Children's Hospital of Philadelphia

Annual Progress Report: 2007 Formula Grant

Reporting Period

July 1, 2010 – June 30, 2011

Formula Grant Overview

The Children's Hospital of Philadelphia received $3,499,480 in formula funds for the grant award period January 1, 2008 through December 31, 2011. Accomplishments for the reporting period are described below.

Research Project 1: Project Title and Purpose

MRI Studies of System Level Brain Dysfunction in Autism Spectrum Disorders - Neuroimaging research is beginning to characterize several important ways in which the brains of persons with an Autism Spectrum Disorder (ASD) develop and function differently. Magnetic resonance imaging (MRI) has been particularly important in this regard. It is also now widely recognized that ASD is a heterogeneous disorder, with individual differences in severity of symptoms, other aspects of behavior and cognitive functioning. This project will use MRI to study the brain anatomy and function, and will correlate neuroanatomy and function with differences in symptomatology and behavior, in order to enhance our understanding of the brain-bases of the ASDs. This research is guided by the belief that understanding the fundamental mechanisms that cause ASDs will allow for the development of better interventions and treatments.

Anticipated Duration of Project

1/1/2008 - 12/31/2011

Project Overview

The autism spectrum disorders (ASDs) involve disruption of the child's ability to interact with others, communicate, play and learn. These are neurobiological disorders that appear to have multiple causes and that involve a good deal of phenotypic heterogeneity. The broad objective of this project is to use magnetic resonance imaging (MRI) to study the ways in which the brains of children with an ASD differ in their neuroanatomical structure and their neural function from typically developing children (TDC). Because the ASDs are heterogenous, large samples are required to discover coherent sets of brain-behavior relationships and to derive coherent subgroups. This project will entail establishing a systematic MRI program of research at Children's Hospital of Philadelphia (CHOP), including the clinical infrastructure necessary to evaluate large numbers of children with an ASD. Using structural MRI (sMRI), functional MRI (fMRI) and diffusion tensor imaging (DTI), we will test for significant brain based differences between ASD and TDC.

The specific aims of this project entail collecting sMRI, DTI and fMRI data from 100 children with an ASD and 100 TDC, matched on age, ethnicity, IQ and gender. While each of these three imaging approaches independently has revealed clues about the underlying neural systems abnormalities in the ASDs, no study to date has combined all 3 sources of imaging data in the same sample. Results from our prior studies using these methods in isolation have converged to suggest that abnormal white matter growth in the temporal lobes may be coupled with abnormalities of long range fiber connections between posterior visual perceptual areas and anterior limbic areas; these anatomical deficits may be responsible for the under-activation of key nodes in the temporal lobes during social perceptual and social cognitive mental activity in research volunteers with an ASD. By combining these three sources of imaging data in a large ASD sample, we expect to find more definitive evidence for the role of temporal lobe neural connectivity problems in causing the primary social deficits in the ASDs. Moreover, because we will collect a large sample, we will be able to characterize differences between individuals with an ASD in terms of different patterns of brain-behavior relationships.

Principal Investigator

Robert T. Schultz, PhD
Professor, Department of Pediatrics
Children's Hospital of Philadelphia
3535 Market Street, 8th floor, Suite 860
Philadelphia, PA 19104

Other Participating Researchers

Sarah J. Paterson, PhD, John D. Herrington, PhD – employed by Children's Hospital of Philadelphia

Expected Research Outcomes and Benefits

This study aims to be the first large-scale effort of its kind to understand and characterize the neuroanatomical bases of the autism spectrum disorders (ASDs). Prior attempts to characterize brain based causal mechanisms have yielded some important clues, but none has attempted the more comprehensive approach to be adopted here. The current project will employ a large sample of children with an ASD and will explicitly attempt to address the heterogeneity in symptoms and levels of functioning. We now believe that there are several and perhaps even many different causal pathways to what we call “autism” or the “ASDs”. Past research has relied on small study samples that likely mixed different types of ASD cases, thereby obscuring our ability to clearly find correlations between brain and behavior. Our strategy will be to create a suitably large clinical infrastructure to enable enrollment of a large sample of children affected with an ASD. Each child will have a broad based expert clinical assessment of their difficulties, as well as of areas of cognitive strengths and behavioral competencies. 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 research volunteers and their families. Scientifically, health benefits include a much deeper understanding of the particular neural systems that are affected by the ASDs. Our research is guided by a fundamental belief that it is difficult and perhaps

impossible to develop effective treatments without first laying out in detail the nature of the problem in terms of causal mechanisms. We expect that not all persons with an ASD will show the exact same neural system abnormalities, and that differences in brain system abnormalities will be mirrored by differences in behavioral and symptomatic features. Better characterization of these brain based mechanisms will facilitate the development of novel and more effective treatments and aid in the process of gene discovery.

Summary of Research Completed

During the current reporting period, the researchers evaluated 106 children for this study. These are full day assessments replete with detailed gold standard diagnostic and neurocognitive evaluation. Of the 106, 18 did not go on to receive the MRI study protocol. Fourteen of the 18 did not continue because they failed to meet research diagnostic criteria for ASD (despite carrying a community based ASD diagnosis) and 4 did not continue for other reasons. Thus, the researchers performed MRI scanning on 88 study participants (56 males, 32 females). Ethnicity was as follows: 3% Hispanic, 5% African-American, 3% Biracial, 82% Caucasian, non-Hispanic, 6% Asian, 1% African-American, Hispanic. Fifty-six of the participants had a diagnosis of autism spectrum disorder (ASD); 32 were typically developing control participants (TDC).

Cumulatively, the study has now scanned 130 participants for this project (82 with ASD; 48 TDC). Most of the study progress has been made in the last year. This research grant has been instrumental in helping launch the Center for Autism Research (CAR) at the Children's Hospital of Philadelphia. The award and the Center began within a few months of each other. CAR had about 10 researchers and research staff at the start of this award and now there are over 100 at the Center. This grant has been instrumental in establishing a major new research center in the Commonwealth of Pennsylvania that already is an international leader in autism research. As the Center has grown, systems for recruiting and evaluating research participants have been established with the assistance of funding from this grant, and operating procedures have matured. Thus, at this point the researchers anticipate continued excellent progress with study recruitment and participant throughput in the final award period.

MRI data analysis is ongoing. Nearly all participants have had their data evaluated at the individual level through processing pipelines. It is a bit premature to focus on group results until all data are collected. Nonetheless, the sample sizes are now sufficient to begin to generate preliminary results. Much progress with the structural MRI (sMRI) analyses was made in the reporting period. Building upon the pattern classification algorithms published previously by the researchers, a computational marker that quantifies the presence of ASD-like abnormality in an individual's brain was derived in preliminary analyses. In this pilot work, the researchers analyzed sMRI scans of 55 participants (26 ASD and 29 TDC).

Using segmented gray matter and white matter (WM) tissue density maps as input to a high-dimensional pattern classification (that works by identifying volumetric patterns that yield optimal classification accuracy with a Support Vector Machines classifier), the researchers found very encouraging preliminary results. The cross-validated accuracy of the classifier in the task of separating ASD from TDCs was 72.7%. The area under the receiver operator curve was also

promising (72%). Figure 1 shows WM volumetric regions that were identified by the pattern recognition approach to be discriminative between ASD and TDC populations. This grant has helped with important methodological developments used in this pilot sample; these methods will be quite valuable for the analyses of final sample.

In order to characterize the correlation between brain anatomy and social competence and other measures of the behavioral phenotype additional new methods were developed during this reporting period. The researchers developed a cross-validated regression methodology that uses a regional feature combining relevance vector machine. This was evaluated in a related study of 98 males with ASD, and the method appears sensitive and valid (e.g., several brain regions of *a priori* interest were strongly correlated with social communication ability as measured with the Autism Diagnostic Observational Schedule). This new pattern regression method is general, and at this point the researchers are performing experiments on the sample from this grant using functional MRI (fMRI) and diffusion tensor imaging (DTI) modalities to finalize its development and to yield insights into the pathobiology of ASD.

Progress has been made developing methods and analyzing DTI data. Fractional Anisotropy (FA) and Mean diffusivity (MD) images were computed from deformably registered diffusion data: 34 ASD (26M/8F) (Mean Age: 11.9±2.7 years) and 26 TDC (26M/8F) (Mean age: 11.6±2.3 years). Two kinds of analysis were performed:

1. Voxel-wise analysis: Voxel based morphology (VBM) was performed on FA and MD images. Preliminary uncorrected group differences are shown in Figure 2. The regions showing decreased FA in ASD (yellow-red) included parts of superior longitudinal fasciculus, inferior longitudinal fasciculus and, some parts of the corpus callosum. MD showed no group differences.

2. Region of interest (ROI) based analyses: ROI classifiers were created for ASD versus TDC with all available DTI data. The researchers obtained 82% cross-validated accuracy (with 0.8 sensitivity, 0.8 specificity, 0.8 area under the cure). The top regions ranked by the feature selection included only FA features (Superior temporal WM, Unicnate fasciculus, Insula, Post-central WM) suggesting no MD contribution (just as observed in the voxel-wise analysis above). The group separation based on classifier scores can be seen in Figure 3. The classifier creation framework is general and these classifiers will be created again, once the entire dataset is acquired.

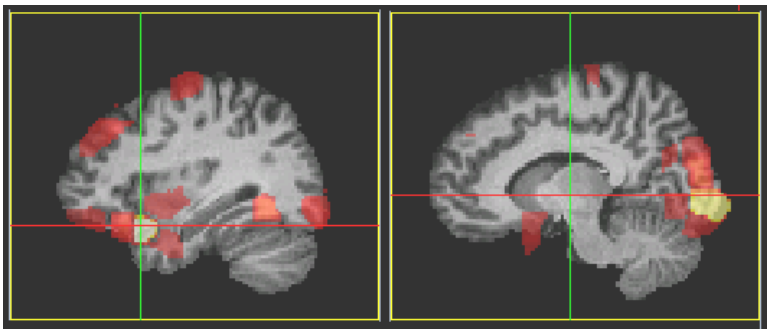

Figure 1. White matter volumetric regions that discriminate ASD and TDC samples

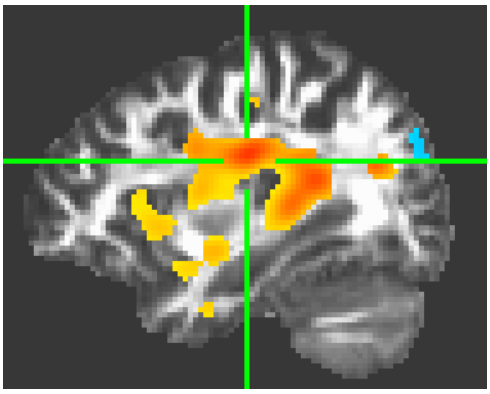

Figure 2. VMB based FA group comparison showing regions where TDC > ASD.

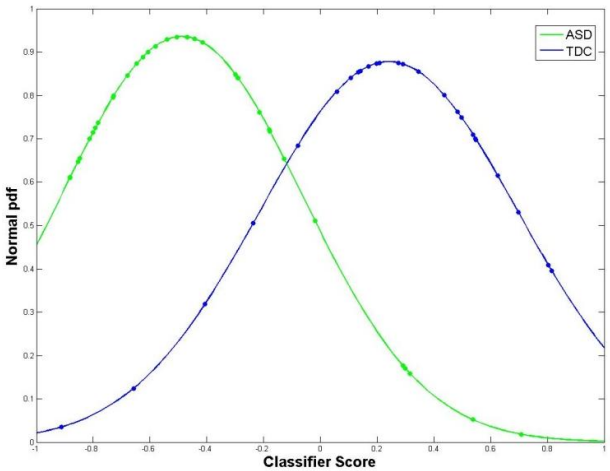

Figure 3. Diagnostic group separation based on classifier scores using DTI data

Research Project 2: Project Title and Purpose

Ethnic Diversity and Gene-Environment Interactions in Childhood Asthma - The purpose of this research is to use a genome-wide association (GWA) approach to uncover the genetic factors that predispose to asthma, the most common chronic disease in children, and to thoroughly examine gene-environment interaction in asthma patients of different ethnic background. The utilization of the GWA approach has excellent potential of identifying the genetic determinants of asthma predisposition and to subsequently dissect out the role of second hand smoke and potentially other specific environmental exposures in the pathogenesis of asthma in subjects from different ethnic background, including Caucasian, African-American and Hispanic Caucasian descent. The study will utilize a well-characterized pediatric asthma cohort from the Children's Hospital of Philadelphia (CHOP) to drive the discovery phase.

Anticipated Duration of Project

1/1/2008 - 12/31/2011

Project Overview

Our primary objectives are directed at unveiling the genes that underlie the pathobiology of asthma and to shed light on how specific environmental factors interact with these genetic factors to trigger the expression of the asthma phenotypes. In our first specific aim we propose to analyze a pediatric asthma cohort from Philadelphia using a GWA approach. We propose to genotype approximately 1000 children with asthma and 1000 non-asthmatic controls and analyze the dataset for single SNP (single nucleotide polymorphism) and SNP- environmental interaction effects on the phenotypic characteristics that will be gathered on the study cohort. In specific aim 2 we propose to select SNPs that show the highest association in specific aim 1 and genotype these selected SNPs in another cohort of 1200 asthma patients and 1200 controls from Philadelphia and analyze the data jointly with the original cohort in specific aim 1 to find single SNPs and SNP-environmental factor interactions that are genome-wide significant after correction for multiple tests. We estimate that approximately 1500 SNPs will have p value of 10^{-5} or lower from where a subset of SNPs (~ 10-15) will be true positives. In specific aim 3 we propose to validate and refine disease-associated SNP alleles and SNP-environmental factor interactions in independent sample sets consisting of 500 Hispanic Caucasian and 500 African-American (AA) subjects. In specific aim 4 we will examine the interplay between genes that harbor the at-risk variants that replicate in specific aims 2 and 3 and specific environmental factors with respect to the expression of the asthma phenotypes and we will re-sequence the most significant genes to identify the functional causative mutations that predispose to asthma.

The institutional resources available to pursue this project are outstanding, most notably with the establishment at CHOP of one of the world's largest genotyping centers (Center for Applied Genomics or CAG) for GWA studies and analyses. CAG has already genotyped over 80,000 individuals and produced over 50 billion genotypes that are being leveraged for this project as control subjects. By incorporating these outstanding resources and expertise available at CHOP we believe that this proposal will effectively address the above specific aims and, thereby, provide critical new insights into gene-environment interactions that underlie asthma in children.

Principal Investigator

Hakon Hakonarson, MD, PhD
Director, Center for Applied Genomics
Children's Hospital of Philadelphia
Abramson Research Institute
3615 Civic Center Boulevard
Philadelphia, PA 19104-4318

Other Participating Researchers

Marcella Devoto, PhD, Struan F.A. Grant, PhD and Patrick Sleiman, PhD –employed by Children's Hospital of Philadelphia

Expected Research Outcomes and Benefits

Asthma is a major health problem in children and an immense economic burden on the health care systems in the US and in the rest of the world. Given the important need for effective translational research in asthma, we believe that recent advances in genotyping technology, coupled to the fundamental information provided by the Human Genome and International HapMap Projects, makes the timing right for this ambitious project. Unique to the project are its cohort study populations, a very large and carefully annotated collection of samples already in hand and a successful strategy for continued collection of samples, and its application of the latest developments in genotyping technology. In view of these strengths of the project, we anticipate that our proposed studies to associate genome-wide individual SNPs and their interactions with environment exposures with varying phenotypes will be successful in identifying genetic variants and gene-environment interactions that underlie the development of asthma in childhood in subjects of different ethnic background. Indeed, we have already observed that exposure to smoking by asthmatic children modifies the predisposition risk of certain genotypes at the ORMDL3 locus and that these effects are differentially expressed in asthmatic children of European vs African ancestry. As such, we have great optimism that other such variants can be identified in pediatric asthma. Thus, we strongly suspect that our findings will ultimately significantly contribute to the health of children with asthma since, by providing an improved understanding of the gene-environment interactions that initiate these interrelated disorders, the findings will serve to identify molecular pathways critical to regulating their phenotypes and, thereby, facilitate the development of more effective new therapeutic and environmental control interventions and preventative measures to alleviate this most common chronic disease in children.

Summary of Research Completed

The successful completion of aim 1 resulted in the identification of a novel asthma predisposition locus by genome wide association (GWA) in both Caucasians and African Americans. The researchers have reported an increase in risk for pediatric asthma as a result of interactions between variants at the ORMDL3 locus and exposure to smoke. As a result, two projects have been initiated in accordance with specific aim 3, to A) use regional capture and next generation sequencing to identify the causal variants at both loci and B) to determine the functional mechanism through which variation in the DENND1B gene predisposes to asthma. Each of the two projects will be briefly outlined below:

<u>Project 1</u>
Genome wide association chips are designed to capture the maximal amount of information throughout the genome by making use of patterns of linkage disequilibrium. In order to identify the causal variants underlying the predisposition to asthma at the DENND1B and ORMDL3 and

loci that are being tagged by the GWAs chips the researchers are resequencing the coding regions of the two genes in several hundred patients. Homologues of the DENND1B gene have been shown to interact with TNFR6, a protein previously implicated in asthma, therefore the TNFR6B gene was also sequenced. Coding sequences were enriched through RainDance Technologies. CAG has acquired an RDT1000 microdroplet system which creates PCR reactions in millions of picoliter-sized droplets. A PCR pre-mix containing sheared genomic DNA is used to generate the droplets containing one-third of a haploid genome. Millions of these droplets are merged with the DENND1B and ORMDL3 primer library which contains five primer pairs per droplet. Amplification takes place in these aqueous drops suspended as an emulsion in a fluorocarbon oil matrix. The instrument can work with a PCR primer library of up to 20,000 amplicons to generate an enrichment of 12 megabases of genomic sequence. The RainDance-enriched DNA is ligated into high molecular weight concatamers and then fragmented by sonication (Covaris). The sheared fragments are then subjected to the library construction process for the Applied Biosystems SOLiD 4 HQ instrument using EZBead automation for the emulsion PCR step. Libraries are constructed in a pooled-barcoded format so that multiple cases can be sequenced in the same flow cell, minimizing cost. The libraries are then sequenced on the SOLiD 4 HQ instrument to achieve depth in excess of 40× per individual. Sequencing reads from the SOLiD instrument are then aligned to a reference genome and variants called using ABI's BioScope software package which is installed on our cluster. The variants are processed through the ANNOVAR annotation pipeline and compared with publically available databases as well as the SeattleSeq project to identify alleles that are rare or which are likely to interfere with gene function. ANNOVAR functionally annotates SNPs and CNVs generated from high-throughput sequencing experiments within exonic regions but it also extends to other functional categories, such as identifying known and putatively functional variants in conserved genomic regions or transcription factor binding sites, as well as functional importance scores, such as SIFT and PolyPhen. Furthermore, ANNOVAR can automatically identify variants that were previously reported in pubic databases, including the 1000 Genomes Project, dbSNP, and the Database for Genomic Variants. When tested on whole human genomes, ANNOVAR requires approximately 5 minutes to annotate 3 million variants using a desktop computer. While reading variants from input files, ANNOVAR scans the gene annotation database stored locally, and identifies intronic, exonic, intergenic, 5/'3'-UTR variants, and those that impact splicing sites. For exonic variants, ANNOVAR scans annotated mRNA sequences to identify and report amino acid changes by a standard nomenclature. The researchers can systematically integrate suites of computational tools into ANNOVAR to classify coding variants that may impact protein function such as POLYPHEN, SIFT, SNPs3D, and Pmut. The researchers have also integrated similar tools to score functionality for consensus transcription factor binding sites (TFSEARCH, MATINSPECTOR and 3'-UTR (ASTRA). These approaches yield a prioritized set of SNPs with composite functional scores for each respective gene and assign a score of "genetic load" for each given individual that can be entered into a generalized linear model (GLM). While the utility of globally assigning *in silico* functional scores obviously comes with uncertainty, the researchers rationalize that it is the most logical unbiased approach to treating such complex datasets. ANNOVAR also offers the ability to scan other types of annotation tracks, such as miRNA and eQTL data that may become available through the community, so that risk and protective alleles can be further categorized based on sequence conservation patterns and predicted functional importance values. Analysis of the sequence data and validation of the mutations is ongoing, however, the researchers can report the identification of two novel

missense mutations at codons 172 and 258 of the TNFR6B gene that are not present in controls and are predicted to result in a loss of function.

To investigate the functional mechanism underlying DENND1B predisposition to asthma, researchers first investigated correlation of the associated markers with gene expression levels. As the SNPs associated with asthma were all in noncoding regions of DENND1B or in intergenic regions, the researchers hypothesized that the SNPs might associate with molecular mechanisms other than coding sequence changes, such as differential expression levels. To examine DENND1B expression, RNA from EBV-transformed lymphoblastoid cell lines (n=37) generated from individuals who were previously genotyped was analyzed on Affymetrix Human Gene 1.0 ST Arrays. Expression information for DENND1B (probeset 7923131) was analyzed versus the genotype at the SNP most strongly associated with asthma (rs2786098). ANOVA shows a statistically significant association of the genotype at this SNP with expression levels ($p=0.044$), with the minor, resistant allele associating with lower DENND1B expression.

To further examine the effects of DENND1B expression levels, plasmids expressing two known splice isoforms of DENND1B were constructed. These plasmids express DENND1B isoforms of 426 or 775 amino acids. Given the homology between DENND1B and the DENN/MADD protein, and the influence that DENN/MADD is known to have on Tumor Necrosis Factor α Receptor (TNFR) signal transduction, the researchers investigated the effect of DENND1B overexpression on TNFR signaling. HEK293T cells were transfected with plasmids expressing the 426 or 775 amino acid isoforms of DENND1B. Cells were co-transfected with plasmids expressing *Renilla* luciferase under the control of the constitutive Thymidine Kinase promoter, and firefly luciferase under the control of an NF-κB promoter. Cells were stimulated with increasing amounts of recombinant, human Tumor Necrosis Factor α (TNFα) for 6 hours. Cells were harvested and lysed, and luciferase activities were determined. Overexpression of the 426 amino acid isoform reduced TNFα-induced luciferase activity, while expression of the 775 amino acid isoform enhanced it. To determine if this effect on TNFR signaling was a result of changes in TNFR surface expression caused by DENND1B expression, the researchers monitored TNFR expression following transfection of HEK293T cells with DENND1B proteins tagged C-terminally with GFP. Transiently transfected cells were stained for surface expression of TNFR-I 48 hours after tranfection, and TNFR-I expression levels in GFP-positive and GFP-negative cell were compared. Expression of GFP-tagged 426 or 775 amino acid isoforms of DENND1B did not alter TNFR-I expression, indicating the altered response to TNFα stimulation is due to DENND1B-induced changes in signal transduction.

In summary, the researchers have completed aims 1 and 2 of the project and are currently engaged in aim 3, to identify the molecular pathways that predispose to asthma, thereby, facilitating the development of more effective therapies.