia Public School system. The Center will engage in a set of closely inter-related studies aimed at improving our understanding of the biological causes of the autism spectrum disorders (ASD), so that more effective treatments can be created. We have new evidence of the first gene for autism, and now need to study this gene in more detail. We will study this gene in a group of 400 children, performing detailed clinical evaluations to understand how this gene affects their brains and behavior. We are also performing a study on the effectiveness of a behavioral treatment, and will test how the gene affects treatment response.

Anticipated Duration of Project

6/1/2008 - 5/31/2013

Project Overview

This funding will support a Center of Excellence for Autism Research located at Children's Hospital of Philadelphia (CHOP), with partnerships at Lincoln and Temple Universities, the University of Pennsylvania, and the Philadelphia Public School System. This project's central theme is understanding the relationship between genetic risk factors, brain and behavior (including treatment response) in persons with an autism spectrum disorder (ASD). We will conduct a series of integrated subprojects that follow upon several recent breakthroughs in the genetics of ASD. We have identified the first common genetic variant for ASD. Our data also suggest that rare sequence and structural variants serve as significant risk factors for ASD. The genes appear to have a common biological function related to synaptic connectivity. We aim to study 400 children with ASD, and to investigate the relationship between genetic risk factors and clinical presentation, brain function, and response to treatment. Individual subprojects in the Center will focus on our common variant finding (Project I), rare genetic variant findings

(Project II), animal knockout models of risk variants replete with social behavioral and brain assays (Project III), comprehensive assessment ("deep phenotyping") of the ASD sample to assess the clinical impact of genetic risk factors, including treatment responsiveness (Project IV), and measures of the "social brain" using magnetic resonance imaging and magnetoencephalography (Project V). Neuroimaging will include several measures of connectivity (including diffusion tensor imaging, functional MRI connectivity, and coherence analysis with magnetoencephalography), as we aim to relate connectivity to genetic risk factors. Participants with an ASD from Project IV will be enrolled in projects I, II, and V, so as to provide a set of tightly inter-related experiments. Our animal models will knock out the rare and common genetic risk factors, and perform behavioral and imaging studies that parallel the human studies.

In addition, we will create undergraduate and graduate level training programs for underrepresented minorities in collaboration with the University of Pennsylvania, Temple University and Lincoln University (Project VI). Students will work within the labs of researchers on this grant, and will receive additional group and individual teaching and mentorship to prepare them for professional work in the area of autism.

Principal Investigator

Robert T. Schultz, PhD
Director, Center for Autism Research
The Children's Hospital of Philadelphia
Joseph Stokes Jr. Research Institute
3615 Civic Center Boulevard
Philadelphia, PA 19104-4318

Other Participating Researchers

Manzar Ashtari, PhD, Peter Doehring, PhD, Hakon Hakonarson, MD, PhD, John Herrington, PhD, Timothy Roberts, PhD - employed by The Children's Hospital of Philadelphia
Edwin Abel, PhD, Edward Brodkin, MD, Maja Bucan, PhD, Christos Davatzikos, PhD, David Mandell, ScD, Steve Marcus, PhD, Ragini Verma, PhD - employed by University of Pennsylvania.
Nicole Stephens, PhD - employed by Lincoln University

Expected Research Outcomes and Benefits

Autism spectrum disorders are caused by a combination of genetic and nongenetic factors. The role of genetics is particularly strong. We have new genetic findings that we will explore in great depth. The integrated set of studies in this project will focus on understanding how these newly discovered genetic risk factors affect the health of children with autism. We will carefully measure the clinical profile of each child in our study, and relate different genetic risk factors to differences in these profiles. We will also perform comprehensive brain imaging studies and relate genetic risk factors to underlying problems in brain structure and function. One immediate health benefit of the research process is that we will give expert clinical evaluations and

feedback free of charge to a large community of individuals with autism and their families. In addition, we expect to make scientific progress that will significantly contribute to the health of children by providing an improved understanding of the genetic factors that initiate this disorder; the findings will serve to identify molecular mechanisms critical to regulating the clinical presentation of ASD, and thereby facilitate the development of effective diagnostic tests. In turn, this could lead to earlier diagnosis and a good possibility for the development of effective new therapeutic interventions. Because a significant proportion of the study subjects will have multiple risk factors that will markedly amplify their genetic risk of developing autism, we will also be able to determine the cumulative risk of autism incurred from having more than one genetic risk factor.

Summary of Research Completed

Considerable progress was made in the first full year of funding for this Center of Excellence for Autism Research. All IRB protocols have been approved. Project PIs and research staff met at Children's Hospital of Philadelphia (CHOP) in early February 2010 for a semi-annual Center retreat. Progress was reviewed systematically on a project-by-project basis and discussion led to further enhancements to ongoing research. This Center is composed of 6 inter-related projects. Progress is reviewed for each Project in order, below.

Project I: Characterizing the Common Intergenic Variant in ASD**.** Sequencing experiments to identify and characterize the causal variant(s) in *the CDH9/CDH10* intergenic region on 5p14 have begun. The preparation phase for these studies was labor intensive and included isolation and amplification of the chromosomal region corresponding to the linkage disequilibrium block harboring the ASD risk locus, plus the two flanking cadherin genes, and the synthesis of the sequencing library for the region. Researchers selected 144 well characterized ASD samples, representing all genotype states, including homozygous risk, homozygous wild type and heterozygous states. Sequences from equal number of controls have been accessed and downloaded from the1000 genome project for comparison. The number of unique amplicons for the 5p14 locus (split into 4 main regions) included: 1) 624 amplicons for 187,063 bps; 2) 312 for 93,531 bps; 3) 526 for 157, 702 bps and 4) 527 for 157 bps. Researchers then included six samples per pool (10ug total DNA) and prepared 24 pools, totaling 144 samples or 288 chromosomes. Additional selection criteria included: cases with transmitted major (risk) allele; cases with large CNVs (for development of CNV algorithm, and for mapping breakpoints); all subjects of European ancestry; sequencing coverage over 30x. These samples are currently being sequenced on the Illumina Genome Analyzer, at 75 bp read lengths using paired-end reads. Researchers have already generated and published a novel algorithm to most optimally align and annotate these sequences (ANNOVAR).

Researchers have taken additional 70 ASD samples and performed focused sequencing of the 500 kb region centered around the most significant association, which includes one of the highest conserved regions in the human genome. Using the Illumina Genome Analyzer, paired-end reads, indexing 10 samples per lane on a flow cell gives us ~360Mb per lane (with usable sequence around ~240Mb per lane), so this effectively achieves ~50X fold coverage of the targeted genomic region. Researchers include complete pedigree for detailed segregation analysis (non-pooled design used in sequencing). In total, there are 212 samples. Researchers are

also completing the sequencing of two samples for a complete genome sequence (3 billion bps), searching for a Mendelian phenocopy in an extended ASD family that has been fully genotyped.

Project II: Understanding the contributions of rare variants. Work has focused on the Specific Aim 2.1. In order to prioritize genomic regions for resequencing, researchers performed comparative sequence analysis on 100 ASD candidate genes, selected based on the family-based genome-wide association analysis of CNVs and published reports on candidate gene approaches: Genomic regions with the strongest associations identified by family-based genome-wide analysis; ASD-associated CNV regions; published data from the SFARI Gene (gene.sfari.org/). Since the majority of ASD-associated CNVs and SNPs have been found in intergenic regions, the entire flanking region (up to the next gene) was included for each of the genes in the list (total of 164 MB) constituting 188 genomic regions of interest. For each gene researchers mapped conserved non-coding elements in intergenic regions, small non-coding and alternatively spliced RNAs, and sequences identified in the analysis of epigenetic modifications. Within highly conserved sequences, researchers annotated potential enhancers and chromatin modification marks. Furthermore, researchers showed that 20% of conserved sequences in non-protein coding regions corresponded to RNA detected in a published RNA-Seq experiment These annotations, combined with the analysis of gene expression (GNF Atlas2: biogps.gnf.org/) resulted in the list of 100 genomic regions (3.5 Mb) to be used for targeted deep resequencing of ASD individuals. Researchers have submitted selected genomic regions to the Agilent and recently received the Sure Select-target enrichment kit (for 3.5 Mb) as well as Agilent exome kit. Researchers have selected 10 ASD samples (individuals from the Autism Genetic Resource Exchange) to test selected prepared panels. Deep sequencing on 3.5 Mb of data, encompassing 100 ASD-associated genes and all exons in 10 individuals will permit identification of combinations of putative rare risk factors that will be tested as a new sample.

In anticipation of DNA sequence data on ASD subjects, researchers used 3.5 MB of sequence around 100 ASD-associated genes to identify rare variants in the 1000 Genome project (pilot 1 that consisted of SNP calls for 60 Cauasians (CEU), 59 Yoruban, 30 Han Chinese and 30 Japanese individuals). This analysis identified approximately 800 SNPs within coding regions. Moreover, approximately 100 variants map within deleted (hemizygous) regions and are predicted to be "damaging" using Polyphen. For example we observed SNPs, unique to CEU population, that map within CNV deletions in genes such as *CDH9*, *CADPS2* and *BZRAP1*. These data will serve as control data in our experiment. Conclusion: Researchers performed functional annotation of 3.5 Mb encompassing 100 ASD-associated genes to design a custom array for resequencing of ASD individuals and identified SNPs predicted to be damaging in key genes.

Project III: Animal Models. The original goal for Specific Aim 1 in Year 1 was to generate a *Cdh9-Cdh10* intergenic region deletion. Because the originally-proposed deletion of a 168 kilobase region is technically very difficult and risky (as pointed out by reviewers of the grant proposal) researchers decided to generate a mouse in which the *Cdh10* gene itself is flanked with loxP sites and then conditionally deleted. Researchers chose to target *Cdh10* because their data suggested that the SNPs in the intergenic region may regulate expression of *Cdh10*. Researchers have contracted with GenOway to carry out floxing of exons 4 and 5 of *Cdh10* to generate a conditional knockout of *Cdh10*. The generation of a targeting vector for exons 4 and 5 has taken

somewhat longer than expected, due to some technical difficulties in the PCR amplification of the long homology arm, but an alternative strategy for amplifying the long homology arm has now succeeded, and therefore all the material for the targeting vector construction is now available. Researchers expect that generation of the conditional *Cdh10* knockout mice will be accomplished in Year 2 of the Award, and phenotyping of the mice will commence in Year 2.

The original goal for Specific Aim 2 in Year 1 was to obtain constitutive heterozygous *Pcdh10* knockout mice, and to begin breeding and phenotyping these mice, has been accomplished. These mutant mice on a pure B6 genetic background were obtained from Dr. Hirano, at Kochi University in Japan. Five male heterozygote mutant mice (het mice) were shipped in August, 2009, and released from quarantine at Penn in October 2009. Breeding of the males with B6 females was initiated. Researchers successfully optimized conditions for PCR genotyping, and procedures are now working reliably. The mice have been breeding well; however, rather than the 1:1 ratio of het:WT expected by Medelian genetics in the het x B6 mating, the ratio of offspring has been approximately 1:3 or 1:4 het:WT, which has necessitated a longer breeding period to generate a sufficient number of het mice to begin phenotyping studies. Nevertheless, 3 generations of offspring were bred in Year 1, generating a sufficient cohort to begin phenotyping studies. As of this report date, 10 mice have undergone behavioral phenotyping.

<u>Project IV. Characterizing the Phenotypes of those carrying rare and common risk alleles.</u> The researchers have recruited, scheduled and completed evaluations on 40 participants (29 male, and 11 female). Ethnic breakdown is as follows: 28 Caucasian, non-Hispanic (70.0%), 5 African American, non-Hispanic (12.5%), 3 Asian, non-Hispanic (7.5%), 1 Biracial (2.5%), 1 Caucasian, Hispanic (2.5%), 2 Hispanic (5.0%). Of the 40 participants evaluated, 8 were typically developing controls (TDCs) and 35 carried a clinical diagnosis of an autism spectrum disorder (ASD). However, 5 did not meet the researchers more rigorous diagnostic definitions and procedures for ASD and will therefore be dropped from the study. All participants have also received comprehensive written feedback describing the evaluation findings, diagnostic results, as well as individually tailored treatment recommendations.

<u>Project V. Characterizing the Brain Endophenotypes of those carrying rare and common risk alleles.</u> Twenty participants from project IV have had magnetic resonance imaging (MRI). Thirteen were male and 7 were female. Ethnicity is as follows: 5 African-American (25%), 1 Biracial (4%), 11Caucasian, non-Hispanic (55%), and 3 Hispanic (15%). Fourteen participants had a confirmed ASD diagnosis and 6 were TDCs. Data analyses are complete at the individual level. Group analyses have not yet been attempted, because of the small and preliminary nature of the current sample. Magneto encephalograph (MEG) evaluation is just beginning. Researchers have successfully piloted the task paradigm on several lab members, and nearly a dozen children to be included in the study are now scheduled for MEG in the summer of 2010.

<u>Project VI, Minority Training Program.</u> During the past year, researchers made good progress towards study aims. The Researchers enrolled one post-baccalaureate student, and are reviewing the application of a second student to start in September 2010. This is consistent with the schedule of activities designed to create a cluster of students who can complete the relevant coursework together. Researchers have also completed the second year of their autism course at CHOP, thereby refining it and having it fully ready for this Project. The autism course will serve

as the foundation for the minority training program. Researchers expanded the course to include a range of topics related to autism identification, intervention, and support, and also added a number of field experiences, including opportunities to participate in parent training workshops, and visits to local community providers. Researchers have also begun discussions with Lincoln University to assist them in developing autism seminars and/or coursework, with the hope that this would increase the likelihood of their students pursuing careers in autism research.

Principle Investigator Talks and Presentations

Schultz, RT "Understanding social deficits in Autism Spectrum Disorders: Converging evidence from genes, brain and behavior." Grand Rounds, Department of Brain and Cognitive Science, Massachusetts Institute of Technology, Cambridge, MA, April 21, 2010