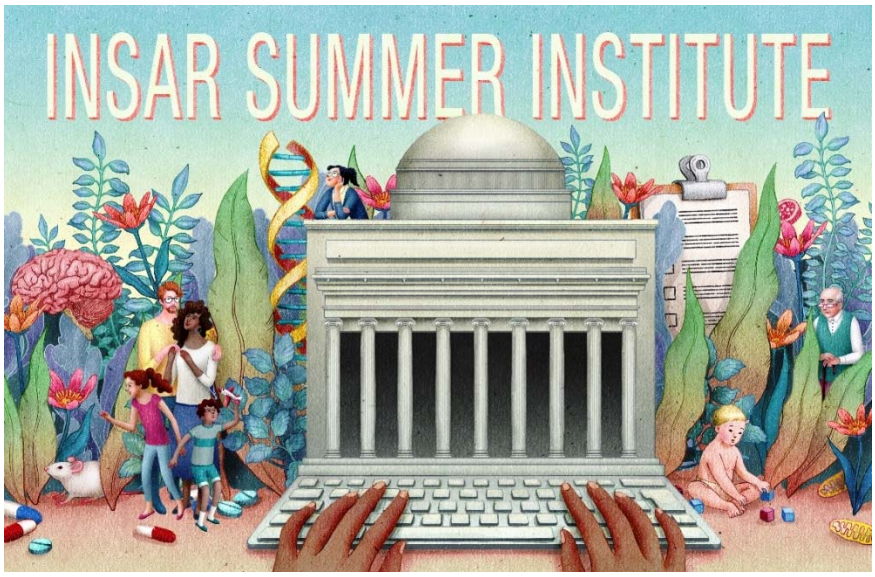
INSAR | International Society for Autism Research



SESSION I: Familial Risk Factors and Comorbidities

Dr. Anthony J. Bailey

Professor and Chair of Child and Adolescent Psychiatry
University of British Columbia

June 23, 2016 // 3:00 pm EDT

Course Materials

The purpose of these materials is to help provide an introduction to the Summer Institute session on familial risk factors and comorbidities. The materials were designed to prepare trainees who are unfamiliar with genetics research with the general background to get the most educational benefit from Dr. Bailey's presentation. Toward this objective, we have prepared the following: (1) learning objectives for this session; (2) some key terms and concepts to become familiar with different methods and concepts in genetic research; (3) some broad review articles that are recommended reading. These materials could be considered "prerequisites" in preparing for Dr. Bailey's presentation.

In collaboration with Dr. Bailey, these materials were developed by **Laura G. Holmes, M.S.** (clinical psychology intern at the Children's Hospital of Philadelphia Center for Autism Research; laura.graham@psych.utah.edu) and **Rachel K. Earl, Ed.S.** (school psychology graduate student at the University of Washington; rkinc78@uw.edu) Feel free to contact us with questions/comments.

Register for this course and other sessions in this series by creating an account at http://eweb.autism-insar.org/eweb/

LEARNING OBJECTIVES

The Summer Institute for Autism Research was established in direct response to requests from early career researchers (graduate students, postdocs, etc.), who asked INSAR for greater training opportunities in multidisciplinary topics. In designing the Summer Institute, the priorities were: (1) to provide a multidisciplinary training platform for young scientists from various backgrounds; (2) allow international participation; and (3) make it freely available. Thus, the second Summer Institute covers broad topics (which are geared to researchers outside the respective topic areas), is offered over a free web platform, and allows researchers from around the world to connect with the presenter. The overarching goal of the Summer Institute is to expose junior scientists to topics they are not currently engaged in, with the hope that basic scientists and clinical scientists could learn from each other to ultimately advance the understanding of autism spectrum disorders.

The current session, Familial Risk Factors and Comorbidities, is lead by Dr. Anthony Bailey and a team of trainees who worked in tandem to prepare these materials and the web presentation. The learning objectives for attendees of this session include:

- **To learn about the different methods that are used to study the genetics and comorbidity patterns of individuals with autism spectrum disorder.**

- **To obtain a broad overview of the history and current research on genetics and comorbidity studies of autism.**

- **To gain the requisite background in genetics and comorbidity in order to read and interpret the related literature more effectively.**

- **To receive a deeper understanding of the areas of genetics and comorbidity that are among the expertise of the presenter, Dr. Anthony Bailey.**

- **To understand the potential clinical benefit that genetic research could have for individuals with autism spectrum disorder.**

GLOSSARY OF KEY TERMS AND CONCEPTS

Basic Terms

Affected relative-pair design: Studies of families with relative pairs who are both affected (usually at least 200 pairs).

Allele: One of two or more alternate forms of a gene, differing in DNA sequence, and affecting the functioning of a single RNA or protein**.**

Amino acids: Organic molecules that make up protein. There are 20 amino acids that are essential for human life to exist, and all proteins in the body are made up of some combination of those 20 amino acids. A sequence of three base pairs of nucleotides codes for a single amino acid. This code can be redundant. Mutations can alter the sequence of the base pairs and lead to a sequence that does not code for an amino acid.

Association studies: Test for a relationship between disease status and genetic variation to identify candidate genes or genome regions that contribute to that disease. For example, a higher frequency of a single-nucleotide polymorphism (SNP) allele or genotype in a series of individuals affected by a disease can be interpreted as meaning that that variant increases the risk for that disease.

Chromatin: The DNA in chromosomes is also associated with transcription factors and other macromolecules. Chromatin in a complex of macromolecules found in cells and consists of DNA, protein, and RNA (the messenger between DNA and proteins). Chromatin is involved in DNA folding and how DNA is accessed in the cell. More specifically, it functions to (1) package DNA into a smaller volume to fit in the cell, (2) reinforce the DNA macromolecule and allow mitosis, (3) prevent DNA damage, and (4) to control gene expression and DNA replication.

Chromosomes: Thread-like structures located inside the nucleus of animal and plant cells. Each chromosome is made of protein and a single molecule of DNA. Rather than being found on its own, it is structured by being wrapped about protein complexes called nucleosomes, which consist of proteins called histones.

Codon: Sequence of three nucleotides that form a unit of genetic code; each codon specifies a specific amino acid to be incorporated into a protein. Ribosomes "read" these 3-nucleotide sequences from a molecule of messenger RNA and assemble chains of amino acids which then fold to become functional proteins.

Copy number variations (CNVs): A genetic variant that is a large gain (duplication) or loss (deletion) of a piece of DNA. CNVs can include a variable number of additional repeats in DNA sequence. Some CNVs appear to have no effect on phenotype, but many have been linked to diseases.

Epistasis: This is when the effect of one gene is dependent on the presence of one or more 'modifier genes.' This may be a situation when one gene masks the expression of another. Epistatic mutations have different, non-additive effects in combination rather than individually.

Gene: A region of DNA (the molecule that carries genetic information) that helps determine a characteristic.

Gene-environment interactions: When different genotypes respond to environmental variation in different ways. For example, genetic sensitivity to environmental risk factors may be inherited rather than the disease itself.

Genetics: The study of heredity and the variation of inherited characteristics.

Genotype: The genetic composition of an organism, as distinguished from its physical appearance; the set of genetic variants, or alleles, carried by an organism.

Heritability: A statistic indicating how much variation in a phenotypic trait in a population is due to genetic variation among individuals in that population. Other causes are characterized as environmental factors. The degree to which a trait is genetically determined.

Histones: The chief protein components of chromatin, acting as spools around which DNA winds, enabling DNA to be compact enough to fit in cell nuclei. Histones can be modified through methylation and other processes and thus play a role in gene regulation (i.e., activating or repressing genes) and DNA repair.

Incomplete penetrance: A case in which clinical symptoms of a disorder are not present in every individual with the disorder-causing mutation(s). In other words, the mutation is not phenotypically expressed. For example, women with the BRCA1 gene have an 80% lifetime risk of developing breast cancer; the penetrance of the condition is therefore 80%.

Inheritance: A quality, characteristic, or predisposition derived genetically from one's parents or ancestors.

Insertion-deletion (indel): a variation in the DNA that involves the insertion or deletion of a small number of nucleotides (the boundary between a small CNV and an in-del is arbitrary).

Locus: Specific location of an allele, or form of a gene, on a chromosome. Each chromosome carries many genes.

Mendelian: A trait or disease transmitted through a single locus. An individual inherits an allele from each parent such that there is a pair of alleles; the dominant form of the gene results in the trait being expressed in an individual's phenotype.

Mutations: A permanent alteration in the DNA sequence that makes up a gene, such that the sequence differs from what is found in most people. Mutations can affect a single DNA building block (base pair) or a large segment of a chromosome that includes multiple genes. Mutations can be inherited from a parent or spontaneous, largely as the consequence of (cosmic) radiation.

Nucleotide: Molecules that form the basic structural units of nucleic acids, DNA and RNA (A, adenine; C, cytosine; G, guanine; T, thymine (DNA); U, uracil (RNA)). Often referred to as "bases," which are paired together to form DNA.

Oligogenic: Oligogenic inheritance is controlled by a few genes responsible for major heritable changes. Oligogenic diseases are produced by combinations of two, three, or four genes; one gene alone is not enough to result in full expression of the trait or disease but requires the presence of other genes.

Penetrance: The proportion of individuals of a specified genotype that show the expected phenotype under a defined set of environmental conditions. For example, if everyone carrying a mutant gene also has the mutant phenotype, the gene is said to have complete penetrance.

Phenotype: The observable properties of an organism, produced by the genotype and environment.

Polygenic: A quantitatively variable phenotype dependent on the interaction of numerous genes. In other words, many genes make a small contribution to the overall outcome, which can be a spectrum or continuum of possible phenotypes. Height and eye color are examples of polygenic traits. Polygenic diseases occur when many mild defects in genes come together to produce chronic diseases; there are hypothesized to be complex multiple interacting defects.

Quantitative Trait Locus (QTL): Quantitative traits are phenotypes (characteristics) that vary in degree and can be attributed to polygenic effects (i.e., the product of two or more genes and their environment). A QTL is a gene that controls the expression of a trait.

Variable expression (expressivity): The range of phenotypes expressed by a given genotype under any given set of environmental conditions or over a range of environmental conditions.

Single-nucleotide polymorphism (SNP): a variation of a single nucleotide at a given position that occurs at a frequency greater than 1% in the genomes of a population; some may be involved in human diseases but most probably are not. SNPs are useful signposts when scientists scan genomes for mutations.

Single-nucleotide variant (SNV): variant that changes a single nucleotide, similar to the SNP, except that it is *rare* (<1%) in the population. AKA: point mutation.

Stop codon: A specific sequence of 3 amino acids that signals to a ribosome to stop translation. (UGA (RNA), TGA (DNA)).

<u>Genetic Methods and Associated Terms</u>

Broader autism phenotype (BAP): An idea, first proposed by Folstein & Rutter (1977), that genetic liability for autism might be expressed in relatives of an individual with autism as characteristics that are milder but qualitatively similar to those seen in autism. Studies with relatives of individuals diagnosed with ASD have observed qualitatively similar deficits in social, stereotyped-repetitive, and communication domains.

Candidate gene and genome screen approaches: In the candidate gene approach, a specific gene of known location is suspected of being involved in a disease. In genome screening, there is an understanding of the pathophysiology of a disorder but no strong hypotheses as to specific genes involved. A genome scan allows scientists to examine biological markers across all of the chromosomes.

Concordance: The presence or absence of a specific trait in both members of a pair of twins.

Exons: Exons are coding sections of an RNA transcript, or the DNA encoding it, that are translated into protein. Exons can be separated by intervening sections of DNA that do not code for proteins, known as introns.

Family history studies: Studying disease status in a particular family using family history interview data (may be problematic due to recall bias).

Introns: Noncoding sections of an RNA transcript that are spliced out before the RNA molecule is translated into a protein

Linkage studies: Examine large families where the disease affects individuals in several generations, and identifies genetic marker(s) that is/are always inherited by family members with the disease but not those without the disease. Linkage studies often start by identifying genetic markers on a section of chromosome (often SNPs) and then narrow this down to a gene or gene variant is identified. Require families with multiple affected members in at least two generations.

Multifactorial threshold model: Assumes that (1) many factors contribute to a disorder, (2) that the effects of each factor are small but are additive or exponentially affect the organism, and (3) that once these factors pass a critical threshold, an individual is affected by the disorder.

Recurrence risk: The risk that a genetic defect that has appeared once in a family will appear in a child born subsequently.

Twin studies method: Twin studies reveal the importance of genetic and environmental influences on phenotypic traits. Twins allow the study of varying family environments and differing genetic make-up. Monozygotic (identical) twins share 100% of genes, suggesting that most differences are due to experiences that only one twin has. Dizygotic (fraternal) twins share only about 50% of genes. In a comparison of identical vs. fraternal twins, if identical twins are much more similar on a given trait, this implies that genes played an important role. Twin studies have shown that almost all studied traits are in part influenced by genetic differences, which some showing a strong genetic influence (e.g., height), others a moderate influence (e.g., personality traits), and some more complex heritabilities, with evidence for different genes affecting different aspects of the trait (e.g., autism).

Whole exome sequencing: A technique for sequencing all of the protein coding regions (exons) in a genome (termed the exome), which make up about 1% of the human genome. Scientists sequence only the DNA that encodes proteins (exons). Lower cost than whole genome sequencing, and especially effective for identifying genetic variants associated with rare Mendelian diseases. Click for a recent NPR article about the application of this technology to help families.

Whole genome scans: Machines look through all of a person's DNA looking for variations that could be associated with disease. It is easy for scientists (and physicians) to find mutations caused by a single mutation (e.g., cystic fibrosis) but more difficult for diseases like heart disease that involve many mutations and environmental factors. It is an approach that identifies genetic linkage using genetic markers as opposed to genes themselves.

Whole genome sequencing: Examines the complete DNA sequence of an organism's genome; this includes chromosomal DNA and DNA contained in mitochondria.

RECOMMENDED BACKGROUND READING

Articles:

Bailey, A., Le Couteur, A., Gottesman, I., Bolton, P., Simonoff, E., Yuzda, E., & Rutter, M. (1995). Autism as a strongly genetic disorder: evidence from a British twin study. *Psychological medicine*, *25*(01), 63-77.

Bolton, P., Macdonald, H., Pickles, A., Rios, P. A., Goode, S., Crowson, M., ... & Rutter, M. (1994). A case-control family history study of autism. *Journal of Child Psychology and Psychiatry*, *35*(5), 877-900.

Jeste, S. S., & Geschwind, D. H. (2014). Disentangling the heterogeneity of autism spectrum disorder through genetic findings. *Nature Reviews Neurology*, *10*(2), 74-81.

Pinto, D., Delaby, E., Merico, D., Barbosa, M., Merikangas, A., Klei, L., ... & Vorstman, J. A. (2014). Convergence of genes and cellular pathways dysregulated in autism spectrum disorders. *The American Journal of Human Genetics*, *94*(5), 677-694.

Useful online tutorials on basic genetic concepts:

http://learn.genetics.utah.edu/

Specific modules

- What is DNA? http://learn.genetics.utah.edu/content/molecules/dna/
- What is a gene? http://learn.genetics.utah.edu/content/molecules/gene/
- Anatomy of a gene http://learn.genetics.utah.edu/content/molecules/geneanatomy/
- How do cells read genes? http://learn.genetics.utah.edu/content/molecules/dnacodes/
- What is a chromosome? http://learn.genetics.utah.edu/content/chromosomes/intro/

Genetic Variation

- Sources of variation http://learn.genetics.utah.edu/content/variation/sources/
- What is mutation? http://learn.genetics.utah.edu/content/variation/mutation/
- Mutation and haplotypes http://learn.genetics.utah.edu/content/variation/haplotype/