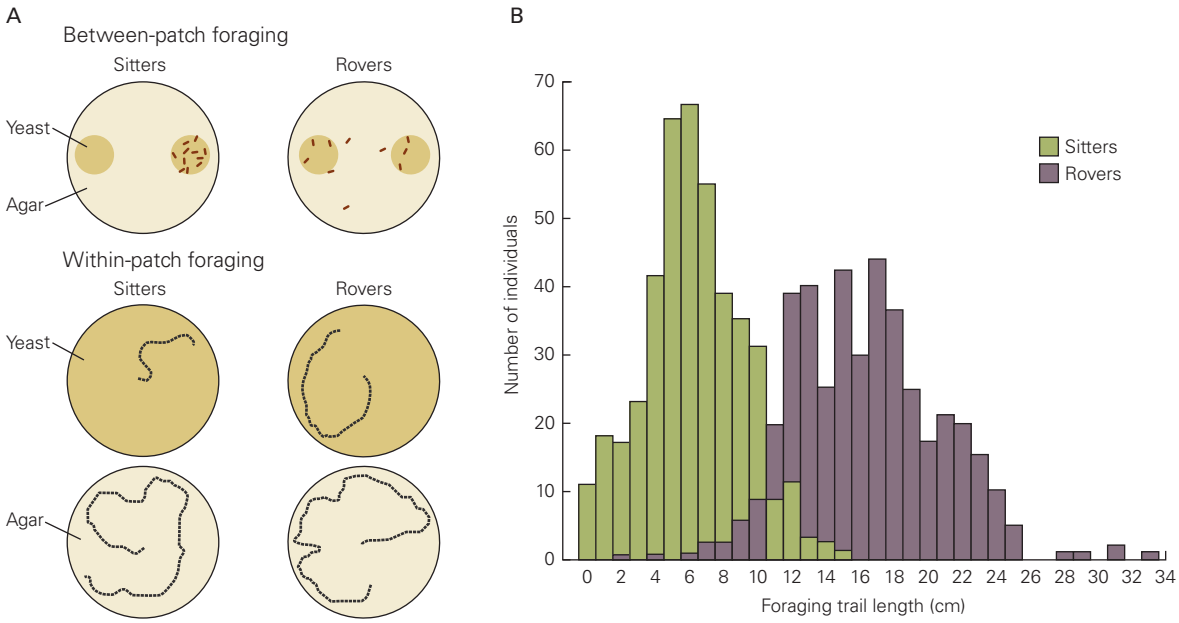

Figure 2–14 Foraging behavior of *Drosophila melanogaster* rover and sitter larvae differs while feasting on patches of yeast. (Reproduced, with permission, from Sokolowski. 2001. Copyright © 2001 Springer Nature.)

A. Rover-type larvae move from patch to patch, whereas sitter-type larvae stay put on a single patch for a long time. When foraging within a single patch, rover larvae move about more than sitter larvae. On agar alone, rover and sitter larvae move about equally.

B. While foraging within a patch of food, rovers have longer trail lengths than sitters (trail lengths were measured over a period of 5 minutes).

This difference in foraging behavior maps to a single protein kinase gene, *for*, which varies in activity in different fly larvae.

Why would variability in signaling enzymes be preserved in wild populations of *Drosophila*, which typically include both rovers and sitters? The answer is that variations in the environment create pressure for balanced selection for alternative behaviors. Crowded environments favor the rover larva, which is more effective at moving to new, unexploited food sources in advance of competitors, whereas sparse environments favor the sitter larva, which exploits the current source more thoroughly.

The *for* gene is also found in honeybees. Honeybees exhibit different behaviors at different stages of their life; in general, young bees are nurses, while older bees become foragers that leave the hive. The *for* gene is expressed at high levels in the brains of active foraging honeybees and at low levels in the younger and more stationary nurse bees. Activation of cGMP signaling in young bees can cause them to enter the forager stage prematurely; this change could be induced by an environmental stimulus or the bee's advancing age.

Thus the same gene controls variation in a behavior in two different insects, but in different ways. In the fruit fly, variations in the behavior are expressed in different individuals, whereas in the honeybee, they are expressed in one individual at different ages. The difference illustrates how an important regulatory gene can be recruited to different behavioral strategies in different species.

Neuropeptide Receptors Regulate the Social Behaviors of Several Species

Many aspects of behavior are associated with an animal's social interactions with other animals. Social behaviors are highly variable between species, yet have a large innate component within a species that is controlled genetically. A simple form of social behavior has been analyzed in the roundworm *Caenorhabditis elegans*. These animals live in soil and eat bacteria.

Different wild-type strains exhibit profound differences in feeding behavior. Animals from the standard laboratory strain are solitary, dispersing across a lawn of bacterial food and failing to interact with each other. Other strains have a social feeding pattern, joining large feeding groups of dozens or hundreds of animals (Figure 2–15). The difference between these strains is genetic, as both feeding patterns are stably inherited.

The difference between social and solitary worms is caused by a single amino acid substitution in a single gene, a member of a large family of genes involved in

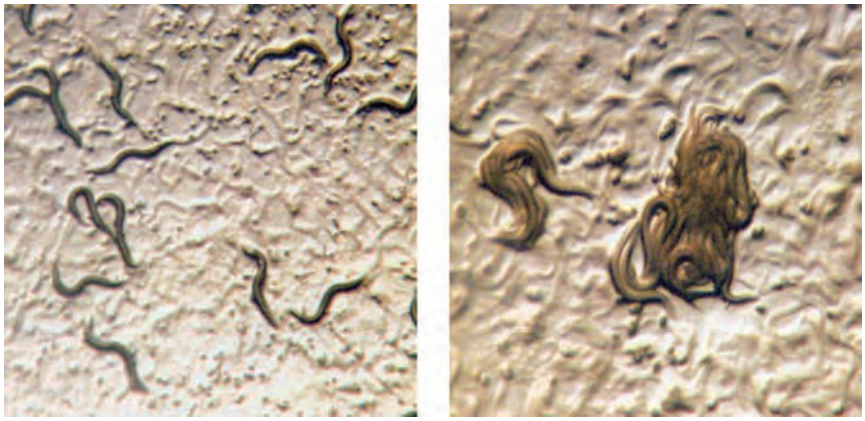

Figure 2–15 Feeding behavior of the roundworm *Caenorhabditis elegans* depends on the level of activity of the gene coding for a neuropeptide receptor. In one strain, individual worms graze in isolation (*left*), whereas in another strain, individuals mass together to feed. The difference is explained by a single amino acid substitution in the neuropeptide receptor gene. (Reproduced, with permission, from De Bono and Bargmann 1998.)

signaling between neurons. This gene, *npr-1*, encodes a neuropeptide receptor. Neuropeptides have long been appreciated for their roles in coordinating behaviors across networks of neurons. For example, a neuropeptide hormone of the marine snail *Aplysia* stimulates a complex set of movements and behavior patterns associated with a single behavior, egg laying. Mammalian neuropeptides have been implicated in feeding behavior, sleep, pain, and many other behaviors and physiological processes. The existence of a mutation in the neuropeptide receptor that alters social behavior suggests that this kind of signaling molecule is important both for generating the behavior and for generating the variation between individuals.

Neuropeptide receptors have also been implicated in the regulation of mammalian social behavior. The neuropeptides oxytocin and vasopressin stimulate mammalian affiliative behaviors such as pair bonding and parental bonding with offspring. In mice, oxytocin is required for social recognition, the ability to identify a familiar individual. Oxytocin and vasopressin have been studied in depth in prairie voles, rodents that form lasting pairs to raise their young. Oxytocin released in the brain of female prairie voles during mating stimulates pair-bond formation. Likewise, vasopressin released in the brain of male prairie voles during mating stimulates pair-bond formation and paternal behavior.

The extent of pair-bonding varies substantially between mammalian species. Male prairie voles form long-lasting pair-bonds with females and help them raise their offspring and are described as monogamous, but the closely related male montane voles breed widely and do not engage in paternal behavior. The difference between the behaviors of males in these species correlates with differences in the expression of the V1a class of vasopressin receptors in the brain. In prairie voles, V1a vasopressin receptors are expressed at high levels in a specific brain region, the ventral pallidum (Figure 2–16). In montane voles, the levels are much lower in this region, although high in other brain regions.

The importance of oxytocin and vasopressin and their receptors has been confirmed and extended by reverse genetic studies in mice, which are easier than voles to manipulate genetically. Introducing the V1a vasopressin receptor gene from prairie voles into male mice, which behave more like montane voles, increases the expression of the V1a vasopressin receptor in the ventral pallidum and increases the affiliative behavior of the male mice toward females. Thus differences between species in the pattern of expression of the vasopressin receptor can contribute to differences in social behaviors.

The analysis of vasopressin receptors in different rodents provides insight into the mechanisms by which genes and behaviors can change during evolution. Thus evolutionary changes in the pattern of expression of the V1a vasopressin receptor in the ventral forebrain have altered the activity of a neural circuit, linking the function of the ventral forebrain to the function of the vasopressin-secreting neurons that are activated by mating. As a result, social behaviors are altered.

The importance of oxytocin and vasopressin in human social behavior is not known, but the central role of pair-bonding and pup rearing in mammalian species suggests that these molecules might play a role in our species as well.

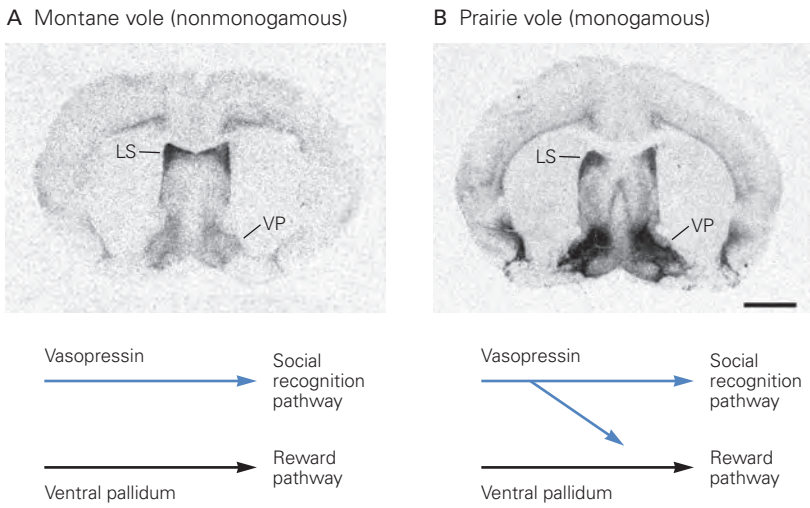

Figure 2–16 Distribution of vasopressin receptors (V1a) in two closely related rodent species. (Adapted, with permission, from Young et al. 2001. Copyright © 2001 Academic Press.)

A. Receptor expression is high in the lateral septum (LS) but low in the ventral pallidum (VP) in the montane vole, which does not form pair bonds.

B. Expression is high in the ventral pallidum of the monogamous prairie vole. Expression of the receptor in the ventral pallidum allows vasopressin to link the social recognition pathway to the reward pathway.

Studies of Human Genetic Syndromes Have Provided Initial Insights Into the Underpinnings of Social Behavior

Brain Disorders in Humans Result From Interactions Between Genes and the Environment

The first gene discovered for a neurological disease in humans clearly illustrates the interaction of genes and environment in determining cognitive and behavioral phenotypes. Phenylketonuria (PKU), described by Asbørn Følling in Norway in 1934, affects one in 15,000 children and results in severe impairment of cognitive function.

Children with this disease have two abnormal copies of the *PKU* gene that codes for phenylalanine hydroxylase, the enzyme that converts the amino acid phenylalanine to tyrosine. The mutation is recessive and heterozygous carrier individuals have no symptoms. Children who lack normal function in both copies of the gene accumulate high blood concentrations of phenylalanine from dietary proteins, which in turn leads to the production of toxic metabolites that interfere with neuronal function. The specific biochemical processes by which phenylalanine adversely affects the brain are still not understood.

The PKU phenotype (intellectual disability) results from the interaction of the genotype (the homozygous *pku* mutation) and the environment (the diet). The treatment for PKU is thus simple and effective: developmental delay can be prevented by a low-protein diet. The molecular and genetic analysis of gene function in PKU has led to dramatic improvements in the life of affected individuals. Since the early 1960s, the United States has instituted mandatory testing for PKU in newborns. Identifying children with the genetic disorder and modifying their diet before the disease appears prevents many aspects of the disorder.

Later chapters of this book describe many examples of single-gene traits that, like PKU, have led to insights into brain function and dysfunction. Certain themes have emerged from these studies. For example, a number of rare neurodegenerative disorders such as Huntington disease and spinocerebellar ataxia result from the pathological, dominant expansion of glutamate residues within proteins. The discovery of these polyglutamine repeat disorders highlighted the danger to the brain of unfolded and aggregated proteins. The discovery that epileptic seizures can be caused by a variety of mutations in ion channels led to the realization that these disorders are primarily disorders of neuronal excitability.

Rare Neurodevelopmental Syndromes Provide Insights Into the Biology of Social Behavior, Perception, and Cognition

Neurological and developmental disorders that manifest themselves in childhood have illuminated the

importance and complexity of genetics in human brain function. Early evidence that genes affect specific cognitive and behavioral circuitry emerged from studies of a rare genetic condition known as Williams syndrome. Individuals with this disorder typically exhibit normal language as well as extreme sociability; early in development, they lack the reticence children typically have in the presence of strangers. At the same time, they are profoundly impaired in spatial processing, show overall intellectual disability, and have very high rates of anxiety (but rarely social anxiety disorder).

The patterns of impairments in Williams syndrome, as compared with for example autism spectrum disorders, suggest that language and social skills can be separated from some other brain functions. Brain areas concerned with language are impaired in children with autism but are active or accentuated in Williams syndrome. By contrast, general and spatial intelligence is more impaired in Williams syndrome than in about half of all children with autism spectrum disorder.

Williams syndrome is caused by a heterozygous deletion of the chromosome region 7q11.23, most often encompassing about 1.5 Mb and 27 genes. The simplest interpretation of this defect is that the level of expression of the genes within the interval is reduced because there is only one copy instead of two of each gene in the region. The precise genes in the interval that influence social communication and spatial processing are not yet known, but they are of great interest because of their potential to provide insight into the genetic regulation of human behavior.

A more recent discovery in studies of autism spectrum disorders has further highlighted the complex relationship between genetic variation and social and intellectual functioning first illuminated by Williams syndrome. Within about the past decade, advances in genomic technology have allowed for high-throughput methods to screen the genome for variations in chromosomal structure, and at much higher resolution than was allowed by the light microscope (see Box 2–1). Seminal studies in 2007 and 2008 demonstrated that individuals with autism spectrum disorder carry new (de novo) copy number variations much more often than unaffected individuals. These findings led to some of the first discoveries of specific genomic intervals contributing to common forms of the syndrome (ie, autism spectrum disorder without evidence of syndromal features, also known as *idiopathic* or *nonsyndromic* autism spectrum disorder).

In 2011, two simultaneous large-scale studies of de novo copy number variations in a very well-characterized cohort found that precisely the same region deleted in Williams syndrome conferred substantial risk for autism spectrum disorder in an individual. However, in these cases, it was rare duplications (one excess copy of the region), and not deletions, that dramatically increased the risk for social disability. These findings, that losses and gains in the identical set of genes may lead to contrasting social phenotypes (while both typically lead to intellectual disability), further support the notion that domains of cognitive and behavioral functioning are separable but may share important molecular mechanisms.

Fragile X syndrome is another neurodevelopmental disorder of childhood that provides insight into the genetics of cognitive function; unlike Williams syndrome, it has been mapped to a single gene on the X chromosome. Fragile X syndrome varies in its presentation. Affected children may have intellectual disability, poor social cognition, high social anxiety, and repetitive behavior; about 30% of boys with fragile X syndrome meet diagnostic criteria for autism spectrum disorder. Fragile X syndrome is also associated with a broader range of traits, including physical characteristics such as an elongated face and protruding ears.

Fragile X syndrome has been shown to result from mutations that reduce expression of a gene called *fragile X mental retardation protein (FMRP)*. Because the gene falls on the X chromosome, males lose all expression of the gene when their one copy is mutated. FMRP protein regulates the translation of mRNAs into proteins in neurons, in a process that is itself regulated by neuronal activity. Regulated translation in neurons is an important component of the synaptic plasticity required for learning. The fragile X defect at the level of translation thus cascades up to affect neuronal function, learning, and higher-order cognitive processes. Interestingly, a large proportion of the other genes associated with increased risk for autism spectrum disorder as well as schizophrenia are regulated by the FMRP protein.

Another Mendelian disorder whose genetic basis is well understood is Rett syndrome (discussed in detail in Chapter 62). Rett syndrome is an X-linked, progressive neurodevelopmental disorder and one of the most common causes of intellectual disability in females. The disorder is almost always confined to females because canonical Rett mutations are very often lethal in the developing male embryo, which has a single X chromosome. Affected girls develop typically until they are 6 to 18 months of age, when they fail to acquire speech, show regression in intellectual functioning, and display compulsive, uncontrolled hand wringing instead of purposeful hand movement. In addition, girls with Rett syndrome often show a

period of markedly impaired social interaction that may be indistinguishable from autism spectrum disorder, although it is thought that social functioning is largely preserved in later life. Huda Zoghbi and her colleagues found that the major cause of this syndrome results from mutations in the *methyl CpG binding protein 2 (MeCP2)* gene. Methylation of specific CpG sequences in DNA alters expression of nearby genes, and one of the established functions of *MeCP2* is that it binds methylated DNA as part of a process that regulates mRNA transcription.

Rare syndromes have also offered some of the first insights into the genetic substrates of schizophrenia (Chapter 60). For example, as first described by Robert Shprintzen and colleagues in 1978, deletions of chromosome 22q11 lead to a wide range of physical and behavioral symptoms, including psychosis, now often referred to as velocardiofacial syndrome (VCFS), DiGeorge syndrome, or 22q11 deletion syndrome. The initial descriptions by Shprintzen were met with some skepticism as a result of the extremely broad range of phenotypes associated with the identical deletion. It is now widely appreciated that the 22q11 deletion is the most common chromosomal abnormality associated with schizophrenia and childhood-onset schizophrenia. Moreover, chromosomal losses of the identical region have been found to be associated with large individual risks for autism. To date, the specific genes within the region responsible for the psychiatric phenotype(s) have not been definitively established. Moreover, recent evidence from the autism literature suggests that it is likely that a combination of multiple genes within this interval, each conferring relatively small individual effects, is responsible for the social disability phenotype.

Psychiatric Disorders Involve Multigenic Traits

As mentioned earlier, single-gene syndromes are rare compared to the total burden of neurodegenerative and psychiatric disease. Consequently, one might question the rationale for studying rare disorders if they represent just a fraction of the total disease burden. The reason is that rare conditions can provide insight into the biological processes involved in more common, complex forms of a disease. For example, among the prominent successes of human genetics has been the discovery of rare genetic variants that lead to early-onset Alzheimer disease or Parkinson disease. Individuals with these severe rare variants represent a tiny subset of all individuals with these conditions, but the identification of rare disease variants uncovered cellular processes that are also disrupted in the larger patient pool, pointing to general therapeutic avenues. Similarly, pursuit of the pathophysiological mechanisms underlying Rett, fragile X, and other neurodevelopmental disorders has already led to some of the first attempts at rational drug development in psychiatric syndromes.

In the remainder of this chapter, we expand the discussion of the genetics of two complex neurodevelopmental and psychiatric phenotypes: autism spectrum disorders and schizophrenia. Compared to the rare Mendelian examples discussed earlier, the genetics of common forms of these conditions are indeed more diverse, varied, and heterogeneous, involving many different genes in different individuals as well as multiple risk genes conferring liability in combinations. Moreover, for both diagnoses, while the support for a genetic contribution is substantial, there is also compelling evidence for a contribution from environmental factors.

Progress in understanding these disorders came from the combination of rapidly advancing genomic technologies and statistical methods, a culture of open data sharing, and the consolidation of very large patient cohorts providing adequate power to detect very rare highly penetrant alleles as well as common genetic variants carrying small increments of risk. Importantly, recent successes in understanding both syndromes have provided a solid foundation for the pursuit of their biological consequences and the molecular, cellular, and circuit-level pathophysiology conveyed by these genetic risk factors.

Advances in Autism Spectrum Disorder Genetics Highlight the Role of Rare and De Novo Mutations in Neurodevelopmental Disorders

Autism spectrum disorders are a collection of developmental syndromes of varying severity affecting approximately 2% to 3% of the population and characterized by impairment in reciprocal social communication as well as stereotyped interests and repetitive behaviors. There is a significant male predominance; on average, three times as many boys as girls are affected. The clinical symptoms of autism spectrum disorders, by definition, emerge in the first 3 years of life, although highly reliable differences between affected and unaffected children are very often identifiable within the first months of life.

There is considerable phenotypic variability between those affected, leading to the development of the quite broad diagnostic classification of autism spectrum disorders. In addition, affected individuals have a higher frequency of seizures and cognitive

problems than the general population and often have serious impairments in adaptive functioning. However, many autistic individuals are not as profoundly affected and lead highly successful lives.

Autism has a very strong heritable component (see Figure 2–1A), which is likely to account for its being among the first genetically complex neuropsychiatric syndromes to yield to modern gene discovery tools and methods. Autism spectrum disorder has broader significance because it provides insight into behaviors that are quintessentially human: language, complex intelligence, and interpersonal interactions. Importantly, the fact that the defects in social communication seen in autism spectrum disorders can coexist with normal intelligence and typical functioning in other domains suggests that the brain is to some degree modular with distinct cognitive functions that can vary independently.

While syndromic forms of autism spectrum disorder account for a small fraction of all cases, the first findings in more common so-called "idiopathic" or "nonsyndromic" forms of the disorder also demonstrated a role for rare mutations carrying large biological effects. For example, in 2003, the sequencing of genes within a region on the X chromosome deleted in a very small number of females with autistic features led to the discovery of rare, loss-of-function mutations in the gene *neuroligin 4X (NLGN4X)*, a gene encoding a synaptic adhesion molecule in excitatory neurons and found in several affected male family members. Soon thereafter, a linkage analysis of a large pedigree with intellectual disability and autism spectrum disorder showed that affected family members all carried a loss-of-function *NLGN4X* mutation.

De novo submicroscopic deletions and duplications in chromosomal structure may dramatically increase an individual's risk for autism spectrum disorder. These copy number variations (CNVs) cluster in specific regions of the genome, identifying specific risk intervals. The earliest reports using this approach showed that the de novo CNVs at chromosome 16p11.2, although present in only about 0.5% to 1% of affected individuals, carried substantial (greater than 10-fold) risk of autism spectrum disorder. Subsequent studies have now identified a dozen or more de novo CNVs that carry risk, including at chromosomes 16p11.2, 1q21, 15q11-13, and 3q29; deletions at 22q11, 22q13 (deleting the gene *SHANK3*), and 2p16 (deleting the gene *NXRN1*); and de novo duplications of 7q11.23 (the Williams syndrome region).

Interestingly, although these CNVs carry large risks for autism spectrum disorder, studies of other psychiatric disorders, including schizophrenia and bipolar disorder, have found that many of the same regions increase the risk for these conditions as well. Moreover, studies of individuals ascertained by genotype (eg, 16p11.2 deletions and duplications) have found a wide variety of associated behavioral phenotypes, ranging from specific language impairment to intellectual disability to schizophrenia. This "one-to-many" phenomenon presents important challenges to illuminating specific pathophysiological mechanisms in psychiatric illness and to conceptualizing the steps from gene discovery to therapies.

The widespread and replicable findings that de novo rare CNVs increase the risk for autism spectrum disorder and other developmental disorders immediately raised the question of whether rare de novo mutations in single genes might carry similar risks. Indeed, the development of technology for low-cost, high-throughput DNA sequencing, initially focused on the coding portion of the genome, led to the identification of an excess of de novo mutations deemed likely to disrupt gene function (LGD mutations) in affected individuals. The repeated occurrence of these mutations in close proximity among unrelated individuals has now been exploited as a means to identify specific risk genes for autism spectrum disorders.

Large-scale studies of de novo mutations in autism spectrum disorders have now identified more than 100 associated genes, with about 45 of these reaching the highest confidence level of statistical significance. These genes have a wide range of known functions, but analyses reveal a statistically significant overrepresentation of genes involved in synaptic formation and function, and in regulation of transcription. Moreover, there are a greater than expected number of risk genes that encode RNAs that are targets of fragile X mental retardation protein and/or proteins that are active in early brain development.

Identification of Genes for Schizophrenia Highlights the Interplay of Rare and Common Risk Variants

Schizophrenia affects about 1% of all young adults, causing a pattern of thought disorders and emotional withdrawal that profoundly impairs life. It is strongly heritable (see Figure 2–1B) and also has a strong environmental component that is associated with stress on a developing fetus. Children born just after the Dutch Hunger Winter famine of World War II had a greatly increased risk of schizophrenia many years later, and children whose mothers were infected with the rubella virus during pregnancy in the 1960s pandemic were also at considerably increased risk.

Genes, as well as the environment, contribute to schizophrenia. As with autism, the sequencing of the human genome, the development of inexpensive methods for genome-wide genotyping of common variants and detection of CNVs, and the consolidation of very large patient cohorts have all resulted in a transformation in the genetics of schizophrenia. First, essentially in parallel with the findings in autism spectrum disorders noted earlier, rare and de novo CNVs began to be implicated in the risk for schizophrenia by the early 2000s. A small percentage of cases are associated with chromosomal abnormalities that carry large risks, including, for example, deletions at chromosome 22q11. These chromosomal abnormalities overlap entirely, or nearly so, with those loci implicated in autism spectrum disorders, but the distribution of risk among deletions and duplications at these loci does not appear to be identical. For instance, although duplications and deletions of the 16p11.2 region are both associated with autism spectrum disorders and schizophrenia, duplications of the region are more likely to lead to schizophrenia, whereas deletions are more likely to be seen with autism spectrum disorders and intellectual disability.

With regard to schizophrenia, the most important development of the last decade and a half has been the emergence of common variant genome-wide association studies (GWASs). In contrast to studies of hypothesis-driven candidate genes described earlier, genome-wide association relies on assaying polymorphisms at every gene in the genome simultaneously. This hypothesis-free approach, when used with well-powered cohorts and appropriate correction for multiple comparisons, has proven to be a highly reliable and reproducible strategy for identifying common risk alleles in common disorders across all of medicine.

GWASs involving nearly 40,000 cases and 113,000 controls have resulted in the identification of 108 risk loci for schizophrenia. The effects attributable to any individual genetic variant in this set have been quite modest, typically accounting for a less than 25% increase in risk. Moreover, many of the genetic polymorphisms assayed in GWASs map to regions outside of the coding segment of the genome. Consequently, although 108 risk loci have been identified, it is not yet entirely clear which genes correspond to all of these risk variants. In some cases, the variations mapped sufficiently close to a single gene that such a relationship could be reasonably inferred; in other cases, this remains to be determined.

The genes implicated in schizophrenia risk provide a starting point for determining the biology underlying the disorder. For example, since the late 1990s, evidence has pointed to the involvement of a region called the major histocompatibility complex (MHC) in schizophrenia risk. Accordingly, the MHC region has the strongest GWAS signal of any part of the human genome in the schizophrenia cohort. Detailed studies made possible by the very large number of patients in the cohort resolved this robust risk association signal in the MHC region into three different loci (and likely three different genes). Among these three loci, one gene, encoding the complement C4 factor, has a strong and definable effect on disease risk. Steven McCarroll and his colleagues showed that the complement C4 locus represents a natural case of CNV, that healthy individuals vary substantially in the number of copies of the gene they have, and that the level of expression of the C4A allele correlates with increasing schizophrenia risk. Subsequent follow-up studies showed that mice with knockouts of the C4 gene had a deficit in synaptic pruning during development, suggesting the hypothesis that excess C4A in humans might cause excessive synaptic pruning, a process that has long been of interest in the schizophrenia literature.

This finding represents an important demonstration of the ability to link genomics to a possible biological mechanism for increased disease risk. Even so, an individual with the highest-risk C4 haplotype who did not have a family history of schizophrenia would on average increase from having a 1% chance of being affected to approximately a 1.3% chance of being affected as a result of this allele. To get a sense of scale, having a first-degree relative with schizophrenia results in an approximately 10-fold increase in risk. This promising start and its limits reflect the challenges that geneticists and neurobiologists now face in moving from successful common variant gene discovery to the elaboration of the specific mechanisms leading to human pathology.

In addition to identifying numerous specific risk loci, GWASs in schizophrenia have repeatedly found that the small individual effects of many common alleles add up to increase risk. These results provide an additional, powerful avenue to study genotype-phenotype relationships in aggregate. Indeed, it is already clear that the number of risk alleles that an individual carries can have a significant (and additive) impact on the risk of developing the disorder. For instance, those in the top decile for a so-called polygenic risk score—a summary statistic relating to the overall amount of additive genetic risk an individual carries—are at 8- to 20-fold increased risk compared to the general population. Although the biology of the cumulative effect is not yet known, the observation nonetheless sets the stage for studying a series

of interesting questions related to disease trajectory and treatment response and will almost certainly reinvigorate studies combining neuroimaing and genomics. These latter types of investigations, similar to early efforts at common variant discovery, have suffered from poor reliability due to the inherent limitations of studying selected, biologically plausible candidate genes.

Finally, high-throughput sequencing methods, similar to those employed in autism spectrum disorders, have begun to yield results in schizophrenia as well. Specifically, exome sequencing in search of rare and de novo risk alleles has been pursued with some success. However, such studies require much larger cohorts to identify statistically significant risks for LGD mutations compared to autism spectrum disorders, suggesting that the overall effect size of these types of variations is likely to be substantially less in schizophrenia. To date, these investigations have identified a handful of risk genes and implicated key neurobiological pathways. In particular, recent exome studies have pointed to the importance of the molecules within the activity-regulated cytoskeleton (ARC) complex, as well as the gene *set domain containing 1A (SETDIA)*, as relevant for schizophrenia pathogenesis.

Perspectives on the Genetic Bases of Neuropsychiatric Disorders

Genes affect many aspects of behavior. There are remarkable similarities in personality traits and psychiatric illnesses in human twins, even those raised separately. Domestic and laboratory animals can be bred for particular, stable behavioral traits; and increasingly, the contributions of a wide range of genetic variations for neurodevelopmental and psychiatric disorders are being discovered.

A series of parallel advances have ushered in an era of remarkable opportunity to understand the relationship between genes, brain, and behavior. The armamentarium available to manipulate and study model systems has been revolutionized. At the same time, progress in defining the genetic risk factors for human neuropsychiatric disorders has advanced considerably. Although the field remains in an early stage in this process, multiple examples of the value of successful gene discovery, and its application to deep biological understanding, have emerged.

Among the many striking findings from recent genetic studies of neurodevelopmental and psychiatric conditions is the overlap in genetic risks across a wide range of diagnostic boundaries. While it may not come as much of a surprise that biology does not hew to categorical diagnostic criteria, it is nonetheless a formidable conceptual challenge to consider how the field will trace these effects and arrive at new therapeutic strategies.

In addition, it is worthwhile noting that for many other psychiatric conditions that have not yet seen the type of progress noted earlier, the calculus is straightforward: greater investment and larger sample sizes will lead to greater insight. For example, recent studies of de novo mutations in Tourette syndrome and obsessive-compulsive disorder clearly demonstrate that the rate-limiting factor for the identification of high confidence risk genes is the availability for sequencing of parent-child trios. In a similar vein, GWAS studies of major depression have only very recently reached sample sizes adequate to confirm statistically significant associated common variants. These studies have included hundreds of thousands of individuals and, not surprisingly, have identified risk alleles carrying very small individual effects.

This last point highlights the idea that one size does not fit all for the genomics of behavioral, developmental, and psychiatric disorders. From the investigations of model systems, to the illumination of rare Mendelian disorders, to the disentangling of both common and rare variants contributing to common disorders, the tools and opportunities available today are unprecedented. The coming years should see deep insights into the biology of psychiatric and neurodevelopmental disorders, and perhaps therapies with the potential to help patients and their families.

Highlights

1. Rare genetic syndromes such as fragile X syndrome, Rett syndrome, and Williams syndrome have provided important insights into the molecular mechanisms of complex human behaviors. Moreover, while considerable work remains to be done, the study of these syndromes has already challenged the notion that associated cognitive and behavioral deficits are immutable and has demonstrated the utility of a wide range of model systems in illuminating conserved biological mechanisms.
2. The sequencing of the human genome, the development of high-throughput genomic assays, and simultaneous computing and methodological advances have led to a profound change in the understanding of the genetics of human behavior and psychiatric illness. Several paradigmatic disorders, including schizophrenia and autism,

have seen dramatic progress, leading to the identification of dozens of definitive risk genes and chromosomal regions.

3. The maturation of the field of psychiatric genetics and genomics over the past decade has revealed the frailty of testing pre-specified candidate genes. These types of studies have now been supplanted by genome-wide scans of both common and rare alleles. Coupled with rigorous statistical frameworks and consensus statistical thresholds, these are yielding highly reliable and reproducible results.
4. At present, the cumulative evidence suggests that the full range of genetic variations underlies complex behavioral syndromes, including common and rare, transmitted and de novo, germline and somatic, and sequence and chromosomal structural variation. However, the relative contributions of these various types of genetic changes vary for a given disorder.
5. A striking finding from recent advances in the genetics of human behavior has been the overlap of genetic risks for syndromes with distinct symptoms and natural histories. Understanding how and why an identical mutation may lead to highly diverse phenotypic outcomes in different individuals will be a major challenge for the future.
6. Findings across common psychiatric disorders point to extremely high rates of genetic heterogeneity. This, coupled with the biological pleiotropy of the risk genes that have been identified to date, as well as the dynamism and complexity of human brain development, all point to important challenges ahead in moving from an understanding of risk genes to an understanding of behavior. Similarly, at present, an important distinction can be made between illuminating the biology of risk genes and unraveling the pathophysiology of behavioral syndromes.

Glossary[1]

Allele. Humans carry two sets of chromosomes, one from each parent. Equivalent genes in the two sets might be different, for example, because of single nucleotide polymorphisms. An allele is one of the two (or more) forms of a particular gene.

Centromere. Chromosomes contain a compact region known as a centromere, where sister chromatids (the two exact copies of each chromosome that are formed after replication) are joined.

Cloning. The process of generating sufficient copies of a particular piece of DNA to allow it to be sequenced or studied in some other way.

Complementary DNA (cDNA). A DNA sequence made from a messenger RNA molecule, using an enzyme called *reverse transcriptase*. cDNAs can be used experimentally to determine the sequence of messenger RNAs after their introns (non–protein-coding sections) have been spliced out.

Conservation of genes. Genes that are present in two distinct species are said to be conserved, and the two genes from the different species are called *orthologous genes*. Conservation can be detected by measuring the similarity of the two sequences at the base (RNA or DNA) or amino acid (protein) level. The more similarities there are, the more highly conserved the two sequences.

Copy number variation (CNV). A deletion or duplication of a limited genetic region that results in an individual having more or fewer than the usual two copies of some genes. Copy number variations are observed in some neurological and psychiatric disorders.

CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats). An enzyme-RNA system in which the enzyme cleaves target sequences that match an RNA guide; the RNA guide can be engineered to recognize a desired gene or sequences within a cell for mutation.

Euchromatin. The gene-rich regions of a genome (see also heterochromatin).

Eukaryote. An organism with cells that have a complex internal structure, including a nucleus. Animals, plants, and fungi are all eukaryotes.

Genome. The complete DNA sequence of an organism.

Genotype. The set of genes that an individual carries; usually refers to the particular pair of alleles (alternative forms of a gene) that a person has at a given region of the genome.

Haplotype. A particular combination of alleles (alternative forms of genes) or sequence variations that are closely linked—that is, are likely to be inherited together—on the same chromosome.

Heterochromatin. Compact, gene-poor regions of a genome, which are enriched in simple sequence repeats.

Introns and exons. Genes are transcribed as continuous sequences, but only some segments of the resulting messenger RNA molecules contain information that encodes a protein product. These segments are called *exons*. The regions between exons are known as *introns* and are spliced from the RNA before the product is made.

[1]Based on Bork P, Copley R. 2001. Genome speak. Nature 409:815.