Attached is an article for dissemination in autism parent group newsletters. It describes the results of the genome screen recently completed by the researchers at the Stanford University School of Medicine (Department of Psychiatry & Behavioral Sciences and Department of Genetics). The article discusses how genetic studies are done, how to interpret results, and compares results of the recently published studies.

We request that the article be published as is without any changes. If that is not possible, please contact Donna Spiker, Ph. D., Clinical Director at (650 723-7809.

# Challenges of Genetic Research in Autism: Results from the Stanford Autism Genetics Project

Autism was once thought to be a sporadic, or non-genetic disorder because most families have only a single affected child. Over the past 20 years, systematic studies of families and twins have provided evidence to the contrary - that genetic factors do play an important role in this disorder. Although the overall recurrence rate of autism in sibships is only around 3-6%, that rate is 50 to 100 times greater than the frequency observed in children from the general population. Also, identical twins, who share 100% of their genetic material, are very frequently concordant for autism (about 75% of the time, both have autism) as compared with fraternal twins, who share, on average, 50% of their genetic material and are concordant only about 3% of the time.

## Methodologies Used in Genetic Studies

The strong circumstantial evidence for genetic causation based on twin and family studies in autism does not provide the identity of any gene or genes that are involved, or how they are involved. Finding genes is the goal of current genetic research efforts in autism. There are essentially two paths that can be taken to try to identify these genes – the "candidate gene" approach and the "genome scanning by linkage analysis" approach. Geneticists studying autism and many other diseases and other traits in humans are using both of these approaches.

In the candidate gene approach, a specific gene of known identity is tested directly for association with autism. Often the sequence of DNA base pairs of the particular gene is already known from other research. Sometime it is determined as part of the genetic study by using a DNA sequencer. The geneticist uses the DNA sequence information differences that are naturally seen in different persons. These differences are referred to as genetic variants or polymorphisms, that is, the different forms a gene can take. Once a variant or variants in a gene have been identified, they are examined for association with autism by one of several designs. Often the simplest design is to examine the frequency of a particular variant in a collection of affected individuals to see if it is increased compared to the frequency in an appropriate "control" group of unaffected subjects. It is particularly important that

the control group be ethnically matched, because it is known that the frequency of genetic polymorphisms can vary from one ethnic group to another.

Another experimental design avoids the need for this type of matching by using parents of affected children as controls. In this case, the DNA sequence variants transmitted to an affected child are compared to the parental variant sequences that are not transmitted. The variant not transmitted is the "control" for the allele that is transmitted, and is automatically "ethnically" matched. If one's guesses about which genes might be involved in susceptibility to autism is correct, this approach is extremely efficient and can quickly lead to a major payoff. The downside of this strategy is that there is somewhere between 50,000 to 100,000 genes in the human genome, and any one of them may be a candidate for susceptibility to autism. Thus, the chance of a correct guess in the first place is low. However, the use of current knowledge of genes involved in human brain structure and function can help us make more educated guesses and improve the odds.

In the second approach for identifying genes involved in genetic traits, called genome scanning by linkage analysis, no specific gene or genes are examined for association with the disease or trait. Rather, random, presumably non-functional DNA markers covering all the chromosomes are examined in an attempt to identify a rough "location" where an autism susceptibility gene or genes may lie. The principle of genetic linkage analysis is based on the fact that a DNA marker that lies on a chromosome near a particular variant of a gene is passed together from parent to offspring most of the time because the marker and the gene variant are physically linked to one another on the chromosome. DNA markers that are further away from the variant gene are not linked because of mixing of the chromosomes, by a process called genetic recombination, that occurs during formation of the sperm and egg prior to fertilization. Thus, instead of testing individual genes that might be involved in autism, DNA markers that are spread throughout all of the chromosomes serve as surrogates that might identify a region of a chromosome that contains a gene that contributes to autism.

It is generally believed that 350-400 DNA markers provides an adequate density to identify involved gene regions that may lie anywhere between these markers. Typically, the initial "scanning" is only the first step, as it gives only preliminary evidence of where a susceptibility gene may lie. Follow-up experiments usually consist of testing many more markers in a given region to help narrow down the

location of the implicated gene or genes. This entire approach is often referred to as "positional cloning", as the identification of a contributing gene is based solely on its location on a chromosome rather than any prior knowledge of its existence or function.

Positional cloning has had many notable successes in leading to the identification of genes for breast cancer, Alzheimer disease, Huntington disease, cystic fibrosis, and numerous other simple genetic disorders. We use the term simple here to indicate that the inherited basis for the disease in these cases was due to a mutation in a single gene. In these cases, the initial stages of the positional cloning, i.e., the linkage analysis, involved examination of DNA markers evenly spaced around the genome for co-inheritance (or similarity of inheritance pattern) with the disease in families with many affected individuals.

Both approaches, candidate gene and genome scanning, have already seen application in autism. To date, the number of candidate gene studies has been quite limited, but we can expect the number to increase rapidly in the near future, as many more research groups have now assembled the materials and scientific resources necessary for such research. Over the last year or two, several research groups have also published the results of linkage analysis with genome scans. Below we discuss the results of these scans, and in particular the results from our own analysis, and the implications of these results for future research.

## Interpreting Genome Scans in Autism

To evaluate the results of linkage studies in autism, some consideration of the genetic basis for the disorder is required. In classic, "Mendelian" disorders, there is a single gene responsible for the disease that is transmitted in families. This "single gene" nature of inheritance makes it relatively simple to detect a gene by linkage analysis, because the transmission pattern in families make a "very loud signal" that can be "heard" even at some distance away on the same chromosome (usually millions of DNA base pairs away). This fact underlies the successful mapping and subsequent identification of many Mendelian disease genes. By contrast, the evidence strongly suggests that autism has no such simple pattern of inheritance, but likely has many different interacting genes leading to susceptibility. This conclusion is supported by the very high ratio of disease concordance for identical

twins versus fraternal twins. That being the case, it is currently impossible to tell the number of genes involved (i.e., whether 10 or 10,000) and/or how large their effects are. If there are a few genes with sizeable effects, we should be able to find them by positional cloning.

When the contribution of a particular gene to a trait (the “gene effect”) is not large, the signal it produces in a genetic linkage study is also not large. This is the primary problem in using genetic linkage for a complex disease such as autism. Such linkage scans have been performed for numerous other disorders with similar complexity, and unfortunately, with only limited success. It is typical that chromosome locations identified as linked in one study appear not to be linked in other studies. When this occurs, it may be difficult, or impossible to know whether the initial finding was simply a chance event (i.e., a "false positive"), or that the gene effect in the region is weak and therefore unlikely to produce positive results in a study of typical size. Examining a large number of families (perhaps gathered together from several sources) may help distinguish between these two explanations.

Nonetheless, replication of initial findings by independent investigators who are using similar methods is a fundamental, essential feature of scientific research. It is the only way scientists can have confidence in the veracity of a hypothesis. However, one also needs to consider possible explanations for non-replicability of initial findings. In genetic studies, some potential reasons for non-replication of results are ethnic differences in study populations and the use of different procedures by different research groups to diagnose subjects. Such concerns are often raised when initial findings are not replicated in other independently collected samples.

## Results of Genome Scans in Autism

With this as background, we can examine some recently published linkage studies of autism. The first such study was performed more than 20 years ago by Dr. Anne Spence and colleagues at UCLA, but this was before the modern DNA era, so the number of genetic markers used was very limited. This research team found no chromosomal locations suggestive of linkage in their analysis. More recently, four linkage genome screens in which families with two or more individuals with autism were studied have been reported -- the one from our group (Stanford), one resulting from an international consortium (IMGSAC), a French consortium

(PARISS), and a U.S. consortium from Tufts University, the University of Iowa, Johns Hopkins University, and Vanderbilt University. The international consortium tested an initial group of 39 families with follow-up in an additional 60, and used a broad definition of autism. They found suggestive evidence of linkage on chromosome 7q, and another positive region on chromosome 16p. However, in neither location were the results formally significant on statistical grounds. The French consortium studied 51 families and obtained their most positive result on chromosome 6q; they also found somewhat positive results on 7q, but again none of the results were formally statistically significant. The U.S. consortium studied 75 families. Their most positive results were on chromosome 13q, but as in the other studies, these fell short of statistical significance. This group also had some positive results on chromosome 7q.

In the fall of 1999, we published the results of our linkage scan, which is the largest and most comprehensive study published to date. In the first stage of our analysis, we examined 360 markers in 90 families containing 97 independent sib pairs with autism. We used narrow diagnostic criteria, excluding all individuals who did not unambiguously meet all the criteria for a diagnosis of autism. In the second stage of our analysis, we examined an additional 49 families with 50 independent sib pairs, and tested an additional 157 markers. In total, we tested 519 markers in the first set of families and 149 in the second set of families. Another novel feature of our study was the inclusion of 30 unaffected siblings as controls. This allowed us to construct 51 affected-unaffected (or discordant) sib pairs to use as a contrast (or control) group when examining sharing of DNA sequence variants in the affected sib pairs. In theory, this group should show a random distribution of DNA sequence variant sharing, where 50% of their variants are in common because all sibs share 50% of their genes.

The first step in our analysis was to examine the extent of DNA sequence variant sharing at the 346 markers we tested in the first set of 90 families. We looked at the proportion of variants shared for each of the 346 markers for the 97 independent concordant sib pairs and 51 discordant sib pairs separately. Over all markers, we found more sharing for the concordant than the discordant pairs (consistent with linkage), but this excess could not be attributed to a small number of genes. In fact, the increased sharing observed in concordant sib pairs was most consistent with a large number (at least 20) of susceptibility genes, none of which has a large effect. Of all the markers we tested, the most significant result we obtained was for a region on

chromosome 1. Our second most noteworthy result was for a marker on chromosome 17. However, neither of these localizations was statistically significant after adjusting for the large number of markers that we tested.

Our follow-up study consisted of analyzing an additional 159 DNA markers on the first set of 90 families and a total of 149 markers on the second set of 49 families. These markers were chosen to lie in chromosomal regions that showed some linkage evidence either in our own study or from other studies. Our evidence for linkage on chromosome 1 increased slightly, but was still short of formal statistical significance. Nonetheless, it was still the strongest result we obtained. The evidence we originally observed on chromosome 17 was diminished but still positive.

We also examined many DNA markers in chromosomal regions identified in other linkage studies, in particular on chromosome 15q (where an inverted chromosomal duplication has been described in some individuals with autism), on chromosome 7q, and chromosome 13q. In none of these locations did we observe an excess of DNA sequence variant sharing that was statistically significant, although we did see small increases in sharing on chromosome 7q and 13q, but at a level very similar to many other regions of the genome in our study.

## Conclusions from Genome Scans to Date

Among all the autism linkage scans published to date, a total of 364 multiplex sibships have been surveyed (139 in our study; 99 in the International Consortium; 51 in the French consortium; 75 in the U.S. consortium). Based on the results of all these studies, it is now apparent that no single region contains a gene with a large effect on the risk of autism in most families. Indeed, the most notable linkage results published to date have been for a different chromosomal location in each of the studies. None of these locations has reached a level considered statistically significant for this type of genetic linkage study.

Although several groups may report positive results on the same chromosome, it is important to examine these results carefully. First of all, human chromosomes, such as chromosome 7, are very large, and positive results in different studies support each other only if they occur at the same location on the chromosome.

Second, the statistical magnitude of any positive result is important. By chance alone, we would expect an excess of sharing of DNA variants, so the amount of sharing must be fairly high to be relevant. Examination of the same chromosomal regions reported by researchers in other published studies (for example, on chromosomes 1, 7, 13 and 15) shows that, for none of the regions, is there consistent statistical replication across studies.

There are efforts currently underway to combine the results of the various linkage projects to determine whether other, more significant localizations can be identified. There are probably now more than 500 informative families collected worldwide that can be used for this effort. However, considering the lack of strong results published to date, it is still possible that no strong linkage results, even from such a large combined collection, will emerge. As described earlier, strong linkage signals are generally required to enable positional cloning of a disease gene.

As a result, we (and others) are currently diversifying our efforts to disentangle the complex genetic basis of autism. Much of our laboratory effort is now devoted to using the candidate gene approach, as this design can be much more powerful to identify a disease gene than can the linkage analysis approach. Although the few studies of this type that have been published recently have not yielded clear answers or replicable findings, we are optimistic that in the long run, this approach will eventually be fruitful. There are several reasons why this candidate gene approach is much more attractive now than when we began our genetic linkage study. First, the technologies for identifying genes and their variants, as well as for scoring variants in a large number of DNA samples, is much more advanced now than even a few years ago. Second, the Human Genome Project is nearing the goal of determining the entire DNA sequence of all of the human chromosomes, and from this effort, it will be possible to identify all or almost all of the 50,000 - 100,000 human genes, as well as the variation that exists in those genes. Thus, a catalog of incredibly valuable information is being generated, and this provides a very rich source of starting material for laboratories selecting genes for candidate gene studies. We are already taking advantage of this remarkable resource, and believe that it is necessary if we are to identify at least some of the genetic contributors to this disabling condition.

In the spirit of collegiality and to move genetic research in autism forward, we have posted downloadable and detailed results of our genome scan study on an internet

ftp site that is available to all autism researchers. We hope that other groups will follow suit as quickly as possible to help answer some of the genetic questions with which we are all currently struggling.

Finally, a growing collaboration of basic and clinical researchers at Stanford are bringing together innovative research methods that will complement the growing genetics database that exists on this disorder. In particular, new magnetic resonance imaging studies are underway that will measure both the structure and function of the brain in persons with autism. When combined with detailed genetic information, these data will provide a stronger framework from which to propose and verify more biologically meaningful subgroups within the autistic spectrum.

Prepared by Drs. Donna Spiker, Linda Lotspeich, Richard M. Myers, and Neil Risch, Stanford University School of Medicine. 2/8/00. For questions, contact Donna Spiker, Ph.D., Clinical Director at (650) 723-7809 or <<spiker@stanford.edu>>

## References for Selected Genetics Studies of Autism


Barrett, S., Beck, J.C., Bernier, R., Bisson, E., Braun, T.A., Casavant, T.L., Childress, D., Folstein, S.E., Garcia, M., Gardiner, M.B., Gilman, S., Haines, J.L., Hopkins, K., Landa, R., Meyer, N.H., Mullane, J.A., Nishimura, D.Y., Palmer, P., Piven, J., Purdy, J., Santangelo, S.L., Searby, C., Sheffield, V., Wang, K., & Winkolosky, B. (1999). An autosomal genomic screen for autism. *American Journal of Medical Genetics (Neuropsychiatric Genetics*), 88, 609-615.

Cook Jr, E.H., Courchesne, R.Y., Lord, C., et.al (1997b). Evidence of linkage between the serotonin transporter and autistic disorder. *Molecular Psychiatry*, 2, 247-250.

Cook Jr, E.H., Lindgren ,V., Leventhal, B.L., et. al (1997a). Autism or atypical autism in maternally but not paternally derived proximal 15q duplication, *American Journal of Human Genetics* , 60, 928-934.

Cook Jr, E.H,, Courchesne, R.Y., Cox, N.J., et. (1998). Linkage-disequilibrium mapping of autistic disorder, with 15q11-13 markers. *American Journal of Human Genetics*, 62, 1077-1083.

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. Archives of General Psychiatry, 53, 985-989.

Hallmayer, J., Spiker, D., Lotspeich, L., McMahon, W.M., Petersen, P.B., Nicholas, P., Pingree, C., Ciaranello, R.D. (1996). Male-to-male transmission in extended pedigrees with multiple cases of autism. American Journal of Medical Genetics: Neuropsychiatric Genetics, 67, 13-18.

Hallmayer, J., Pintado, E., Lin, A., Hebert, J., Lotspeich, L., Spiker, D., McMahon,W.M., Petersen, P.B., Nicholas, P., Pingree, C., Kraemer, H.C., Wong, D., Ritvo, E.R., Ciaranello, R.D., Cavalli-Sforza, L.L. (1994). Molecular analysis and test for linkage between the FMR-1 gene and infantile autism in multiplex families. American Journal of Human Genetics, 55, 951-959.

International Molecular Genetic Study of Autism Consortium (1998). A full genome screen for autism with evidence of linkage to a region of chromosome 7p. *Human Molecular Genetics*, 7, 571-578.

Lotspeich, L. J. & Ciaranello, R. D. (1993). The neurobiology and genetics of infantile autism. *International Review of Neurobiology*, 35, 87-129.

Pericak-Vance MA, Wolpert CM, Menold MM, et. (1997). Linkage evidence supports the involvement of chromosome 15 in autistic disorder (AUT). *American Journal of Human Genetics Supplement*, A40: 208.

Philippe, A., Martinez, M., Guilloud-Bataille, M., Gillberg, C., Rastam, M., Sponheim, E., Coleman, M., Zapella. M., Aschauer, H., van Malldergerme, L., Penet, C., Feingold, J., Brice, A., Leboyer, M, & the Paris Autism Research International Sibpair Study. (1999). Genome-wide scan for autism susceptibility genes. *Human Molecular Genetics*, 8, 805-812.

Risch, R., Spiker, D., Lotspeich, L., Nouri, N., Hinds, D., Hallmayer, J., Kalaydjieva, L., McCague, P., Dimiceli, S., Pitts, T., Nguyen, L., Yang, J., Harper, C., Thorp, D., Vermeer, S., Young, H., Hebert, J., Lin, A., Ferguson, J., Chiotti, C., Wiese-Slater, S., Rogers, T., Salmon, B., Nicholas, P., Petersen, P.B., Pingree, C., McMahon, W., Wong, D.L., Cavalli-Sforza, L.L., Kraemer, H.C., & Myers, R.M. (1999). A genomic screen of autism: Evidence for a multi-locus etiology. *American Journal of Human Genetics*, 65, 493-507.

Spence, M.A., Ritvo, E.R., Marazita, M.L., Funderburk, S.L., Sparkes, R.S. , Freeman, B.J. (1985). Gene mapping studies with the syndrome of autism. *Behavior Genetics*, 15, 1-13.

Spiker, D., Lotspeich, L., Kraemer, H.C., Hallmayer, J., McMahon, W.M., Petersen, P.B., Nicholas, P., Pingree, C., Wiese-Slater, S., Chiotti, C., Wong, D.L., Dimiceli, S., Ritvo, E.R.,Cavalli-Sforza,L.L., & Ciaranello, R.D. (1994). Genetics of autism: Characteristics of affected and unaffected children from 37 multiplex families. American Journal of Medical Genetics: Neuropsychiatric Genetics, 54, 27-35.