07 - 1.7 Neurogenetics

1.7 Neurogenetics

have a beneficial effect, not only in terms of mitigating the effect of stress on immune functioning but perhaps also on diminishing emotional disturbances that arise in the context of immune system dysregulation. Two factors that have been repeatedly identified as protective against stress-induced immune alterations are social support and the ability to see stressors as being to some degree under the individual's control. In this regard, a recent study that conducted a genome-wide scan to assess gene expression activity in socially isolated versus nonisolated individuals found that social isolation was associated with increased activation of a number of proinflammatory, cytokine-related pathways and reduced activity in anti-inflammatory cytokine pathways, as well as in the glucocorticoid receptor, which plays an important role in neuroendocrine control of inflammatory processes. Of interest, the two types of psychotherapy most often examined in illnesses associated with immune dysregulation are group therapy, which provides social support, and cognitive behavioral therapy, which provides cognitive reframing techniques aimed at enhancing one's sense of agency (and hence control). REFERENCES Bajramovic J. Regulation of innate immune responses in the central nervous system. CNS Neurol Disord Drug Targets. 2011;10:4. Capuron L, Miller AH. Immune system to brain signaling: Neuropsychopharmacological implications. Pharmacol Ther. 2011;130(2):226. Danese A, Moffitt TE, Pariante CM, Ambler A, Poulton R. Elevated inflammation levels in depressed adults with a history of childhood maltreatment. Arch Gen Psychiatry. 2008;65:409. Dantzer R, O'Connor JC, Freund GG, Johnson RW, Kelley KW. From inflammation to sickness and depression: When the immune system subjugates the brain. Nat Rev Neurosci. 2008;9:46. Raison CL, Borisov AS, Woolwine BJ, Massung B, Vogt G, Miller AH. Interferon-α effects on diurnal hypothalamicpituitary-adrenal axis activity: Relationship with proinflammatory cytokines and behavior. Mol Psychiatry. 2010;15:535. Raison CL, Cowles MK, Miller AH. Immune system and central nervous system interactions. In: Sadock BJ, Sadock VA, Ruiz P, eds. Kaplan & Sadock's Comprehensive Textbook of Psychiatry. 9th edition. Philadelphia: Lippincott Williams & Wilkins; 2009:175. Ransohoff RM, Brown MA. Innate immunity in the central nervous system. J Clin Invest. 2012;122(4):1164. Steiner J, Bernstein HG, Schiltz K, Müller UJ, Westphal S, Drexhage HA, Bogerts B. Immune system and glucose metabolism interaction in schizophrenia: A chicken–egg dilemma. Prog Neuropsychopharmacol Biol Psychiatry. 2014;48:287–294. Wilson EH, Weninger W, Hunter CA. Trafficking of immune cells in the central nervous system. J Clin Invest. 2010;120(5):1368. Yousef S, Planas R, Chakroun K, Hoffmeister-Ullerich S, Binder TM, Eiermann TH, Martin R, Sospedra M. TCR bias and HLA cross-restriction are strategies of human brain-infiltrating JC virus-specific CD4+ T cells during viral infection. J Immunol. 2012;189(7):3618.

Starting from the rediscovery of Gregor Mendel's basic concepts at the turn of the 20th century, the field of genetics has matured into an essential cornerstone not only of the biological sciences but of all of medicine. The discovery of the basic structure and properties of deoxyribonucleic acid (DNA) in the middle of the century led to an exponential acceleration in our understanding of all aspects of the life sciences, including deciphering the complete sequence of the human genome, and those of myriad other species. Massive databases of such sequences now provide 21st-century biologists with the task of decoding the functional significance of all this information. In particular, attention has turned to determining how sequence variations contribute to the phenotypic variation between species and between individuals within a species; in humans it is hoped that discoveries about the relationship between genotypes and phenotypes will revolutionize our understanding of why and how some individuals but not others develop common diseases. This hope is particularly strong for psychiatry, as our knowledge of the pathogenic mechanisms of psychiatric disease remains sparse. Genetic mapping studies aim to identify the genes implicated in heritable diseases, based on their chromosomal location. These studies are carried out by investigating affected individuals and their families through two approaches, linkage and association (Fig. 1.7-1). It is now straightforward to genetically map Mendelian traits (traits for which a specific genotype at one particular locus is both necessary and sufficient to cause the trait). Psychiatric diseases, however, do not follow simple Mendelian inheritance patterns but rather are examples of etiologically complex traits. Etiological complexity may be due to many factors, including incomplete penetrance (expression of the phenotype in only some of the individuals carrying the disease-related genotype), the presence of phenocopies (forms of the disease that are not caused by genetic factors), locus heterogeneity (different genes associated with the same disease in different families or populations), or polygenic inheritance (risk for disease increases only if susceptibility variants at multiple genes act in concert). Mapping a complex disorder involves several component steps, including definition of the phenotype to be studied, epidemiological studies to determine the evidence for genetic transmission of that phenotype, choice of an informative study population, and determination of the appropriate experimental and statistical approaches.

FIGURE 1.7-1 Comparison of gene-mapping strategies. Genetic mapping approaches can be divided into those that rely on linkage analysis and those that rely on association analysis. Linkage studies can be further categorized as either focused on investigation of pedigrees or focused on investigation of sib pairs. Association studies can be categorized as either case–control or family-based. Some of the key features as well as advantages and disadvantages of these different approaches are shown. (From Sadock BJ, Sadock VA, Ruiz P. Kaplan & Sadock's Comprehensive Textbook of Psychiatry. 9th ed. Philadelphia: Lippincott Williams & Wilkins; 2009:321.) GENETIC EPIDEMIOLOGICAL APPROACHES Genetic epidemiological investigations provide quantitative evidence regarding the degree to which a given trait aggregates in families and, furthermore, can suggest to what degree such aggregation reflects a genetic contribution to the etiology of the trait. Family studies compare the aggregation of disease among the relatives of affected individuals compared to control samples. Because these studies do not differentiate between genetic and environmental contributions to such familial aggregation, they provide only indirect evidence regarding the heritability of a trait. Often these studies

measure the relative risk (λ), defined as the rate of occurrence of a disease among specified categories of relatives of an affected individual divided by the rate of occurrence of the disease for

the general population. A relative risk of >1 suggests a genetic etiology, and the magnitude of the measure gives an estimate of the genetic contribution to the disease. Relative risks can be calculated for sibling pairs, parent– offspring pairs, and various other types of family relationships. Likely modes of transmission can be assessed by comparing the degree of relative risk for each type of relationship. Multiple family studies have been carried out for many of the major psychiatric disorders, including major depression, bipolar disorder, schizophrenia, and obsessive-compulsive disorder (OCD). Although these studies have consistently reported familial aggregation for all of these disorders, the degree of such aggregation has varied substantially across studies, largely reflecting differences in phenotype definition and how study samples were ascertained and assessed. Twin studies examine the concordance rates of a particular disorder (the percentage of twin pairs where both twins have the disorder) in monozygotic (MZ) and dizygotic (DZ) twins. For a disorder that is strictly determined by genetic factors, the concordance rate should be 100 percent in MZ twin pairs (who share 100 percent of their genetic material) and 25 or 50 percent in DZ twin pairs (who are no more closely related than any siblings), depending on whether the disease is recessive or dominant, respectively. For a disorder where genetic factors play a role in disease causation but are not the exclusive cause of disease, the concordance rates should be greater for MZ twins than those for DZ twins. The higher the degree of concordance of MZ twins, the higher the trait heritability or the evidence for a genetic contribution to disease risk. When genetic factors do not play a role, the concordance rates should not differ between the twin pairs, under the simplifying assumption that the environment for MZ twin pairs is no more similar than that for DZ twin pairs. The several twin studies that have been conducted for traits such as autism, bipolar disorder, and schizophrenia have consistently suggested high heritability and have therefore spurred efforts to genetically map loci for each of these conditions. Different twin studies may however generate varying point estimates for the heritability of any given disorder. When evaluating the results of twin studies, it is therefore important to scrutinize how the phenotype was ascertained because, as with family studies, the different heritability estimates are likely due to differences in the mode of assessing and defining phenotypes. For example, early twin studies of psychiatric disorders often relied for their phenotypes on unstructured interviews by a single clinician. In contrast, modern studies generally utilize standardized assessments and review of diagnostic material by a panel of expert clinicians. Similarly, part of the apparent variation in heritability between different twin studies can be attributed to the fact that some studies employ narrow definitions of affectedness for a given phenotype, while other studies employ broader phenotype definitions (e.g., considering a twin with major depressive disorder to be phenotypically concordant with a co-twin diagnosed with bipolar disorder). Because of such differences in approach across studies it is usually prudent to view such investigations as providing a rough estimate of the genetic contribution to trait variability. Nevertheless, even such estimates are useful in deciding which traits are likely to be mappable. BASIC CONCEPTS OF GENE MAPPING Recombination and Linkage Once genetic epidemiological studies of particular phenotypes have suggested that these phenotypes are heritable, genetic mapping studies are conducted to identify the specific genetic variants that contribute to the risk of the disorder. All genetic mapping methods

aim to identify disease-associated variants based on their chromosomal position and the principle of genetic linkage. All cells contain two copies of each chromosome (called homologs), one inherited from the mother and one inherited from the father. During meiosis, the parental homologs cross over, or recombine, creating unique new chromosomes that are then passed on to

the progeny. Genes that are physically close to one another on a chromosome are genetically linked, and those that are farther apart or are on different chromosomes are genetically unlinked. Genes that are unlinked will recombine at random (i.e., there is a 50 percent chance of recombination with each meiosis). Genetic loci that are linked will recombine less frequently than expected by random segregation, with the degree of recombination proportional to the physical distance between them. The principle of linkage underlies the use of genetic markers, segments of DNA of known chromosomal location that contain variations or polymorphisms (described in more detail later). Strategies to map disease genes are based on identifying genetic marker alleles that are shared—to a greater extent than expected by chance—by affected individuals. It is presumed that such sharing reflects linkage between a disease locus and a marker locus, that is, the alleles at both loci are inherited "identical by descent" (IBD), from a common ancestor, and, furthermore, that this linkage pinpoints the chromosomal site of the disease locus. The evidence for linkage between two loci depends on the recombination frequency between them. Recombination frequency is measured by the recombination fraction (Θ) and is equal to the genetic distance between the two loci (1 percent recombination equals 1 centimorgan [cM] in genetic distance and, on average, covers a physical distance of about 1 megabase [mB] of DNA). A recombination fraction of 0.5 or 50 percent indicates that two loci are not linked but rather that they are segregating independently. A LOD (logarithm of the odds) score is calculated to determine the likelihood that two loci are linked at any particular genetic distance. The LOD score is calculated by dividing the likelihood of acquiring the data if the loci are linked at a given recombination fraction by the likelihood of acquiring the data if the loci are unlinked (Θ = 0.5). This step gives an odds ratio, and the log (base 10) of this odds ratio is the LOD score. A LOD score can be obtained for various values of the recombination fraction, from Θ = 0 (completely linked) to Θ = 0.5 (unlinked). The value of Θ that gives the largest LOD score is considered to be the best estimate of the recombination fraction between the disease locus and the marker locus. This recombination fraction can then be converted into a genetic map distance between the two loci. Linkage Disequilibrium Linkage disequilibrium (LD) is a phenomenon that is used to evaluate the genetic distance between loci in populations rather than in families. When alleles at two loci occur together in the population more often than would be expected given the allele frequencies at the two loci, those alleles are said to be in LD. When strong LD is observed between two loci it usually indicates that the two loci are sited in very close physical proximity to one another on a given chromosome, and is useful in mapping disease susceptibility loci because one locus can be used to predict the presence of another locus. This predictability is important because current gene-mapping strategies are able to sample only a subset of the estimated 10 million common human

polymorphisms. Because of the existence of LD, one can use data from a subset of genotyped polymorphisms to infer genotypes at nearby loci. Clusters of alleles that are in LD and inherited as a single unit are termed haplotypes. Thus LD mapping "consolidates" genomic information by identifying haplotypes in populations that can then be used to infer IBD sharing among unrelated individuals. There are several methods to measure the extent of LD. One of the most commonly used measures of LD is r2, a measure of the difference between observed and expected haplotype probabilities. Unlike D′, another widely used measure of LD, r2 values do not depend on the allele frequencies of the loci being assessed. A large r2 value indicates that the observed frequency of association between two alleles is greater than that expected by chance; that is, the alleles are in LD. LD studies have traditionally been used to complement traditional pedigree analyses, for example, to hone in on a locus that has been mapped by linkage analysis. However, LD-based

association analysis has become the method of choice for whole genome screens, particularly for diseases where traditional linkage studies have been unsuccessful. These studies have one great advantage over a traditional family analysis: because affected individuals are chosen from an entire population rather than from one or a few pedigrees, the number of potential subjects is limited only by the size of the population and the frequency of the disease. Maximizing the potential number of affected individuals that can be included in the analysis is extremely important for disorders where genetic heterogeneity or incomplete penetrance is likely to be a factor. Genetic Markers Mapping studies, regardless of their type, depend on the availability of genetic markers. The most widely used markers are microsatellite markers (also called simple tandem repeats [STRs], or simple sequence length polymorphisms [SSLPs]) and single nucleotide polymorphisms (SNPs). SSLPs are stretches of variable numbers of repeated nucleotides two to four base pairs in length. These markers are highly polymorphic, as the number of repeat units at any given STR locus varies substantially between individuals. SNPs, as the name implies, are single base pair changes at a specific nucleotide; they are the most common form of sequence variation in the genome. SNPs are widely used for genetic mapping studies because they are distributed so widely across the genome and because they can be assessed in a high-throughput, automated fashion. Other forms of genetic variation that have been investigated for use as genetic markers include small insertion or deletion polymorphisms, termed indels, that generally range between 1 and 30 base pairs and copy number variations (CNVs), which can refer to either deletions or duplications. Recent genomewide surveys have revealed that CNVs are common and can range in length from several base pairs to several million base pairs. CNVs may contribute to chromosomal recombination and rearrangements, thereby playing an important role in generating genetic diversity, and also, as many of these variants are sizable, it is hypothesized that they may significantly influence the expression of genes that encompass or are adjacent to the variant. MAPPING STRATEGIES The genetic variants that contribute to disease susceptibility can be roughly categorized

into those that are highly penetrant and those that are of low penetrance. Highpenetrance variants by definition have a large effect on phenotype, and therefore identifying these variants usually provides fundamental insights into pathobiology. Because individuals carrying high-penetrance variants have a high probability of expressing a disease phenotype, such variants tend to be rare and to segregate in families and are generally most powerfully mapped using pedigree-based approaches (see Fig. 1.7-1). In contrast, low-penetrance variants have a relatively weak effect on phenotype, and therefore identification of individual low-penetrance variants may, at least initially, provide relatively little new biological knowledge. However, because of their small effects, such variants are typically common in the population, and therefore identifying them may add to our understanding of disease risk in the population as a whole. Because we do not expect these variants to segregate strongly with the disease phenotype in pedigrees, efforts to identify them focus on population samples. Pedigree Analysis A pedigree analysis, which is conducted in multigenerational families, consists of scanning the genome or a portion of the genome with a series of markers in one or more affected pedigrees, calculating a LOD score at each marker position, and identifying the chromosomal regions that show a significant deviation from what would be expected under independent assortment. The primary goal of pedigree analysis is to determine if two or more genetic loci (i.e., a genetic marker of known location and the unknown disease loci) are cosegregating within a pedigree. Following the successful application of pedigree analysis to map Mendelian disorders such as Huntington’s disease, many investigators adopted this

strategy for mapping psychiatric disease genes with, at best, mixed success. In the late 1980s and mid-1990s, several pedigree-based studies reported the mapping of susceptibility loci for Alzheimer's disease, bipolar disorder, and schizophrenia. Although the linkage findings for three Alzheimer's disease loci were relatively quickly replicated, the findings reported for bipolar disorder and schizophrenia were ultimately determined to have been false positives. A number of different explanations have been proposed for the failure of pedigree-based approaches to map psychiatric loci; however, most investigators now recognize that these studies were generally drastically underpowered considering the apparent etiological complexity of psychiatric disorders. Pedigree analysis in psychiatry has increasingly turned toward an application that is more appropriately powered, namely, the mapping of quantitative trait loci (QTLs). QTLs are defined as genetic loci that contribute to the variation in continuously varying traits (as opposed to categorical traits such as disease diagnoses). QTLs are typically loci of small effect that only contribute to a portion of the observed variance of a trait in the population. It is now generally accepted that, using analytical methods developed in the late 1990s, it may be possible to use pedigree studies to map a wide range of quantitative traits that are relevant for understanding psychiatric disorders. Several such studies are now being undertaken, typically with multiple phenotypes being assessed in

each individual in the pedigree. Sib Pair Analysis Affected sib pair (ASP) analysis became widely used during the 1990s for the genetic mapping of complex traits, including many psychiatric disorders. Sib pair analysis examines the frequency with which sibling pairs concordant for a trait share a particular region of the genome compared with the frequency that is expected under random segregation. Sib pair analysis is based on the fact that siblings share approximately 50 percent of their genomes IBD. Therefore, if a set of unrelated sib pairs affected with a given trait shares a particular area of the genome at a frequency significantly greater than 50 percent (the proportion of sharing expected under conditions of random segregation), then that area of the genome is likely to be linked to the trait in question. In this method, siblings are genotyped, and population frequencies and parental genotypes are used to estimate the proportion of genes shared IBD at each site for each sib pair. The linkage analysis then compares those pairs concordant and discordant for each locus. Like pedigree studies, ASP studies have more power to locate genes of large effect than genes of small effect. This limitation can be partially addressed by a two-tiered design that incorporates additional markers or family members after an initial linkage study in affected siblings or by increased sample size. It generally requires less effort to identify and assess even large sets of affected sibs than to identify and assess all members of extended pedigrees, particularly when investigators can take advantage of data repositories that include samples and phenotype data from sib pairs ascertained from multiple sites. For example, the U.S. National Institute of Mental Health (NIMH) maintains such repositories for sizable collections of sib pairs affected with schizophrenia, bipolar disorder, autism, and Alzheimer's disease. An additional benefit of the ASP design is that it allows for the incorporation of epidemiological information, permitting the simultaneous examination of environmental and gene–environment interactions. Association Studies In the last few years, there has been increasing acceptance of the notion that association studies are more powerful than linkage approaches for mapping the loci of relatively small effect that are thought to underlie much of the risk for complex disorders. Whereas linkage studies attempt to find cosegregation of a genetic marker and a disease locus within a family or families, association studies examine whether a particular allele occurs more frequently than expected in affected individuals within a population. As noted previously in this chapter, mapping of genes using association studies is based on the idea that certain alleles at markers closely

surrounding a disease gene will be in LD with the gene; that is, these alleles will be carried in affected individuals more often than expected by random segregation, because they are inherited IBD. There are two common approaches to association studies (see Fig. 1.7-1), case–control designs and family-based designs, which typically investigate trios (mother, father, and an affected offspring). In a case–control study, allele frequencies are compared between a group of unrelated affected individuals and a matched control sample. This design is

generally more powerful than a family-based design, because large samples of cases and controls are easier to collect than trios and are less expensive, since they require the genotyping of fewer individuals. Case–control samples may be the only practical design for traits with a late age of onset (such as Alzheimer's disease) for which parents of affected individuals are typically unavailable. The main drawback of the case–control approach is the potential problem of population stratification; if the cases and controls are not carefully matched demographically, then they may display substantial differences in allele frequency that reflect population differences rather than associations with the disease. Family-based association studies are designed to ameliorate the problem of population stratification. In this design, the nontransmitted chromosomes (the copy of each chromosome that is not passed from parent to child) are used as control chromosomes, and differences between allele frequencies in the transmitted and nontransmitted chromosomes are examined, eliminating the problem of stratification, as the comparison group is by definition genetically similar to the case group. Although more robust to population stratification than a case–control study, family-based studies are only about two-thirds as powerful using the same number of affected individuals, as noted previously. Until recently, it was not practical to conduct association studies on a genomewide basis, as relatively few SNPs were available. Therefore, association studies focused on testing one or a few markers in candidate genes chosen on the basis of their hypothesized function in relation to a given disease. Recently, however, as a result of international efforts that have identified millions of SNPs distributed relatively evenly across the genome and that have developed technology for genotyping them relatively inexpensively, genomewide association (GWA) studies are now a reality. Such studies hold much promise for the identification of common variants contributing to common diseases. While few GWA studies of psychiatric disorders have been completed, such studies have already reported remarkable findings for complex traits such as rheumatoid arthritis, inflammatory bowel disease, and type 2 diabetes. The successful studies of these diseases have made use of very large samples (in some cases up to several thousand cases and controls), providing further support for the hypothesis that underpowered study designs bear much of the responsibility for the disappointing results to date of psychiatric genetic investigations. Statistical Considerations Scientists in other biomedical research fields are often surprised by the apparently high level of statistical evidence that geneticists require to consider a linkage or association result to be significant. Most simply, this requirement can be thought of in terms of the very low expectation that any two loci selected from the genome are either linked or associated with one another. The likelihood that any two given loci are linked (i.e., the prior probability of linkage) is expected to be approximately 1:50, based on the genetic length of the genome. To compensate for this low prior probability of linkage and bring the posterior (or overall) probability of linkage to about 1:20, which corresponds to the commonly accepted significance level of $P = .05$, a conditional probability of 1,000:1 odds in favor of linkage is required, corresponding to the traditionally accepted LOD score threshold of 3. This generally provides an acceptable false-positive rate (Fig. 1.72), but some false-positive findings have exceeded even this threshold.

FIGURE 1.7-2 Number of false positives expected in a whole genome scan for a given threshold of logarithm of odds (LOD) score. Solid line represents the expectation for a perfect genetic map. Symbols represent the results for 100 sib pairs using genetic maps with markers spaced every .1 cM (circles), every 1 cM (squares), and every 10 cM (triangles). The dotted line indicates the 5 percent genomewide significance level. (Courtesy of Dr. Eric Lander). Geneticists generally assume that the expectation that any two loci in the genome are associated with one another is even lower than that of their being in linkage, and typically a P value of less than about 10−7 is considered to indicate “genomewide significance.” This standard essentially discounts the prior probability that some investigators assign to variants in candidate genes chosen on the basis of their hypothesized functional relevance to a given disorder or trait. GWA studies are now replicating associations with very low P values for a wide range of complex traits, whereas most candidate gene associations (which usually report as significant much higher P values) remain unreplicated. It is therefore increasingly apparent that genomewide levels of significance are appropriately applied to all initial association studies for a given trait. DEFINING PHENOTYPES FOR MAPPING STUDIES The generally disappointing results of psychiatric genetic mapping studies have focused increasing attention on the problem of defining and assessing phenotypes for such studies. Most psychiatric mapping studies to date have relied on categorical disease diagnoses, as exemplified by the Diagnostic and Statistical Manual (DSM-5) classification scheme. Criticisms of this approach rest on two arguments. First, diagnosis of psychiatric disease depends on subjective clinical evaluation, a fact that underscores the difficulty in ascertaining individuals who can be considered definitely affected with a given disease.

Second, even when a psychiatric diagnosis can be established unambiguously, the menubased system used for psychiatric classification provides the possibility that any two individuals affected with a given disorder may display largely nonoverlapping sets of symptoms, likely reflecting distinct etiologies. Concern that the diagnosis-based approach to phenotyping may represent one of the chief obstacles to the genetic mapping of psychiatric phenotypes has generated considerable interest in mapping heritable traits known to demonstrate continuous variation in the population. Continuous measures that are hypothesized to be related to psychiatric disorders include biochemical measures (e.g., serum or CSF levels of neurotransmitter metabolites or hormones), cognitive measures, personality assessments, structural or functional brain images, biophysical markers such as responses to evoked potentials, or molecular assays such as gene expression profiles. Key features of categorical and continuous phenotyping strategies are shown in Figure 1.7-3, and each is discussed in more detail below. FIGURE 1.7-3 Two alternate schemes for conceptualizing psychiatric phenotypes. A. Categorical Traits as conceptualized by the Diagnostic and Statistical Manual (DSM-5) represent a “menubased” approach to psychiatric disorders. Individuals are assessed for a checklist of signs and symptoms that are then used to categorize the individual as “affected” according to a specific diagnosis. Not all symptoms are present in samples of individuals who carry a particular DSM diagnosis, and many of these symptoms occur across diagnostic boundaries, as illustrated in this Venn diagram. DSM phenotypes therefore probably represent etiologically heterogeneous categories, and this fact may help to explain the

limited progress thus far of genetic mapping investigations focused on these phenotypes. B. Alternatively, in the Continuous Traits model, “affectedness” can be conceptualized in terms of an expectation that an individual will demonstrate extreme values on a set of continuous measures

that correlate with psychopathology and thus are hypothesized to underlie the disorder (as illustrated by examples of six different types of measures shown in the hexagon). Such measures may also be associated with particular components of categorical phenotypes, such as those depicted in the Venn diagram in Figure 19-3A. The justification for using continuous measures as the phenotypes for genetic mapping studies is that they are considered etiologically simpler and more reliably assessed compared to categorical phenotypes. In addition, mapping such traits combines information from all members of the study population (affected and unaffected individuals alike), which adds considerably to power. (From Sadock BJ, Sadock VA, Ruiz P. Kaplan & Sadock's Comprehensive Textbook of Psychiatry. 9th ed. Philadelphia: Lippincott Williams & Wilkins; 2009:325.) Categorical Phenotypes The most commonly used categorical phenotypes in psychiatry are DSM diagnoses. Some studies focus on a single DSM diagnosis, whereas other studies include individuals with a range of different diagnoses. The latter approach is typically used for disorders that are hypothesized to represent a single disease spectrum, such as mood disorders. Using the categorical approach, it is important to be able to classify subjects as unambiguously as possible. Several strategies are used to accomplish this goal. The first strategy involves deciding on the appropriate diagnostic criteria for the study in question and deciding how these criteria will be applied to individuals in the study. One way of standardizing the procedures used to identify and assess potential study subjects is to use only experienced clinicians in the diagnostic process and to train them in the administration of the instruments and the diagnostic criteria to be employed. In addition, a "best estimate" procedure and/or a consensus diagnosis is frequently used. The best estimate process involves making use of every piece of available information, including medical records, interviews, and videotapes, to arrive at a diagnosis. For a consensus diagnosis, two or more diagnosticians independently review the material and make a diagnosis for each individual. The diagnoses are then compared, and individuals for whom an agreement in diagnosis cannot be reached are not entered as "affected" into the study. A well-designed study makes use of all available information about the genetic epidemiology of the disorder to choose a sample of affected individuals to study. It is often the case that a subset of families carries the disorder in what appears to be a simple Mendelian pattern, whereas the inheritance pattern is less clear for other families or groups. In a disorder where there are likely to be multiple genes contributing to the phenotype, it makes sense to begin with a study sample where there may be major loci. Redefining the disease phenotype can often simplify the mapping process by identifying such groups or families. For example, in the search for a genetic defect for

Alzheimer's disease, the process was advanced enormously by limiting the study population to those individuals who had early age of onset (before age 65); the early onset trait segregated in an autosomal dominant fashion. Other ways of redefining the phenotype include focusing on factors such as ethnic background, age of onset, treatment response, symptom severity, or the presence of comorbid disorders. Narrowing the phenotype using the approaches discussed earlier may increase the chances of finding a genetic defect in complex diseases, but it can also greatly reduce the power of the study by limiting the number of available affected individuals. For this reason, it has been argued that for some disorders broadening the phenotype is an appropriate strategy. The suggestion is that for some complex diseases the phenotype of interest may represent the extreme end of a spectrum and that to have enough power to map genes other phenotypes within the spectrum must also be included. For example, mapping studies of bipolar disorder might include as affected individuals with major depressive disorder as well as those individuals diagnosed with bipolar disorder. Although the two approaches of narrowing the disease phenotype and broadening

the disease phenotype may seem to be mutually exclusive, many groups studying complex disorders have incorporated both approaches into their study designs. One way to do this is to create stratified diagnostic categories, ranging from a narrow diagnostic category to a broad diagnostic category, and test for genetic linkage under each of these schemas. Some investigators argue that for complex diseases that are part of a spectrum, this strategy decreases the rate of false negatives, that is, of missing an existing linkage because of misspecification. Others argue that using several models and picking the one that gives the highest scores greatly increases the rates of false positives, that is, of identifying an area of linkage where none exists. One problem that clearly exists with the use of multiple diagnostic categories is that as more models are used (and therefore more statistical tests are performed), increasingly stringent levels of evidence are required to consider a result significant. While categorical phenotypes remain the mainstay of psychiatric genetic studies, the limitations of DSM nosology as the basis of phenotyping for genetic studies are becoming clear. Genetic investigations are focusing increasingly on traits that may be components of one or more DSM diagnostic categories. For example, there is growing evidence that genetic susceptibility to psychosis, broadly defined, contributes to both severe bipolar disorder and schizophrenia, and a number of investigative approaches are being employed to attempt to identify genes that underlie such susceptibility and even to explore possible etiological relationships between psychiatric and nonpsychiatric disorders. For example, bioinformatics models have been employed to investigate medical records databases and have uncovered extensive pairwise correlations among a diverse list of psychiatric disorders, neurological disorders, autoimmune disorders, and infectious diseases. Eventually, the results of such model-fitting experiments may provide a framework to design more powerful linkage and association studies that can search for alleles that contribute to susceptibility to multiple disorders. Continuous Phenotypes Because of the difficulties experienced in genetic mapping of categorical diagnoses, neurobehavioral geneticists are increasingly focused on investigating quantitative traits that are hypothesized to underlie a particular psychiatric diagnosis and that may be simpler to genetically map. The rationale for efforts to map such alternative phenotypes, or endophenotypes, is that the genes identified through such efforts may provide clues regarding the biological pathways that are relevant to understanding a particular disorder. Several features characterize useful endophenotypes. First, they should be state-independent; that is, they should not fluctuate as a function of the

disease course or medication treatment and should show adequate test–retest stability. Second, they should be heritable; that is, there should be evidence that genetic factors are responsible for a substantial proportion of the variability of the trait within the population. Third, the endophenotype should be correlated with the disease under investigation; that is, different values of the trait measure are observed in patients compared to unrelated control subjects. Measures of brain structure and function provide most of the traits now under investigation as endophenotypes for psychiatric disorders. For example, several features of brain morphometry (as assessed by magnetic resonance imaging [MRI]) are highly heritable (in the range of 60 to 95 percent) including total brain volume, cerebellar volume, gray and white matter density, amygdala and hippocampal volume, and regional cortical volume. Several studies show that brain structural features that are correlated in clinical samples with disorders such as schizophrenia or bipolar disorder are also abnormal in relatives of affected individuals. Physiological measures of brain activity that have been employed as candidate endophenotypes for psychiatric disorders include electroencephalography (EEG) patterns. Several “pencil and paper” assessments have been

employed to measure endophenotypes relating to neurocognitive function and temperament. Animal Models In contrast to categorical phenotypes, endophenotypes can be more straightforwardly related to phenotypes that can be assessed in animal models. Studies of genetic variations that affect circadian rhythms provide a good example. Variations in circadian rhythms have long been recognized as important features of mood disorders, and quantitative assessments of activity patterns have been proposed as endophenotypes for such disorders. Numerous studies in animal models have demonstrated that genetically controlled biological clocks determine circadian activity and that variations in clock genes are associated with variations in such activity from bacteria to humans. Genetic mapping efforts in fruit flies starting in the early 1970s resulted in the identification of at least seven “clock genes,” beginning with period. Subsequent studies showed that the homologs of several of these genes play essential roles in regulating mammalian circadian rhythms. Genetic mapping studies in mice also have identified previously unknown circadian rhythm genes, beginning with the discovery and characterization in the early 1990s of clock. These genetic discoveries have not only explicated the cellular networks and neurophysiological circuits responsible for the control of mammalian circadian rhythms but have also generated animal models that may shed light on the pathobiology of psychiatric syndromes such as bipolar disorder. For example, mice carrying a targeted mutation in clock demonstrate abnormal activity patterns, such as hyperactivity and decreased sleep, which are apparently modified by administration of lithium. PROGRESS IN THE GENETICS OF SPECIFIC DISORDERS Taken as a whole, the progress in identifying susceptibility genes for psychiatric disorders has been disappointing compared to that observed for nonpsychiatric disorders. Alzheimer’s disease represents the most successful application of gene-

mapping strategies to complex neurobehavioral disorders, and the section on this disease provides an example of how genetic linkage studies add to understanding of the pathogenesis of a complex trait. An overview section on autism describes genetic investigations of syndromes that have features of autism but have relatively simple inheritance patterns and discusses how these studies have provided starting points for investigations of more complex autism spectrum disorders. Finally, the frustrating search for unequivocal gene findings for bipolar disorder and schizophrenia is used to illustrate the challenges that are motivating new approaches in the field of neurobehavioral genetics. ALZHEIMER’S DISEASE Alzheimer’s disease provides an excellent example of the power of genetics to elucidate the complex biology of a neuropsychiatric disorder. Alzheimer’s disease is a well-defined form of dementia characterized by progressive impairment of memory and intellectual functioning. The clinical signs and symptoms, although characteristic, are not limited to Alzheimer’s disease; they are also found in several other types of dementia. For this reason, the diagnosis of Alzheimer’s disease can only be confirmed histopathologically at autopsy. The presence of senile plaques (made up of a core of β-amyloid fibrils surrounded by dystrophic neurites), tau-rich neurofibrillary tangles, and congophilic angiopathy in the brain parenchyma and associated blood vessels are pathognomonic for Alzheimer’s disease. A variable age of onset has been noted for Alzheimer’s disease, ranging from as early as age 35 to as late as age 95. The concordance rate for Alzheimer’s disease in MZ twin pairs is about 50 percent, indicating a moderately strong genetic contribution to disease risk. It is now evident from a wide range of genetic studies that Alzheimer’s disease can be divided into two broad categories: familial forms, which account for a tiny minority of Alzheimer’s disease cases and are characterized by early onset and autosomal dominant inheritance with high penetrance; and sporadic forms, in which the genetic contribution is hypothesized to be similar to that characterizing other common

neuropsychiatric diseases. The search for the genetic basis of familial Alzheimer's disease began with traditional linkage studies. First, an investigation of a candidate locus on chromosome 21 in humans identified mutations in the amyloid precursor protein (APP) gene in a small number of families in which significant linkage had previously been observed to markers from this region. Transgenic mice with different APP mutations were created and have been shown to produce β-amyloid deposits and senile plaques as well as to show synapse loss, astrocytosis, and microgliosis, all part of the pathology of Alzheimer's disease. Mutations in the genes that encode β-APP all lead to an increase in the extracellular concentration of longer fragments of β-amyloid (Aβ42). Most of the strains of transgenic mice with mutations in APP exhibit increased rates of behavioral changes and impairment in several memory tasks, indicating dysfunction in object-recognition memory and working memory among others. These findings represent striking evidence that mutations in the β-amyloid gene are indeed responsible for at least some of the histopathological elements of Alzheimer's disease. Even as the preceding findings were being reported, it was clear that mutations in the β-amyloid gene could not

completely explain the etiology and pathology of Alzheimer's disease, not least because it was shown that linkage to chromosome 21 was excluded in most early onset Alzheimer's disease families. In addition, no neurofibrillary tangles are observed in most of the different β-amyloid transgenic mice. The subsequent search for the genetic underpinnings of Alzheimer's disease using genomewide linkage analysis of early onset Alzheimer's disease families resulted in the identification of two additional Alzheimer's disease susceptibility genes: presenilin-1 (PS-1) on chromosome 14q24.3 and presenilin-2 (PS-2) on chromosome 1q. PS-1 and PS-2 are integral transmembrane proteins with at least seven transmembrane domains. Although their function has not yet been completely elucidated, they are clearly involved in the pathogenesis of Alzheimer's disease. Inactivation of presenilins in mice leads to neurodegeneration and behavioral manifestations of memory loss. Biochemical and cellular studies have implicated presenilins in several important pathways, including apoptosis (programmed cell death) and protein processing in the endoplasmic reticulum. These findings emphasize one of the strengths of using family-based linkage analysis. Pedigree-based studies are especially suited to identify highly penetrant disease genes that serve important roles in important biological processes. Although mutations in APP and presenilin are rare, research into the biology of the expressed proteins has provided key insights into the pathophysiology of dementia. Because these highly penetrant mutations elucidate important biological functions, they also provide a firm ground to design therapeutic interventions. For example, amyloid-β "vaccines" designed to induce an immunogenic response to pathogenic amyloid are now in advanced clinical trials. Unlike the current psychopharmacological treatments for Alzheimer's disease that nonspecifically target cholinergic and glutaminergic neuronal systems, the amyloid-β vaccines specifically treat the causes of Alzheimer's disease by generating an immune response that may actually reverse the deposition of senile plaques. Sporadic and Late-Onset Alzheimer's Disease Mutations in APP, PS-1, or PS-2 are present in a majority of familial cases of early-onset Alzheimer's disease but do not account for sporadic or familial late-onset Alzheimer's disease. For this reason, investigators turned to other approaches to search for evidence of linkage in a large number of small families with late-onset Alzheimer's disease. In 1991, the results of a nonparametric linkage study using 36 markers in late-onset Alzheimer's disease families provided evidence for a susceptibility gene on the long arm of chromosome 19. In 1993, association studies revealed that the e4 allele of the apolipoprotein E gene was strongly associated with late-onset Alzheimer's disease and that this association almost certainly was

responsible for the previously observed linkage signal on chromosome 19. There are three known alleles of this gene— e2, e3, and e4. In most populations, the e3 allele is the most common. However, in familial late-onset Alzheimer's disease the incidence of e4 is approximately 50 percent, and in sporadic late-onset Alzheimer's disease it is 40 percent, compared with about 16 percent in normal controls. Epidemiological studies suggest that between 30 and 60 percent of late-onset Alzheimer's disease cases have at least one apoE-e4 allele. The e4 genotype appears to be a more important risk factor for Alzheimer's disease in populations of European and Asian origin when compared with populations of African origin. Overall, the association of apoE-e4 with Alzheimer's disease remains probably the strongest association yet identified for a common human disease. The establishment of apoE-e4 as a susceptibility allele for late-onset Alzheimer's disease has led to the search for additional alleles that might interact with apoE-e4 to

modify disease risk. In 2007, investigators used genomewide association strategies (in histologically confirmed cases and controls) to identify GAB2 (GRB-associated binding protein 2) as an additional risk allele in apoE-e4 carriers (but not in Alzheimer's disease patients who were not e4 carriers). Initial studies suggest that carriers of both apoE-e4 and GAB2 risk alleles have an almost 25-fold greater risk for Alzheimer's disease than individuals who do not carry either risk allele. Larger-scale GWA studies of Alzheimer's disease are in progress and will likely yield further associations; however, it is unlikely that any will have as strong an effect as apoE. AUTISM Autism is a severe neurodevelopmental disorder that is characterized by three primary features: impaired language and communication; abnormal or impaired social interaction; and restricted, repetitive, and stereotyped patterns of behavior. Understanding of the etiology of autism has proceeded slowly, but there is now convincing evidence that alterations in specific cellular and molecular neurodevelopmental pathways are important in its etiology. In comparison with other neuropsychiatric disorders, there is particularly strong evidence for a genetic contribution to the risk of autism and autism spectrum disorders (ASDs). The sibling recurrence risk for autism and/or ASD is between 2 and 6 percent. Given a population prevalence of about 1 in 2,000 (.04 percent), this means that the siblings of autistic individuals are approximately 50 to 100 times more likely to develop autism than a person in the general population. Twin studies of autism show an extraordinarily high heritability (as demonstrated by MZ twin concordance of 80 to 92 percent) but also demonstrate the genetic complexity of these disorders, with the DZ twin concordance rate of 1 to 10 percent suggesting a highly multigenic mode of inheritance. Increasing interest is now focused on the possibility that individuals affected with autism may display larger numbers of large-scale chromosomal aberrations (5 to 10 percent in some studies) than unaffected individuals. In addition to such gross abnormalities, several recent studies have suggested that autism is associated with an unusually high prevalence of submicroscopic CNVs. For example, in 2007, the Autism Genome Project Consortium applied microarray strategies to almost 8,000 individuals from about 1,500 families, each with at least two affected family members, and found that about 10 percent of the ASD families carried CNVs, with an average size of more than 3 million base pairs, mostly consisting of duplications rather than deletions. Although the design of this study did not permit assessment of whether the frequency of CNVs is greater in patients with autism than that in controls, another study found a de novo CNV incidence of 10 percent in sporadic (no family history) cases of autism compared to an incidence of 1 percent in controls. These results, while exciting, are still considered preliminary. Even before the demonstration of high rates of de novo mutations in autism, epidemiological studies had strongly suggested that the genetic basis of this disorder is likely complex. For example, although the risk of autism in firstdegree relatives of autistic probands

is high, there is a substantial falloff for second-

degree and third-degree relatives of such probands, suggesting that multiple genetic variants must interact to increase susceptibility to this syndrome. Segregation analyses of autism also support the hypothesis that it is a heterogeneous disorder that reflects the actions of multiple genetic variants of small effect. A latent class analysis performed to study possible modes of transmission suggested an epistatic model with up to about 10 interacting loci, whereas other studies have estimated that as many as 15 such loci may be involved. Genetic studies of autism have included whole genome screens, candidate gene studies, chromosome rearrangement studies, mutation analyses, and, most recently, comparative genomic hybridization studies. Taken together and recognizing that most findings still await adequate replication, these studies have contributed to an emerging picture of autism susceptibility that includes genes involved in three major systems: those involving synapse formation and maintenance, those involving cell migration, and those involving the excitatory/inhibitory neurotransmitter networks. Figure 1.7-4 shows a schematic of the currently known potential candidate genes for autism and their molecular relationships with one another. FIGURE 1.7-4 Schematic of the cell biology of proteins expressed from genes identified through mapping studies of autism spectrum disorders. The function of each gene product falls into three broad functional categories. Proteins involved in synapse formation and

maintenance include FMR1, TSC1, TSC2, MeCP2, NLGN 3 and 4, and SHANK3. Another set of proteins is involved in neuronal migration and cell fate including REELIN, WNT2, LAMB1, and NrCAM. Proteins involved in neurotransmitter systems are also altered in some individuals with autism and include 5-HTT (serotonin transporter encoded by SLC6A4), GABAR, and the NMDA subunit encoded by GRIN2A. See text for details. (From Persico AM, Bourgeron T. Searching for ways out of the autism maze: Genetic, epigenetic and environmental clues. Trends Neurosci. 2006;29:349, with permission.) Synapse Formation and Maintenance Perhaps the biggest breakthroughs in identifying susceptibility genes for autism have come from studies of disorders that display clinical features associated with autism or ASDs but with simpler inheritance patterns, including fragile X syndrome, tuberous sclerosis, and Rett syndrome. In general, the genetic defects associated with these disorders affect synapse formation and maintenance. Fragile X, which accounts for 3 to 4 percent of autism cases, is caused by an unstable trinucleotide repeat in the 5′ region of the fragile X mental retardation 1 (FMR1) gene at Xq27.3. This repeat expands as it is transmitted to succeeding generations, resulting in abnormal methylation and inhibition of expression of FMR1. FMR1 produces a ribonucleic acid (RNA)-binding protein that acts as a chaperone for the transport of RNA from the nucleus to the cytoplasm and is involved in messenger RNA (mRNA) translation at the synapse. Abnormalities in dendritic spine density (increased over normal) and anatomy (longer and thinner than normal) have been reported in individuals with fragile X as well as in mouse models of this disorder. Tuberous sclerosis, which accounts for perhaps 2 to 10 percent of autism cases (the rate of tuberous sclerosis is higher among autistic individuals with seizure disorders), results from mutations in one of two tumor suppressor genes, TSC1 on 9q34, and TSC2 on 16p13, both of which are involved in guanosine triphosphatase (GTPase) inactivation. Loss of a single copy of TSC1 in mice has been shown to disrupt cytoskeletal dynamics and dendritic spine structure. Although somewhat less well understood, the genetics of Rett syndrome, an X-linked pervasive developmental disorder (the first with a known genetic etiology) that occurs only in girls and is associated with normal early development followed by loss of skills—particularly social engagement and purposeful hand skills by age 4—also point to

abnormalities in synapse formation and maintenance in ASD and ASD-like disorders. Rett syndrome is caused by mutations in MeCP2, which makes a methylated-DNA-binding protein that regulates gene expression and chromatin structure. Although little is known about the exact role of MeCP2 in the development of Rett syndrome, the pattern of normal early development and later regression suggests that this gene is more likely to be involved in synapse maintenance and remodeling than in synapse development. Neuroligin (NLGN) 3 and 4 and SHANK3, additional genes that appear to play a role in synapse formation, may be affected by chromosomal rearrangements observed in some individuals affected with autism. The neuroligin genes, sited on the X chromosome, produce cell adhesion molecules that are located on postsynaptic glutamatergic neurons. When

mutated in rodents, these genes show defective trafficking and synapse induction. In nonmutated form, their expression induces the formation of normal, presynaptic terminals in axons. SHANK3 is a binding partner of the neuroligins and regulates the structural organization of dendritic spines. Mutations in SHANK3 have been identified in ASD-affected members of at least three families to date, and a comparative genomic hybridization study of autistic individuals, their family members, and controls recently identified a large deletion in chromosome 22q13, the region containing SHANK3, in at least one individual with autism. Cell Migration Of the regions highlighted by a genome screen in autism families, chromosome 7q has provided the most consistent evidence for linkage, albeit over a very broad region. Known chromosomal rearrangements in this region in individuals affected with autism add to its interest. The linkage region on chromosome 7q contains several genes that are strong candidates for autism, most notably RELN, which maps to chromosome 7q22. RELN codes for reelin, a signaling protein secreted by Cajal-Retzius cells located in the marginal zone of the developing brain. It plays an important role in neuronal migration as well as in the development of neural connections. Reeler mice, which have spontaneous deletions of RELN, have cytoarchitectonic alterations in their brains during development that are similar to those that have been described in autistic brains. The complete absence of RELN in humans leads to a more severe phenotype with lissencephaly and severe mental retardation but not autism. Individuals with autism show reduced levels of reelin mRNA and protein in brain and blood serum, suggesting that mutations leading to reduced expression of RELN rather than its absence may be important in ASD. Genetic association studies with RELN have been equivocal, suggesting that if RELN does contribute to the development of autism, then it may play such a role in a small subset of affected individuals. WNT2 (wingless-type MMTV integration site family member 2) is another gene identified as a potential candidate for autism based on linkage studies. WNT2 is located on 7q31 and is part of a family of genes that encode secreted signaling proteins implicated in several developmental processes, including the regulation of cell fate and patterning during embryogenesis. At least two families have been identified in which nonconservative coding sequence variants in WNT2 segregate with autism. LD between a SNP in the 3′ untranslated region of WNT2 and autism is also present in families with severe language abnormalities that accounted for most of the evidence for linkage on chromosome 7q in one of the original genome screens. Excitatory/Inhibitory Neurotransmitter Systems Although there is little current evidence that mutations in genes encoding neurotransmitter transporters and/or receptors are directly responsible for the development of autism, there is some evidence that such genes might act as modifiers or susceptibility factors for an autism spectrum phenotype. The evidence is perhaps strongest for the role of the γ-aminobutyric acid (GABA) receptors in the development and expression of autistic disorders. These receptors occur in a cluster on chromosome

15q11–13, and duplications of this region are the most common cytogenetic abnormalities seen in autism cases (up to 6 percent of cases). GABA is an important inhibitory neurotransmitter in the central nervous system and is responsible for controlling excitability in mature brains. Chromosome 15q11–13 is one of the most complex regions of the genome. It has a high rate of genomic instability, including frequent duplication and deletion events, and imprinting plays an important role in the expression of genes in this region. The 15q11–13 region is the critical region for Angelman and Prader-Willi syndromes, neurological disorders due to deletions or mutations in this region that occur on maternally and paternally inherited chromosomes, respectively. Despite the high rate of duplications of 15q11–13 among autistic individuals, genome screens have not shown strong support for linkage or association to this region. Candidate gene studies continue, however, in part because a rate of 6 percent of autistic individuals with duplications in this region is hard to ignore. BIPOLAR DISORDER The search for the genetic basis of bipolar affective disorder has been fraught with missteps and partial answers. The history of genetic mapping attempts for bipolar disorder illustrates not only the extreme complexity of psychiatric disorders but also the evolution of genetic approaches to such diseases. Bipolar disorder is an episodic illness characterized by recurrent periods of both mania and depression. Psychotic symptoms are often a part of the clinical picture, particularly in more severely affected individuals. Numerous genetic epidemiological investigations conducted over several decades have strongly supported a genetic contribution to risk for bipolar disorder. As with other psychiatric disorders, however, the definition of the bipolar disorder phenotype in these studies has varied substantially, and this in turn has resulted in a wide range in estimates of its heritability. For example, many early studies into the genetic basis of mood disorders did not distinguish between unipolar and bipolar mood disorders. Furthermore, the diagnostic methodology used in such early studies differs substantially from that employed in current-day genetic studies. For example, a Danish twin study that suggested a very high heritability for bipolar disorder and thereby had a heavy influence on the design of initial genetic mapping studies of mood disorders employed only unstructured diagnostic interviews by a single clinician rather than the structured assessments used in current studies, which have suggested somewhat lower heritabilities. Current estimates of concordance for bipolar disorder range between 65 and 100 percent in MZ twins and between 10 and 30 percent in DZ twins, indicating that the disorder is highly heritable (between about 60 and 80 percent). Several studies have shown that bipolar disorder is substantially more heritable than unipolar major depression, which has an estimated heritability between 30 and 40 percent. Early family studies suggested that bipolar disorder segregation patterns were compatible with single gene inheritance of a locus of major effect. However, although it is possible that some bipolar disorder pedigrees segregate such a locus,

mounting evidence indicates that if such pedigrees exist they must be quite rare. Furthermore, the fact that genetic linkage studies have failed to uncover such a locus with unequivocal evidence in any pedigrees argues against this possibility. The observed rapid decrease in recurrence risk for bipolar disorder from monozygotic co-twins to first-degree relatives is also not consistent with single gene inheritance models but rather suggests models of multiple interacting genes. Early Linkage Studies Tremendous excitement followed the first reports of linkage to bipolar disorder on chromosomes X and 11 in 1987. Investigators noted that in several families, bipolar disorder and other affective disorders appeared to be inherited in an X-linked fashion. Likewise, these disorders appeared to cosegregate in several Israeli families with color blindness and G6PD deficiency, which map to the X chromosome. Linkage studies in these pedigrees, using color blindness or G6PD

deficiency as marker loci, gave LOD scores between 4 and 9. Early studies of chromosome 11 were similar to those for chromosome X in that they reported significant linkage after testing only a few markers in a single region, in this case in an extended Old Order Amish pedigree heavily loaded for bipolar disorder. Not surprisingly, these findings generated a great deal of interest. Both studies showed high LOD scores and seemed to provide clear evidence for linkage. However, replication studies in other populations failed to produce positive results for either the X chromosome or chromosome 11, and evidence for linkage essentially disappeared in both chromosomal regions in the samples in which linkage was originally reported when the pedigrees were extended to include additional affected individuals and when additional markers were typed in the putative linkage regions. The most likely explanation in each case is that the original linkage results were false-positive findings and may have reflected overoptimistic interpretation of evidence that, in retrospect, was relatively scanty. Genomewide Screens The early linkage studies of bipolar disorder evaluated only a few markers because they were all that were available. With the construction of genetic linkage maps of the genome in the 1990s, linkage studies of most complex traits, including bipolar disorder, began to search genomewide. The advantage of genomewide mapping studies is that they do not require a priori knowledge of the biological underpinnings of a particular phenotype. Complete genome screens provide an opportunity to evaluate the evidence of linkage at all points in the genome without bias (see Color Plate 1.7-5). Although genomewide studies clearly had greater power to detect true linkage than studies focused on only a few markers in arbitrary locations or around a few candidate genes, these investigations have also generally had disappointing results. The challenge of achieving replicated significant linkage results for bipolar disorder and other complex traits is apparent when one reviews the many gene-mapping studies that have suggested —but not demonstrated unequivocally—bipolar disorder susceptibility loci on chromosome 18.

Chromosome 18 The first report of linkage came from a partial genome screen that examined 11 markers on chromosome 18 and identified suggestive linkage near the centromere. Because the inheritance patterns for bipolar disorder are unknown, the results were analyzed using both recessive and dominant models. Some of the markers were positive under a recessive model in some families, some were positive under a dominant model in other families, and some markers gave positive LOD scores in a subset of families under both models. Attempts to replicate this finding in other populations have been mixed. So far at least two groups have found no evidence for linkage to the pericentromeric region of chromosome 18 in their samples, although one other group has found evidence to support linkage to this region. Other studies have found suggestive evidence for linkage on chromosome 18, including a complete genome screen in two large Costa Rican pedigrees that gave evidence for linkage on chromosome 18q22–23 as well as in an area on 18p. The combined evidence of these several studies, although somewhat contradictory and confusing, points to at least two different susceptibility loci on chromosome 18: one on 18p and one on 18q. Improving Study Power The equivocal findings represented by the attempts to pinpoint susceptibility loci on chromosome 18 have led investigators to implement several new strategies to map bipolar disorder genes. One such strategy is meta-analysis, which involves combining data across multiple individual investigations to increase statistical power, and in some cases the combined analysis points to loci not originally found in the individual studies. Several meta-analytical techniques have been used to explore gene-mapping studies for bipolar disorder. The multiple scan probability (MSP) and genome scan meta-analysis (GSMA) methods require only linkage statistics and P-values from each study to examine combined data. MSP was used to

combine chromosomal regions with P-values less than .01 from 11 independent bipolar disorder studies and provided evidence for susceptibility loci on chromosomes 13q and 22q. Although the MSP and GSMA methods have the advantage of requiring only linkage significance data, they are not able to account for study-specific issues that will limit the extent to which multiple studies can be compared. Combining original genotype data from multiple studies can circumvent this problem. With this method, the largest meta-analysis to date combined 11 bipolar disorder genomewide linkage scans consisting of 5,179 individuals from 1,067 families. Access to the original genotype data allowed the construction of a standardized genetic map in which the markers of each respective study were mapped onto one common gender-averaged map. The results of this meta-analysis identified two susceptibility loci with genomewide significance on 6q and 8q. Another strategy that has been used to increase the power of gene-mapping studies is the formation of consortia that combine data across multiple clinical sites. A consortium combining data from the UK and Ireland led to support for linkage at 9p21 and 10p14–21. Likewise, combining data from Spanish, Romanian, and Bulgarian families provided

additional support for findings on chromosomes 4q31 and 6q24. Investigators can also increase power by standardizing marker sets and clinical evaluation protocols between independent studies to permit direct comparisons between such studies. This approach was used to identify a bipolar disorder susceptibility locus on chromosome 5q31–33. The region showed suggestive nonparametric linkage results in pedigrees from the Central Valley of Costa Rica. With identical genetic markers and diagnostic criteria, the same region was highlighted in an independent analysis of a set of Colombian families who have a genetic background similar to that of the Costa Rican families. A follow-up study using additional markers in an expanded set of Colombian and Costa Rican families confirmed genomewide significant evidence to a candidate region of 10 cM in 5q31–33. This finding is especially interesting given that the linkage peak in the bipolar studies overlaps with linkage regions for schizophrenia and psychosis, identified in a previous study of 40 families from the Portuguese Islands. These results contribute to a growing opinion that there may be substantial genetic overlap between different DSM disorders. SCHIZOPHRENIA As with bipolar disorder, investigations of the genetic basis of schizophrenia exemplify the frustrations still characteristic of psychiatric genetics, and the field still struggles to interpret the significance of initially promising linkage and association results that began to emerge over a decade ago. Unlike with bipolar disorder, however, candidate genes have emerged from each of the regions highlighted from these studies. Thus, although none of these findings have been validated unequivocally, they have spawned a diverse range of basic and clinical investigations aiming to elucidate their functional significance, for example, using mouse gene targeting and functional MRI. Here we discuss some of the more extensively investigated loci for purposes of illustration; it could be argued that roughly equivalent evidence supports schizophrenia candidate loci that we do not discuss in detail, for example, AKT1 on chromosome 14 or COMT on chromosome 22. Chromosome 6p24–22 was among the first regions to be implicated by a complete genome screen for schizophrenia, in this case from a study of Irish families heavily loaded for schizophrenia. The linkage results were strongest under a broad diagnostic definition that included schizophrenia spectrum disorders, such as schizotypal personality disorder. Six additional linkage studies have shown positive results over approximately the same region, but at least three studies have found no linkage to the region. Fine-scale mapping of this region using association analysis in the original Irish kindreds led to the proposal of Dysbindin (DTNB1) as a candidate gene for schizophrenia. Additional association studies of Dysbindin have been equivocal. Although multiple association

studies in a variety of populations have shown positive results, interpretation of the results has been difficult. Different association studies have not used the same SNP marker sets. Meta-analysis of five “positive” association studies using a high-resolution haplotype map designed to compare the five studies showed significant inconsistencies with regard to the identified disease-associated Dysbindin allele. Although it is possible that several different variants in the same gene could each contribute to disease susceptibility in different families or populations, this possibility does not explain the inconsistencies between the several Dysbindin association studies.

Linkage studies subsequently pointed to a region on chromosome 1 containing the candidate genes DISC 1 and DISC 2 (disrupted in schizophrenia 1 and 2) located on chromosome 1q21–22 and 1q32–42. These genes were initially identified in a large Scottish pedigree in the early 1990s. A balanced translocation between chromosomes 1 and 11 segregated in this pedigree and was possibly associated with serious mental illness. DISC 1 and 2 were identified in the original Scottish family because of their location near the chromosomal translocation breakpoint. As with Dysbindin, follow-up studies of DISC 1 and 2 have been equivocal. Genome screens, including a screen focused on extended Icelandic kindreds, have identified a schizophrenia candidate region on chromosome 8p21–22. Fine mapping of the region narrowed the search and eventually led to the proposal of neuregulin 1 (NRG1) as a schizophrenia candidate gene. Association studies again provided equivocal and difficult-tointerpret results. Meta-analysis of 14 separate studies using the SNP marker that demonstrated an association in the original study showed significant heterogeneity between the follow-up studies. It also showed that there is no consistent association between the specific risk allele “tagged” by the marker SNP and schizophrenia in different populations. However, after taking account of the statistical power of each association study, the meta-analysis showed a positive association between NRG1 at the level of the gene (as opposed to the SNP or haplotype level). Despite the equivocal genetic studies, significant resources have been channeled into molecular and neurophysiological investigations of the functional products of dysbindin, DISC 1 and 2, and neuregulin. Mutant mice for each of the three genes are now available and have been used to demonstrate interesting biological findings. For example, dysbindin is expressed in the hippocampus and dorsolateral prefrontal cortex. The dysbindin protein binds to B-dystrobrevin and has been implicated in synaptic structure and signaling. DISC 1 has been shown to influence neurite formation in cellular studies, and mutant mice for DISC 1 show impairments in a wide variety of tests including learning, memory, and sociability. Neuregulin belongs to a family of growth factors that mediate numerous functions including synapse formation, neuronal migration, and neurotransmission. Targeted disruption of erbB4, the postsynaptic target of neuregulin, leads to synaptic glutamatergic hypofunction. Despite the interesting biology uncovered, it remains unclear whether and to what extent any of these genes contribute to the etiology of schizophrenia in humans, and many geneticists have been cautious in their endorsement of the legitimacy of the mutant mice generated from the current list of candidate genes as models of psychiatric disorders. As with bipolar disorder, the genetic mapping findings for schizophrenia are promising but equivocal. Unlike for bipolar disorder, these mapping studies have generated a set of candidate genes that have stimulated a wide range of functional investigations, many of which have biologically interesting findings. As with bipolar disorder and other psychiatric disorders, the primary challenge in elucidating the genetic basis of schizophrenia is assembling adequate richly phenotyped samples for wellpowered genomewide mapping studies. REFERENCES Craddock N, O’Donovan MC, Owen MJ. Phenotypic and genetic complexity of psychosis. Invited commentary on Schizophrenia: A common disease caused by multiple rare alleles. Br J Psychiatry. 2007;190:200. De Luca V, Tharmalingam S, Zai C, Potapova N, Strauss J,

Vincent J, Kennedy JL. Association of HPA axis genes with

Revision #1
Created 2026-01-04 19:50:28 UTC by Omar Ayman
Updated 2026-01-04 19:50:28 UTC by Omar Ayman