Neuropsychiatric genomics in precision medicine: diagnostics, gene discovery, and translation

Anna C. Need, PhD; David B. Goldstein, PhD

Only a few years after its development, next-generation sequencing is rapidly becoming an essential part of clinical care for patients with serious neurological conditions, especially in the diagnosis of early-onset and severe presentations. Beyond this diagnostic role, there has been an explosion in definitive gene discovery in a range of neuropsychiatric diseases. This is providing new pointers to underlying disease biology and is beginning to outline a new framework for genetic stratification of neuropsychiatric disease, with clear relevance to both individual treatment optimization and clinical trial design. Here, we outline these developments and chart the expected impact on the treatment of neurological, neurodevelopmental, and psychiatric disease.

Author affiliations: Division of Brain Sciences, Department of Medicine, Imperial College London, London, W12 0NN, UK (Anna C. Need); Institute for Genomic Medicine, Columbia University, New York, NY, 10032, USA (David B. Goldstein).

Address for correspondence: Anna C. Need, Imperial College Hammersmith Campus, 405 Burlington Danes, Du Cane Road, London, W12 0NN, UK (email: a.need@imperial.ac.uk)

Introduction

The field of human genetics is currently in a process of transformation due to technological advances in next-generation sequencing (NGS) that allow nearly comprehensive sequencing of large numbers of human genomes. This technology has emerged as by far the most successful general genetic diagnostic test, with diagnostic yields ranging from 25% to 60% in the setting of pediatric genetic disease. It has also led to a remarkable acceleration in overall gene discovery. The Centers for Mendelian Genomics (CMGs) established by the US National Institutes of Health in 2011 to perform genetic analyses across the Mendelian disease population, has discovered 956 genes, including 375 novel disease genes.[1] Beyond facilitating diagnosis and gene discovery, there are also qualitative differences in the types of mutations that can be discovered. For example, until recently, it has been very difficult to identify the cause of dominant conditions that compromise reproduction because linkage generally cannot be used. However, such conditions are now easily resolved by comparing the genomic sequences of multiple affected individuals. NGS has also led to a remarkable number of "phenotype expansions" in which already known genes are implicated in unexpected presentations.

Of particular importance is the fact that the successes in gene discovery are by no means limited to Mendelian disease. For example, a recent burden analysis implemented on a genome-wide basis implicated a new

State of the art

gene in amyotrophic lateral sclerosis (ALS).[2] Strikingly, such analyses allow a direct investigation of the underlying architecture of disease, and the cases of ALS, schizophrenia,[3] and unpublished studies of epilepsy and other conditions show that it is the very rarest variants that carry the strongest signal of risk. These results illustrate two critical features about complex human diseases. First, they are tractable through NGS ; second, there is a less sharp divide between at least some complex diseases and Mendelian diseases than had previously been assumed. Together with a now nearly comprehensive assessment of common variation through genome-wide association studies (GWAS), these results provide an emerging picture of the genetic architectures of neuropsychiatric diseases. NGS has therefore already transformed the diagnosis of rare neurological conditions and has opened the door to systematic gene discovery in more complex cases. These developments make it clear that NGS already provides concrete clinical deliverables. The next phase of research will focus on the biological modeling of disease-causing mutations in order to allow the development of personalized treatments that target the precise underlying cause of disease. The work on both rare and common forms shows that this paradigm may be applicable to both rare, severe, pediatric neurological disease and later-onset, genetically more complex neuropsychiatric illness. Here, we review these developments in the following sections: (i) genetic diagnoses of undiagnosed and unresolved diseases; (ii) common variants in complex neurological and psychiatric disorders: genome-wide association studies; (iii) rare variants in complex neurological and psychiatric disorders; and (iv) clinical impact.

Genetic diagnosis of undiagnosed and unresolved diseases

While NGS is proving invaluable to new gene discovery, a striking and largely unanticipated application is in the accurate diagnosis of patients with undiagnosed or unresolved presumed-genetic diseases. Mendelian disorders are a common cause of intellectual disability,

At present, the majority of next-generation sequencing (NGS) studies applied to complex disorders use exome sequencing and restrict analyses to loss-of-function (nonsense, frameshift, and splice) and damaging missense variants to predict the effects of missense variants. This is facilitated by the availability of large publically available data sets of exome sequence data that can be used to assess the frequency of particular variants and the likelihood of particular classes of variants appearing in specific genes. The largest publically available control cohort is the Exome Aggregation Consortium (ExAC) database (exac.broadinstitute.org), a catalog of rare coding variants comprising summary data from the exome sequencing of 60 706 unrelated individuals. A gene search in ExAC will show the position, quality, consequence, and frequency of all the variants found in that gene in the cohort, as well as metrics to indicate the gene's tolerance to damaging mutation. Individuals with severe pediatric disease have (as far as possible) been excluded from ExAC, which has facilitated the search for genes underlying severe early-onset Mendelian disorders. For complex neurological and psychiatric disorders, one has to be more careful with ExAC as the cohort is enriched for patients with stroke, schizophrenia, bipolar disorder, and Tourette syndrome. It also probably includes genetic variants that cause Mendelian disorders with onset in adulthood and controls with disease-associated variants that are not fully penetrant. Nonetheless, given that complex neuropsychiatric disorders are extremely genetically heterogeneous and that overall patients with neuropsychiatric illnesses make up a minority of the ExAC population, it is expected that disease-associated genes would have a lower burden of pathogenic variants in ExAC than they would in case cohorts and that most highly penetrant disease-associated variants would be missing or occur at a very low frequency in ExAC.

Box 1. Publicly available controls.

microcephaly, and brain malformations, and they account for a substantial fraction of neurodevelopmental disorders, seizure disorders, ataxias, and neurodegenerative disorders. There are currently approximately 7500 known monogenic Mendelian disorders in the OMIM (Online Mendelian Inheritance in Man) database (www.omim.org) and the number increases steadily each year as newly characterized disorders are added. The underlying genes have been identified for about half of these disorders.[1]

Individuals suspected to have monogenic neurodevelopmental or neurocognitive disorders are usually referred to a clinical geneticist, who will first establish that the probable cause is genetic and then attempt to

---

Chromosomal nomenclature. Copy number variants (CNVs) are usually described in terms of their chromosomal location, eg, 22q11.2 indicates the variant is on chromosome 22, located on the long arm (p represents the short arm, q the long), in the first cytogenetic band from the telomere, in the first sub-band of that band, in the second sub-sub-band of the sub-band.

Collapsing analysis/burden analysis. An emerging method for analyzing exome sequence data in a case-control framework. For each gene, cases and controls are evaluated on the basis of whether they carry a "qualifying" variant in a particular gene (defined on the basis of the variant's allele frequency and predicted function). A comparison of counts of qualifying variants in cases and controls is made for each gene in order to identify genes with a significant excess or deficit of qualifying variants.

Complex disease. A disease judged to have a more complex causation than Mendelian diseases.

Compound heterozygous. When both maternal and paternal copies of a gene are affected by a recessive genetic variant, but the variants are different from each other.

Copy number variant. A deletion or duplication of part of a chromosome, ranging in size from several kilobases to several megabases, but submicroscopic.

Exons. The parts of genes that are transcribed into messenger RNA (mRNA) and translated into protein.

Functional annotation. Information on the effect or possible effect of a genetic variant on the encoded protein, eg, a missense or nonsynonymous variant changes the protein sequence, a 3'–untranslated region (UTR) variant is located in the UTR of the messenger RNA (mRNA) after the last exon and may or may not affect the stability or translation efficiency of the mRNA.

Genome-wide association study. When hundreds of thousands of common single-nucleotide polymorphisms (SNPs) are genotyped in cases and controls to search for SNPs that have a significant ($P < 5 \times 10^{-8}$) difference in frequency between cases and controls. Genome-wide association studies are considered to simultaneously analyze all common variations in the genome.

Heritability. An estimate of the amount of the population variability of a particular trait (eg, disease susceptibility) that can be attributed to genetic variation between individuals (as opposed to environmental factors).

Homozygous variant. When both maternal and paternal copies of the gene carry the same mutation.

Loss-of-function variant (nonsense, frameshift, and splice). A genetic variant that probably truncates or alters the protein sequence to the extent that it can no longer perform its normal function.

Box 2. Glossary.

identify the causal mutation(s). Traditionally, this has been done by carefully assessing the child's physical appearance, development, intellectual function, and medical history, and by performing any necessary laboratory studies and then comparing all this information to the presentation of known Mendelian disorders. Once a provisional diagnosis has been made, the causal gene(s), if known, may be sequenced to confirm the diagnosis. This approach relies on the patient presenting with a pattern of symptoms close enough to the "standard" to be recognized as having that disorder. It results in a diagnosis in about 50% of cases, most of which are diagnosed at the patient's first visit.[4] Only about 10% of patients who do not receive a diagnosis by the second visit will go on to receive a genetic diagnosis with the traditional method.[4]

In 2012, we published a pilot study testing the use of NGS to find genetic diagnoses for patients with a disease that remains undiagnosed after the traditional diagnostic approach has failed. A key feature of this early effort was an attempt to estimate the diagnostic yield in a "typical" clinical genetics setting. Out of 12 patients in the study, we identified a genetic diagnosis for six,[5]

and with subsequent reanalysis with updated variant identification software and newly discovered disease genes, this rate has now risen to 10 out of 12 receiving a diagnosis,[6] providing evidence for the effectiveness of this approach in the clinical setting. Since then, thousands of patients have undergone NGS for genetic diagnosis, resulting in the discovery of hundreds of new disease genes and expanding the phenotype associated with many known genes.[1] Importantly, every patient receiving a disease diagnosis in this manner is considered as an individual, with his or her genome and the clinical features weighed together to identify the probable causal genes. The key point to emphasize is that when a clear genetic diagnosis is obtained for an individual patient's disease, the diagnosis affords the opportunity to target a treatment to the precise underlying cause of disease, which is the basis of precision medicine. Even when such targeted treatments are not immediately available, simply having a correct genetic diagnosis sets the stage for trialing new targeted treatments as they do become available.

Currently, the diagnostic yield of NGS is higher when a trio design is used than when only probands are

Study	Year published	Sample size*	Patient population	Diagnosis yield†
Deciphering Developmental Disorders project, UK[7]	2015	1133	Previously investigated but undiagnosed children with developmental disorders	31%
Centers for Mendelian Genomics, USA[1]	2015	8838	Known and novel Mendelian phenotypes	31%
Clinical Diagnostic Sequencing at Ambry Genetics Laboratory, USA[8]	2015	500	Diverse; primarily pediatric; 65% pediatric-onset neurological disorders	37% trios 21% singleton
Research NGS diagnostic program at Duke University Medical Center, USA[6]	2015	119	Diverse; primarily pediatric; previously investigated but undiagnosed	24%
Clinical Diagnostic Sequencing at Baylor College of Medicine, USA[9]	2015	486	Various; adult at time of referral	17.5%
Clinical Diagnostic Sequencing at the Hamad Medical Corporation, Qatar[10]	2015	149	Various undiagnosed, suspected Mendelian disorders; predominantly neurocognitive; high consanguinity	60%
Clinical Diagnostic Sequencing at Baylor College of Medicine, USA[11]	2014	2000	Diverse; primarily pediatric; 90% with neurological disorders or developmental delay	25%
Clinical Diagnostic Sequencing at the University of California, USA[12]	2014	814	Diverse; 64% pediatric, of which 53% had DD; ataxia most common in adults (26%)	31% trios 22% singleton 41% DD

Table I. Diagnostic yield in large (n>100) diagnostic next-generation sequencing studies. †Diagnostic yield refers to the percentage of patients for whom a genetic variant was identified as probably causing all or some of their symptoms *not including sequenced relatives. DD, developmental delay; NGS, next-generation sequencing.

sequenced. The trio design, which includes both biological parents, permits the exclusion of many rare variants that are observed in the parents but not in population controls, making it much easier to zero in on candidate disease variants. Of particular importance is the identification of de novo mutations, which are now known to explain the majority of early-onset, severe genetic conditions. After the trio is sequenced, bioinformatics filters are used to identify rare genotypes that are present in the affected proband and absent in the parents. This approach can be used to identify dominant de novo causal variants, as well as recessive homozygous, compound heterozygous, or X-linked genotypes. When this approach is applied to general populations of people referred for clinical diagnostic sequencing, the diagnostic yield is usually around 20% to 40% (*Table I*).[1,6-12] Several groups have begun to explore the diagnostic yield for individual neurological and neurodevelopmental disorders—predominantly seizure disorders, ataxias, and intellectual disability (*Table II*).[13-25]

Diagnostic yields vary depending on several factors, including the following: (i) the sequencing paradigm—higher yields are achieved when genes of parents or other family members have also been sequenced, and trio sequencing is particularly important for detecting de novo and compound heterozygous disease variants; (ii) the genetic model—eg, highest yields are obtained for suspected monogenic recessive disorders; and (iii) how comprehensively the patients have been screened beforehand for mutations in genes associated with their condition and whether or not they have been prescreened for pathogenic variants that would not be identified by NGS, eg, repeat expansions. Diagnostic yields will increase as capture kits, alignment, and variant identification software improve and, perhaps most importantly, as more genes are implicated in disease and in new phenotypes. However, despite the expectation that some of the cases that now have negative results will resolve as current genes of unknown significance are proven to be pathogenic, groups performing diagnostic exome sequencing have all identified cases with no good candidate variants at all. Possible explanations for these unresolved cases include the following: (i) inherited mutations outside of the exome that cause disease through an effect on gene expression or splicing; (ii) mutations in the exome that have been ignored, perhaps because they act only in combination with mutations at one or more other loci and are thus filtered out because of the appearance of the individual variants in controls; and (iii) postzygotic somatic changes.

Phenotype expansion

When a genetic diagnosis is made by NGS in patients that have been very intensely investigated beforehand, it is often because they are presenting with a known Mendelian disorder in an atypical manner. They may be showing symptoms that have not previously been

Study	Sample size	Disorder	Diagnosis yield
Helbig et al,[13] 2016	293	Epilepsy	38%
Dimassi et al,[14] 2016	10	Infantile spasm syndrome	40%
Veeramah et al,[15] 2013	10	Syndromic epilepsy	40%
Thevenon et al,[16] 2016	43	Intellectual disability or epileptic encephalopathy	33%
Keogh et al,[17] 2015	12	Late-onset cerebellar ataxia	33%
Pyle et al,[18] 2015	22	Undiagnosed ataxias	64%
Fogel et al,[19] 2014	76	Cerebellar ataxias	21%
Sawyer et al,[20] 2014	28	Pediatric-onset ataxia	46%
Ohba et al,[21] 2013	23	Cerebellar and/or vermis atrophy	39%
Tammimies et al,[22] 2015	95	Autism spectrum disorder	8%
Rump et al,[23] 2016	35	Intellectual disability and microcephaly	29%
Monroe et al,[24] 2016	17	Syndromic and nonsyndromic intellectual disability	29%
de Ligt et al,[25] 2012	100	Severe intellectual disability	16%

Table II. Diagnostic whole-genome and exome sequencing studies focused on specific neurological disorders (sample size ≥10).

associated with that disorder, and/or missing key diagnostic features. For example, in our original study,[5] we identified a pathogenic mutation in the *TCF4* (transcription factor 4) gene—which causes Pitt-Hopkins syndrome—in a girl who had neither seizures nor periods of hyperventilation, common and differentiating features of the disorder. The Finding of Rare Disease Genes (FORGE) Canada Consortium recently investigated 264 rare pediatric-onset Mendelian disorders of unknown cause and identified genes for 146 of them.[26] Of these, 95 were already known disease-associated genes, many representing expansion of the known phenotype. Similarly, of 956 genes identified by the US National Institutes of Health CMGs, 198, or approximately 1 in 5, represent phenotype expansion.[1] This is a very important development in human genetics, made possible only by the NGS diagnostic paradigm, as it is by definition impossible to significantly expand the range of clinical features associated with a gene when the diagnosis is being made based on a defined phenotype. This will not only be important for identifying genes associated with Mendelian disorders, but may also be critical to our understanding of complex disorders. It seems probable that many Mendelian diseases have a sufficiently broad phenotypic spectrum that a fraction of affected individuals will end up classified as having a complex disease. In other words, some so-called common complex diseases may in fact be, at least in part, a heterogeneous collection of genetically simpler conditions. Within neuropsychiatric diseases, epilepsy appears quite clearly to fit this category, and evidence is building for autism and schizophrenia.

Gene discovery

Approximately 30% of the 486 genetic diagnoses made by the Baylor NGS diagnostic team were in disease genes that have been discovered since 2011,[9] and 23% of the positive findings from the 500 cases reported by Ambry were within genes characterized within the past 2 years.[8] Of the 146 genes discovered by FORGE to be underlying rare Mendelian disorders, 67 had not previously been associated with human disease, 41 of which have been genetically or functionally validated. The CMG identified 375 genes not previously associated with human disease (or 128 by more conservative criteria), and the DDD project (Deciphering Developmental Disorders) and Ambry Genetics respectively identified 12 and 31 novel disease genes.[7,8] One key lesson of this rapid rate of discovery is the critical importance of regular reanalysis of clinical exomes. A further interesting finding from diagnostic sequencing is the apparent commonness of more than one pathogenic mutation. Such a combination would of course be expected to result in an undiagnosed condition because the presentation would not match any single Mendelian disease. It may be that the effects of the mutations blend to cause the major clinical features, or it may be that they have two different nonoverlapping disorders. This was observed in 7% of cases with a positive finding reported by Ambry, 5% of the Baylor pediatric patients, 7% of the Baylor adult patients, and 5% of the DDD cohort.[7-9,11]

Common variants in complex neurological and psychiatric disorders: genome-wide association studies

GWAS are designed to identify common genetic variants that individually confer a small increased risk of illness but that added together may account for a substantial fraction of the heritability of a particular condition. GWAS are used to investigate common disorders where family history does not suggest a single underlying causal gene. Large panels of single-nucleotide polymorphisms (SNPs; usually between 0.5 and 2.5 million) are used to represent the majority of common variants in the human genome, and to be declared genome-wide significant, an associated variant needs to achieve a *P* value of less than 5×10^{-8}. Because SNPs associated with complex neuropsychiatric traits may have very small effect sizes, very large numbers are often needed to have the power to identify real associations. Because GWAS use a standardized set of SNPs to represent a much larger set of common variants, they can only identify disease-associated loci, and further work is needed to track down the causal gene or variant. The identification of variants underlying GWAS associations has proven to be much harder than expected and is a major impediment to the movement from GWAS hit to clinical utility. The reasons for this remain unresolved, but there are a number of possible explanations. One possibility is that the causal variants themselves are common (like the actual associated SNPs) and have subtle functional effects on gene expression or splicing that are hard to discern. Another potential explanation is that the causal variants are rare, and distributed over a

broad genomic region, creating a "synthetic" signal of risk difficult to track down to individual causal variants. The balance between these two possible explanations is impossible to determine, and both probably contribute to the difficulty of fine mapping of GWAS signals. However, whatever the explanation, the reality is that the majority of GWAS signals remain entirely unexplained. *Table III* displays the results of the most powerful GWAS to date for a set of common neurological, neurodevelopmental, and psychiatric illnesses.[27-45] What is clear from the pattern of discovery is that any given condition seems to have a threshold sample size above which real discoveries increase regularly with increasing sample size. The exact reasons underlying the various thresholds for different diseases remain unclear in that multiple different underlying architectures are theoretically consistent with the GWAS findings. Not until a reasonable proportion of causal variants responsible for GWAS signals are tracked down will the reasons for these patterns become clear.

Here, we outline the latest GWAS findings for adult-onset neurodegenerative disorders, neurological disorders, neurodevelopmental disorders, and psychiatric disorders.

Adult-onset neurodegenerative disorders

The most common adult-onset neurodegenerative disorders—Alzheimer disease (AD), frontotemporal dementia (FTD), Parkinson disease (PD), and ALS—all have earlier-onset Mendelian forms that account for a minority of the cases (approximately 5%, 20%, 10%, and 10%, respectively) and later-onset, more genetically complex forms. Linkage analysis has led to the identification of major Mendelian disease genes, and in the case of AD, a common variant with major effect in the *ApoE4* (apolipoprotein E4) gene,[46] but the majority of cases remain unexplained, and many will not be caused by a single underlying genetic variant. GWAS have been well powered for AD and PD (*Table III*), and interesting recent findings include a signal within the *SORL1* (sortilin-related receptor, L[DLR Class] A repeats containing) gene, which had previously been identified through candidate gene studies of AD that

Disorder	Largest GWAS sample size (unrelated cases)	Number of genome-wide significant risk loci ($<5\times10^{-8}$)
Schizophrenia[38]	36 989	108
Multiple sclerosis[31]	9772	52
Parkinson disease[29]	13 708	24
Alzheimer disease[27]	25 586	20
Migraine[33]	23 285	12
Stroke[32]	24 164	8
Bipolar disorder[39]	9747	5
Epilepsy[34]	8696	3
Frontotemporal dementia[28]	3526	2
Amyotrophic lateral sclerosis[30]	2323	2
Autism[36]	5842	2
Major depressive disorder*[40]	5303	2
Anxiety disorders[43]	7016	2
Tourette syndrome[35]	1285	0
Attention-deficit/hyperactivity disorder[37]	2960	0
Anorexia[41]	5551	0
Obsessive-compulsive disorder[42]	1865	0
Posttraumatic stress disorder[44]	1708	0

Table III. Genome-wide association study results for common neurological and psychiatric illnesses. A larger study of 16 023 cases (9240 discovery and 6783 replication) failed to find any genome-wide significant hit.[45] GWAS, genome-wide association study.

investigated processing of the *APP* (amyloid precursor protein) gene, a Mendelian AD gene.[47] An NGS study investigating rare variants in AD GWAS-associated genes identified an excess of missense and loss-of-function variants in the *ABCA7* (ATP-binding cassette, sub-family A [ABC1], member 7) gene in patients with AD.[48] This has since been replicated,[49,50] although this did not account for the GWAS signal at this locus,[48] suggesting that both common and rare variants in this gene may be associated with AD. Association testing of rare variants identified by NGS and imputed onto AD GWAS data identified a rare missense variant (R47H) in the *TREM2* (triggering receptor expressed on myeloid cells 2) gene in AD.[51,52] A common GWAS signal in the human leukocyte antigen (HLA) region has been observed in AD and PD.[27,29] GWAS of ALS and FTD remain underpowered, and larger studies will probably identify more associated variants. The most common genetic cause of both ALS and FTD is an intronic repeat polymorphism in the *C9orf72* (chromosome 9 open reading frame 72) gene, accounting for approximately 40% and 25% of familial ALS and FTD respectively, as well as about 7% of sporadic ALS.[53] A combined GWAS meta-analysis of the two conditions including 4377 ALS patients and 435 FTD patients identified shared susceptibility in the *UNC13A* (unc-13 homolog A) gene as well as *C9orf72*.[54]

Neurological disorders

Epilepsy

There are many Mendelian epilepsies, both syndromic (associated with other features such as developmental delay/intellectual disability, physical abnormalities, or other clinical symptoms) and nonsyndromic. However, as is the case for other neurological and psychiatric disorders, the genetic basis of common, nonfamilial epilepsies remains largely unknown. The largest GWAS study for epilepsy combined data from 12 cohorts and searched for associations with genetic generalized epilepsy, focal epilepsy, and unclassified epilepsy.[34] They identified two GWAS signals for "all epilepsy," one centered on the *SCN1A/SCN9A* (sodium voltage-gated channel α subunit 1/sodium voltage-gated channel α subunit 9) gene and the other on the *PCDH7* (protocadherin 7) gene; a third association was seen in the generalized epilepsy analysis, close

to the *VRK2* (vaccinia-related kinase 2) and *FANCL* (Fanconi anemia complementation group L) genes. Interestingly, the epilepsy-associated allele close to *VRK2* confers protection from epilepsy,[34] but susceptibility to schizophrenia.[38,55] Mutations in *SCN1A* are seen in familial epilepsy with phenotypes ranging from severe myoclonic epilepsy in infancy to benign febrile seizures; mutations are seen less frequently in *SCN9A*, with some suggestion of interaction between variants in *SCN1A* and *SCN9A* in mouse models of epilepsy.[56] It is not yet clear whether the GWAS signal is a result of common variants with very small effect or rare variants in *SCN1A* and/or *SCN9A* causing synthetic associations.[57]

Multiple sclerosis

GWAS have had an unusually large yield of signals for the sample sizes in comparison with other neurological conditions.[31] It is tempting to infer from this a relatively more important role of common variation. This might reflect the fact that some of the variants that confer risk for multiple sclerosis (MS) have been selected in response to pathogens, whereas the variants underlying conditions such as epilepsy and schizophrenia seem generally more likely to reflect a mutation selection balance (damaging variants constantly enter the population through mutational events but are maintained at low frequency by purifying selection). The majority of MS-associated variants are related to immune function rather than neurological processes.[31]

Stroke and migraine

Both stroke and migraine have had well-powered GWAS that have had relatively few hits for the sample size. For stroke, this may be partly an ascertainment bias, as those with the most severe strokes with early mortality are less likely to be included in current GWAS. Interestingly, there is substantial overlap between common variants associated with stroke and migraine.[58]

Neurodevelopmental disorders: autism and attention-deficit/hyperactivity disorder

The role of rare, highly penetrant genetic variants in autism has been the focus of very significant genetic investigation (see later), but the GWAS signals await repli-

cation and finer mapping.[36] GWAS of attention-deficit/hyperactivity disorder (ADHD) have yet to identify a signal.

Psychiatric disorders

Psychiatric disorders are complex traits with strong environmental influences, which complicates their genetic investigation, despite high heritability. Unlike the adult-onset neurodegenerative disorders, there are no (known) Mendelian forms of psychiatric disorders; it is not clear why. Large-scale GWAS of schizophrenia,[38] depression,[40,45] bipolar disorder,[39] anorexia,[41] and anxiety[43] suggest that hundreds or even thousands of common variants of very small effect will be needed to explain a reasonable fraction of the heritability of these disorders. Schizophrenia has one of the highest heritabilities of all psychiatric disorders, and is leading the field of psychiatry in terms of systematic large-scale genetic analyses and positive findings. GWAS have identified 108 associated loci[38]; however, much of the heritability remains to be explained, and larger studies will probably identify many more associated variants.[59] Impressively, the strongest signal, in the major histocompatibility complex (MHC) locus, has been traced to its causal gene, *C4* (complement component 4), and it has been suggested that increased C4 activity in the brain of people with schizophrenia causes excessive synaptic pruning during postnatal brain development.[60] If this is supported by further work, it would represent one of very few times that a GWAS hit has informed about a specific underlying biological process.

A "mega-GWAS" of major depressive disorder showed no genome-wide association signals[45]; however, two risk loci were identified in a study focusing on Chinese women,[40] possibly due to the more homogeneous sample of females with severe, recurrent depression or possibly reflecting population-specific associations. Other disorders, such as bipolar disorder,[39] have shown some interesting patterns but without as much definitive discovery. An obvious question for the community of researchers is whether there is any utility at all in expanding GWAS sample sizes. The answer to this question depends on one's view about the primary reasons to be investigating the genetics of neuropsychiatric diseases. If the primary reason is to identify genetic leads that can elucidate the underlying molecular pathways, the answer would seem to be that there is little value in fur-

ther expanding GWAS sample sizes. It seems clear that GWAS signals are a rather poor guide to the underlying causal variants because there are now many thousands of independent GWAS signals and only a handful that have been tracked down to their underlying cause. From this perspective, it is rather hard to imagine that moving the number of independent GWAS signals across neuropsychiatric diseases from 500 to 5000 would offer a great advantage in the identification of key disrupted genes and pathways. The one counter to this might be an effort to achieve a reasonable threshold size for the most important conditions, recognizing that some leading GWAS signals have apparently resolved to causal variants. For example, variants in *C4* were discovered to underlie the strong association at the MHC locus in schizophrenia GWAS. One relatively new approach in following-up GWAS findings is to perform targeted NGS of regions surrounding GWAS hits in order to directly identify both rare and common associated variants. This has been successfully used to further explore, for example, bipolar disorder[61] and AD[50] GWAS hits.

Rare variants in complex neurological and psychiatric disorders

Since so few GWAS hits have been tracked to causal variants, we have an exceptional paucity of implicated common causal variants for neuropsychiatric illness. What we do know is that, in most cases, common causal variants underlying GWAS hits have only a very modest impact on risk. The total cumulative effect of associated common variants in GWAS also clearly leaves significant heritability unexplained for all common conditions. This makes it very difficult to apply GWAS findings at the individual level. However, in contrast to the paucity of common causal variants, a large number of distinct rare variants have been definitively implicated in neuropsychiatric disease. Rare copy number variants (CNVs) with high penetrance have already been shown to be associated with many neuropsychiatric disorders (see below), and NGS studies are underway to explore the role of rare sequence variants in these disorders. The identification of rare, highly penetrant sequence variants offers an unprecedented opportunity to learn about the etiology of neuropsychiatric conditions, as—unlike with GWAS and CNV analyses—we discover the exact base change associated with the disorder. It is much more straightforward to move from these find-

ings to biological interpretation and functional modeling of the mutations. When sequence variants are implicated in risk, the variants themselves become reagents to probe the underlying causes of disease. Thus, a critical and largely new focus in genetic studies of neuropsychiatric diseases is the systematic functional characterization of the mutations, implicated through sequencing studies, that confer risk of disease.

Copy number variants

CNVs are deletions and duplications of large sections of the genome (approximately 10 kilobases to megabases). In 2007, it was reported that de novo CNVs were strongly associated with autism,[62] and CNVs at 1q21, 17p12, and the *NRXN1* (neurexin 1) gene were reported in autism patients for the first time.[63] Other studies rapidly confirmed and extended these findings to other CNV regions[64,65] and other conditions, including intellectual disability, seizure disorders, and schizophrenia.[66-70] There are now 11 CNVs statistically associated with neuropsychiatric/neurodevelopmental disorders; these CNVs include 1q21.1, *NRXN1*, 2q37, 3q29, 7q11.23, 15q11.2, 15q13.3, 16p11.2, 16p13.1, 17q12, and 22q11.2; they vary considerably in their presentation, with most appearing at low frequencies in controls, some presenting with specific phenotypes such as obesity, and some in a severe syndromic manner with facial dysmorphism, physical abnormalities, and intellectual disability (ID).[71,72] Interestingly, 22q11 deletions have very recently been found in sporadic PD patients too.[73] With the exception of *NRXN1*, these CNVs affect multiple genes, and so they have not offered mechanistic insight into the conditions they associate with; it seems probable that in most cases the disorder is due to the combined effects of multiple genes being deleted or duplicated.

Because the phenotypic effects of these rare CNVs are unpredictable, genetic counseling for personal and familial risk of neuropsychiatric disease in cases where the associated variant has been identified before any symptoms have emerged is complicated. However, if a patient with ID, epilepsy, autism, or schizophrenia is found to have one of these variants, it is very probably relevant to the disorder, and the patient and family should be counseled accordingly. Unlike in patients with seizures, ID, or developmental delay, genetic testing is currently very unusual in psychiatry as traditionally there are no obvious genetic tests such as those that exist for Mendelian epilepsies and intellectual disability syndromes. However, we now know that approximately 2.5% of schizophrenia patients will carry one of the associated CNVs, and that many more genes will probably be associated through more powerful exome sequencing studies in the near future. Genetic testing should therefore be considered as part of routine clinical care for schizophrenia and potentially for other severe or early-onset psychiatric illnesses.[74]

Exome sequencing for gene discovery

In terms of studying multiple individuals, all classified as having a similar neuropsychiatric condition (distinct from what we refer to as the "genetic diagnosis" above, which is applied to individual cases), there are two principal discovery frameworks that have now been successfully applied; these include (i) testing for excess de novo mutations in patients; and (ii) testing for burden of rare functional variants in case-control studies.

The de novo–mutation approach uses sample sets in which parents are available and identifies and confirms de novo mutations in trios or quads. In this framework, the focus is on testing whether there is a significant excess of de novo mutations either overall or in individual genes across the cohort studies, taking account of both the total mutation rate of genes and of the number of genes that are considered (about 19 000 in most exome studies). In addition to the identification of specific genes, analyses have been performed showing that de novo mutations are differentially drawn from specific groups of genes, for example, targets of the fragile X mental retardation protein (FMRP[75]) and genes that have been shown to be intolerant to damaging mutation.[76] For diseases that strongly compromise reproduction, de novo mutations are often of particular importance, and their rarity in the genome means that it is easier to implicate disease genes when information about whether variants are de novo or not is available.

The primary analysis approach currently being used in case-control studies of whole-exome sequence data sets in neuropsychiatric diseases is to use some variant of a "collapsing" analysis. In this framework, criteria such as functional annotation (eg, missense, splice-affecting) and allele frequency are used to define "qualifying variants" (in the terminology of Cirulli et al[2]).

Once qualifying variants are defined, each gene is assessed to see if there is a qualifying variant present or not in each case and each control. The resulting counts of qualifying variants in cases and controls are then compared. If there is a significant difference after accounting for all genes in the genome, then the gene can be considered implicated in risk, assuming artifactual signals have been controlled. This framework has only been applied a handful of times so far, but it looks very probable that it will emerge as the primary discovery engine for neuropsychiatric disease genetics.

NGS for gene discovery in complex neuropsychiatric illness to date has largely focused on schizophrenia, autism, ALS, epileptic encephalopathy, stroke, and obsessive-compulsive disorder (OCD).

Schizophrenia

The first trio NGS study to look for de novo mutations in complex neuropsychiatric illness sequenced the exomes of 53 schizophrenia patients without a family history of schizophrenia, and 22 control trios. No excess of de novo variants in general was found in these patients; however, an excess of damaging de novo missense variants was reported.[77] An excess of gene-disrupting de novo variants was reported in a study of 231 trios the following year.[78] Two papers published in *Nature* in 2014, including analysis of de novo variants in 604 trios[79] and a case-control comparison of 2536 cases and 2543 controls[80] failed to identify specific associated alleles or genes, but both showed an excess of damaging variants in glutamatergic postsynaptic proteins and proteins whose messenger RNAs are targets of the FMRP. Reanalysis of the 604 trios in 2015 did not find any evidence of recessive genotypes playing a role in schizophrenia.[81] Most recently, a combined analysis of 4264 cases, 9343 controls, and 1077 trios (which included the exomes sequenced in the two 2014 papers) reported a significant excess of very rare (including de novo) gene-disrupting variants in the *SETD1A* (SET domain containing 1A) gene in schizophrenia patients, finding 10 in 5341 unrelated cases (0.19%) and none in 9343 controls.[3] This represents the first time that NGS has shown a statistically significant association between schizophrenia and a single candidate gene. SETD1A is involved in histone methylation and the association with schizophrenia further substantiates the report that common risk variants for psychiatric disorders may ag-

gregate on histone methylation pathways.[82] Like the schizophrenia-associated CNVs, *SETD1A*-disrupting mutations are present at a low frequency in controls (2 out of 45 375 Exome Aggregation Consortium (ExAC) exomes without schizophrenia) and are also seen in patients with a more severe range of developmental and physical abnormalities.

Autism spectrum disorders

Four NGS studies were published in 2012 that investigated de novo variants in autism spectrum disorders (ASDs).[75,83-85] Using either a trio or quad design, these studied showed, as for schizophrenia, that many of the disrupted genes were connected to the FMRP network and to chromatin-remodeling networks; the studies demonstrated that the majority of the mutations are paternal in origin and reported several genes with recurrent mutations. These findings were confirmed in 2014 in a reanalysis of small frameshift insertions and deletions (INDELs) in the same trios.[86] In 2013, rare inherited variants were examined for the first time: Yu et al[87] looked at 277 children with ASD from consanguineous families from the Middle East and identified recessive mutations in Mendelian disease genes in patients with familial ASD, reporting a generally milder presentation and a lack of specific diagnostic features of the diseases previously associated with those genes. Lim et al[88] searched for homozygous or compound heterozygous variants in 933 cases and 869 controls and identified in ASD patients an excess of complete knockouts in genes that generally have low rates of loss-of-function variation. Two large-scale analyses were reported in 2014: the first compared de novo mutations in affected and unaffected siblings from 2500 quads.[89] They reported that 13% of de novo missense variants and 43% of gene-disrupting variants contribute to ASD diagnoses, and that, combining CNVs and sequence variants, de novo mutations account for about 30% of all simplex cases of autism and about 45% of all female cases. A case-control analysis of 3871 cases and 9937 controls used gene-burden tests for the first time in autism to search for a statistical association of different classes of gene variants in individual genes. They reported a set of 22 genes with a false discovery rate of less than 0.05. Both papers implicate genes in synaptic formation and chromatin-remodeling pathways.

Overall, considering genes, regions, and pathways implicated through GWAS, CNV, and NGS, there is a great deal of overlap between the findings of autism and schizophrenia genomics, which would not necessarily have been predicted from their clinical presentation and treatment. It is clear for both disorders that many hundreds of genes are probably enriched for rare, damaging variants in cases and that the genes that have been identified already represent only the tip of the iceberg. Very large sample sizes will be needed to be able to begin to systematically characterize the genetic basis of these disorders, and it is not yet clear if they will be distinguishable from each other or even from seizure disorders and severe developmental disorders at the genetic level.

Epileptic encephalopathies

Epileptic encephalopathies are genetically heterogeneous, severe epilepsy syndromes. A study of 264 trios in 2012 implicated two new diseases genes—*ALG13* (UDP-N-acetylglucosamine transferase subunit ALG13 homolog) and *GABRB3* (γ-aminobutyric acid (GABA) A receptor, β3).[90] A more highly powered follow-up in 2014 analyzed 356 trios (including those previously analyzed) and, unlike in schizophrenia and ASD, showed an overall excess of exonic de novo mutations in cases compared with control trios, but in common with neuropsychiatric illnesses, also showed an excess of damaging de novo variants in synaptic genes. They also provided statistical evidence for de novo missense mutations in the synaptic gene *DNM1* (dynamin 1) in epileptic encephalopathy; *DNM1* was implicated in their previous study.[91] Several other genes had recurrent damaging mutations and will probably show statistical significance in larger cohorts.

Amyotrophic lateral sclerosis

In March 2015, Cirulli and colleagues published an exome-sequencing study of 2900 ALS patients and 6400 controls. They performed gene-wise collapsing association testing and found nonbenign coding mutations in the *TNK1* (TANK-binding kinase 1) gene in 1.1% of patients and 0.2% of controls, thus establishing *TBK1* as an ALS gene. This finding was confirmed soon afterwards by Axel Freischmidt and colleagues in a study of 252 familial ALS cases.

Stroke

Part of the National Heart, Lung, and Blood Institute (NHLBI) Exome Sequencing Project (ESP), Auer et al reported last year on a discovery sample of 365 cases and 809 controls, followed-up by sequencing and targeted genotyping of 1800 cases and 4800 controls[91]; they identified two novel stroke-associated genes: *PDE-4DIP* (phosphodiesterase 4D interacting protein) and *ACOT4* (acyl-coenzyme A thioesterase 4). Analysis of previously implicated candidate genes provided further evidence for a role for *ABCA1* (ATP binding cassette subfamily A member 1) and *ZFHX3* (zinc finger homeobox 3) in stroke susceptibility. These findings implicate novel molecular pathways in ischemic stroke.

Obsessive-compulsive disorder

OCD is the most recent neuropsychiatric disorder to have de novo mutations investigated with exome sequencing of trios. The small pilot study of 17 sporadic cases reported a slight excess of de novo mutations in patients compared with unaffected siblings of ASD patients and suggested that larger studies may begin to implicate associated genes for OCD as they have for other neuropsychiatric conditions.[92]

Whole-genome sequencing for gene discovery

At present, there remain very few studies of whole-genome sequence data in neuropsychiatric illness. Whole-genome sequencing (WGS) is more expensive to generate and store and more difficult to analyze. We also lack large publically available databases like ExAC for noncoding regions. However, WGS not only provides data on noncoding regions, which are probably important in complex disease, but can also provide sequence data on coding regions that are currently missed by exome sequencing due to high GC content. A recent autism study used WGS for a set of 53 ASD patients, most of whom had been previously investigated by exome sequencing and CNV analysis without causal genes being identified.[93] The authors investigated rare genetic variants in noncoding regulatory regions and reported a significant excess of mutations in cases in DNase I–hypersensitive sites. They also reported the generation of sequence data for an additional 1854 genes that are commonly missed by exome sequencing and the reli-

able detection of small CNVs affecting single exons that cannot be identified with exome sequencing. As the cost of WGS decreases, it seems probable that high-coverage WGS will gradually overtake exome sequencing as the primary gene discovery method. However, while the collapsing framework is already working very well for genes, it is clear that it will prove a significant challenge to implement on a genome-wide basis. While expansion of the rules for qualifying variants should easily permit inclusion of other classes of variants very clearly associated with genes (eg, synonymous variants), implementation outside of genic regions poses significant challenges for the fundamental reason that it is unclear how to define the regions within which variants should be "collapsed." This is a key area of necessary development. One direction that will probably prove profitable for defining such regions is to develop genome-wide annotations that integrate information about regulatory potential with information about regions of the genome under purifying selection.

Clinical impact

Although exome-sequencing approaches remain quite recent, there is already a growing body of evidence of clear clinical relevance. The most immediate clinical relevance is in providing an individual diagnosis that would not have otherwise been obtained and that directly determines the appropriate treatment. A particularly clear example of this is a rare neurological condition called Brown-Vialetto-Van Laere (BVVL) syndrome, which results from loss-of-function mutations in a riboflavin transporter. Petrovski et al[94] report the analysis of the exome of an 18-month-old girl with an undiagnosed progressive neurological condition. Once her sequence data were available, the previously unsuspected diagnosis of BVVL was immediately clear, and a probably life-saving treatment of riboflavin supplementation was initiated.[95] Although this is one of the more extreme examples of personalized therapeutic benefits, the proportion of cases that derive some direct benefit has been estimated at between 19% and 49%, depending on how liberal the definition of therapeutic consequence.[96]

Beyond direct changes in clinical management of individual patients, exome sequencing has already helped to identify disease-causing pathways that are now the focus of drug development efforts in those conditions.

For example, the exome-sequencing work identifying *TBK1* highlights autophagy in ALS,[2] and sequencing of patients with early-onset epilepsies has emphasized the importance of synaptic vesicle trafficking and other biological processes in epilepsy.[97] These results not only identify pathways as therapeutic targets for intervention, but also begin to develop a new molecular stratification of these diseases, which will have relevance to both drug development and prognosis.

It is worth emphasizing explicitly that in order to use a genetic diagnosis to provide an effective treatment, the key requirement is to understand how the responsible mutations cause disease. Once this is determined, it becomes possible to screen for compounds that act specifically against the underlying cause of disease. As one clear illustration of this paradigm, mutations in the *KCNT1* (potassium sodium-activated channel subfamily T member 1) gene have been shown to cause a number of epilepsies and appear to do so through a straightforward increase in activity of the channel.[98] This suggests the possibility that inhibitors of the channel could provide an effective targeted therapy, and quinidine has indeed already been tried in *KCNT1*-positive epilepsies, so far with limited benefit presumably due to dose-limiting toxicities.[99] However, it does appear probable that a more selective and potent inhibitor will eventually be found, opening the way to more effective targeted treatments for this form of epilepsy. For this experience to become more common, it is essential that we develop modeling systems that can be used for any kind of mutant protein because many genes that cause neuropsychiatric diseases encode proteins that cannot be modeled well with traditional electrophysiology. One exciting paradigm in this context is the use of multielectrode arrays and other monitoring approaches to characterize the behavior of in vitro neuronal networks engineered to carry the mutations of interest. If phenotypes can be observed in these mutant neuronal networks, it becomes possible to screen for candidate molecules that revert those phenotypes, which in turn become candidate targeted treatments. Although this work is in its infancy and there are as yet few successes, there is growing enthusiasm to try this kind of approach in a range of different neuropsychiatric approaches, and there is already a compound being tested in ALS that was identified based on an excitability profile in engineered neurons monitored with multielectrode arrays.[100] Despite the justified enthusiasm for precision

State of the art

medicine based on examples such as these, it is important to recognize that in the vast majority of cases where a genetic diagnosis is obtained, there will currently not be any available treatment to choose on the basis of the diagnosis. The reality of this is clear from a simple consideration that among the nearly 5000 genes now implicated in Mendelian disease, only a tiny fraction have clear treatments associated with them. Despite this sobering reality, it seems reasonable to hope that advances in the biological modeling of disease mutations will steadily increase the proportion of genetic diagnoses that are associated with targeted treatments. As this work progresses, the proportion of genetic diagnoses associated with changes in management will only grow. ❑

REFERENCES

1. Chong JX, Buckingham KJ, Jhangiani SN, et al. The genetic basis of Mendelian phenotypes: discoveries, challenges, and opportunities. *Am J Hum Genet.* 2015;97(2):199-215.
2. Cirulli ET, Lasseigne BN, Petrovski S, et al. Exome sequencing in amyotrophic lateral sclerosis identifies risk genes and pathways. *Science.* 2015;347(6229):1436-1441.
3. Singh T, Kurki MI, Curtis D, et al. Rare loss-of-function variants in *SET-D1A* are associated with schizophrenia and developmental disorders. *Nat Neurosci.* 2016;19(4):571-577.
4. Shashi V, McConkie-Rosell A, Rosell B, et al. The utility of the traditional medical genetics diagnostic evaluation in the context of next-generation sequencing for undiagnosed genetic disorders. *Genet Med.* 2014;16(2):176-182.
5. Need AC, Shashi V, Hitomi Y, et al. Clinical application of exome sequencing in undiagnosed genetic conditions. *J Med Genet.* 2012;49(6):353-361.
6. Zhu X, Petrovski S, Xie P, et al. Whole-exome sequencing in undiagnosed genetic diseases: interpreting 119 trios. *Genet Med.* 2015;17(10):774-781.
7. Deciphering Developmental Disorders Study. Large-scale discovery of novel genetic causes of developmental disorders. *Nature.* 2015;519(7542):223-228.
8. Farwell KD, Shahmirzadi L, El-Khechen D, et al. Enhanced utility of family-centered diagnostic exome sequencing with inheritance model-based analysis: results from 500 unselected families with undiagnosed genetic conditions. *Genet Med.* 2015;17(7):578-586.
9. Posey JE, Rosenfeld JA, James RA, et al. Molecular diagnostic experience of whole-exome sequencing in adult patients. *Genet Med.* 2015 Dec 3. Epub ahead of print. doi:10.1038/gim.2015.142.
10. Yavarna T, Al-Dewik N, Al-Mureikhi M, et al. High diagnostic yield of clinical exome sequencing in Middle Eastern patients with Mendelian disorders. *Hum Genet.* 2015;134(9):967-980.
11. Yang Y, Muzny DM, Xia F, et al. Molecular findings among patients referred for clinical whole-exome sequencing. *JAMA.* 2014;312(18):1870-1879.
12. Lee H, Deignan JL, Dorrani N, et al. Clinical exome sequencing for genetic identification of rare Mendelian disorders. *JAMA.* 2014;312(18):1880-1887.
13. Helbig KL, Farwell Hagman KD, Shinde DN, et al. Diagnostic exome sequencing provides a molecular diagnosis for a significant proportion of patients with epilepsy. *Genet Med.* 2016 Jan 21. Epub ahead of print. doi:10.1038/gim.2015.186.
14. Dimassi S, Labalme A, Ville D, et al. Whole-exome sequencing improves the diagnosis yield in sporadic infantile spasm syndrome. *Clin Genet.* 2016;89(2):198-204.
15. Veeramah KR, Johnstone L, Karafet TM, et al. Exome sequencing reveals new causal mutations in children with epileptic encephalopathies. *Epilepsia.* 2013;54(7):1270-1281.
16. Thevenon J, Duffourd Y, Masurel-Paulet A, et al. Diagnostic odyssey in severe neurodevelopmental disorders: toward clinical whole-exome sequencing as a first-line diagnostic test. *Clin Genet.* 2016;89(6):700-707.
17. Keogh MJ, Steele H, Douroudis K, et al. Frequency of rare recessive mutations in unexplained late onset cerebellar ataxia. *J Neurol.* 2015;262(8):1822-1827.
18. Pyle A, Smertenko T, Bargiela D, et al. Exome sequencing in undiagnosed inherited and sporadic ataxias. *Brain.* 2015;138(pt 2):276-283.
19. Fogel BL, Lee H, Deignan JL, et al. Exome sequencing in the clinical diagnosis of sporadic or familial cerebellar ataxia. *JAMA Neurol.* 2014;71(10):1237-1246.
20. Sawyer SL, Schwartzentruber J, Beaulieu CL, et al. Exome sequencing as a diagnostic tool for pediatric-onset ataxia. *Hum Mutat.* 2014;35(1):45-49.
21. Ohba C, Osaka H, Iai M, et al. Diagnostic utility of whole exome sequencing in patients showing cerebellar and/or vermis atrophy in childhood. *Neurogenetics.* 2013;14(3-4):225-232.
22. Tammimies K, Marshall CR, Walker S, et al. Molecular diagnostic yield of chromosomal microarray analysis and whole-exome sequencing in children with autism spectrum disorder. *JAMA.* 2015;314(9):895-903.
23. Rump P, Jazayeri O, van Dijk-Bos KK, et al. Whole-exome sequencing is a powerful approach for establishing the etiological diagnosis in patients with intellectual disability and microcephaly. *BMC Med Genomics.* 2016;9:7.
24. Monroe GR, Frederix GW, Savelberg SM, et al. Effectiveness of whole-exome sequencing and costs of the traditional diagnostic trajectory in children with intellectual disability. *Genet Med.* 2016 Feb 4. Epub ahead of print. doi:10.1038/gim.2015.200.
25. de Ligt J, Willemsen MH, van Bon BW, et al. Diagnostic exome sequencing in persons with severe intellectual disability. *N Engl J Med.* 2012;367(20):1921-1929.
26. Beaulieu CL, Majewski J, Schwartzentruber J, et al. FORGE Canada Consortium: outcomes of a 2-year national rare-disease gene-discovery project. *Am J Hum Genet.* 2014;94(6):809-817.
27. Lambert JC, Ibrahim-Verbaas CA, Harold D, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet*. 2013;45(12):1452-1458.
28. Ferrari R, Hernandez DG, Nalls MA, et al. Frontotemporal dementia and its subtypes: a genome-wide association study. *Lancet Neurol*. 2014;13(7):686-699.
29. Nalls MA, Pankratz N, Lill CM, et al. Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. *Nat Genet*. 2014;46(9):989-993.
30. van Es MA, Veldink JH, Saris CG, et al. Genome-wide association study identifies 19p13.3 (*UNC13A*) and 9p21.2 as susceptibility loci for sporadic amyotrophic lateral sclerosis. *Nat Genet*. 2009;41(10):1083-1087.
31. International Multiple Sclerosis Genetics Consortium; Wellcome Trust Case Control Consortium 2; Sawcer S, Hallenthal G, Prirnen M, et al. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature.* 2011;476(7359):214-219.
32. Neurology Working Group of the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium; Stroke Genetics Network (SiGN); International Stroke Genetics Consortium (ISGC). Identification of additional risk loci for stroke and small vessel disease: a meta-analysis of genome-wide association studies. *Lancet Neurol.* 2016;15(7):695-707.

Genómica neuropsiquiátrica en la medicina de precisión: diagnóstico, descubrimiento de genes y traducción clínica

Solo unos pocos años después de su descubrimiento, la secuenciación de nueva generación se está constituyendo rápidamente en una parte esencial del cuidado clínico de los pacientes con patologías neurológicas serias, especialmente en el diagnóstico de cuadros de aparición precoz y grave. Más allá de este papel diagnóstico, ha constituido una explosión en el descubrimiento de genes definitivos en una amplia variedad de enfermedades neuropsiquiátricas. Esto está aportando nuevos indicadores para sustentar la biología de la enfermedad y se está empezando a perfilar un nuevo marco de referencia para la estratificación genética de las enfermedades neuropsiquiátricas, con especial relevancia tanto en la optimización del tratamiento individual como en el diseño de ensayos clínicos. En este artículo se resumen estos desarrollos y se proyecta el impacto esperado en el tratamiento de enfermedades neurológicas, del neurodesarrollo y psiquiátricas.

Génomique neuropsychiatrique dans la médecine de précision : diagnostic, découverte génique et traduction clinique

Quelques années seulement après son développement, la nouvelle génération de séquençage est rapidement devenue une part essentielle des traitements pour les patients atteints de maladies neurologiques graves, surtout dans le diagnostic des tableaux sévères et d'installation précoce. Au-delà de ce rôle diagnostique, de très nombreux gènes ont été formellement impliqués dans plusieurs maladies neuropsychiatriques. L'explosion de ces découvertes a fourni de nouvelles indications sur les mécanismes sous-tendant ces maladies et a permis d'établir une ébauche de leur stratification génétique, ce qui est très utile à la fois pour optimiser les traitements individuels et pour planifier les études cliniques. Dans cet article, nous soulignons ces évolutions et retraçons l'impact attendu sur le traitement de la maladie neurologique, neurodéveloppementale et psychiatrique.

33. Anttila V, Winsvold BS, Gormley P, et al. Genome-wide meta-analysis identifies new susceptibility loci for migraine. *Nat Genet*. 2013;45(8):912-917.
34. International League Against Epilepsy Consortium on Complex Epilepsies. Genetic determinants of common epilepsies: a meta-analysis of genome-wide association studies. *Lancet Neurol*. 2014;13(9):893-903.
35. Scharf JM, Yu D, Mathews CA, et al. Genome-wide association study of Tourette's syndrome. *Mol Psychiatry*. 2013;18(6):721-728.
36. Xia K, Guo H, Hu Z, et al. Common genetic variants on 1p13.2 associate with risk of autism. *Mol Psychiatry*. 2014;19(11):1212-1219.
37. Neale BM, Medland SE, Ripke S, et al. Meta-analysis of genome-wide association studies of attention-deficit/hyperactivity disorder. *J Am Acad Child Adolesc Psychiatry*. 2010;49(9):884-897.
38. Schizophrenia Working Group of the Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature*. 2014;511(7510):421-427.
39. Muhleisen TW, Leber M, Schulze TG, et al. Genome-wide association study reveals two new risk loci for bipolar disorder. *Nat Commun*. 2014;5:3339.
40. CONVERGE Consortium. Sparse whole-genome sequencing identifies two loci for major depressive disorder. *Nature*. 2015;523(7562):588-591.
41. Boraska V, Franklin CS, Floyd JA, et al. A genome-wide association study of anorexia nervosa. *Mol Psychiatry*. 2014;19(10):1085-1094.
42. Stewart SE, Yu D, Scharf JM, et al. Genome-wide association study of obsessive-compulsive disorder. *Mol Psychiatry*. 2013;18(7):788-798.
43. Otowa T, Hek K, Lee M, et al. Meta-analysis of genome-wide association studies of anxiety disorders. *Mol Psychiatry*. 2016 Feb 9. Epub ahead of print. doi:10.1038/mp.2016.11.
44. Ashley-Koch AE, Garrett ME, Gibson J, et al. Genome-wide association study of posttraumatic stress disorder in a cohort of Iraq-Afghanistan era veterans. *J Affect Disord*. 2015;184:225-234.
45. Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium; Ripke S, Wray NR, Lewis CM, et al. A mega-analysis of genome-wide association studies for major depressive disorder. *Mol Psychiatry*. 2013;18(4):497-511.

46. Strittmatter WJ, Saunders AM, Schmechel D, et al. Apolipoprotein E: high-avidity binding to β-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer disease. *Proc Natl Acad Sci U S A*. 1993;90(5):1977-1981.
47. Rogaeva E, Meng Y, Lee JH, et al. The neuronal sortilin-related receptor SORL1 is genetically associated with Alzheimer disease. *Nat Genet*. 2007;39(2):168-177.
48. Steinberg S, Stefansson H, Jonsson T, et al. Loss-of-function variants in *ABCA7* confer risk of Alzheimer's disease. *Nat Genet*. 2015;47(5):445-447.
49. Le Guennec K, Nicolas G, Quenez O, et al. *ABCA7* rare variants and Alzheimer disease risk. *Neurology*. 2016;86(23):2134-2137.
50. Cuyvers E, De Roeck A, Van den Bossche T, et al. Mutations in *ABCA7* in a Belgian cohort of Alzheimer's disease patients: a targeted resequencing study. *Lancet Neurol*. 2015;14(8):814-822.
51. Jonsson T, Stefansson H, Steinberg S, et al. Variant of *TREM2* associated with the risk of Alzheimer's disease. *N Engl J Med*. 2013;368(2):107-116.
52. Guerreiro R, Wojtas A, Bras J, et al. *TREM2* variants in Alzheimer's disease. *N Engl J Med*. 2013;368(2):117-127.
53. Majounie E, Renton AE, Mok K, et al. Frequency of the *C9orf72* hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. *Lancet Neurol*. 2012;11(4):323-330.
54. Diekstra FP, Van Deerlin VM, van Swieten JC, et al. *C9orf72* and *UNC13A* are shared risk loci for amyotrophic lateral sclerosis and frontotemporal dementia: a genome-wide meta-analysis. *Ann Neurol*. 2014;76(1):120-133.
55. Steinberg S, de Jong S; Irish Schizophrenia Genomics Consortium; Andreassen OA, et al. Common variants at VRK2 and TCF4 conferring risk of schizophrenia. *Hum Mol Genet*. 2011;20(20):4076-4081.
56. Meisler MH, O'Brien JE, Sharkey LM. Sodium channel gene family: epilepsy mutations, gene interactions and modifier effects. *J Physiol*. 2010;588(pt 11):1841-1848.
57. Dickson SP, Wang K, Krantz I, Hakonarson H, Goldstein DB. Rare variants create synthetic genome-wide associations. *PLoS Biol*. 2010;8(1):e1000294.

State of the art

58. Malik R, Freilinger T, Winsvold BS, et al. Shared genetic basis for migraine and ischemic stroke: a genome-wide analysis of common variants. *Neurology*. 2015;84(21):2132-2145.

59. Ripke S, O'Dushlaine C, Chambert K, et al. Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nat Genet*. 2013;45(10):1150-1159.

60. Sekar A, Bialas AR, de Rivera H, et al. Schizophrenia risk from complex variation of complement component 4. *Nature*. 2016;530(7589):177-183.

61. Ament SA, Szelinger S, Glusman G, et al. Rare variants in neuronal excitability genes influence risk for bipolar disorder. *Proc Natl Acad Sci U S A*. 2015;112(11):3576-3581.

62. Sebat J, Lakshmi B, Malhotra D, et al. Strong association of de novo copy number mutations with autism. *Science*. 2007;316(5823):445-449.

63. Autism Genome Project Consortium; Szatmari P, Paterson AD, Zwaigenbaum L, et al. Mapping autism risk loci using genetic linkage and chromosomal rearrangements. *Nat Genet*. 2007;39(3):319-328.

64. Marshall CR, Noor A, Vincent JB, et al. Structural variation of chromosomes in autism spectrum disorder. *Am J Hum Genet*. 2008;82(2):477-488.

65. Christian SL, Brune CW, Sudi J, et al. Novel submicroscopic chromosomal abnormalities detected in autism spectrum disorder. *Biol Psychiatry*. 2008;63(12):1111-1117.

66. International Schizophrenia Consortium. Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature*. 2008;455(7210):237-241.

67. Stefansson H, Rujescu D, Cichon S, et al. Large recurrent microdeletions associated with schizophrenia. *Nature*. 2008;455(7210):232-236.

68. Sharp AJ, Mefford HC, Li K, et al. A recurrent 15q13.3 microdeletion syndrome associated with mental retardation and seizures. *Nat Genet*. 2008;40(3):322-328.

69. Brunetti-Pierri N, Berg JS, Scaglia F, et al. Recurrent reciprocal 1q21.1 deletions and duplications associated with microcephaly or macrocephaly and developmental and behavioral abnormalities. *Nat Genet*. 2008;40(12):1466-1471.

70. de Kovel CG, Trucks H, Helbig I, et al. Recurrent microdeletions at 15q11.2 and 16p13.11 predispose to idiopathic generalized epilepsies. *Brain*. 2010;133(pt 1):23-32.

71. Grayton HM, Fernandes C, Rujescu D, Collier DA. Copy number variations in neurodevelopmental disorders. *Prog Neurobiol*. 2012;99(1):81-91.

72. Torres F, Barbosa M, Maciel P. Recurrent copy number variations as risk factors for neurodevelopmental disorders: critical overview and analysis of clinical implications. *J Med Genet*. 2016;53(2):73-90.

73. Mok KY, Sheerin U, Simon-Sanchez J, et al. Deletions at 22q11.2 in idiopathic Parkinson's disease: a combined analysis of genome-wide association data. *Lancet Neurol*. 2016;15(6):585-596.

74. Brand H, Pillalamarri V, Collins RL, et al. Cryptic and complex chromosomal aberrations in early-onset neuropsychiatric disorders. *Am J Hum Genet*. 2014;95(4):454-461.

75. Iossifov I, Ronemus M, Levy D, et al. De novo gene disruptions in children on the autistic spectrum. *Neuron*. 2012;74(2):285-299.

76. Petrovski S, Wang Q, Heinzen EL, Allen AS, Goldstein DB. Genic intolerance to functional variation and the interpretation of personal genomes. *PLoS Genet*. 2013;9(8):e1003709.

77. Xu B, Roos JL, Dexheimer P, et al. Exome sequencing supports a de novo mutational paradigm for schizophrenia. *Nat Genet*. 2011;43(9):864-868.

78. Xu B, Ionita-Laza I, Roos JL, et al. De novo gene mutations highlight patterns of genetic and neural complexity in schizophrenia. *Nat Genet*. 2012;44(12):1365-1369.

79. Fromer M, Pocklington AJ, Kavanagh DH, et al. De novo mutations in schizophrenia implicate synaptic networks. *Nature*. 2014;506(7487):179-184.

80. Purcell SM, Moran JL, Fromer M, et al. A polygenic burden of rare disruptive mutations in schizophrenia. *Nature*. 2014;506(7487):185-190.

81. Rees E, Kirov G, Walters JT, et al. Analysis of exome sequence in 604 trios for recessive genotypes in schizophrenia. *Transl Psychiatry*. 2015;5:e607.

82. Network and Pathway Analysis Subgroup of Psychiatric Genomics Consortium. Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. *Nat Neurosci*. 2015;18(2):199-209.

83. Sanders SJ, Murtha MT, Gupta AR, et al. De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature*. 2012;485(7397):237-241.

84. O'Roak BJ, Vives L, Girirajan S, et al. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature*. 2012;485(7397):246-250.

85. Neale BM, Kou Y, Liu L, et al. Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature*. 2012;485(7397):242-245.

86. Dong S, Walker MF, Carriero NJ, et al. De novo insertions and deletions of predominantly paternal origin are associated with autism spectrum disorder. *Cell Rep*. 2014;9(1):16-23.

87. Yu TW, Chahrour MH, Coulter ME, et al. Using whole-exome sequencing to identify inherited causes of autism. *Neuron*. 2013;77(2):259-273.

88. Lim ET, Raychaudhuri S, Sanders SJ, et al. Rare complete knockouts in humans: population distribution and significant role in autism spectrum disorders. *Neuron*. 2013;77(2):235-242.

89. Iossifov I, O'Roak BJ, Sanders SJ, et al. The contribution of de novo coding mutations to autism spectrum disorder. *Nature*. 2014;515(7526):216-221.

90. Epi4K Consortium; Epilepsy Phenome/Genome Project; Allen AS, Berkovic SF, Cossette P, et al. De novo mutations in epileptic encephalopathies. *Nature*. 2013;501(7466):217-221.

91. EuroEPINOMICS-RES Consortium; Epilepsy Phenome/Genome Project; Epi4K Consortium. De novo mutations in synaptic transmission genes including *DNM1* cause epileptic encephalopathies. *Am J Hum Genet*. 2014;95(4):360-370.

92. Cappi C, Brentani H, Lima L, et al. Whole-exome sequencing in obsessive-compulsive disorder identifies rare mutations in immunological and neurodevelopmental pathways. *Transl Psychiatry*. 2016;6:e764.

93. Turner TN, Hormozdiari F, Duyzend MH, et al. Genome sequencing of autism-affected families reveals disruption of putative noncoding regulatory DNA. *Am J Hum Genet*. 2016;98(1):58-74.

94. Petrovski S, Shashi V, Petrou S, et al. Exome sequencing results in successful riboflavin treatment of a rapidly progressive neurological condition. *Cold Spring Harb Mol Case Stud*. 2015;1(1):a000257.

95. Shashi V, Petrovski S, Schoch K, et al. Sustained therapeutic response to riboflavin in a child with a progressive neurological condition, diagnosed by whole-exome sequencing. *Cold Spring Harb Mol Case Stud*. 2015;1(1):a000265.

96. Soden SE, Saunders CJ, Willig LK, et al. Effectiveness of exome and genome sequencing guided by acuity of illness for diagnosis of neurodevelopmental disorders. *Sci Transl Med*. 2014;6(265):265ra168.

97. Dhindsa RS, Bradrick SS, Yao X, et al. Epileptic encephalopathy-causing mutations in *DNM1* impair synaptic vesicle endocytosis. *Neurol Genet*. 2015;1(1):e4.

98. Milligan CJ, Li M, Gazina EV, et al. *KCNT1* gain of function in 2 epilepsy phenotypes is reversed by quinidine. *Ann Neurol*. 2014;75(4):581-590.

99. Mikati MA, Jiang YH, Carboni M, et al. Quinidine in the treatment of KCNT1-positive epilepsies. *Ann Neurol*. 2015;78(6):995-999.

100. Wainger BJ, Kiskinis E, Mellin C, et al. Intrinsic membrane hyperexcitability of amyotrophic lateral sclerosis patient-derived motor neurons. *Cell Rep*. 2014;7(1):1-11.