Unraveling Autism

Dietrich A. Stephan[1],*

In this issue of *AJHG*, Alarcón et al.,[1] Arking et al.,[2] and Bakkaloglu et al.[3] identify a series of functional variants in the *CNTNAP2* gene that unequivocally implicate this gene as causing Type 1 autism in the general population.

Autism Spectrum Disorder (ASD) is a catch-all diagnosis for a set of poorly understood neurodevelopmental disorders that are clinically heterogeneous, with a spectrum of severity, characterized by repetitive self-stimulatory behaviors and communication and socialization deficits. ASD is traditionally diagnosed by the age of 3 years and the severe forms can be accompanied by language regression, seizures, and low measured IQ. The more strict diagnosis of "autism" is made through behavioral testing on the ADOS and/or ADI-R rating systems. The umbrella diagnosis of ASD is approaching 1% of all births, likely through a combination of ascertainment bias and increased population exposure to unknown environmental risk factors or mutagens. Epidemiologic studies of ASD have failed to identify definitive evidence of exposures correlating to increased risk of the disorder. It will be critical to dissect the genetic subclasses of ASD prior to embarking on a new wave of epidemiology to identify environmental risk factors, because it is reasonable to assume that individual environmental exposures act differently on each genetic subclass of the disorder.

Twin studies (both with strictly defined autism and the broader ASD phenotype) as well as familial clustering indicate a strong genetic component to predisposition. The clinical heterogeneity of this disorder will likely be explained in large part by genetic heterogeneity; a multitude of linkage studies (including the large multinational Autism Genome Project) have reproducibly identified several loci by studying ASD phenotypes segregating through rare pedigrees with multiple affecteds.[4] Because the more common forms of ASD are sporadic and ASD individuals are less likely to reproduce, we can assume that this majority of ASD predisposition is caused by either SNP variants segregating through the population or by a high new mutation rate in predisposition genes.

Until recently, there existed only three genes with limited evidence (often only in a few probands) implicating them as causative of ADOS/ADI-R-defined autism, when mutated. Mutations of the SH3 and multiple ankyrin repeat domains 3 (*SHANK3*) gene have recently been reported to be associated with ASD in a small number of individuals, and mutations in this gene have been found in less than 1% of probands tested.[5,6] *SHANK3* codes for an adaptor protein in the postsynaptic density of the excitatory neuron and likely plays a role in the functional and structural organization of the dendritic spine and the synapse. Mutations in a second gene, neuroligin-3 (*NLGN3*), have been shown to cause ASD when mutated, but have not been replicated widely, and there are reports of lack of replication.[7,8] Finally, the contactin associated protein-like 2 (*CNTNAP2*) gene at 7q35, a member of the neurexin superfamily, was described by our group in 2006 to cause severe autism with medication-insensitive temporal lobe seizures, language regression, and low IQ when the carboxy terminal of the protein product

was truncated through a homozygous loss-of-function mutation in a single family.[9] The mechanism of action of the mutation is likely altered attachment of the axon to the glia via the TAG-1 protein and mislocalization of ion channels at the juxtaparanodal junction leading to cortical dysplasia. This finding is now replicated in a large sampling of the autism population by three groups in this issue of *AJHG* and places the *CNTNAP2* gene as the first widely replicated autism-predisposition gene. Alarcón et al.,[1] Arking et al.,[2] and Bakkaloglu et al.[3] all describe functional variants (both common and rare) that predispose to autism in the general population. It is reasonable at this point to define *CNTNAP2* mutation-positive autistic cases as having "Type 1 autism."

Bakkaloglu and colleagues[3] report a de novo 7q35 inversion that disrupts *CNTNAP2* between exons 10 and 13 in a child with autistic features. This evidence led them to resequence all 24 exons of the gene in 635 affecteds and 942 controls. Thirteen rare variants were identified in cases, and of these eight were predicted to have negative consequences on gene function because they occurred in evolutionarily conserved regions of the gene. Although predicted deleterious variants were also found in the control cohort, there were roughly twice as many variants found in the cases. One predicted deleterious variant (I869T) was found to be present in four affected individuals from three different families, but not present in >4000 chromosomes from unaffecteds. This study illustrated a large structural

[1]Neurogenomics Division, Translational Genomics Research Institute, Phoenix, AZ 85004, USA
*Correspondence: dstephan@tgen.org

event that likely leads to ASD in a single individual and characterizes a series of rare variants in probands that are predicted to have functional consequences on gene function.

Alarcón et al.[1] also replicate the *CNTNAP2* locus as causative of autism. Predicated on their previous work identifying a 10 cM linkage peak at 7q35 in families with language deficits and ASD, the authors genotyped 172 parent-child trios at 2758 SNPs across the linkage peak and narrowed the region to 4 genes (5 haplotype blocks), including *CNTNAP2*. Eight SNPs covering these 5 loci were genotyped in an independent 304 ASD families, and only SNP rs2710102, which is in *CNTNAP2*, remained significant, and only in male probands. The significant association was specifically with "age of first word" in the ASD probands. The authors go on to describe a microdeletion in the gene that is carried by a proband and his father, but never seen in 1000 control chromosomes.

A common SNP variant (rs7794745) within the *CNTNAP2* gene was identified by Arking and colleagues[2] under the established 7q35 linkage peak by association mapping with strictly defined autism cases, and confirmed in an independent replication population (with broader diagnostic inclusion criteria) as enriched ("T" allele) in probands with autism. Of particular interest is that the T allele is more strongly associated with the phenotype when inherited from the mother, although studies examining parent-of-origin imprinting at this locus have yet to be done. The finding that a common variant is associated with increased risk of autism was made possible by careful clinical assessment and reduction of clinical heterogeneity to the greatest extent possible followed by focused association analysis within a linkage region, which allowed this subtle association to be detected without genome-wide background noise.

These three publications reiterate the point that rare monogenic forms of common complex genetic disorders yield critical insight into disease processes, and these studies should continue to be carefully evaluated in combination with careful phenotyping of "sporadic" disease and recent whole-genome association approaches. To address the overarching hypothesis that common variants predispose to any subtype or all ASD, a series of whole-genome association studies at high SNP density have been performed. The largest of these data sets was generated on more than 3000 probands and family members with approximately 500,000 SNPs and deposited into the Autism Genetic Resource Exchange (AGRE) database for public access. Preliminary analysis of this elegant data set indicates that extremely significant p values that exceed multiple testing correction are not present. This is possibly due to the clinical and genetic heterogeneity that we know exists within ASD. It is reasonable to think, based on the three studies presented herein, that a priori phenotypic stratification or linkage positivity to a certain locus will enhance the ability to detect associations between common variants and these subgroups.

The three studies herein[1–3] have moderate sample overlap and use different strategies to narrow the phenotype of the ASD cohorts to relative homogeneity before performing genotyping/resequencing across the *CNTNAP2* gene locus. Nevertheless, the three studies together identify a set of common and rare variants that provide unequivocal evidence that the *CNTNAP2* gene, when disrupted, leads to a subtype of ASD. This genetic subtype can be clinically characterized by ADOS/ADI-R-defined autism with language deficits and potential gender bias and parent-of-origin effects. Type 1 autism may also be associated with seizures. It will be important to begin to characterize the genotype-phenotype correlations across this gene so that we may begin to use the *CNTNAP2* as a diagnostic and prognostic tool. This gene is a very large target (~2.3 Mb genomic locus) for mutations. ASD individuals have reduced reproduction and thus it is reasonable to assume a high new mutation rate in this gene. Now that we have definitive evidence from several perspectives that integrity of the neuroligin-neurexin axis is critical for normal development, we must launch into a candidate gene-resequencing effort to fully describe mutations in the other members of these gene families in ASD. Finally, our collaborative group has been able to identify individuals from our initial family who carry the homozygous loss-of-function mutation that we originally described in 2006. Several of these children were given the anti-seizure medication valproic acid prior to onset of symptoms in an attempt to control seizures and other aspects of the disorder, and the preliminary results are promising (K. Strauss, personal communication). These preliminary findings lead one to speculate whether early detection of *CNTNAP2* mutation carriers coupled with early intervention could coax children through a critical period in development (12–24 months of age) and allow them to emerge undamaged and continue to develop normally thereafter. The modern technologies and strategies derived from the Human Genome Project, coupled with the elegant sample banking, phenotyping, and data dissemination resources of groups like AGRE, are resulting, finally, in the unraveling of Autism Spectrum Disorder.

References

1. Alarcón, 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., et al. (2008). Am. J. Hum. Genet. *82*, this issue, 150–159.
2. Arking, D.E., Cutler, D.J., Brune, C.W., Teslovich, T.M., West, K., Ikeda, M., Rea, A., Guy, M., Lin, S., Cook, E.H. Jr., and Chakravarti, A. (2008). Am. J. Hum. Genet. *82*, this issue, 160–164.
3. Bakkaloglu, B., O'Roak, B.J., Louvi, A., Gupta, A.R., Abelson, J.F., Morgan, T.M., Chawarska, K., Klin, A., Ercan-Sencicek, A.G., Stillman, A.A., et al. (2008). Am. J. Hum. Genet. *82*, this issue, 165–173.
4. Hu-Lince, D., Craig, D.W., Huentelman, M.J., and Stephan, D.A. (2005). Am. J. Pharmacogenomics *5*, 233–246.
5. Durand, C.M., Betancur, C., Boeckers, T.M., Bockmann, J., Chaste, P.,

Fauchereau, F., Nygren, G., Rastam, M., Gillberg, I.C., Anckarsäter, H., et al. (2007). Nat. Genet. *39*, 25–27.

6. Moessner, R., Marshall, C.R., Sutcliffe, J.S., Skaug, J., Pinto, D., Vincent, J., Zwaigenbaum, L., Fernandez, B., Roberts, W., Szatmari, P., and Scherer, S.W. (2007). Am. J. Hum. Genet. *81*, 1289–1297.

7. Jamain, S., Quach, H., Betancur, C., Råstam, M., Colineaux, C., Gillberg, I.C., Soderstrom, H., Giros, B., Leboyer, M., Gillberg, C., Bourgeron, T., and Paris Autism Research International Sibpair Study. (2003). Nat. Genet. *34*, 27–29.

8. Blasi, F., Bacchelli, E., Pesaresi, G., Carone, S., Bailey, A.J., Maestrini, E., and International Molecular Genetic Study of Autism Consortium (IMGSAC). (2006). Am. J. Med. Genet. B. Neuropsychiatr. Genet. *141*, 220–221.

9. Strauss, K.A., Puffenberger, E.G., Huentelman, M.J., Gottlieb, S., Dobrin, S.E., Parod, J.M., Stephan, D.A., and Morton, D.H. (2006). N. Engl. J. Med. *354*, 1370–1377.