Proposed Research for NIDA Genetics and Epigenetics Portfolio

Contributed by

NIDA Genetics Coordinating Committee and NIDA Genetics Workgroup*

January 23, 2015

*Will Aklin, Beth Babecki, Mark Caulder, Naresh Chand, Colin Hodgkinson, Meyer Glantz, Harold Gordon, Roger Little, Falk Lohoff, Gerald McLaughlin, Jonathan Pollock, Tanya Ramey, Jagadeesh Rao, Joni Rutter, John Satterlee, Shelley Su, Kevin Walton, Naimah Weinberg, and Da-Yu Wu

NIDA Genetics and Epigenetics Research Proposal

Mission Statement:

Elucidate the genetic and epigenetic mechanisms of substance abuse and addiction to facilitate diagnosis, treatment, and prevention.

Background to strategic goals:

NIDA's efforts in gene discovery have been highly successful in identifying gene variants for nicotine addiction, caffeine consumption, clearance of HCV, and host response to HIV. The discovery of the variants on 15q25 for chrna5-chrnb3-chrna3 has led to important insights into the role that the habenula plays in mediating the aversive component of nicotine. Ongoing gene discovery studies have provided promising results for opioid, cocaine, amphetamine, and cannabis dependence. Pharmacogenomic studies supported by NIDA suggest that variants in CYP2A6 can be used to predict the response to pharmacological aids for smoking cessation. Gene discovery efforts provide the foundation for identification of drug targets, tailoring treatments by genotype, and assist in identifying both environment and genetic factors that contribute to substance abuse. By comparing substance abuse gene discovery data sets with other genome wide association studies it possible to identify gene variants that are co-morbid with other disorders. Human genetic data will be used to inform preclinical genetic studies and vice versa, so that animal genetic studies are informative of human addiction.

Proposed Priorities for NIDA Genetics Research:

1. Discover additional common and rare variants for addictions to nicotine, cocaine, amphetamines, opioids, and cannabis, and identify the common genetic mechanisms underlying substance abuse.

2. Epigenetic analysis in animal models of addiction and epigenetic analysis of post-mortem human brain tissue

3. Functional characterization of genetic variants

4. Translation strategies for improved diagnosis, treatment, and prevention

1. Discover additional common and rare variants for addictions to nicotine, cocaine, amphetamines, opioids, and cannabis and identify the common genetic mechanisms underlying substance abuse.

a) Complete the genetic analysis of samples collected over the past 20 years.

More than 100,000 samples collected over the past 20 years remain to be genotyped. Genotyping these samples on a standardized chip platform will assist in identifying gene variants for opioids, psycho-stimulants, and cannabis. It may also provide sufficient power to identify gene variants that are common to all addictions. Sequencing of all these samples is not practical. The cost would exceed $100,000,000 and validated data analytic strategies do not exist to handle such quantities of data. In contrast, the cost with chip technology is about a tenth the cost, $10,000,000.

Over the next 2-4 years NIDA proposes to support genotyping of 50,000 samples through the NIDA Center for Genetic Studies (NCGS) at Rutgers University using the Smokescreen microarray. Funding for this initiative will come from both an SBIR contract to BioRealm for microarray purchasing and through the pre-existing NCGS contract for genotyping. Data will be made publicly available through dbGAP to assist investigators over the next five years to

- Identify additional common variants associated with nicotine, cocaine, amphetamine, and opioid addiction
- Personalize pharmacological treatment of smoking cessation
- Integrate substance abuse genetic data with data from the Psychiatric Genetics Consortium to identify genetic variants that co-occur with both substance abuse and psychiatric disorders, and identify likely risk and protective markers
- Genetically define the different stages of drug abuse and addiction

b) Recruit and genotype additional individuals with opioid, cocaine, or cannabis dependence

The identification of genetic variants for nicotine dependence in part resulted from the large number of samples recruited by NIDA investigators around the world. The number of samples for cocaine dependence, opioid dependence, and for cannabis dependence in the NIDA Center for Genetic Studies remains small and is under-powered. The small sample size makes replication of findings problematic. To address this deficiency, NIDA proposes to recruit additional participants suffering from opioid, cocaine, and cannabis addiction over the next five years by

- Leveraging genetic analysis of samples that contain large numbers of individuals with substance use disorders such as patients being treated for pain, infected with HIV, or other medical conditions associated with substance abuse.

- Accessing populations and studies available through the NIDA CTN to recruit additional participants and associated phenotypes and genetic/epigenetic markers. Encouraging interdisciplinary studies to confirm initial correlations and to identify associated mechanisms.

Developing and applying bioinformatic and computational approaches to provide novel insights into genomic and epigenomic data sets, including data in dbGAP and other databases as a control data set for smaller cohorts of opioid, cocaine, and cannabis dependence, to increase the statistical power and scope of the studies.

c) Develop and implement methods to identify the missing heritability. Genome wide association studies of complex traits have identified only a small fraction of heritability estimated by twin studies. Possible explanations for the missing heritability is an overestimation of heritability by twin studies arising from shared common environment, poorly defined phenotypes, interaction of heritabilities of phenotypes of different cell types that constitute a complex trait, epistasis, poor chip coverage of rare variants, gene-environment interactions and somatic mutation.

To address these issues, NIDA proposes in the next five years to:

Increase emphasis on continuous quantitative phenotypes: Having a large population on which to do genetic analysis does not provide results if the wrong phenotype is selected. This has been a particular problem in searching for genes for mental illness where symptoms were used as the basis for searches for gene variants, e.g. schizophrenia. However, when a more defined phenotype is used, e.g. psychosis, then genes associated with that emerged and were found across DSM diagnoses (e.g. schizophrenia, bipolar disorder). Thus, carefully selected, measurable phenotypes, or symptom cluster domains based on biomarkers or traits shown to be heritable in animals (e.g. impulsivity) that more closely underlie the biological mechanism, provide increased power and reduce the need for ever increasing sample sizes. For example, the use of carbon monoxide as a continuous quantitative trait significantly reduced the sample size needed to identify genes on 15q25 that are associated with nicotine dependence. Thus, future genetic studies should end reliance on DSM criteria and use quantitative measures, enabling integration of human and animal genetic studies. Over the next five years NIDA proposes to encourage studies that will identify gene variants in model genetic organisms and in humans using continuous quantitative phenotypes such as:

- Amount of drug ingested
- Frequency of use
- Length of abstinence
- Somatic and affective symptoms of drug withdrawal
- Novelty preference or novelty seeking

- Cue reactivity as a marker of craving
- Impulsivity
- Inhibitory control
- Reward processing
- Stress reactivity
- Working memory
- Emotional regulation
- Motivation
- Disruption of circadian rhythms
- Structural, functional, and chemical neurochemical correlates
- Metabolomic markers of substance use
- Neuroimaging of neuronal activity

Identify Rare Variants

Another factor contributing to the missing heritability is the failure of gene chips to cover rare variants and the extraordinarily large sample sizes needed in an association study to capture information from these variants. One approach to identify rare variants is to conduct targeted next generation sequencing in a region that has been shown to be associated with a substance abuse phenotype. This has proven to be successful for identifying rare variants for smoking. An alternative approach to identify rare variants is to carry out deep sequencing in families and pedigrees at risk for substance abuse. In the next five years NIDA proposes to

- Leverage the "The Adolescent Brain and Cognitive Development" initiative and collect biosamples from parents and children for a family- based or trio design for subsequent genotyping and identification of rare variants for substance abuse related phenotypes.
- Leverage the family-based COGA study to identify rare variants for cocaine addiction comorbid with alcoholism
- Leverage well characterized lineages, i.e., Iceland, Finland, etc. to identify rare variants in human populations

Discover gene-gene, gene-development, and gene-environment interactions

Genetic epidemiologic and molecular genetic approaches have contributed increasingly significant advances to our understanding of the causes of use, abuse, and dependence of addictive substances. These studies have established that SUDs are complex developmental disorders, with high heritability that are also strongly influenced by environment; particularly early in adolescence and at the point of drug use initiation. New genetic methodologies are needed to elucidate the complex interplay of genetic and environmental factors in developmental trajectories of SUDs and comorbid conditions.

Epistatic interactions (gene-gene), gene-development, and gene-environment interactions are likely to contribute to the missing heritability that is currently lacking in the field. To uncover these interactions from whole genome association studies will require methods to reduce data dimensions to overcome the problem of multiple comparisons. Mice with defined genetic backgrounds (inbred strains, recombinant inbred strains, congenic, consomic strains, and mouse strains carrying defined naturally and induced genetic variations) provide a way to test gene-environment, and gene-development interactions under controlled experimental conditions.

In the next five years NIDA proposes to encourage research that

- Examines interactions among multiple gene variants and the environment and development instead of single gene interactions.
- Develops statistical methods to reduce data dimensions of whole genome wide analysis to overcome problems of multiple comparisons.
- Identifies mechanisms of epistasis, gene-environment, gene-development interactions in humans and in recombinant inbred animals and across phenotypically characterized different inbred strains.

d) Gene Discovery in Model Organisms

Genetic mapping of substance abuse-related phenotypes in model organisms may also suggest candidate genes to test in human populations and provide powerful insights into the mechanisms of substance use disorders. For instance, the identification of the *ob* gene in mice as encoding leptin gave great insight into metabolism but is a rare variant in the human population. In the next five years NIDA proposes to encourage studies to

- Discover genetic variants in inbred and outbred animals using selective breeding strategies, haplotype associated mapping, and QTL mapping for identification and refinement of heritable phenotypes (i.e. impulsivity, novelty seeking, etc.)
- Identify phenotypic and genetic correlations
- Create genomic resources in non-human primates

2. Epigenetic analysis in animals models of addiction and epigenetic analysis of post-mortem human brain tissue

Epigenetic mechanisms (e.g. DNA methylation, histone modification, expression of certain non-coding RNAs, splicing of retrotransposons, miRNAs) are well known mediators of organismal development and appear to play a significant role in mediating organismal responses to environmental stimuli. Mouse genetic studies may provide an opportunity to explicitly investigate the importance of epigenetic phenomena in mediating gene-development or gene-environment effects related to substance use

disorders. For example it may be possible to identify mouse genetic variants that influence epigenetic regulation of addiction-related phenotypes.

Validate relevance of epigenetic findings in animals to the human condition by investing in epigenetic studies of post-mortem brain tissue of substance abusers

NIDA has invested in research that examined epigenetic modifications of gene expression in animals yet these findings have yet to be validated in post-mortem human tissue. Using human tissue to validate findings in animal models and identify differences between the two is key to maximizing the translation of animal genetic findings to humans. Epigenetic studies in post-mortem human brain are particularly needed because epigenetic changes in peripheral tissue are not necessarily correlated with epigenetic changes in the brain. These studies are expected to reveal the role of epigenetic mechanisms in substance use disorders and to elucidate the role that somatic mutation, non-coding RNAs, retrotransposon hopping, retrotransposon splicing, imprinting, and other variables play in processes that underlie substance abuse.

In the next five years, NIDA proposes to:

- Focus efforts related to basic research on epigenetics and non-coding RNAs and encourage research in the following areas of opportunity:
  - Non-coding RNA regulation in neuroplasticity and addictive processes
  - Basic research on structure and function of RNA modifications in the nervous system
  - Investigate the role of higher order chromatic structure in the nervous system in response to neuroplastic changes and/or drugs of abuse
  - Develop small molecular epigenetic modulators that target epigenetic regulators with expression or function only in the nervous system, particularly in regions and cell types relevant to substance use disorders
  - Examine the stability and dynamics of somatic epigenomic changes in response to drugs of abuse and potential drug abuse therapeutics
  - Epigenetic, gene expression, behavioral, neuroscience, and transgenerational effects of prescription opioid exposure
  - Discover epigenetic modifications that modify the expression or activation of retrotransposons
  - Identify epigenetic changes in post-mortem human brain tissue and validate epigenetic changes seen in animals

- Improve methods to map the epigenome

- Determine how the microbiome, diet, exercise, family environment, and/or environmental toxins affect substance use disorders through the epigenome.

3. Functional Analysis of Genetic Variants.

Association of gene variants with substance use disorders is insufficient to show causality or demonstrate function. Knockouts of chrna5 and knockins of the A118G have successfully linked nicotine aversion to these genetic loci and identified signaling pathways to focus drug discovery efforts for nicotine addiction. In the next five years NIDA proposes to encourage studies that

- Characterize the function of genes in whole organisms including non-human primates using the CRISPr/Cas9 system.
- Characterize the function of genes in whole organism by leveraging the KOMP2 resource.
- Differentiate iPSC cells derived from lymphocytes of affected individuals stored at the NIDA Center for Genetics and other repositories, as well as from KOMP2 mouse ES cells into the appropriate cell type to analyze the function of a gene.
- Use other cell culture systems to measure the impact of gene variants on cell physiology.

Addictome

In the next five years, the NIDA genetic program proposes to expand addictome-related research by supporting studies that

- Utilize single cell analysis and heterogeneity studies for developing brain, with emphasis on learning, memory and emotional control
- Leverage the BRAIN initiative to identify unique cell types within the nervous system
- Identify genetic variants that regulate variation in neuronal circuitry
- Generate or assemble a diverse, interoperable collection of multi-scale data sets that can be mined by the scientific community and visualized in a user-friendly, 4D-framework to discover novel relationships and scientific knowledge gaps for the stages of addiction

4. Translation strategies for improved diagnosis, treatment and prevention

Recent genetic studies suggest that genetic variants can be used to predict treatment response in smoking cessation. For example, different allelic variants of Catechol-O-

methyltransferase VAL158MET polymorphism (COMT VAL/MET), *CHRNB2, CHRNA5*, and *CHRNA4, CYP2A6, CYP2B6*, and *$5HT_3$* have been shown to affect risk for smoking (some protective, some higher risk) and duration of relapse to smoking in nicotine replacement trials as well as response to buproprion and nicotine replacement. The VAL158MET polymorphism also has been associated with response to to methylphenidate response in ADHD children with VAL/VAL homozygotes showing significantly less symptoms than MET/MET homozygotes following methylphenidate treatment. In addition, there are data to suggest that specific biomarkers, particularly modulation of cortical-limbic connectivity (or reduced metabolism in several prefrontal regions), are associated with response to cognitive behavior therapy or antidepressant medications among individuals with major depressive disorder, a condition often comorbid with many substance use disorders. Exploratory research to identify potential genetic variants and SNPs (mainly focused in the dopaminergic pathway) that might be associated with response to substance use disorder treatment is essential.

Improved SUD therapy relies heavily on studies that can integrate specific genetic polymorphism data with specific treatment protocols. Additionally, it is important to determine whether characteristics that are already known to moderate treatment outcome (e.g., drug dependence severity, gender, ethnicity, or concurrent alcoholism) are associated with specific polymorphisms. For example, opioid analgesics are widely used to treat pain; however, they have a narrow therapeutic window, are accompanied by many adverse effects, and often do not provide adequate pain relief. In addition, genetic factors affecting pain pathways, pain perception, analgesic metabolism (pharmacokinetics), transport and receptor signaling account for large individual differences in opioid analgesic efficacy and adverse effects. Exploratory analyses can begin to determine whether various polymorphisms differentially predict success in different types of SUD treatments, including pharmacological approaches, by conducting large prospective trials that track treatment success along with genetic polymorphism data.

NIDA researchers have tested models showing that genetic markers may influence the efficacy of universal preventive interventions. Additional studies are needed to substantiate this work and elucidate how genetic factors impact preventive interventions for at risk youth and youth who have used drugs but do not have a diagnosis/disorder and whether the efficacy of the intervention is also related to when the intervention is delivered during the course of development of substance use disorders. Understanding how genetic differences influence the efficacy of interventions is critical for the development and refinement of effective approaches that prevent SUDs.

In the next five years, the NIDA Genetic program proposes to support

- A prospective pharmacogenomics trial of smoking cessation
- Discovering the genetic variants affecting the efficacy of opioid analgesics and associated adverse effects to maximize analgesia and decrease side effects

- Elucidating the genetics of vaccine response against substance abuse (i.e., antibody titer, free and bound drug in tissue) utilizing diversity outcross (DO) in mice and in human populations
- Studies examining the genetics of drug toxicity
    - Drugs of abuse
    - Therapeutics
- Studies examining the genetics of treatment response in model organisms
    - Utilize DO and inbred strains to identify and test genetic/syntenic loci
- Developing CRISPR/CAS9 for human gene therapy methods.
- Improving substance abuse treatment efficacy by honing in on the specific biologic processes that leaves one susceptible to addiction and/or responsivity to treatment
- Genome wide association studies (GWAS) to develop personalized and targeted treatments

HIV/HCV

NIDA has had considerable success in identifying gene variants associated with HIV/HCV infectivity. A key challenge now is the development of technologies to delete the integrated provirus from the genomes of somatic cells infected with HIV. These technologies include nanoparticles to deliver RNAi and CASP/CAS9. In the next five years NIDA proposes to encourage studies that support

- Systems biology,
- Epigenetics, and
- Biomarkers of HIV and HCV progression

SBIR/STTR

The development and commercialization of technologies that treat SUDs can have a large impact on public health as well as on the ability to conduct research. The NIDA Genetics program proposes to encourage the development of technologies that treat SUDs by releasing RFAs utilizing the SBIR, STTR, and other contract mechanisms to support technology development in the following areas:

- Protein capture
- Metabolomics
- Gene delivery
- Analysis of fixed tissue specimens
- shRNA, microRNA
- Biomarkers
- iPS cells for therapeutic development
- Data analysis, integration, visualization, etc.
- Computerized methods for phenotyping individuals on active tasks.