



We need a large sample size to have enough statistical power to detect the associations, because we don't expect any individual allele to have a large effect on the phenotype. Instead, we expect that there are many alleles contributing only a small fraction of phenotypic variability that we observe. The challenge is balancing the need for sample size without sacrificing the quality of the phenotyping.

**Q: How do you recruit and test subjects?**

**AP:** In the human impulsivity study, we screened the subjects to make sure that they weren't heavy drug users and that they didn't suffer from any major psychiatric disorders. They completed questionnaires online, as well as performed behavioral tasks in the lab to measure their degree of impulsivity quantitatively. The subjects provided a blood or saliva sample as a source of genomic DNA. We performed genotyping to assess alleles at hundreds of thousands of sites across the genome. We measured variable sites where both alleles are common in human populations. At each of those alleles, we determined whether or not a variation at that site influences variation in the phenotype that we're interested in.

We take a similar approach in rats and mice. We obtain a quantitative measure for a number of different phenotypes and then isolate genomic DNA from the tail tip or spleen, and used that sample to genotype hundreds of thousands of sites across the genome. One of the added benefits of working in animals is that we can measure genome-wide gene expression in trait-relevant brain regions. We can associate phenotypes with genotypes and gene expression, and determine if a particular locus influences a behavioral trait through its influence on the gene expression trait. We can test this hypothesis by manipulating gene expression in the model organisms, and then assessing how the implicated gene affects behavior.

**Q: What Illumina technologies have enabled these studies?**

**AP:** We've been working with Illumina technology for the past decade. In our early mouse genotyping studies, we customized a GoldenGate® array. More recently, we've relied entirely on NGS with the HiSeq 2500 System for mouse and rat studies. We use a genotyping by sequencing (GBS) protocol. Similarly, we used to use microarrays to measure gene expression in mice and rats. In the last three or four years, we've transitioned to using NGS to perform RNA sequencing (RNA-Seq) to obtain measures of gene expression in the brain quantitatively. In human genetic studies, we continue to use single nucleotide polymorphism (SNP) arrays for genotyping. We're really excited about the Illumina PsychArray.

“More recently, we've relied entirely on NGS with the HiSeq 2500 System for mouse and rat studies.”

**Q: What are the advantages of GBS over genotyping arrays?**

**AP:** GBS is a good fit for us because there aren't any SNP genotyping array products available for mice and rats with coverage for all the applications we're interested in. GBS is well suited for mice because they have a greater density of polymorphisms, with common alleles occurring at 1 per 100 sites. If you randomly sequenced regions of the mouse genome, you're more likely to come across an informative marker than if you did the same thing in humans.

GBS provides an efficient way to capture polymorphic alleles without having to design an array or select a subset of SNPs in advance. The selection of SNPs carries biases that can confound analysis. GBS gives us a relatively unbiased sampling of SNPs that exist within a particular population. Some SNPs are already known. The ones we're discovering *de novo* are either recent mutations or relatively rare in the mouse strains that have been sequenced to date. Working with Illumina, we've modified GBS to allow us to obtain 100,000 or more SNP genotypes per animal at a low cost.

“Working with Illumina, we've modified GBS to allow us to obtain 100,000 or more SNP genotypes per animal at a low cost.”

**Q: How are you using the Illumina Infinium® PsychArray BeadChip?**

**AP:** The PsychArray is perfect for our human studies. It was designed through a collaboration between Illumina and leading scientists in the field and includes customized content informative for psychiatric traits. When we're looking at intermediate phenotypes it's not clear *a priori* whether content relevant to schizophrenia, smoking, or drug abuse will be of interest. We'd much rather have a product that has all the SNPs previously implicated in psychiatric and behavioral traits represented. Then we don't have to impute key SNPs.



