Introduction

Many papers on the genetics of common human diseases start with the following statement: 'Disease X is a complex, multifactorial disorder'. This rubric has been applied to schizophrenia, autism, depression, asthma, epilepsy, diabetes, rheumatoid arthritis, hypertension, coronary artery disease, obesity, Crohn's disease, Alzheimer's and Parkinson's disease, multiple sclerosis and probably hundreds of other conditions - even dandruff! But what does it mean? It means that the disease is influenced by multiple genetic and environmental factors. Diseases may earn this label if they are clearly heritable and also influenced by environmental factors (as in the case of diabetes), or if the inheritance of genetic liability is not sufficient to predict whether a person will actually develop the disease; that is, there is some probabilistic element to the emergence of the disease state itself (as in many inherited cancers or psychiatric disorders).

Quite apart from whether environmental or stochastic factors are at play, the terms complex and multifactorial are also commonly used to describe the architecture of just the genetic component of disease liability. In these cases, these terms are usually at least implicitly equated with the trait being polygenic; in fact, 'complex', 'multifactorial' and 'polygenic' are commonly used as synonyms. It is important here to make a distinction in how the term polygenic is used: the implication is that the disorder arises in each individual due to the combined effects of a large number of genetic variants [1]. This definition is distinct from a model of genetic heterogeneity, in which many different variants are involved across the population, but where each case is caused by a single variant (or a few variants).

Arguments that the inheritance of a disorder is polygenic usually derive from the observation that, while the disorder might aggregate in families, it does not tend to segregate in ways that are consistent with simple Mendelian inheritance. This is indeed true for the disorders referred to above, for which risk of disease is increased if an individual has a relative with the disease (and increased more, the closer the relative) but where sporadic cases are also common, sometimes forming the majority of cases. The precise values of the relative risks to family members of different degrees of relatedness can be fed into mathematical models of genetic architecture, and are, in many cases, consistent with polygenic inheritance (for example, for schizophrenia [2-4]). In some cases, these kinds of analyses have even been taken as proof that Mendelian inheritance with genetic heterogeneity can be rejected unequivocally as a model of the genetic architecture of the disorder (for example, [3-6]).

The conclusion that many common disorders are polygenic and caused by the aggregate effects in individuals of many common variants [1], along with the development of the HapMap Project [7], laid the foundation for genome-wide association studies (GWAS) designed to identify loci that harbor such common variants [8-10]. Unfortunately, this conclusion is based entirely on circular logic and a number of unfounded assumptions. The most fundamental of these is that the disorder in question represents a single, biologically valid category.

When is a disorder not a disorder?

Analyses of the segregation patterns of a ‘disorder’ across a population necessarily assume that the disorder in question is a valid category; otherwise, there would be no point in lumping all cases together and calculating things such as heritability and relative risks across the entire population. If the clinical diagnosis of a specific disorder is based on superficial criteria, then this assumption is unlikely to hold.

For example, ‘blindness’ is not a very informative diagnosis - genetic forms can be caused by cataracts, corneal defects, optic nerve atrophy and various forms of photoreceptor degeneration, such as retinitis pigmentosa.
over 100 different loci have been found [16-19]. For such as autism and schizophrenia, where mutations in complex conditions.

degree of genetic heterogeneity underlying common, genes in different people clearly demonstrates a very high individually causal variants in a large number of different disorders. By contrast, the identification of rare, individually causal variants in a large number of different disorders.

vidually causal variants in a large number of different disorders. By contrast, the identification of rare variants conferring disease risk.

The theoretical foundation of genome-wide association studies

GWAS are founded on the polygenic model of disease liability, which itself arises from an assertion of breathtaking audacity by the godfather of quantitative genetics, DS Falconer. In an attempt to demonstrate the relevance of quantitative genetics to the study of human disease, Falconer, based on work of others before him (for example, [24]), came up with a nifty solution [25]. Even though disease states are typically all-or-nothing, and even though the actual risk of disease is clearly very discontinuously distributed in the population (being dramatically higher in relatives of affected people, for example), he claimed that it was reasonable to assume that there was something called the underlying liability to the disorder that was actually continuously distributed. This could be converted to a discontinuous distribution by further assuming that only individuals whose burden of genetic variants passed an imagined threshold actually got the disease. To transform discontinuous incidence data (that is, mean rates of disease in various groups, such as people with different levels of genetic relatedness to affected individuals) into mean liability on a continuous scale, it was necessary to further assume that...
this liability was normally distributed in the population. The corollary is that liability is affected by many genetic variants, each having a small effect [1,26].

This declaration meant that the statistical techniques developed in animal breeding could supposedly be applied legitimately to the study of human disease. Unfortunately, there is no basis for the underlying assumption of normally distributed liability, nor for the invoked threshold of genetic burden (Box 1; Figure 1).

Nevertheless, GWAS have gone ahead on a very large scale for many complex disorders and have produced statistically significant findings. Do these findings validate the assumptions I have claimed are flawed? They do not, at least not necessarily.

**What have we learned about complex disorders from GWAS?**

GWAS follow a simple design: compare allele frequencies for hundreds of thousands of common variants spread across the genome between large samples of disease cases and controls [10]. If a variant predisposes to the disease, even slightly, this should be apparent as an increased allelic frequency in cases versus controls, given a large enough sample size. This design has been applied to many different complex disorders, with varying degrees of success.

It is important to note that these studies use a sample of single nucleotide polymorphisms (SNPs) to tag variation across the genome (on the basis of blocks of low recombination or linkage disequilibrium (LD)). If a single SNP shows an association with the disease, this does not necessarily mean that the SNP itself is involved, the association could be due to any of the other variants in LD with it. As we will see, these can include tagged rare variants [27]. The goal of these studies is thus not necessarily to identify causal alleles but to point to loci that might harbor them.

By that criterion, GWAS have been extremely successful for many complex disorders. They have pointed to numerous candidate loci for type 1 and type 2 diabetes, Crohn's disease, coronary artery disease, schizophrenia, bipolar disorder, multiple sclerosis and many other diseases. Some of these loci were previously implicated as sites of known rare mutations that cause obviously Mendelian forms of the disease in question, whereas others are novel findings that implicate new genes in disease processes. For some disorders, the findings converge on particular biochemical processes or pathways, such as beta-cell dysfunction or insulin action in type 2 diabetes [28,29]; natriuretic peptide signaling in high blood pressure and cardiovascular disease risk [30]; immune system genes in multiple sclerosis [31]; innate immunity and inflammation in Crohn's disease [32]; and neural development in schizophrenia [33,34].

Studies have also revealed some shared genetic risk across multiple disorders, including various autoimmune disorders (type 1 diabetes, Crohn's disease, multiple sclerosis and others) [35] and between schizophrenia and bipolar disorder [33,34,36].

The general trend across these studies is that the SNPs that give statistically significant association signals have tiny effects on disease risk, with odds ratios typically in the region of 1.05 to 1.2 (which means that if you carry such an allele, your risk of disease is increased 1.05- to 1.2-fold). This is exactly as predicted under a polygenic model: individual variants are not expected to have large effects. The aggregate risk caused by all the identified variants considered together is, however, still relatively small. In most cases, the SNPs that meet the criteria for genome-wide significance can collectively explain only a small percentage of the genetic variance of the disorder [37] and hardly any of the familial risk [38]. There is also no statistical evidence for
the kind of epistatic interactions that might be expected: combinations of alleles simply increase the overall effect additively (for example, for height [39,40] and body-mass index [41]).

There are several, not mutually exclusive ways to interpret the positive signals that have emerged from GWAS of complex disorders (Figure 2). First, they represent the actual effects of either the genotyped SNP or a common variant that is in LD with it. Given the power of the studies and the lack of overall variance explained by the identified variants, one could further conclude that, while real, the overall effect of such common variants is very modest [42,43]. This would be consistent with a model in which such variants act as modifiers of rare mutations, but, even in aggregate, do not cause disease in the absence of some such mutation.

Second, the common variants found represent merely the tip of the iceberg. Many other variants exist that have even smaller effect sizes, which current studies are underpowered to detect. Collectively, these could explain a sizeable fraction of the overall genetic variance – much more than actually observed, due to incomplete LD with the causal variants – leaving little need to invoke rare mutations (for example, [33]). Simulations exploring this possibility are discussed in Box 2.

Third, the common SNPs that show association signals are actually tagging rare mutations that segregate in the sampled populations [23,27]. Whenever a rare mutation arises, it necessarily occurs on the background of some ancient haplotype. If such a mutation predisposes to disease with relatively high penetrance, explaining even a small percentage of cases in a population, then the common haplotype will be slightly increased in frequency among cases when compared to controls. The odds ratio of the associated common SNP, which suggests a very modest increase in risk, could thus actually signal a highly penetrant variant on a fraction of the chromosomes with that haplotype. This kind of effect will be especially prevalent in studies from small, defined populations. Though one might expect it to be diluted out when multiple populations are combined (because different rare mutations in the same gene will occur on different haplotypes), it has been argued that synthetic associations with a single SNP allele can arise by chance due to multiple rare variants in the same locus [23,27], though others contend this is unlikely [21,22].

It is not possible to determine definitively which of these interpretations is correct from the GWAS data themselves. In particular, the strongly worded claim that GWAS signals provide strong support for a polygenic architecture of complex disorders, involving large numbers of common variants of small effect in each individual [33], is not justified [18,27,44] (Box 2). GWAS simply cannot determine whether the alleles responsible for the positive associations are common or rare, nor can aggregate scores or genome-partitioning models [45,46] tell how many alleles are involved, either across the population or in each affected individual (especially if the assumption that the signals from the SNPs in question are independent is not valid (D Goldstein, personal communication)).
Fortunately, we do not have to rely on statistical simulations to answer these questions. There is direct empirical evidence that rare mutations play the predominant role in the inheritance of such disorders. This evidence includes the nature of the spectrum of genetic variation in humans and the growing number of examples of identified, rare, disease-causing mutations.

The spectrum of human genetic variation
There is a common view that the human genome can be divided into bases that are pretty much constant across all people and those that are polymorphic. The logical extension of this idea is that heritable phenotypic differences between people must be caused by the particular combinations of polymorphisms that they inherit at the variable sites (for example, [47]). Recent data from whole-genome sequencing efforts show just how wrong this view is.

Far from most of the genome being effectively constant, it seems that every position in the genome has been mutated many, many times over in the human population [48-51]. Each of us carries thousands of very rare variants, including hundreds of novel mutations [52-56]. Recent, rare mutations are far more likely to have a deleterious effect on protein production or function and much more likely to cause disease than common variants [54-59].

New mutations may spread in the pedigree or population in which they arise for some time, depending largely on whether they have a deleterious effect on fitness or not [51,60,61]. Mutations that do have a deleterious effect will be quickly selected against, though the recent human population explosion could allow less penetrant or recessive alleles to persist for some time at low frequency [50,60]. It is, however, highly paradoxical to suppose that variants that predispose to serious...
Box 2. Aggregate polygenic scores

GWAS for various disorders have identified a number of SNPs that are above the threshold for genome-wide statistical significance as disease risk factors. Collectively, these loci explain a small fraction of risk. It remains possible, however, that additional SNPs are also associated with risk but that their signals are buried in the noise and cannot meet the burden of multiple testing correction with current sample sizes. To attempt to measure whether additional associated SNPs exist, various researchers have generated aggregate scores for the top $p$ percentage of SNPs based on a ranking of $p$-values in a discovery sample. The scores for each individual in a replication sample are then used to see if they can distinguish cases and controls. In a study on schizophrenia, such an aggregate score, involving thousands of SNPs, was indeed significantly increased in cases versus controls. However, this score accounted for only 3% of the variance [33]. This effect has since been replicated in a family study, where it explained 5% of the variance [129], ruling out population stratification concerns, and in a much larger GWAS, where it explained 6% of the variance [34].

These data indicate that there probably are additional SNPs that are associated with schizophrenia risk that could be detected with larger samples. Indeed, the most recent GWAS mega-analysis for schizophrenia [34] reports additional significant SNPs that were not found in earlier studies. Taken at face value, however, they also suggest that the overall contribution of common variants to the genetic variance that affects the disorder is very modest (less than 10%). But based on the idea that the real signals might be swamped out by non-associated SNPs in these aggregate scores, and that the linkage between the genotyped and the putative causal SNPs was probably imperfect, the authors performed simulations where they suggest that the overall impact of common SNPs was in fact much higher (as much as 33% of the total variance). Despite the claims of convergence onto a narrow range of values, seven models (out of 560 tested), involving combinations of multiple parameters such as allele frequency, effect size, number of SNPs involved and others, actually give widely different estimates of the total true variance explained (from 34% to 98%) and the number of SNPs contributing (from 6% to 100%).

The actual lack of convergence does not provoke much confidence in the overall claim that these results provide molecular genetic evidence for a substantial polygenic component to the risk of schizophrenia involving thousands of common alleles of very small effect.' That claim assumes (circularly) that the associations of common SNPs actually reflect the biological effects of common variants, as opposed to associations due to tagged rare variants [23,27]. It also assumes that the signals carried by the SNPs used in the score are independent: if not, then no information can be deduced as to the actual number of underlying causal variants (D Goldstein, personal communication). For now, the conclusion that many loci are involved in the genetics of schizophrenia across the population is uncontested. The conclusion that thousands of loci are causally involved in the inheritance of the disorder in each individual is not justified.

diseases would ever rise to a high frequency [51,62,63]. The casual invocation of balancing selection as a mechanism to maintain disease-causing alleles at high frequency is not supported by any evidence [64]. If a disorder is associated with reduced fitness, then the distribution of alleles affecting it will be expected to shift towards very rare ones that are highly deleterious [65] (and not, as previously concluded for disorders for which it was assumed that no correlation with fitness exists, towards slightly deleterious alleles at moderate frequency [66]).

Identifying rare mutations

The best evidence that so-called common disorders really encompass many distinct genetic disorders is the growing numbers of rare, highly penetrant mutations causing such disorders that are now being identified. Examples of single mutations causing disorders such as autism, schizophrenia, diabetes, epilepsy and many other common diseases have long been known. While these could be identified in only a small proportion of cases, they could, however, be disregarded as exceptions to the generality of the disease (for example, [67]): they did not cause ‘real schizophrenia’ or ‘real autism’. But what if there is no such thing? What if all cases are due to some rare mutation? The growing number of cases explained by such examples makes this view more and more difficult to argue against.

Such cases include copy number variants (CNVs; that is, deletions or duplications of sections of chromosomes, which often affect more than one gene), as well as point mutations. CNVs have become more easily detected, using genomic microarray and sequencing technologies, and have been found to contribute significantly to the total number of cases of a range of psychiatric and neurological disorders, including schizophrenia [68-76], autism [73,77-80], attention deficit-hyperactivity disorder [81-84], Tourette syndrome [85], developmental delay and mental retardation [14], and epilepsy [86].

Whole-genome or whole-exome sequencing strategies are now also identifying many point mutations that predispose with high penetrance to various disorders. Studies on psychiatric disorders have again led the way here [13,87-93], but recent reports have also identified single mutations causing neonatal diabetes mellitus [94], coronary artery disease [95] and Crohn’s disease [96].

Real sources of complexity in linking genotype to phenotype

If complex disorders really arise due to rare mutations, then why is their inheritance not more obviously Mendelian? There are a number of factors that contribute to the complexity of inheritance of these disorders. I have
argued that much of the complexity is simply apparent, due to lumping together what are actually distinct disorders under one umbrella term. Certain pathophysiological states could arise due to mutation in any of a large number of different genes. There are clearly many ways to cause psychosis or seizures or poor control of blood sugar level, just as there are many different ways, genetically, to cause blindness or deafness or mental retardation. More common diseases are likely to reflect a larger mutational target (that is, more genes will be involved in the affected underlying processes) [97].

This can only be part of the answer, however. Even within single families, the inheritance patterns of these phenotypes are usually not simple. What other factors might contribute to this complexity? First, the definition of the phenotype is probably very imprecise. A major finding that has emerged from recent studies is that specific mutations do not respect the boundaries of diagnostic categories - their effects can manifest in many different ways, leading to different symptoms and diagnoses in different carriers [15,17,98-103]. Analyzing segregation patterns on the basis of overly specific diagnostic categories could thus be highly misleading.

Second, any particular mutation could be required but not sufficient to cause disease in individual carriers [17]. It is extremely common, the norm actually, for Mendelian mutations to be modified by additional variants in the genetic background [104,105]. Some highly penetrant mutations will be clearly responsible for disease, by themselves. For diseases that have a strong effect on fitness, these will be almost immediately selected against and therefore enriched in sporadic cases that are caused by a novel occurrence of the mutation. Mutations that have lower penetrance might require additional 'hits' to result in the disease phenotype. There are numerous examples of compound heterozygosity or digenic inheritance underlying complex disorders [51,106-110]. Oligogenic interactions, involving several mutations, might also be important in some cases. Epistatic interactions between multiple mutations could be highly complex and unpredictable [111-114], or even paradoxical, as in cases where two disease mutations suppress each other's effects [115,116].

Finally, non-genetic factors must also be important in the emergence of disease phenotypes, given the incomplete concordance in the phenotypes of monozygotic twins for many of these diseases. Incomplete penetrance and variable expressivity could result from environmental factors and also from intrinsic developmental variation. The latter is especially important for neurodevelopmental disorders, where a certain probability that a pathogenic route will be followed could be inherited but the actual outcome of development would be strongly influenced by stochastic events [117].

Concluding remarks
The apparent complexity of common disorders arises to a large extent because of our poor ability to discriminate between what are in reality many distinct genetic disorders. Most cases of such disorders are likely to be the result of a rare, recent mutation that has a strong biological effect, or of interactions between a small number of such mutations. GWAS point to loci that might be involved in complex disorders but the population-based metrics that they provide say little about the number or type of causal alleles in individuals. Modifying effects of common alleles are certainly possible, though the evidence for these remains indirect. This change in paradigm is already having an impact in the clinic, as more and more cases of complex disorders are clinically defined on the basis of a genetic diagnosis, indexing the primary cause of the disease and not merely the surface symptoms [94-96,118,119].

Abbreviations
CNV, copy number variant; GWAS, genome-wide association study; LD, linkage disequilibrium; RP, retinitis pigmentosa; SNP, single nucleotide polymorphism.

Competing interests
The author has no competing interests.

Published: 23 January 2012

References
1. Plomin R, Haworth CM, Davis OS: Common disorders are quantitative traits. Nat Rev Genet 2009, 10:872-878.
2. Gottesman II, Shields J: A polygenic theory of schizophrenia. Proc Natl Acad Sci U S A 1967, 58:199-205.
3. Mcgue M, Gottesman II, Rao DC: Resolving genetic models for the transmission of schizophrenia. Genet Epidemiol 1985, 2:99-110.
4. Risch N: Genetic linkage and complex diseases, with special reference to psychiatric disorders. Genet Epidemiol 1990, 73:16; discussion 17-45.
5. O'Rourke DH, Gottesman II, Suarez BK, Rice J, Reich T: Refutation of the general single-locus model for the etiology of schizophrenia. Am J Hum Genet 1982, 34:630-649.
6. Risch N: Linkage strategies for genetically complex traits. I. Multilocus models. Am J Hum Genet 1990, 46:222-228.
7. International HapMap Consortium: The International HapMap Project. Nature 2003, 426:789-796.
8. Risch N, Merikangas K: The future of genetic studies of complex human diseases. Science 1996, 273:1516-1517.
9. Reich DE, Lander ES: On the allelic spectrum of human disease. Trends Genet 2001, 17:502-510.
10. Hirschhorn JN, Daly MJ: Genome-wide association studies for common diseases and complex traits. Nat Rev Genet 2005, 6:95-108.
11. Wright AF, Hastie ND: Complex genetic diseases: controversy over the Croesus code. Genome Biol 2001, 2:COMMENT2007.
12. Wright AF, Chakrava S, Abd El-Aziz MM, Bhattacharya SS: Photoreceptor degeneration: genetic and mechanistic dissection of a complex trait. Nat Rev Genet 2010, 11:273-284.
13. Najmabadi H, Hu H, Garshasi M, Zemotjel T, Abedini SS, Chen W, Hosseini M, Behjati F, Haas S, Jamali P, Zeca A, Mohseni M, Püttmann L, Vahid LN, Jensen C, Mohajer LA, Bienen M, Lani F, Mueller I, Weissmann R, Darvish H, Wrognemann K, Hadavi V, Lipkowitz B, Esmaeeli-Nieh S, Weiczorek D, Karimnejad F, Firoozabadgi SG, Cohen M, Fatatzi Z, et al. Deep sequencing reveals 50 novel genes for recessive cognitive disorders. Nature 2011, 478:57-63.
14. Cooper GM, Coe BP, Girirajan S, Rosenfeld JA, Yu TH, Baker C, Williams C, Stalker H, Hamid R, Hannig V, Abdel-Hamid H, Bader P, McCracken E, Niyazov D, Leppig K, Thiese H, Hummel M, Alexander N, Gorski J, Kussmann J, Shashi V,
69. Xu B, Roos JL, Levy S, van Rensburg EJ, Gogos JA, Karayiorgou M: Strong association of de novo copy number variants with sporadic schizophrenia. *Nat Genet* 2008, 40:880-885.

70. Stefansson H, Rujescu D, Cichon S, Pietiläinen OP, Ingason A, Steinberg S, Fossdal R, Sigurdsson E, Sigurdsson T, Buizer-Vorrips J, Hansen T, Jakobsen RD, Muglia P, Fratiglioni L, Svedberg T, Thorsdottir I, Ingason A, Jonasson A, Jonasson A, Bjornsson A, Mattiasdottir S, Blondal T, Haraldsson M, Magnusdottir BB, Geiling I, Möller HI, Hartmann A, et al. *Large recurrent microdeletions associated with schizophrenia*. *Nature* 2008, 455:232-236.

71. International Schizophrenia Consortium: *Rare chromosomal deletions and duplications increase risk of schizophrenia*. *Nature* 2008, 455:237-241.

72. Shi YY, He G, Zhang Z, Tang W, Zhang J Jr, Zhao Q, Zhang J, Sr., Li XW, Xi ZR, Fang C, Zhao XZ, Feng GY, He L. *A study of rare structural variants in schizophrenia patients and normal controls from Chinese Han population*. *Mol Psychiatry* 2013, 8:1911-1111.

73. Awadalla P, Gauthier J, Myers RA, Casals F, Hamdan FF, Griffin AR, Côté M, Henrion EI, Spiegelman D, Tarabichi J, Putton A, Yang Y, Boyko A, Bystance M, Xiong L, Rapoport JL, Addington AM, Deiuliis LB, Krieger AC, Jobber M, Miller B, Fombonne E, Mottron L, Zilversmit M, Keebler L, Daoust H, Marineau C, Roy-Gagnon M-H, Dubé MP, Eise-Walker A, et al. *Direct measure of the de novo mutation rate in autism and schizophrenia cohorts*. *Am J Hum Genet* 2010, 87:316-324.

74. Dinto P, Pagnamenta AT, Klei L, Arney NM, Merico D, Regan R, Conroy J, Magalhaes TR, Correia C, Abrahams BS, Almeida J, Baccelli E, Bader GD, Bailey A, Baird G, Battaglia M, Berney T, Boshkova N, Boti H, Bolton FP, Bourgeron T, Brenan S, Brian J, Bryson SE, Carson AR, Casullo G, Casey J, Chung BH, Cochrane L, Corsell L, et al. *Functional impact of global rare copy number variation in autism spectrum disorders*. *Nature* 2010, 466:368-372.

75. Levinson DF, Duan J, Oh S, Wang K, Sanders AR, Shi Z, Zhang M, Mouw BJ, Ollinie A, Amin F, Cloning CR, Silverman JM, Buckla NG, Byerley WF, Black DW, Kendler KS, Freedman RP, Budtrudic F, Pelet I, Hakonarson H, Bergen SE, Fanous AH, Holmans PA, Gejman PV. *Copy number variants in schizophrenia: confirmation of five previous findings and new evidence for 3q29 microdeletions and VIPR2 duplications*. *Am J Psychiatry* 2011, 168:802-816.

76. Vacov C, McCarthy S, Malhotra D, Murray F, Chou HJ, Peoples A, Makarav O, Yoon S, Bhanandari A, Corominas I, Lakoucheva LM, Krassotkhevy O, Krause V, Larach-Walsh V, Welsh DK, Craig D, Kelsoe JR, Gershon ES, Leal SM, Dell Aquila M, Morris DW, Gill M, Corvin A, Insel PA, McClellan J, King MC, Karayiorgou M, Levy DL, Dei-Les L, Sebat J. *Duplications of the neuropeptide receptor gene VIPR2 confer significant risk for schizophrenia*. *Nature* 2011, 471:499-503.

77. Slab J, Lakschmi B, Malhotra D, Troje J, Lese-Martin C, Walsh T, Yarnom B, Yoon S, Krassat I, Jonasson J, Leotta A, Pa D, Tsai D, Zhang R, Lee H, Hicks J, Spence SJ, Lee AT, Puura K, Lehtimakki T, Lederbrett D, Greskewen P, Bregman J. *Sutcliffe JS, Jobanputra V, Chung W, Warburton D, King MC, Skuse D, Geschwind DH, Gilliam TC, et al. Strong association of de novo copy number mutations with autism*. *Science* 2007, 316:445-449.

78. Marshall CR, Noor A, Vincent JB, Lionel AC, Feuk L, Rujescu D, Pietiläinen OP, Ingason A, Jonasson A, Jonasson A, Bjornsson A, Mattiasdottir S, Blondal T, Haraldsson M, Magnusdottir BB, Geiling I, Möller HI, Hartmann A, et al. *Large recurrent microdeletions associated with schizophrenia*. *Nature* 2008, 455:237-241.

79. Sanders SJ, Ercan-Sencicek AG, Hus V, Luo R, Murtha MT, Moreno-De-Luca D, Chu SH, Moreau MP, Gupta AR, Thomson SA, Mason CE, Bilguvar K, Celestino-Soper PB, Choi M, Crawford EL, Davis L, Wright NR, Dhodapkar RM, DiCola M, Dilullo JM, Candor G, Hanvill T, Fishman DO, Framh S, Garagalay R, Goh GS, Kamel S, Klei L, Love J, Kende V, et al. *Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism*. *Neuron* 2011, 70:863-867.

80. Elia J, Gai X, Xie HM, Perin JC, Geiger E, Glessner JT, D'arcy M, deBerardinis R,
93. O'Roak BJ, Deriziotis P, Lee C, Vives L, Schwartz JJ, Girirajan S, Karakoc E, Myers RA, Casals F, Gauthier J, Hamdan FF, Keebler J, Boyko AR, Bustamante CD, Girard SL, Gauthier J, Noreau A, Xiong L, Zhou S, Jouan L, Dionne-Laporte A, Glessner JT, Reilly MP, Kim CE, Takahashi N, Albano A, Hou C, Bradfield JP, Vissers LE, de Ligt J, Gilissen C, Janssen I, Steehouwer M, de Vries P, van Lier B, Piton A, Gauthier J, Hamdan FF, Lafrenière RG, Yang Y, Henrion E, Laurent S, Mefford HC, Muhle H, Ostertag P, von Spiczak S, Buysse K, Baker C, Franke A, Gazzolino M, Carson AR, Lantieri F, Muganga BM, Wang L, Takeda T, Rappaport EF, Grant SF, Berrettoni W, Devoto M, Shaikh TH, Hakonarson H, White PS. Rare structural variants found in attention-deficit hyperactivity disorder are preferentially associated with neurodevelopmental genes. Mol Psychiatry 2010, 15:637-646.

82. Williams NM, Zaharieva I, Martin A, Langley K, Mantriopagada K, Fossdal R, Stefansson H, Stefansson K, Magnusson P, Gudmundsson OQ, Gustafsson O, Holmans P, Owen MJ, O'Donovan M, Thapar A. Rare chromosomal deletions and duplications in attention-deficit hyperactivity disorder: a genome-wide analysis. Lancet 2010, 376:1401-1408.

81. Lionel AC, Crabtree JB, Barbosa N, Goddard T, Trithurariprapan R, Rickaby J, Gazzolino M, Carson AR, Howe JL, Wang Z, Wei J, Stewart AF, Roberts R, McPherron R, Freiberg A, Franke A, Schreiber S, Zwangenaub L, Fernandez BA, Roberts W, Arnold PD, Szatmari P, Marshall CR, Schachar R, Scherer SW. Rare copy number variation discovery and cross-disorder comparisons identify risk genes for ADHD. Sci Trans Med 2011, 3:95a75.

80. Lesch KP, Selch S, Renner TJ, Caijie Q, Nguyen TT, Hahn T, Romans M, Walita S, Shochet S, Denpifie AE, Heine M, Boreath-Hummer A, Romans A, Gross-Leusch S, Zerlaht H, Wulschtl T, Heinzelt S, Fassnacht M, Hallgarter A, Allolio B, Schäfer H, Warnke A, Reif A, Ropers HH, Ullmann R. Genome-wide copy number variation analysis in attention-deficit/hyperactivity disorder: association with neuropsychiatric Y gene dosage in an extended pedigree. Mol Psychiatry 2011, 16:491-503.

79. Sundaram SK, Hug AM, Wilson BI, Chuangyi HT. Tourette syndrome is associated with recurrent exonic copy number variant markers. Neurology 2010, 74:1536-1540.

78. Mefford HC, Muenke M, Eichler EE: Shendure J, Eichler EE: Mackenzie AP, Ng SB, Baker C, Rieder MJ, Nickerson DA, Bernier R, Fisher SE, Glessner JT, Reilly MP, Kim CE, Takahashi N, Albano A, Hou C, Bradfield JP, Vissers LE, de Ligt J, Gilissen C, Janssen I, Steehouwer M, de Vries P, van Lier B, Piton A, Gauthier J, Hamdan FF, Lafrenière RG, Griffing AR, Stone EA, Rouleau GA, Awadalla P: Rare copy number variations in schizophrenia. Neurology 2010, 75:867-880.

77. Funayama M, Li Y, Tsoi TH, Lam CW, Ohi T, Yazawa S, Uyama E, Djaldetti R, Bolszak M, Anttonen AK, Komulainen T, Hinttala R, Pakanen S, Sormunen R, Hohmann-Seele B, Brzustowicz L, Seretis M, Palsania P, Amoozegar A, Valsesia C, Ophoff RA, Veltman JA. CNNTAP2 gene dosage variation is associated with schizophrenia and epilepsy. Mol Psychiatry 2008, 13:261-266.

76. Mitochondrial disorders identifies severe diabetes mellitus using next-generation sequencing of the whole exome. Genome Biol 2013, 14:R92.

75. Kim S, Tzschachen MR, Veith RL, Basehouse MJ, Broeckel UT, Tomita-Mitchell A, Arca MJ, Casper JT, Margolis DA, Bick DP, Hessner MJ, Routes JM, Verbsky JW, Jacob HJ, Dimmock DP: Making a definitive diagnosis: successful clinical application of whole exome sequencing in a child with intractable inflammatory bowel disease. Genet Med 2011, 13:255-262.

74. Glessner JT, Reilly MP, Kim CE, Takahashi N, Albano A, Hou C, Bradfield JP, Zhang H, Sloman PM, Flory JH, Imelinsmisc M, Frackelton EC, Chivacchiott R, Thomas KA, Carins M, Otomo FG, Davidson M, Weiser M, Reichenberg A, Davis KL, Friedman JL, Cappola TP, Margulies KB, Rader DJ, Grant SF, Busbaum JD, Gur RE, Hakonarson H. Strong synaptic transmission impact by copy number variations in schizophrenia. Proc Natl Acad Sci USA 2010, 107:10584-10589.

73. Xu B, Roos JL, Delemer P, Boone B, Plummer B, Levy S, Gagos JA, Karayorgou M. Exome sequencing supports a de novo mutational paradigm for schizophrenia. Nat Genet 2011, 43:864-868.

72. Myamya M, Kakkula J, Hamdan FF, Lafrenière RG, Yang Y, Hennion E, Laurent S, Noreau A, Thibeudeau P, Karemara L, Spiegelman D, Kukuf D, Duguy I. Dystroismaisons J, Pollet V, Çetin M, Lapachelle K, Daillo O, Raymond A, Marineau C, Champhagne N, Xiong L, Gaspar C, Rivière JB, Tarabeux J, Cossette P, Krebs MO, Rappaport JL, Addington A, Delisi LE, et al.: Systematic resequencing of X-chromosome synaptic genes in autism spectrum disorder and schizophrenia. Mol Psychiatry 2011, 16:867-880.

71. Williams NM, Zaharieva I, Martin A, Langley K, Mantriopagada K, Fossdal R, Stefansson H, Stefansson K, Magnusson P, Gudmundsson OQ, Gustafsson O, Holmans P, Owen MJ, O’Donovan M, Thapar A. Rare chromosomal deletions and duplications in attention-deficit hyperactivity disorder: a genome-wide analysis. Lancet 2010, 376:1401-1408.

70. Lionel AC, Crabtree JB, Barbosa N, Goddard T, Trithurariprapan R, Rickaby J, Gazzolino M, Carson AR, Howe JL, Wang Z, Wei J, Stewart AF, Roberts R, McPherron R, Freiberg A, Franke A, Schreiber S, Zwangenaub L, Fernandez BA, Roberts W, Arnold PD, Szatmari P, Marshall CR, Schachar R, Scherer SW. Rare copy number variation discovery and cross-disorder comparisons identify risk genes for ADHD. Sci Trans Med 2011, 3:95a75.

69. Lesch KP, Selch S, Renner TJ, Caijie Q, Nguyen TT, Hahn T, Romans M, Walita S, Shochet S, Denpifie AE, Heine M, Boreath-Hummer A, Romans A, Gross-Leusch S, Zerlaht H, Wulschtl T, Heinzelt S, Fassnacht M, Hallgarter A, Allolio B, Schäfer H, Warnke A, Reif A, Ropers HH, Ullmann R. Genome-wide copy number variation analysis in attention-deficit/hyperactivity disorder: association with neuropsychiatric Y gene dosage in an extended pedigree. Mol Psychiatry 2011, 16:491-503.
113. Zwarts L, Magwire MM, Carbone MA, Versteven M, Herteleer L, Anholt RR, Callaerts P, Mackay TF: Complex genetic architecture of Drosophila aggressive behavior. *Proc Natl Acad Sci U S A* 2011, 108:17070-17075.

114. Schaaf CP, Sabo A, Sakai Y, Crosby J, Muzny D, Hawes A, Lewis L, Akbar H, Varghese R, Boervinkle E, Gibbs RA, Zoghbi HY: Oligogenic heterozygosity in individuals with high-functioning autism spectrum disorders. *Hum Mol Genet* 2011, 20:3366-3375.

115. Glasscock E, Qian J, Yoo JW, Noebels JL: Masking epilepsy by combining two epilepsy genes. *Nat Neurosci* 2007, 10:1554-1558.

116. Klassen T, Davis C, Goldman A, Burgess D, Chen T, Wheeler D, McPherson J, Bourquin T, Lewis L, Villasana D, Morgan M, Muzny D, Gibbs R, Noebels J: Exome sequencing of ion channel genes reveals complex profiles confounding personal risk assessment in epilepsy. *Cell* 2011, 145:1036-1048.

117. Mitchell KJ: The genetics of brain wiring: from molecule to mind. *PLoS Biol* 2007, 5:e113.

118. Vissers LE, de Vries BB, Veltman JA: Genomic microarrays in mental retardation: from copy number variation to gene, from research to diagnosis. *J Med Genet* 2010, 47:826-837.

119. Miller DT, Adam MP, Aadhiya S, Biesecker LG, Brothman AR, Carter NP, Church DM, Crolla JA, Eichler EE, Epstein CJ, Faucett WA, Feuk L, Friedman JM, Hamosh A, Jackson L, Kaminski EB, Kuk K, Krantz ID, Kuhn RM, Lee C, Ostell JM, Rosenberg C, Scherer SW, Srinivas SN, Stavropoulos DJ, Tepperberg JH, Thorland EC, Vermeesch JR, Wagggoner DJ, Watson MS, et al: Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am J Hum Genet* 2010, 86:749-764.

120. Visscher PM, Goddard ME, Derks EM, Wray NR: Evidence-based psychiatric genetics, AKA the false dichotomy between common and rare variant hypotheses. *Mol Psychiatry* 2011. doi: 10.1038/mp.2011.65

121. Chahal HS, Stals K, Unterländer M, Balding DJ, Thomas MG, Kumar AV, Besser GM, Atkinson AB, Morrison PJ, Howlett TA, Levy MJ, Orme SM, Akker SA, Abel RL, Grossman AB, Burger J, Eildard I, Korbonits M: AIP mutation in pituitary adenomas in the 18th century and today. *N Engl J Med* 2011, 364:93-95.

122. Laederich MB, Horton WA: Achondroplasia: pathogenesis and implications for future treatment. *Curr Opin Pediatr* 2010, 22:516-523.

123. Klingseisen A, Jackson AP: Mechanisms and pathways of growth failure in primordial dwarfs. *Genes Dev* 2011, 25:2011-2024.

124. Kitano H: Biological robustness. *Nat Rev Genet* 2004, 5:826-837.

125. Raj A, van Oudenaarden A: Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* 2008, 135:216-226.

126. Eldar A, Elowitz MB: Functional roles for noise in genetic circuits. *Nature* 2010, 467:167-173.

127. Albert R, Jeong H, Barabasi AL: Error and attack tolerance of complex networks. *Nature* 2000, 406:378-382.

128. Ruderfer DM, Kirov G, Chambert K, Moran JL, Owen MJ, O’Donovan MC, Sklar P, Purcell SM: A family-based study of common polygenic variation and risk of schizophrenia. *Mol Psychiatry* 2011, 16:887-888.

Cite this article as: Mitchell KJ: What is complex about complex disorders? *Genome Biology* 2012, 13:237.