One gene, many neuropsychiatric disorders: lessons from Mendelian diseases

Xiaolin Zhu[1], Anna C Need[1,2], Slavé Petrovski[1,3] & David B Goldstein[1]

Recent human genetic studies have consistently shown that mutations in the same gene or same genomic region can increase the risk of a broad range of complex neuropsychiatric disorders. Despite the steadily increasing number of examples of such nonspecific effects on risk, the underlying biological causes remain mysterious. Here we investigate the phenomenon of such nonspecific risk by identifying Mendelian disease genes that are associated with multiple diseases and explore what is known about the underlying mechanisms in these more 'simple' examples. Our analyses make clear that there are a variety of mechanisms at work, emphasizing how challenging it will be to elucidate the causes of nonspecific risk in complex disease. Ultimately, we conclude that functional approaches will be critical for explaining the causes of nonspecific risk factors discovered by human genetic studies of neuropsychiatric disorders.

Neurological and neuropsychiatric disorders are among the major contributors to human morbidity and mortality[1]. There is broad agreement that genetics represents one of the most promising avenues for identifying new therapeutic directions to treat this substantial unmet clinical need. Moreover, elucidating the genetic basis of neurological and neuropsychiatric disorders provides invaluable insights into how our nervous system functions and dysfunctions.

After many years of stagnation, genetic studies of complex forms of neurological and neuropsychiatric disorders have finally entered an era of systematic and definitive discovery, thanks in large part to technological advances permitting genome-wide interrogation. One of the most striking features of the recent findings is that many genes, and even mutations, that have been implicated in one neurological/neuropsychiatric disorder have also been implicated in other disorders (http://omim.org/). Genome-wide association studies (GWASs) have repeatedly shown a shared risk for schizophrenia and bipolar disorder both for specific loci[2–4] and through analysis of polygenic risk scores[2,5] and related approaches[6]. Relationships between schizophrenia, bipolar disorder, depression, attention deficit hyperactivity disorder (ADHD) and autism have also been demonstrated through studies of common single nucleotide polymorphisms (SNPs)[2,6].

There remains a debate as to whether such relationships are truly a result of shared genetic loci between different disorders or are simply an artifact of inaccuracies in psychiatric diagnosis. The latter may indeed contribute to the phenomenon when considering closely related disorders, such as schizophrenia and bipolar disorder, or bipolar disorder and depression, or autism with low cognitive ability and intellectual disability, where individual patients are often diagnosed with both at varying points in their lives. Even in these cases, however, there are strong arguments against misdiagnosis being the primary explanation; for example, patients with bipolar and schizophrenia clearly tend to respond best to different drugs. In addition, theoretical modeling shows that although misdiagnosis can inflate estimates of genetic correlation among disorders, it is unlikely to explain the high genetic correlation observed between bipolar disorder and schizophrenia[6,7]. In addition, the shared genetic risk of multiple neuropsychiatric disorders is reminiscent of what is observed for autoimmune disorders[2].

Moreover, misdiagnosis is an even less persuasive argument when applied to variants associated with very distinct neuropsychiatric disorders. Virtually without exception, copy number variants (CNVs) associated with one neuropsychiatric disorder have been found to predispose to other neurological or neuropsychiatric phenotypes[8–12]. This includes sharing of risk loci between conditions that could not reasonably reflect diagnostic ambiguity, such as 16p13.11 deletions, which are seen in multiple neuropsychiatric conditions, including a range of epilepsies without psychiatric illness or intellectual disability[13], as well as in profound intellectual disability without seizures[14]. In further support of the case for 'one variant, multiple disorders' is the observation that many of these CNVs were found not only to confer risk of multiple different neurodevelopmental disorders, but also to confer risk of other disorders, such as obesity, cardiac abnormalities and cataracts[8,9,15]. For instance, the 1q21.1 deletion associated with schizophrenia[16] and other neuropsychiatric disorders has been reported to associate with variable pediatric phenotypes such as cardiac abnormalities and cataracts, usually including intellectual disability or developmental delay, but occasionally with no obvious neuropsychiatric phenotype[8]. It has also been reported in at least one adult with cardiac abnormalities and no neuropsychiatric disorder[17], as well as in several patients with multiple severe congenital abnormalities[18].

Most of these CNVs contain a very large number of genes, but, in some cases, specific genes have been associated with neuropsychiatric disorders through CNV analysis, including deletions in

[1]Center for Human Genome Variation, Duke University School of Medicine, Durham, North Carolina, USA. [2]Division of Brain Sciences, Department of Medicine, Imperial College London, London, UK. [3]Department of Medicine, The University of Melbourne, Austin Health and Royal Melbourne Hospital, Melbourne, Victoria, Australia. Correspondence should be addressed to D.B.G. (d.goldstein@duke.edu).

Received 12 November 2013; accepted 31 March 2014; published online 27 May 2014; doi:10.1038/nn.3713

NRXN1 (refs. 19,20) and *CNTNAP2* (ref. 21). Notably, even these specific CNVs cause multiple, highly variable, disorders[22,23]. More recently, this model of shared causality across multiple complex neurodevelopmental disorders has also been observed at the level of mutations in individual genes. Exome-sequencing studies, often focused on trios (proband and both parents), have implicated the same genes in different diseases, including autism spectrum disorder (ASD)[24–27], intellectual disability (ID)[28,29], schizophrenia[30] and epileptic encephalopathies (EE)[31]. The most obvious explanation for this is that, across these complex neurodevelopmental disorders, the children are ascertained for genetic abnormalities in genes that have a global neurodevelopmental effect. The complexity of neurodevelopment makes it easy to imagine that a disruption in the amount or the timing of expression of a single gene could differentially affect multiple pathways and the resulting phenotype may be influenced by other rare and common genetic variants in these pathways, as well as by environmental factors such as nutrition and general health. Most notably across the aforementioned studies, *de novo* mutations in previously known neurodevelopmental disorder genes, such as *STXBP1*, *SCN2A* and *TCF4*, are observed multiple times across patients with broader clinical diagnoses of ASD, EE and severe ID[24–29,31].

Despite the definitive emerging pattern of shared causality among complex neuropsychiatric disorders, we need to be cautious of circumstances in which this phenomenon might be overstated because of a lack of conclusive evidence. The complex genetic architecture of neuropsychiatric disorders poses daunting challenges for gene mapping efforts[32]. Appropriate standards for statistically valid identification of genes that confer risk at the population level are maturing rapidly and will permit unambiguous identification of disease genes when sample sizes are sufficiently large. Nevertheless, implicating a gene as conferring significant risk in the population as a whole is different from assigning causation in individual patients. In the case of complex disease, causation in individual patients represents a substantial challenge for fine-scale genotype-phenotype correlation analyses. Nevertheless, we share the view often expressed that human genetics is the best general framework for providing statistically convincing identification of disease genes[33].

Although nonspecificity of risk genes contributing to complex forms of neuropsychiatric disorders has been the subject of much attention[9,34], there has been little consideration of what the examples of genes causing Mendelian neurological diseases can tell us about the causes of such nonspecific effects on risk of disease. We should clarify that nonspecificity refers simply to the phenomenon of a gene associated with more than one disease, and we make no claims regarding the underlying causes. Thus, nonspecific risk does not necessarily imply nonspecific biological effects, and may or may not apply to the underlying nature of diseases. For example, mutations in a gene might cause disruption of a critical neurodevelopmental process that ultimately contributes to a diversity of neuropsychiatric phenotypes across current diagnostic boundaries. In this case, the biological effect is specific, but we observe nonspecific risk. Keeping this in mind, we review what is known about Mendelian disease genes that cause distinct diseases with a view to using that experience to help understand the recent observations related to complex neuropsychiatric disorders.

We first summarize the known examples of a single gene leading to different diseases. Next, we discuss the possibility that already known Mendelian disease genes may, in some cases, carry mutations that confer risk of more complex forms of neuropsychiatric disorders. Finally, we discuss experimental procedures that may allow elucidation of exactly how mutations in single genes can cause a variety of disease presentations. In particular, we emphasize that the combination of rapidly maturing genome editing and stem cell technologies presents a powerful platform for modeling the effects of disease-causing mutations in cellular models of disease. Because of the throughput afforded by this approach, it will be particularly valuable in efforts to assess the biological consequences of the many mutations identified in individual genes and how those mutations may interact with the genetic background to confer risk.

Lessons from Mendelian disease genes

As the name implies, monogenic disease is caused by mutation(s) in a single gene. However, the reverse is not necessarily true. In fact,

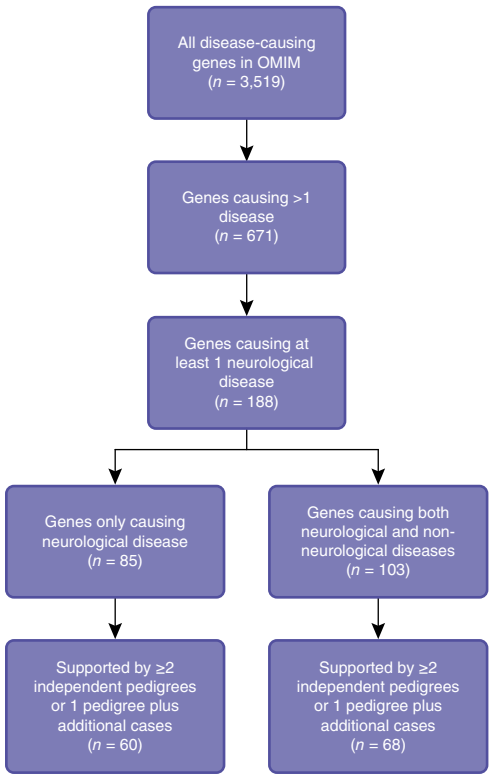

Figure 1 Nonspecific disease-causing genes that cause at least one Mendelian neurological disease. Disease means an independent phenotype entry in OMIM with a unique OMIM number and without brackets, braces or a question mark[38]. Neurological diseases were included by manual curation, excluding metabolic diseases not primarily affecting nervous system, ophthalmologic diseases and deafness, cranial skeletal diseases, brain vascular diseases, neuromuscular junction diseases, primary muscular diseases, neuroendocrine diseases, and non-familial nervous system tumors. We only considered germline mutations. To enhance reliability, we further required each gene-disease correlation to be supported by at least two independent pedigrees or one pedigree plus additional cases as documented in OMIM. We found 128 genes causing at least one Mendelian neurological disease, for which OMIM and PubMed were examined. For genes for which relevant literature clearly indicates a mechanism, we recorded that mechanism. Doing this over all 128 genes with informative literature suggested the presence of three different mechanism groups: distinct locations of mutations, extent of functional change and qualitatively different effects, although these groups are not necessarily mutually exclusive. Examples of genes (**Table 1**) with supporting literature are shown in **Supplementary Table 1**. A few of these genes can also harbor identical mutations associated with more than one disease; these genes are indicated in **Table 1**. There is a unique group of genes with identical mutations linked to different diseases. This is particularly interesting because it clearly implicates something other than the mutation alone. We identified these genes by checking OMIM 'Allelic Variants' for any mutations linked to more than one disease.

one gene–one disease might be the exception rather than the rule. The phenomenon of one gene or genomic region conferring risk of multiple diseases is closely associated with the concept of an allelic series. As formulated by Victor A. McKusick in 1973, the term allelic series describes disease phenotypes that are seemingly caused by different genes on clinical grounds, but with genetic and/or biochemical evidence supporting mutations in the same gene[35]. Similarly, nonspecific risk is also closely related to the concept of pleiotropy, defined as "the phenomenon in which a single locus affects two or more distinct phenotypic traits"[36]. We do not refer to an allelic series or pleiotropy, however, as they represent only two possible mechanisms by which mutation in the same gene may cause different diseases. In fact, our fundamental aim is to try to determine how often the known examples follow the concept of an allelic series or pleiotropy, or some other mechanism (such as interactions of the same class of mutations in a gene with genetic background, or epistasis[37]).

To obtain an overview of known genes that cause more than one disease, we referred to OMIM (Online Mendelian Inheritance in Man) (http://omim.org/), which hosts up-to-date, curated information about human Mendelian disorders and their causal genes[38], focusing particularly on neurological illness (**Fig. 1**). To enhance the reliability of our analysis, we further required each gene-disease correlation to be supported by at least two independent pedigrees or one pedigree plus additional cases as documented in OMIM. We next studied what is known about precisely which mutations in each gene cause which disease based on the literature. In particular, we focused on examples in which the causal mutations for different diseases fall in different locations of a protein, differ in their magnitudes of effect on protein function, have different biochemical/functional effects on the protein, and are clearly the same type of mutations, or indeed the same exact mutation, which suggests genetic or environmental modification or other mechanisms (**Table 1** and **Supplementary Table 1**). These four groups of mechanisms are not meant to be exhaustive or mutually exclusive, but represent an initial effort to interpret the complexity of mutational effects in the context of pathogenesis of neurological disorders.

Distinct locations of mutations

One mechanism of nonspecificity is that mutations in different parts of a protein are associated with different effects and thereby associated with different presentations. For genes encoding proteins that are structurally resolved experimentally or computationally, disease-specific mutations can be observed to be in different functional domains of the protein, allowing an assessment of specificity at the functional level within a protein. For example, among the broad spectrum of X-linked mental retardation syndromes caused by *ATRX* mutations, major urogenital abnormalities are found to be associated with truncating mutations at the C-terminal part of the protein[39]. Further clinical genetic studies show that mutations in the plant homeodomain (PHD)-like domain are associated with more severe phenotypes than mutations in the helicase domain[40]. As more disease-causing mutations are discovered and mapped onto functional substructures of proteins, a clearer picture of such genotype-phenotype correlations is likely to emerge.

Even when such a one domain–one disease pattern is not clear, an uneven distribution of disease-specific mutations across a gene can hint at different causes of pathogenicity. A well-understood case is the amyloid beta precursor protein gene *APP* that, when mutated, can cause familial Alzheimer disease or cerebral amyloid angiopathy[38]. Notably, Alzheimer disease–causing mutations are more likely to be localized near beta-secretase or gamma-secretase cleavage sites, with amino acid changes flanking beta-amyloid sequence, whereas *APP* mutations associated with cerebral amyloid angiopathy tend to cluster in the beta-amyloid sequence; this phenomenon can be explained by different physicochemical properties of beta-amyloid fragments, which in turn determines which type of tissue (brain parenchyma versus vascular wall) they are more likely to deposit in[41]. In many cases, the unusual distribution of disease-specific mutations points to distinct pathobiology that awaits further investigation. For example, *ATP1A3* (encoding a Na^+/K^+ pump) mutations, resulting in alternating hemiplegia of childhood (AHC), are concentrated in transmembrane domains of the protein, whereas *ATP1A3* mutations, resulting in rapid-onset dystonia-parkinsonism (RDP), tend to be distributed throughout the protein[42]. Notably, germline mutations throughout the tumor suppressor gene *SMAD4* cause familial cancer syndromes, including juvenile polyposis with hereditary hemorrhagic telangiectasia, and three different substitutions at a single site in the same protein (I500V, I500M and I500T) cause Myhre syndrome, a rare mental retardation syndrome with dysmorphic facial features and skeletal anomalies[43,44]. Such a restricted mutational spectrum of Myhre syndrome–causing mutations (versus a broad spectrum of cancer syndrome–causing mutations) clearly indicates a specific association between genetic lesion and pathogenic mechanism.

Tissue-specific alternative splicing is one of the key mechanisms for cellular specialization[45]. This tissue specificity can also lead to disease specificity for mutations in different parts of a gene. Known examples are limited, but instructive. For example, mutations in *WNK1* can cause either hereditary sensory and autonomic neuropathy (HSAN) or pseudohypoaldosteronism (PHA), a disorder of the adrenal cortex[38]. It was found that HSAN-causing recessive mutations truncate the isoform exclusively expressed in the nervous system[46], whereas the PHA-causing mutations are heterozygous intronic deletions that increase gene expression[47].

Indeed, recent analyses of temporally and spatially resolved human brain transcriptomes and the *de novo* mutations identified by sequencing studies have led to important new hypotheses about the

Table 1 Examples of genes causing more than one Mendelian disease

	Causing only neurological diseases	Causing both neurological and non-neurological diseases
Distinct locations of mutations	*APP*, ATP1A3, PLP1*	*ATRX, INF2, SMAD4*, TUBB3, WNK1*
Extent of functional change	*CASK, RPGRIP1L, RPS6KA3, SCN1A*, SCN2A*	*ASAH1, ATP7A, GCH1, OFD1, OPA3, POMGNT1, POMT1, POMT2*, TMEM67**
Qualitatively different effects	*CACNA1A*, FUS*, MPZ*, NFIX, SCN9A*	*AR*, DNM2, FLNA*, PRPS1, PTCH1, SMARCB1, SOX10**
Same mutation, different diseases	*GARS, KCNQ2, L1CAM, MAPT, MECP2, PMP22, PRNP, PRRT2, SIX3, SMN1, TBCE, TMEM216, TSEN54, VAPB, VCP*	*ADAR, LMNA, NF1, NPHP1, PAX2, POLG, PSAP, PSEN2, PTEN, SCO2, SHH, SLC2A1, TREX1, TRPV4, TSC1, WFS1*

Examples of genes are represented in a way that reflects the underlying biological mechanism based on the literature (**Supplementary Table 1**). In some cases, there may be multiple potential mechanisms linking a gene to multiple diseases. For each gene (and its associated diseases), the corresponding mechanism is the best established but not necessarily the only mechanism involved.
*These genes can also harbor mutations causing different diseases. However, they are shown not as examples of 'same mutation, different diseases', but as examples highlighting other mechanisms.

neuropathogenesis of schizophrenia[48] and autism[49,50]. Interestingly, Parikshak *et al.* found that ASD genes converged to specific brain circuitry, whereas ID genes lacked developmental or anatomical specificity, in spite of substantial overlap between the two sets of genes[50]. As transcriptome databases grow, we expect that an increasing fraction of disease-specific mutational consequences can be linked to transcript isoforms that are specific to a cell type and/or a developmental stage.

Extent of functional change

Even if the biochemical or functional changes caused by mutations in a gene are qualitatively similar, quantitative differences between mutant proteins could translate into different clinical diagnoses. Illustratively, when a metabolic enzyme is mutated, null alleles tend to cause a congenital, severe form of systemic enzyme deficiency syndrome (including severe neurological phenotypes), whereas hypomorphic alleles might be more likely to generate a later-onset, milder form of disease with restricted, distinct nervous system manifestation. For instance, mutations in *GCH1* (encoding GTP cyclohydrolase I, the rate-limiting enzyme in tetrahydrobiopterin (BH4) biosynthesis) can cause either BH4-deficient hyperphenylalaninemia B (HPABH4B) or dopa-responsive dystonia (DRD)[38]. When homozygous or compound heterozygous mutations cause a loss of enzymatic activity, HPABH4B results and presents as hyperphenylalaninemia with severe neurological conditions, including dystonia; when heterozygous mutations disrupt the protein, but residual enzymatic activity is preserved, DRD results and presents as a primarily neurological disorder without systemic hyperphenylalaninemia[51]. Not surprisingly, there is also a transitional, recessive form of DRD that can present with or without hyperphenylalaninemia[52]. Similarly, *ASAH1* (encoding acid ceramidase) mutations that eliminate most or all of the lysosomal acid ceramidase activity cause Farber lipogranulomatosis, an early-onset lysosomal storage disorder with severe systemic manifestations[53], whereas mutations in the same gene retaining higher residual activity cause spinal muscular atrophy with progressive myoclonic epilepsy (SMAPME), a milder, later-onset disease confined to neurons[54].

Similar examples can be found in genes that do not encode enzymes. For example, among the four clinically discernible ciliopathies with overlapping features caused by *TMEM67* mutations, nephronophthisis with liver fibrosis (NPHP11) is distinguished from the other three by no or only mild neurological involvement, which can be explained by the hypomorphic nature of NPHP11-causing mutations[55]. Such observations indicate that different quantitative effects of mutations could result in either the presence or absence of specific symptoms or differing degrees of severity. Relevant examples for the latter might include mutations in enzymes required for O-mannosylation of proteins (for example, *POMGNT1*, *POMT1* and *POMT2*) causing related types of muscular dystrophy-dystroglycanopathies distinguished by clinical severity[56], and possibly *SCN1A* mutations causing familial febrile seizures (FEB, less severe), generalized epilepsy with febrile seizures plus (GEFS+, more severe) or Dravet syndrome (most severe)[57]. Although a nearly perfect genotype-phenotype correlation is extremely rare, it is also clear that we lack the appropriate assays to assess the quantitative effect of mutations for many kinds of proteins. This may be the reason that there are more clear-cut examples for enzymes, where quantification of the effect of mutations is more straightforward.

Qualitatively different effects of mutations

In some cases, mutations can affect protein function in ways that are clearly qualitatively distinct. In some cases, this category will overlap with the first category described above, where mutations fall in defined regions of genes. The most obvious example of mutations with clearly different effects would be loss-of-function (LoF) versus gain-of-function (GoF) mutations, although there may of course be different GoF mutations that have distinct gains of function. For instance, GoF mutations in *SCN9A* cause primary erythermalgia characterized by episodic vasodilation and burning pain of lower limbs[58], whereas LoF mutations in the same gene cause insensitivity to pain[59]. Conversely, sometimes apparently dissimilar biochemical properties of mutations can have more similar pathogenic effects, demonstrating the complexity of genotype-phenotype relationships. For instance, among the four diseases caused by mutations in *PRPS1* (encoding phosphoribosylpyrophosphate synthetase)[38], at least two (Arts syndrome and X-linked recessive Charcot-Marie-Tooth disease 5) are associated with decreased or loss of enzyme activity and one (phosphoribosylpyrophosphate synthetase superactivity) is associated with enzyme hyperactivity[60]. However, all four *PRPS1* spectrum diseases (the three aforementioned plus X-linked deafness 1) have neurological phenotypes that could be explained by a unifying biochemical mechanism[60].

In most situations in which LoF or GoF is associated with different pathologies, it seems that nonsense-mediated decay (NMD), a quality-control mechanism of the cell to selectively eliminate transcripts carrying truncating mutations[61], has a pivotal role in determining the severity of the clinical outcome. Typically, if NMD is in effect, haploinsufficiency (dominant LoF) could be the disease-causing mechanism; if NMD is not able to eliminate an abnormal transcript, the mutant protein itself may be pathogenic, as a result of either toxicity or a dominant-negative mechanism. This could explain why sometimes more severe diseases are associated with missense rather than truncating mutations of the same gene[38]. For example, *de novo NFIX* mutations can cause either Sotos syndrome-2 (SOTOS2), an overgrowth syndrome, or Marshall-Smith syndrome (MRSHSS), a malformation syndrome that is more severe than SOTOS2 (ref. 2). It is believed that the milder SOTOS2 is a result of haploinsufficiency when NMD is working, whereas MRSHSS is the result of a dominant-negative effect of the mutant protein that escapes NMD[62]. Other well-understood cases include *MPZ* mutations causing several distinct peripheral neuropathies[63] and *SOX10* mutations causing different neurological syndromes[64]. The more severe disease phenotypes being associated more often with NMD failure not only highlights the extensive and critical roles of NMD in maintaining the stability of our functional genome, but also raises the possibility of manipulating NMD as a potential treatment for human diseases[61].

LoF mutations knockdown or knockout a protein; however, the consequence of a GoF mutation can be difficult to predict on the basis of the known protein function. A conspicuous group of such GoF mutations, which are major culprits of inherited neurodegenerative diseases, is trinucleotide repeat expansion[65]. Such GoF mutations can have toxic effects at both the protein and RNA level and their pathogenetic mechanism (causing neurodegeneration) can be unrelated to the original function of the affected gene. Kennedy disease, for example, is a neurodegenerative disorder characterized by spinal and bulbar muscular atrophy associated with CAG trinucleotide expansion in the coding region of androgen receptor (*AR*) gene[66]. Most other *AR* mutations, however, cause androgen insensitivity[38].

Same mutation, different diseases

In some cases, it is clear that the same mutation can cause different diseases[38]. Such a situation can occur across families or within single pedigrees. For example, a homozygous 12-bp deletion in exon 2 of *TBCE* causes hypoparathyroidism-retardation-dysmorphism (HRD)

syndrome or Kenny-Caffey syndrome (similar to HRD, but with additional features of osteosclerosis and recurrent bacterial infections) across different Middle Eastern pedigrees[67]. In a single pedigree, a heterozygous R155H substitution in *VCP* reportedly manifested either as inclusion body myopathy with Paget disease and frontotemporal dementia (IBMPFD) or as classic amyotrophic lateral sclerosis (ALS)[68]. Such multiple disease–causing mutations highlight the extensive variability of phenotypic expression and are uniquely informative for potential underlying mechanisms, including genetic background, environment, stochastic events and epigenetics.

In fact, there is a large and growing list of mutations that cause more than one disease (**Table 1**), and some have been characterized mechanistically. First, a mutation can cause different diseases depending on zygosity (that is, heterozygosity, homozygosity or compound heterozygosity). For example, two mutations (Q53X and E140K) in *SCO2* (encoding a copper chaperone that is important for normal function of cytochrome *c* oxidase (COX)) can individually cause familial high-grade myopia in heterozygous form, whereas compound heterozygosity (with each other) causes cardioencephalomyopathy as a result of COX deficiency-1 (CEMCOX1)[69].

Second, interaction between a disease-causing genotype and individual genetic background, a form of epistasis[37], can result in different phenotypes linked to the same mutation. Some pedigree-specific variable phenotypic expression of mutations, such as the *TBCE* deletion aforementioned, might indicate the effect of genetic background, if relevant pedigree-specific environmental exposure can be ruled out. In a well-understood example, copy number polymorphism has a role, in which *SMN2* gene copy number modulates the severity of spinal muscular atrophy associated with mutations in *SMN1* (ref. 70). Although such genetic background effects tend to be attributed to a disease modifier gene, we note that the nature and presentation of genetic background can be very complex and might involve multiple loci and even a whole spectrum of personal genetic variation[71], similar to strain-specific effects of engineered mutations in mouse models[72].

In addition to genetic and environmental factors, stochastic events in cellular and developmental processes can also contribute to phenotypic variability. Important lines of evidence include monozygotic twins raised in presumably very similar environments who are discordant for disease phenotype (or susceptibility to disease)[73], and incomplete penetrance of Mendelian disease–causing mutations within pedigrees[38]. Despite a paucity of mechanistic understanding of these epidemiological observations in humans, recent investigation in model organisms has elucidated interesting principles that might be relevant to variable presentations of disease-causing mutations[74]. These studies found that stochasticity is a hallmark of biological systems and can cause substantial variation in gene expression, enzyme-substrate and protein-protein interaction, and cell-cell interaction, shaping development and finally leading to variation of organismal-level phenotypes[75].

Furthermore, epigenetic regulation can contribute to different outcomes of the 'same' genetic defect[38]. This mechanism is particularly important for neuropsychiatric disorders because epigenetic regulation is fundamental to brain development and function[76]. A classic example is Prader-Willi syndrome and Angelman syndrome, which are associated with disruption of imprinted genes at 15q11–15q13 (ref. 77). In genomic imprinting, the active allele is solely determined by its parent of origin[78], whereas, in other cases, allele-specific expression can be random, including X-inactivation and autosomal random monoallelic expression[78]. Regardless of the underlying mechanisms, all these forms of monoallelic expression can potentially complicate genotype-phenotype correlations[78].

Sometimes mutations linked to multiple diseases may largely be a matter of clinical resolution, for example, in cases of different diagnoses with overlapping features. As an illustration, two phenotypically similar, but clinically different, diseases distinguished by severity, Charcot-Marie-Tooth disease type 2D (CMT2D) and distal hereditary motor neuronopathy type VA (dHMN5A), can coexist in a family and are caused by a single segregating mutation (E71G or D500N) in the *GARS* gene[79]. Multiple diagnoses associated with exactly the same genetic defect do accentuate the phenotypic variability that might benefit clinical management, but an umbrella diagnosis could help to understand the common disease biology. However, mechanism-oriented diagnosis can be extremely challenging in reality, especially when a mutation has multiple direct effects culminating in apparently distinct disorders. Such pleiotropic effects of mutations are widespread in mouse models[80] and might contribute to nonspecific risk observed in both Mendelian neurological and complex neuropsychiatric disorders by adding complexity to ascertaining phenotypes unambiguously. On the other hand, it is interesting to see examples of unexpected genotype-phenotype correlations that challenge the prevailing biological paradigm or diagnostic nosology, such as the identical *WDR62* frameshift mutation contributing to microcephaly with a broad spectrum of cortical abnormalities that have been generally considered as distinct in etiology and pathogenesis[81].

Mendelian disease genes causing complex forms of neuropsychiatric disorders

Until recently, the primary mechanism for identifying disease-causing genes relied on patients being recognized as having similar conditions on the basis of strictly clinical criteria. Once a syndrome had been recognized clinically, it was then possible to track down the responsible mutations, initially through linkage and, more recently, by next generation sequencing[33]. Critically, when patients are organized first clinically and then subsequently studied genetically, we are virtually guaranteed to miss examples in which mutation in a given gene results in an atypical presentation. For this simple reason, it is possible, and even likely, that many already 'known' Mendelian disease–causing genes cause other conditions that have not yet been recognized.

Overlap between genes associated with complex disease and those linked to Mendelian disease is well documented in the human genetics literature[38]. In some cases, such a link between monogenic and complex forms of disease can shed light on common pathobiology, where rare, highly penetrant, often protein-damaging mutations cause the Mendelian form of disease, whereas common, less penetrant, often regulatory variants are associated with susceptibility[71]. Such a pattern is also likely to emerge for complex neuropsychiatric disorders[82]. Notably, Blair *et al.* analyzed over 110 million medical records and found consistent associations between Mendelian and complex diseases[83]. These included association of multiple complex neuropsychiatric disorders with neurological Mendelian conditions such as spinocerebellar ataxia and degenerative diseases of the basal ganglia, and, perhaps more surprisingly, with non-neurological conditions such as hemophilia and long QT syndrome[83]. As sporadic cases of neuropsychiatric disorder yield to investigation, we can also begin to understand the potential contribution that *de novo* mutations make to more classically defined complex disorders. This paradigm permits the identification of strong risk factors at the point of introduction into the population.

The primary and most reliable analysis framework for implication of *de novo* mutations in risk of complex disease is to statistically assess whether individual genes carry significantly more *de novo* mutations than expected by chance[25,31]. Properly implemented, this analysis

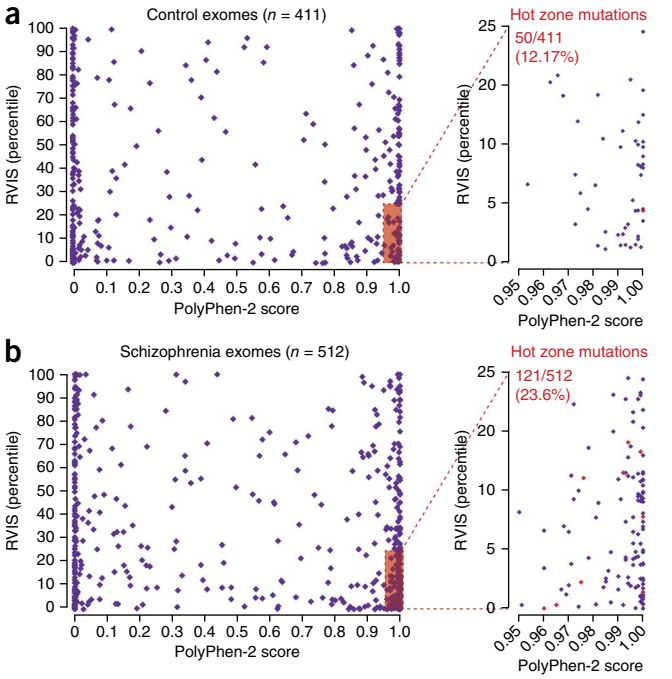

Figure 2 Distribution of *de novo* mutations across controls and schizophrenia patients. (**a**,**b**) *De novo* mutations (DNMs) in controls (**a**) and schizophrenia patients (**b**) are plotted across genic-level residual variation intolerance score (RVIS) and quantitative variant-level PolyPhen-2 score axes. The previously defined hot zone (RVIS ≤25th percentile and PolyPhen-2 ≥0.95)[86] is highlighted at bottom right. Nonsense and canonical splice mutations are recoded with a PolyPhen-2 score of 1 and synonymous with 0. We took control DNM data from seven independent studies[24,26,27,29,30,48,88], three of which were not leveraged by Petrovski *et al.*[86]. We assessed enrichment of hot zone mutations among cases using the collection of DNMs from four schizophrenia studies[30,48,85,89]. For cases and controls, we began with the 830 and 674 CCDS single nucleotide substitution DNMs, respectively. All DNMs were annotated together using Variant Effect Predictor[86]. We excluded 79 case and 66 control DNMs overlapping ESP6500SI variants (http://evs.gs.washington.edu/EVS/) and 34 case and 20 control DNMs in RVIS unassessed genes[86]. Cases and controls with multiple DNMs only contributed the single most-damaging DNM for the hot zone illustration. The most-damaging DNM was defined as the DNM with the shortest Euclidean distance from coordinate (1,0). This reflects 512 cases and 411 controls that reported at least one eligible DNM. Hot zone mutations were significantly enriched in cases versus controls (23.6 versus 12.17%, two-tail Fisher's exact test, $P = 8.2 \times 10^{-6}$). This reflects an excess of 49% or 59 DNMs in the hot zone among cases. The insets show amplification of the hot zone region. In our previous publication in EE, we found an estimated relative risk of 81 for the risk factors present among the 25th percentile most intolerant genes[31]; DNMs in the hot zones overlapping genes with DNMs among the EE study[31] are colored red.

must account for the size and mutability of genes and the number of cases that have been sequenced, and must correct for all the genes in the genome. This framework has now definitively implicated genes in multiple neuropsychiatric disorders, including EE[31].

Beyond implicating individual genes in this way, patterns in the data can also highlight the overall presence of additional risk factors, and this can provide important pointers to the nature of those risk factors. For example, a number of different studies have shown that, among the genes carrying *de novo* mutations in patient genomes, there is a clear excess of genes that are regulated by the Fragile-X protein[84], after appropriately correcting for the gene size and mutability of the FMRP-regulated set of genes[24,31,85].

Similarly, Petrovski *et al.* recently introduced an approach that combines gene- and variant-level prioritization scores to test for the presence of *de novo* mutation that confers risk in patient genomes. Briefly, a two-dimensional plot of *de novo* mutations is constructed. One dimension reflects the corresponding residual variation intolerance score (RVIS) percentile for each gene, reflecting the degree to which the gene tolerates mutation in the human population[86]. The other dimension is PolyPhen-2 (ref. 87), which assesses the likelihood a given mutation alters protein function. In this two-dimensional space, Petrovski *et al.* defined a hot zone as the region occupied by an RVIS genic percentile score ≤25th percentile and a PolyPhen-2 quantitative score ≥0.95 (ref. 86). They showed that, compared with *de novo* mutations from control individuals, this hot zone was significantly enriched for *de novo* mutations among ASD, EE and ID sequenced trios[86]. By expanding the control data set used by Petrovski *et al.* to consist of seven studies that published control *de novo* mutation data, we observed the rate of *de novo* mutations in the hot zone is 12.2% based on these control individuals (**Fig. 2a**)[24,26,27,29,30,48,88].

This framework allows us to evaluate whether there are *de novo* mutation risk factors present across four recent schizophrenia trio sequencing studies[30,48,85,89]. We found that, of the 512 schizophrenia sequenced probands who had at least one protein-coding *de novo* single nucleotide substitution, 121 probands (23.6%) had a *de novo* mutation in the most damaging hot zone (**Fig. 2b** and **Supplementary Table 2**). This is a significant excess over the 12.2% observed among controls ($P = 8.2 \times 10^{-6}$). Moreover, this not only proves the presence of risk factors in the published schizophrenia *de novo* mutation data, but also shows us where to find some of them. Specifically, these results indicate that, among the 121 hot zone *de novo* mutations, there are an excess of 59 (49%) *de novo* mutations compared with what we would expect from control trios. This further translates to approximately 59 (6.2%) of the combined 951 schizophrenia sequenced probands[30,48,85,89] being affected by one of the excess hot zone *de novo* risk mutations. It is therefore of particular interest to investigate the genes carrying hot zone *de novo* mutations. In particular, we observed a striking overlap of genes with specifically hot zone mutations observed across the multiple neuropsychiatric disorders (**Fig. 3**).

We also found that, of the 121 *de novo* mutations in the hot zone, 51 occurred in genes that are known or suspected to carry mutations that confer significant risk of disease, through either already known Mendelian disease genes ($n = 27$) or genes observed to have a *de novo* mutation among recent exome-sequencing studies of neuropsychiatric ascertained trios ($n = 34$). For example, specifically for EE, we had previously demonstrated that the risk factors present among the 4,264 genes that are in the 25th percentile for intolerance had an estimated relative risk of 81 (ref. 31). In this schizophrenia data set, we observed that, among the 121 hot zone mutations, nine *de novo* mutations overlapped intolerant (RVIS genic percentile score ≤25th percentile) genes from the previous EE study (*ALS2CL*, *CHD4*, *GNAO1*, *ITPR1*, *KIAA2018*, *PAQR8*, *SCN2A*, *TRRAP* and *ZBTB40*)[31].

Resolving nonspecific disease-causing mechanisms

Ultimately, a fuller understanding of nonspecificity of risk will depend not only on genetic association, but also on functional modeling. In particular, it will be critical to assess the functional effect of allelic series in genes defined on clinical criteria and how the effects of those alleles depend on genetic background. Notably, it is well known that nearly all known genetic diseases show marked variation in presentation, and this is true even given the pronounced ascertainment biases that have caused the community to systematically miss patients that do not show typical presentations, as described above. Despite the clear evidence that something is modifying presentation, we have

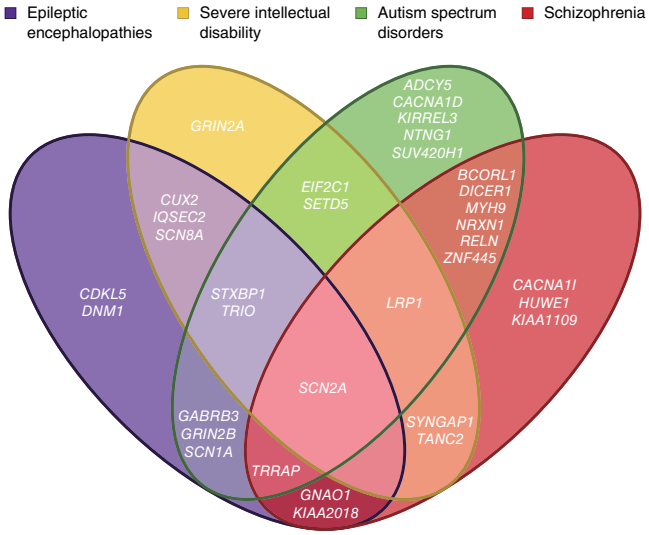

Figure 3 Venn diagram representing the overlap of genes affected by hot zone *de novo* mutations across four neuropsychiatric disorders. Hot zone *de novo* mutations are *de novo* mutations that occur in the 25th percentile most-intolerant genes, genome-wide (RVIS), and are either putatively LoF (nonsense or canonical splice) or PolyPhen-2 (HumVar) ≥ 0.95 *in silico* prediction[86].

only a handful of cases in which such modifiers have been identified, among which cystic fibrosis (CF) is prominent in that several genes have been found to modify several different clinical manifestations, including lung function, *Pseudomonas aeruginosa* infection, meconium ileus, CF-related diabetes and probably liver disease[90].

Until now, systematic assessment of disease-causing mutations and how they interact with the genetic background has been beyond reach. Recent technological advances, however, provide a clear direction for systematic study of how mutations interact to confer disease risk. In particular, the combination of rapid and flexible genome editing[91] and stem cell technologies[92] makes possible the systematic assessment of the effects of specific mutations in controlled genetic backgrounds and even specific combinations of mutations. This is not to say that cellular modeling will entirely replace animal models, but cellular models permit a scale of investigation that will be difficult to achieve in any other system.

Of particular importance in this regard are multi-electrode arrays, which permit the effect of mutations and combinations of mutations on neuronal network behavior to be assessed[93]. More recent optogenetic approaches will even allow such networks to be monitored in much finer detail, with exquisite spatial and temporal resolution of electrical impulses[94]. Because many neurological and neuropsychiatric disorders are fundamentally diseases that involve networks of neurons, this level of complexity may be required to fully realize the effect of the mutations, as opposed to studies, for example, in oocytes or in single neurons.

Conclusions

Nonspecific effects of pathogenic mutations have emerged as one of the key themes of recent studies of neuropsychiatric disorders. Currently, almost nothing is known about the underlying biological causes of this nonspecificity of risk. The sobering conclusion from our analysis of nonspecific effects of risk mutations in Mendelian disease genes is that there is no obvious predominant mechanism responsible. Indeed, Mendelian disease genes offer examples of virtually all possible mechanisms. There are clear examples of nonspecificity emerging from mutations of different effects in a single gene, and these different effects can be either quantitative or qualitative. There are also, however, numerous examples in which the same exact mutation, in the same zygosity, can cause different diseases, clearly implicating modifiers outside of the gene. This complexity, even in the case of Mendelian diseases, strongly suggests that only careful integration of genetic data with experimental investigation will reveal the underlying mechanisms of nonspecificity. Fortunately, recent advances in experimental technologies provide a powerful approach to investigate interactions of genetic variants in individual patients. As more disease-causing genes are discovered by sequencing studies, we anticipate the cellular models in particular may be the most flexible experimental procedure for dissecting the ways in which mutations in a single gene can cause very different clinical presentations. Once specific hypotheses are identified in the higher throughput setting of cellular models, these hypotheses can be further refined and evaluated in appropriate animal models.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

1. Insel, T.R. & Landis, S.C. Twenty-five years of progress: the view from NIMH and NINDS. *Neuron* **80**, 561–567 (2013).
2. Cross-Disorder Group of the Psychiatric Genomics Consoritum & Genetic Risk Outcome of Psychosis Consortium. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *Lancet* **381**, 1371–1379 (2013).
This is the first study showing that specific genetic variants are significantly associated with risk of different psychiatric disorders. By integrating genome-wide association study (GWAS) data of five different psychiatric disorders from multiple resources, this study uses statistical modeling strategies to convincingly demonstrate presence of nonspecific genetic risk factors in terms of both individual alleles and polygenic risk scores.
3. Ruderfer, D.M. *et al.* Polygenic dissection of diagnosis and clinical dimensions of bipolar disorder and schizophrenia. *Mol. Psychiatry* published online, doi:10.1038/mp.2013.138 (November 26 2013).
4. Williams, H.J. *et al.* Fine mapping of ZNF804A and genome-wide significant evidence for its involvement in schizophrenia and bipolar disorder. *Mol. Psychiatry* **16**, 429–441 (2011).
5. International Schizophrenia Consortium. *et al.* Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* **460**, 748–752 (2009).
6. Cross-Disorder Group of the Psychiatric Genomics Consortium. *et al.* Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat. Genet.* **45**, 984–994 (2013).
7. Wray, N.R., Lee, S.H. & Kendler, K.S. Impact of diagnostic misclassification on estimation of genetic correlations using genome-wide genotypes. *Eur. J. Hum. Genet.* **20**, 668–674 (2012).
8. Mefford, H.C. *et al.* Recurrent rearrangements of chromosome 1q21.1 and variable pediatric phenotypes. *N. Engl. J. Med.* **359**, 1685–1699 (2008).
This paper presents an excellent example of a rare, recurrent microdeletion associated with a wide spectrum of phenotypes including intellectual disability, microcephaly, cardiac abnormalities, and cataracts, emphasizing the importance of a genotype-based approach to clinical management of relevant patients.
9. Grayton, H.M., Fernandes, C., Rujescu, D. & Collier, D.A. Copy number variations in neurodevelopmental disorders. *Prog. Neurobiol.* **99**, 81–91 (2012).
10. Bijlsma, E.K. *et al.* Extending the phenotype of recurrent rearrangements of 16p11.2: deletions in mentally retarded patients without autism and in normal individuals. *Eur. J. Med. Genet.* **52**, 77–87 (2009).
11. Ben-Shachar, S. *et al.* Microdeletion 15q13.3: a locus with incomplete penetrance for autism, mental retardation, and psychiatric disorders. *J. Med. Genet.* **46**, 382–388 (2009).
12. Morrow, E.M. Genomic copy number variation in disorders of cognitive development. *J. Am. Acad. Child Adolesc. Psychiatry* **49**, 1091–1104 (2010).
13. Heinzen, E.L. *et al.* Rare deletions at 16p13.11 predispose to a diverse spectrum of sporadic epilepsy syndromes. *Am. J. Hum. Genet.* **86**, 707–718 (2010).
14. Ullmann, R. *et al.* Array CGH identifies reciprocal 16p13.1 duplications and deletions that predispose to autism and/or mental retardation. *Hum. Mutat.* **28**, 674–682 (2007).

15. Bachmann-Gagescu, R. *et al.* Recurrent 200-kb deletions of 16p11.2 that include the SH2B1 gene are associated with developmental delay and obesity. *Genet. Med.* **12**, 641–647 (2010).
16. Stefansson, H. *et al.* Large recurrent microdeletions associated with schizophrenia. *Nature* **455**, 232–236 (2008).
17. Digilio, M.C. *et al.* Congenital heart defects in recurrent reciprocal 1q21.1 deletion and duplication syndromes: rare association with pulmonary valve stenosis. *Eur. J. Med. Genet.* **56**, 144–149 (2013).
18. Rosenfeld, J.A. *et al.* Proximal microdeletions and microduplications of 1q21.1 contribute to variable abnormal phenotypes. *Eur. J. Hum. Genet.* **20**, 754–761 (2012).
19. Kirov, G. *et al.* Neurexin 1 (NRXN1) deletions in schizophrenia. *Schizophr. Bull.* **35**, 851–854 (2009).
20. Rujescu, D. *et al.* Disruption of the neurexin 1 gene is associated with schizophrenia. *Hum. Mol. Genet.* **18**, 988–996 (2009).
21. Friedman, J.I. *et al.* CNTNAP2 gene dosage variation is associated with schizophrenia and epilepsy. *Mol. Psychiatry* **13**, 261–266 (2008).
22. Ching, M.S. *et al.* Deletions of NRXN1 (neurexin-1) predispose to a wide spectrum of developmental disorders. *Am. J. Med. Genet. B. Neuropsychiatr. Genet.* **153B**, 937–947 (2010).
23. Gregor, A. *et al.* Expanding the clinical spectrum associated with defects in CNTNAP2 and NRXN1. *BMC Med. Genet.* **12**, 106 (2011).
24. Iossifov, I. *et al. De novo* gene disruptions in children on the autistic spectrum. *Neuron* **74**, 285–299 (2012).
25. Neale, B.M. *et al.* Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature* **485**, 242–245 (2012).
This is one of the first studies using a trio design to investigate the contribution of *de novo* mutations to a complex neuropsychiatric disorder (autism spectrum disorder, ASD). Importantly, this study establishes an analysis framework to statistically assess whether individual genes harbor significantly more *de novo* mutation than expected by chance. This study indicates an important, but limited, role of *de novo* mutations in the pathogenesis of ASD.
26. O'Roak, B.J. *et al.* Sporadic autism exomes reveal a highly interconnected protein network of *de novo* mutations. *Nature* **485**, 246–250 (2012).
27. Sanders, S.J. *et al. De novo* mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* **485**, 237–241 (2012).
28. de Ligt, J. *et al.* Diagnostic exome sequencing in persons with severe intellectual disability. *N. Engl. J. Med.* **367**, 1921–1929 (2012).
29. Rauch, A. *et al.* Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. *Lancet* **380**, 1674–1682 (2012).
30. Xu, B. *et al.* De novo gene mutations highlight patterns of genetic and neural complexity in schizophrenia. *Nat. Genet.* **44**, 1365–1369 (2012).
31. Epi4K Consortium. *et al. De novo* mutations in epileptic encephalopathies. *Nature* **501**, 217–221 (2013).
This is the first study that comprehensively investigates the contribution of *de novo* mutations to epileptic encephalopathies (EE), a spectrum of devastating neurodevelopmental disorders that can overlap with ASD and/or intellectual disability (ID) in clinical presentation. Using statistical modeling, this study clearly implicates *de novo* mutations as genetic risk factors of EE in both individual genes and the group of ~4,000 genes that are most intolerant to functional genetic variation.
32. Sullivan, P.F., Daly, M.J. & O'Donovan, M. Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nat. Rev. Genet.* **13**, 537–551 (2012).
This paper provides an elegant review on genetic architectures of nine major neuropsychiatric disorders, based on discoveries from GWASs and structural variation studies.
33. Goldstein, D.B. *et al.* Sequencing studies in human genetics: design and interpretation. *Nat. Rev. Genet.* **14**, 460–470 (2013).
This paper provides an overview of implementation issues for sequencing study design, sequence data generation, variant discovery and prioritization, statistical analysis and interpretation, and functional evaluation of candidate variants along with genetic evidence to assess pathogenicity. This review introduces the idea of the 'narrative potential' inherent in sequence data and advocates for the importance of appropriate statistical criteria to secure genetic discoveries.
34. Stessman, H.A., Bernier, R. & Eichler, E.E. A genotype-first approach to defining the subtypes of a complex disease. *Cell* **156**, 872–877 (2014).
35. McKusick, V.A. Phenotypic diversity of human diseases resulting from allelic series. *Am. J. Hum. Genet.* **25**, 446–456 (1973).
36. Stearns, F.W. One hundred years of pleiotropy: a retrospective. *Genetics* **186**, 767–773 (2010).
37. Mackay, T.F. Epistasis and quantitative traits: using model organisms to study gene-gene interactions. *Nat. Rev. Genet.* **15**, 22–33 (2014).
38. Online Mendelian Inheritance in Man. McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University http://omim.org/ (accessed 2 September 2013).
39. Gibbons, R.J. & Higgs, D.R. Molecular-clinical spectrum of the ATR-X syndrome. *Am. J. Med. Genet.* **97**, 204–212 (2000).
40. Badens, C. *et al.* Mutations in PHD-like domain of the ATRX gene correlate with severe psychomotor impairment and severe urogenital abnormalities in patients with ATRX syndrome. *Clin. Genet.* **70**, 57–62 (2006).
41. Revesz, T. *et al.* Genetics and molecular pathogenesis of sporadic and hereditary cerebral amyloid angiopathies. *Acta Neuropathol.* **118**, 115–130 (2009).
42. Heinzen, E.L. *et al. De novo* mutations in ATP1A3 cause alternating hemiplegia of childhood. *Nat. Genet.* **44**, 1030–1034 (2012).
43. Caputo, V. *et al.* A restricted spectrum of mutations in the SMAD4 tumor-suppressor gene underlies Myhre syndrome. *Am. J. Hum. Genet.* **90**, 161–169 (2012).
44. Le Goff, C. *et al.* Mutations at a single codon in Mad homology 2 domain of SMAD4 cause Myhre syndrome. *Nat. Genet.* **44**, 85–88 (2012).
45. Blencowe, B.J. Alternative splicing: new insights from global analyses. *Cell* **126**, 37–47 (2006).
46. Shekarabi, M. *et al.* Mutations in the nervous system–specific HSN2 exon of WNK1 cause hereditary sensory neuropathy type II. *J. Clin. Invest.* **118**, 2496–2505 (2008).
47. Wilson, F.H. *et al.* Human hypertension caused by mutations in WNK kinases. *Science* **293**, 1107–1112 (2001).
48. Gulsuner, S. *et al.* Spatial and temporal mapping of *de novo* mutations in schizophrenia to a fetal prefrontal cortical network. *Cell* **154**, 518–529 (2013).
49. Willsey, A.J. *et al.* Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. *Cell* **155**, 997–1007 (2013).
50. Parikshak, N.N. *et al.* Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. *Cell* **155**, 1008–1021 (2013).
This is one of the most successful studies mapping genes implicated by exome-based *de novo* mutation screens onto temporally and spatially resolved human brain transcriptomes. Notably, this study shows strong connections among the genes implicated in ASD in terms of developmental and spatial specificity, indicating involvement of specific brain circuits in the pathogenesis of ASD.
51. Ichinose, H. *et al.* Hereditary progressive dystonia with marked diurnal fluctuation caused by mutations in the GTP cyclohydrolase I gene. *Nat. Genet.* **8**, 236–242 (1994).
52. Furukawa, Y. *et al.* Dystonia with motor delay in compound heterozygotes for GTP-cyclohydrolase I gene mutations. *Ann. Neurol.* **44**, 10–16 (1998).
53. Koch, J. *et al.* Molecular cloning and characterization of a full-length complementary DNA encoding human acid ceramidase. Identification Of the first molecular lesion causing Farber disease. *J. Biol. Chem.* **271**, 33110–33115 (1996).
54. Zhou, J. *et al.* Spinal muscular atrophy associated with progressive myoclonic epilepsy is caused by mutations in ASAH1. *Am. J. Hum. Genet.* **91**, 5–14 (2012).
55. Otto, E.A. *et al.* Hypomorphic mutations in meckelin (MKS3/TMEM67) cause nephronophthisis with liver fibrosis (NPHP11). *J. Med. Genet.* **46**, 663–670 (2009).
56. Mercuri, E. *et al.* Congenital muscular dystrophies with defective glycosylation of dystroglycan: a population study. *Neurology* **72**, 1802–1809 (2009).
57. Kanai, K. *et al.* Physicochemical property changes of amino acid residues that accompany missense mutations in SCN1A affect epilepsy phenotype severity. *J. Med. Genet.* **46**, 671–679 (2009).
58. Waxman, S.G. & Dib-Hajj, S.D. Erythromelalgia: a hereditary pain syndrome enters the molecular era. *Ann. Neurol.* **57**, 785–788 (2005).
59. Cox, J.J. *et al.* An SCN9A channelopathy causes congenital inability to experience pain. *Nature* **444**, 894–898 (2006).
60. de Brouwer, A.P. *et al.* PRPS1 mutations: four distinct syndromes and potential treatment. *Am. J. Hum. Genet.* **86**, 506–518 (2010).
61. Frischmeyer, P.A. & Dietz, H.C. Nonsense-mediated mRNA decay in health and disease. *Hum. Mol. Genet.* **8**, 1893–1900 (1999).
62. Malan, V. *et al.* Distinct effects of allelic NFIX mutations on nonsense-mediated mRNA decay engender either a Sotos-like or a Marshall-Smith syndrome. *Am. J. Hum. Genet.* **87**, 189–198 (2010).
63. Khajavi, M. *et al.* Curcumin treatment abrogates endoplasmic reticulum retention and aggregation-induced apoptosis associated with neuropathy-causing myelin protein zero-truncating mutants. *Am. J. Hum. Genet.* **77**, 841–850 (2005).
64. Inoue, K. *et al.* Molecular mechanism for distinct neurological phenotypes conveyed by allelic truncating mutations. *Nat. Genet.* **36**, 361–369 (2004).
65. Nelson, D.L., Orr, H.T. & Warren, S.T. The unstable repeats–three evolving faces of neurological disease. *Neuron* **77**, 825–843 (2013).
66. La Spada, A.R., Wilson, E.M., Lubahn, D.B., Harding, A.E. & Fischbeck, K.H. Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature* **352**, 77–79 (1991).
67. Parvari, R. *et al.* Mutation of TBCE causes hypoparathyroidism-retardation-dysmorphism and autosomal recessive Kenny-Caffey syndrome. *Nat. Genet.* **32**, 448–452 (2002).
68. Johnson, J.O. *et al.* Exome sequencing reveals VCP mutations as a cause of familial ALS. *Neuron* **68**, 857–864 (2010).
69. Tran-Viet, K.N. *et al.* Mutations in SCO2 are associated with autosomal-dominant high-grade myopia. *Am. J. Hum. Genet.* **92**, 820–826 (2013).
70. Parsons, D.W. *et al.* Intragenic telSMN mutations: frequency, distribution, evidence of a founder effect, and modification of the spinal muscular atrophy phenotype by cenSMN copy number. *Am. J. Hum. Genet.* **63**, 1712–1723 (1998).
71. Lupski, J.R., Belmont, J.W., Boerwinkle, E. & Gibbs, R.A. Clan genomics and the complex architecture of human disease. *Cell* **147**, 32–43 (2011).
This paper proposes a unified genetic model for human morbidities, highlighting the role of recent mutations. There is an instructive discussion on genetic relationships between Mendelian and complex diseases.
72. Nadeau, J.H. Modifier genes in mice and humans. *Nat. Rev. Genet.* **2**, 165–174 (2001).
73. Wong, A.H., Gottesman, I.I. & Petronis, A. Phenotypic differences in genetically identical organisms: the epigenetic perspective. *Hum. Mol. Genet.* **14**, R11–R18 (2005).

74. Raj, A., Rifkin, S.A., Andersen, E. & van Oudenaarden, A. Variability in gene expression underlies incomplete penetrance. *Nature* **463**, 913–918 (2010).
75. Burga, A. & Lehner, B. Beyond genotype to phenotype: why the phenotype of an individual cannot always be predicted from their genome sequence and the environment that they experience. *FEBS J.* **279**, 3765–3775 (2012).
76. Tsankova, N., Renthal, W., Kumar, A. & Nestler, E.J. Epigenetic regulation in psychiatric disorders. *Nat. Rev. Neurosci.* **8**, 355–367 (2007).
77. Jiang, Y.H., Bressler, J. & Beaudet, A.L. Epigenetics and human disease. *Annu. Rev. Genomics Hum. Genet.* **5**, 479–510 (2004).
78. Chess, A. Random and non-random monoallelic expression. *Neuropsychopharmacology* **38**, 55–61 (2013).
79. Antonellis, A. *et al.* Glycyl tRNA synthetase mutations in Charcot-Marie-Tooth disease type 2D and distal spinal muscular atrophy type V. *Am. J. Hum. Genet.* **72**, 1293–1299 (2003).
80. White, J.K. *et al.* Genome-wide generation and systematic phenotyping of knockout mice reveals new roles for many genes. *Cell* **154**, 452–464 (2013).
81. Bilgüvar, K. *et al.* Whole-exome sequencing identifies recessive WDR62 mutations in severe brain malformations. *Nature* **467**, 207–210 (2010).
This paper demonstrates a mutation associated with a broad spectrum of malformations of cortical development that have been recognized as distinct in pathogenesis.
82. McCarroll, S.A. & Hyman, S.E. Progress in the genetics of polygenic brain disorders: significant new challenges for neurobiology. *Neuron* **80**, 578–587 (2013).
83. Blair, D.R. *et al.* A nondegenerate code of deleterious variants in mendelian Loci contributes to complex disease risk. *Cell* **155**, 70–80 (2013).
This is the first study that systematically examines the correlation between Mendelian and complex diseases. By mining medical records over 110 million patients, the study identifies widespread comorbidity between Mendelian-Mendelian and Mendelian-complex disease pairs. The analysis also highlights the effect of genetic interactions in both Mendelian and complex diseases.
84. Darnell, J.C. *et al.* FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. *Cell* **146**, 247–261 (2011).
85. Fromer, M. *et al. De novo* mutations in schizophrenia implicate synaptic networks. *Nature* **506**, 179–184 (2014).
86. Petrovski, S., Wang, Q., Heinzen, E.L., Allen, A.S. & Goldstein, D.B. Genic intolerance to functional variation and the interpretation of personal genomes. *PLoS Genet.* **9**, e1003709 (2013).
This study provides the first comprehensive gene prioritization scoring system based on human population genetic data. The basic approach is to predict common functional variation using total variation in a regression framework, thereby identifying genes that are intolerant to functional genetic variation and therefore likely under purifying selection. The authors showed that intolerant genes are much more likely to cause neurodevelopmental disorders than tolerant genes.
87. Adzhubei, I.A. *et al.* A method and server for predicting damaging missense mutations. *Nat. Methods* **7**, 248–249 (2010).
88. Zaidi, S. *et al. De novo* mutations in histone-modifying genes in congenital heart disease. *Nature* **498**, 220–223 (2013).
89. Girard, S.L. *et al.* Increased exonic *de novo* mutation rate in individuals with schizophrenia. *Nat. Genet.* **43**, 860–863 (2011).
90. Cutting, G.R. Modifier genes in Mendelian disorders: the example of cystic fibrosis. *Ann. NY Acad. Sci.* **1214**, 57–69 (2010).
This paper provides an elegant review on genetic modifiers in Mendelian disorders, based on the well-studied example of cystic fibrosis.
91. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).
92. Hou, Z. *et al.* Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis. *Proc. Natl. Acad. Sci. USA* **110**, 15644–15649 (2013).
93. Spira, M.E. & Hai, A. Multi-electrode array technologies for neuroscience and cardiology. *Nat. Nanotechnol.* **8**, 83–94 (2013).
94. Venkatachalam, V. *et al.* Flash memory: photochemical imprinting of neuronal action potentials onto a microbial rhodopsin. *J. Am. Chem. Soc.* **136**, 2529–2537 (2014).