Editor's choice
Scan to access more
free content

From the periphery to centre stage: de novo single nucleotide variants play a key role in human genetic disease

Chee-Seng Ku,[1] Eng King Tan,[2,3] David N Cooper[4]

▸ Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/jmedgenet-2013-101519)

[1]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden
[2]Department of Neurology, National Neuroscience Institute, Singapore, Singapore
[3]Duke-National University of Singapore Graduate Medical School, Singapore General Hospital, Singapore, Singapore
[4]School of Medicine, Institute of Medical Genetics, Cardiff University, Cardiff, UK

Correspondence to
Dr Chee-Seng Ku, Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm 17177, Sweden; cheeseng.ku@ki.se
and
Professor David N Cooper, Institute of Medical Genetics, School of Medicine, Cardiff University, Cardiff CF14 4XN, UK; cooperDN@cardiff.ac.uk

Received 4 January 2013
Accepted 9 January 2013
Published Online First
9 February 2013

To cite: Ku C-S, Tan EK, Cooper DN. *J Med Genet* 2013;**50**:203–211.

ABSTRACT

Human germline mutations arise anew during meiosis in every generation. Such spontaneously occurring genetic variants are termed de novo mutations. Although the introduction of microarray based approaches led to the discovery of numerous de novo copy number variants underlying a range of human genetic conditions, de novo single nucleotide variants (SNVs) remained refractory to analysis at the whole genome level until the advent of next generation sequencing technologies such as whole genome sequencing and whole exome sequencing. These approaches have recently allowed the estimation of the mutation rate of de novo SNVs and greatly increased our understanding of their contribution to human genetic disease. Indeed, de novo SNVs have been found to underlie various common human neurodevelopmental conditions such as schizophrenia, autism and intellectual disability, as well as sporadic cases of rare Mendelian disorders. In many cases, however, confirmation of the pathogenicity of identified de novo SNVs remains a major challenge.

INTRODUCTION

Germline mutations arise anew during meiosis in every generation. Such spontaneously occurring genetic alterations are termed de novo mutations and this term serves to describe those heritable mutations that neither parent possessed or transmitted. Thus, de novo mutations are mutations that arose in the gametes of an individual's parents as distinct from post-zygotic somatic mutations that may have arisen post-fertilisation.[1–3]

De novo mutations are evident in the context of a range of different types of lesion, from single nucleotide variants (SNVs) to small indels of multiple bases, to larger structural variations including deletions, duplications, and other chromosomal rearrangements. Studies of human de novo mutations on a genome wide scale were, until comparatively recently, extremely challenging owing to the technological limitations of available screening methods. Microarray based technologies have been deployed very successfully in the context of de novo copy number variants (CNVs), most notably in the identification of submicroscopic deletions and duplications underlying schizophrenia and intellectual disability.[4–9] However, these microarray based methodologies were inadequate to the task of investigating de novo SNVs and small indels on a genome wide scale. Such studies required large scale sequencing which was simply not feasible by means of Sanger sequencing. Thus, the extent to which de novo SNVs and small indels contribute to the burden of human genetic disease has remained largely unexplored. The advent of high throughput, next generation sequencing (NGS) technologies has now ushered in a new era in the study of these types of de novo mutation in the human genome.[10]

Whole exome sequencing (WES) and whole genome sequencing (WGS) can now be rapidly performed on parent–offspring trios to identify de novo SNVs residing either within the protein coding regions or the entire genome.[1–3] However, the majority of studies reported to date have adopted WES (see online supplementary table S1[11–36]) owing to: (1) the fact that WES is cheaper and analytically less challenging than WGS; and (2) that most of the sought after disease causing and/or deleterious mutations are expected to be found within the protein coding regions. In this article, we review and summarise recent discoveries of de novo SNVs (ie, single nucleotide substitutions) causing human genetic disease. We focus on those studies which have applied WES to the investigation of de novo SNVs and their implications for our understanding of the aetiology of common complex diseases as well as rare Mendelian diseases. We also discuss the potential clinical applications of NGS in the genetic diagnosis of those diseases characterised by a high frequency of de novo SNVs.

TECHNOLOGICAL ADVANCES

By definition, WES focuses on sequencing the entire set of exons (protein coding regions) in the genome.[37] As such, and in contrast to WGS, it requires exome enrichment before massively parallel sequencing. The development of a battery of whole exome enrichment methods was therefore a prerequisite for the success of this approach.[38] During the sequence enrichment steps, the genomic regions of interest (ie, all exons) are captured through hybridisation selected DNA fragments using oligonucleotide probes, whereas the unwanted DNA sequences (ie, the non-coding regions) are removed before sequencing. Different commercial exome enrichment methods vary markedly in terms of the size of their targeted exomes and the specific genomic regions being enriched.[39–42] Genomic enrichment leads to a significant reduction in the proportion of the genome that needs to be sequenced; hence WES can achieve higher sequencing coverage. Thus, since the exome comprises only ∼30 Mb of genomic DNA sequence[43] (the actual targeted size being dependent on the precise choice of exome enrichment method adopted), approximately 3 Gb of filtered sequence reads or sequencing data are sufficient to achieve an average 100× depth of coverage. Achieving sufficient depth of coverage is particularly important for the detection of heterozygous de novo mutations which are characteristic of dominant conditions, in order to ensure that both alleles are sequenced adequately. These sequence enrichment methods, coupled with the high throughput NGS technologies such as Illumina HiSeq and Life Technologies SOLiD sequencing platforms, have together heralded a veritable explosion of WES applications in delineating the genetic causes of rare Mendelian disorders with new insights even for some common complex diseases (eg, schizophrenia, autism, and intellectual disability).[44–49]

DE NOVO MUTATIONS IN HUMAN GENETIC DISEASE

Owing to technological limitations, the contribution of de novo SNVs to human genetic disease (both common complex diseases and rare Mendelian disorders) has until recently remained largely unexplored, at least at the whole genome level. De novo SNVs are nevertheless likely to have profound clinical and/or phenotypic consequences when they impact functionally important nucleotides in the genome (eg, nonsense and splice site mutations or missense mutations in evolutionarily conserved sequences). However, only a relatively small proportion of the de novo SNVs occurring at each meiosis are nonsense and splice site mutations or will alter functionally important nucleotides, and hence are likely to be disease causing (see online supplementary table S1). Although de novo CNVs were shown quite early on, by means of whole genome microarrays, to be associated with several common neurodevelopmental diseases,[4–9] it was not until comparatively recently that de novo SNVs were also implicated in the aetiology of schizophrenia, autism/autism spectrum disorder (ASD) and intellectual disability through the WES of affected trios.[50–56] In similar vein, de novo SNVs have also been found to be responsible for sporadic cases of various rare dominant Mendelian disorders such as Kabuki syndrome (MIM 147920), Schinzel–Giedion syndrome (MIM 269150), and Bohring–Opitz syndrome (MIM 605039).[11–13] These Mendelian disorders are also characterised by multiple neurodevelopmental defects; thus, for example, Bohring–Opitz syndrome is a clinically recognisable syndrome characterised by severe intellectual disability, distinctive facial features, and multiple congenital malformations.[13]

COMMON DISEASES

It has been postulated that the occurrence of de novo mutations might explain why diseases characterised by dramatically reduced fecundity such as schizophrenia, autism, and intellectual disability nevertheless remain fairly common in the general population. If this were to be the case, the de novo mutations occurring in these diseases would serve to replenish the number of highly penetrant disease mutations in every generation despite continual negative selection against the disease alleles. Consistent with this hypothesis, de novo CNVs have been shown to be a common cause of schizophrenia, autism, and intellectual disability.[57 58]

De novo SNVs have also been identified in schizophrenia through the WES of affected case–parent trios.[52] The exomes of 53 sporadic schizophrenia cases (with no history of the disease in a first or second degree relative), 22 unaffected controls and their parents were sequenced. This study implemented a set of filters to eliminate false positive de novo variants (ie, these would appear to occur de novo either as a consequence of non-detection in the parents or due to systematic false positive calls in the offspring), including validation by Sanger sequencing. As such, the study identified a total of 34 de novo point mutations (33 SNVs and one dinucleotide substitution) and four de novo indels (microdeletions or microinsertions) in the affected trios throughout the exomes. This study found an excess of non-synonymous changes among the identified de novo SNVs in schizophrenia cases, as well as an increased likelihood of missense variants affecting protein structure and function when compared with the rare inherited exonic variants identified in the same study. Of the 34 de novo point mutations identified, 32 were missense mutations, 19 of which affected evolutionarily conserved positions and were predicted bioinformatically (by PolyPhen-2 analysis) to alter protein function. In addition, three of the de novo indels were predicted to give rise to protein truncations whereas one resulted in a single amino acid deletion. Overall, 27 of the 53 cases of schizophrenia were found to carry at least one de novo mutation (SNVs or/and indels). However, it is noteworthy that several exonic de novo SNVs (four non-synonymous and three synonymous) were also identified in seven of the 22 control subjects.[52] This highlights the fact that distinguishing pathogenic and/or deleterious de novo mutations from non-disease-associated de novo mutations can be a tricky task; indeed, it has become clear that de novo non-synonymous SNVs detected in controls need not necessarily be pathogenic.

By contrast, sequencing of complete genomes from two ‘apparently healthy’ parent-offspring trios (from the HapMap CEU (Utah residents with Northern and Western European ancestry) and YRI (Yoruba in Ibadan, Nigeria)) identified a total of 49 and 35 germline de novo mutations in two parent offspring (CEU and YRI) trios, respectively.[2] However, none of these was non-synonymous SNVs, and only one synonymous SNV was detected in the CEU trio. A larger study on apparent healthy trios would be needed to determine the occurrence of de novo non-synonymous SNVs occurred per meiosis. Thus, additional evidence would generally be required to demonstrate the pathogenicity of a given de novo non-synonymous SNV—for example, prediction of protein damage, evolutionary conservation of the affected nucleotide or amino acid residue—and whether the SNV is located in a gene already known to be responsible for the disease in question, or at least a gene which is a biologically plausible candidate for disease involvement (this is discussed in more detail below). Girard *et al*[53] also applied WES to 14 trios (each trio comprising an individual with schizophrenia plus his or her parents) and identified 15 de novo SNVs in different genes in eight of the schizophrenia probands. Of these 15 validated de novo SNVs, four were nonsense SNVs (the remaining were missense) predicted to lead to premature termination of translation. The predicted effect on protein function of the 11 missense SNVs includes benign, possibly damaging, and probably damaging. It has been well known from the inherited variants as well, that the prediction on protein function alone is inadequate to determine the pathogenicity of de novo missense SNVs, and additional lines of evidence are required.

In a parallel development, de novo SNVs were also found at elevated frequencies in ASD/autism.[51 54–56] This phenomenon was first noted in a WES study of 20 sporadic ASD cases and their parents that identified 18 de novo SNVs within coding sequence, 11 of which were protein altering (missense and nonsense).[51] Potentially causative de novo SNVs were identified in the *FOXP1*, *GRIN2B*, *SCN1A*, and *LAMC3* genes, respectively, in four unrelated ASD probands who were among the most severely affected individuals studied.[51] Interestingly, de novo mutations in some of these genes had already been found in association with a variety of different neurodevelopmental phenotypes. For example, four de novo mutations in the *GRIN2B* gene were identified in a study of 468 individuals with intellectual disability.[59] Similarly, de novo mutations in the *FOXP1* gene had been identified in individuals with intellectual disability, autism, and language impairment.[60 61] A pleiotropic effect for these genes in different neurodevelopmental disorders is therefore to be suspected—it follows that de novo mutations in these genes could give rise to variable phenotypic expressivity or even a spectrum of neurodevelopmental phenotypes, depending upon mutation severity and modulation by other genetic and/or environmental factors.

A much larger WES study was performed in 238 families from a comprehensively phenotyped ASD cohort, comprising pedigrees with two unaffected parents and an affected proband (and in 200 ASD families, with an unaffected sibling).[55] This study found that highly deleterious (nonsense and splice site) de novo mutations in brain expressed genes were associated with ASD, but the total number of non-synonymous de novo SNVs was significantly higher in ASD probands than in the unaffected siblings. In addition, the odds ratio of de novo non-synonymous to silent mutations in probands versus unaffected siblings was 1.93. When the analysis was confined to de novo nonsense and splice site mutations in brain expressed genes, this substantially increased the estimate of the effect size and demonstrated a significant difference in cases versus controls based either on an analysis of mutational burden or an evaluation of the odds ratio of nonsense and splice site mutations to silent SNVs (OR=5.65).[55] These results demonstrated that non-synonymous de novo mutations, and particularly highly deleterious nonsense and splice site de novo mutations in brain expressed genes, are associated with ASD.

Several new insights have been obtained from the recent WES studies of de novo mutations in ASD.[54–56] First, the number of de novo SNVs correlates positively with increasing paternal age; this is consistent with the postulate that the increased risk for children of older fathers to develop ASD is the result of an increased mutation rate with paternal age,[56] an observation also made by Sanders *et al*.[55] Further, new biologically plausible disease genes, such as *NTNG1*, were also identified on the basis that *NTNG1* harboured recurrent, protein disruptive mutations. *NTNG1* is a strong biological candidate given its role in the laminar organisation of dendrites and axonal guidance; both de novo SNVs identified were missense located at highly conserved positions and predicted to disrupt protein function.[56] Another gene of interest was *SCN2A* since two probands carrying de novo nonsense SNVs in *SCN2A* were identified; gain-of-function mutations in *SCN2A* were already known to be associated with a range of epilepsy phenotypes.[56 62–64] The presence of two or more de novo nonsense and/or splice site mutations in the same gene in unrelated affected individuals is unlikely to have occurred by chance alone, and hence provides strong evidence to implicate *SCN2A* as a disease gene in ASD. In addition, *KATNAL2* and *CHD8* were each found to carry two highly deleterious mutations in different individuals (by combining data with another study), again strongly implicating this gene in the aetiology of ASD.[55]

Taken together, these studies identified a variety of de novo SNVs that are predicted to disrupt gene function in ASD, thereby strengthening the hypothesis that the occurrence of de novo SNVs underlies at least a proportion of ASD. The de novo nature of the underlying lesions may account for the high prevalence of this disease which is clearly associated with a notable reduction in reproduction fitness. Further, these studies highlighted the extreme locus heterogeneity of ASD. This might explain why de novo mutations that are individually very rare could play a role in the causation of common diseases, because many different genes may harbour such mutations in different individuals. In support of this view, for example, it has been estimated that mutations in more than 1000 different genes may cause intellectual disability[65]; this large number of disease 'targets' could account for the prevalence of intellectual disability in the general population, as demonstrated by the findings of de Ligt *et al*[66] who implicated a total of 79 de novo mutations in 77 genes in causing intellectual disability (discussed below).

Comparable progress in identifying de novo SNVs has also been achieved in individuals with hitherto unexplained causes of intellectual disability. The first report of its kind employed the WES approach to identify de novo SNVs in children with a normal karyotype in whom array based genome profiling had excluded a potential contribution from de novo CNVs.[50] Thus, WES was performed on 10 children with unexplained intellectual disability and their parents, leading to the identification of a total of nine de novo SNVs. These mutations occurred in different genes (including *RAB39B* and *SYNGAP1*), some of which had previously been implicated in causing intellectual disability,[67 68] thereby lending further support to the results obtained by exome sequencing. Six of these nine de novo SNVs were predicted to be pathogenic on the basis of gene function, evolutionary conservation, and likely mutational impact.[50]

A recent WES study, performed on a larger cohort of 100 people with intellectual disability and their unaffected parents, identified 79 de novo SNVs in 53 patients.[66] Of these, 10 de novo SNVs identified in 10 patients were predicted to compromise the structure or function of known intellectual disability genes. Potentially causative de novo SNVs in novel candidate genes were also detected in 22 patients. Taken together, this study identified the underlying genetic cause in 10 patients with de novo mutations in known intellectual disability genes and in three male patients with severely disruptive, maternally inherited mutations in known X linked intellectual disability genes, giving a diagnostic yield of 13%. This study once again conveyed the message that de novo SNVs represent an important cause of intellectual disability.[66] Given the extensive locus heterogeneity of intellectual disability (ie, it is caused by mutations in a very large number of different genes), WES represents a promising diagnostic tool to detect the underlying mutations. The diagnostic yield of 13% in patients with intellectual disability using WES is most encouraging. It should be appreciated that this is a conservative estimate because a further 24 new candidate genes affected by de novo mutations could also be potentially pathogenic, but their actual clinical significance would require further investigation in additional patient cohorts to be confirmed. Indeed, a pathogenic role for three of these genes (DYNC1H1, GATAD2B, and CTNNB1) was supported by the identification of additional patients with intellectual disability and severely disruptive mutations in these genes.[66]

Since CNVs constitute a significant cause of intellectual disability, improvements in the detection of genomic deletions and duplications from WES data should further increase the diagnostic yield (while eliminating the need for concomitant microarray screening), as different types of mutations would then be detectable in a single experiment.[69–71] Although promising, various challenges remain in relation to the implementation of NGS methods as diagnostic tools in the clinical setting (eg, with reference to the development of a standardised pipeline and protocol from the experiment to interpretation of results), data analysis, interpretation of the results (eg, variants of uncertain significance), communication of the results to patients and their families, and other ethical concerns such as the disclosure of incidental findings. Challenges are also to be faced in the application of NGS methods to the genetic diagnosis of rare Mendelian diseases as discussed below.[72–74]

RARE MENDELIAN DISORDERS

Over the last 2 years, WES has been widely applied to the identification of new inherited causal mutations for a range of dominant and recessive Mendelian disorders.[48 75–78] However, our focus here has been placed firmly on the utility of WES in the context of identifying de novo SNVs for rare Mendelian disorders. Heterozygous de novo mutations are believed to be a common cause of sporadic instances of rare diseases characterised by multiple congenital malformations or anomalies, developmental delay and intellectual disability such as Schinzel–Giedion syndrome (MIM 269150),[12] Bohring–Opitz syndrome (MIM 605039),[13] and Coffin–Siris syndrome (MIM 135900),[22 23] just to name a few, whose genetic bases had previously remained elusive.

Among recent WES studies, of particular interest is Coffin–Siris syndrome, characterised by various anomalies such as developmental delay and severe speech impairment. It is a rare congenital anomaly syndrome in which the majority of affected individuals are sporadic cases, strongly implying a dominant genetic basis for the disorder with underlying de novo mutations. This has now been confirmed to be the case by molecular analysis. WES, performed on five affected individuals with Coffin–Siris syndrome, identified 51 plausible variants under the hypothesis that an abnormality in a causal gene would be shared by at least two individuals.[23] These variants were then validated by targeted Sanger sequencing of genomic DNA from the five affected individuals and their parents. This led to the identification of two de novo heterozygous SNVs in *SMARCB1* in two (unrelated) affected individuals. Since the a priori probability that two different coding sequence mutations in unrelated individuals would be found to occur de novo within the same gene is extremely low, this provided strong support for the disease gene candidacy of *SMARCB1* in Coffin–Siris syndrome. Further, in another study that performed WES on Coffin–Siris syndrome patients (including one case–parents trio and two sporadic cases), a second disease gene was identified[22]; this time, three de novo mutations were found to truncate the *ARID1B* reading frame (one frameshift and two nonsense mutations). It would be interesting to reanalyse the data from Tsurusaki *et al*[23] to confirm if de novo mutations in this gene were missed in the remaining three patients (who did not harbour de novo mutations in *SMARCB1*) in their earlier analysis, or if a third gene for these three patients should be suspected. Intriguingly, both *SMARCB1* and *ARID1B* encode components of the SWI/SNF (SWItch/Sucrose NonFermentable) chromatin remodelling complex, which acts as an epigenetic modifier modulating the accessibility of transcription factors to DNA by altering chromatin structure. This was the first evidence to show that germline de novo mutations in SWI/SNF complex genes are associated with a multiple congenital anomaly syndrome. WES has clearly already proved itself to be a powerful discovery tool.

Among other new discoveries has been the detection of de novo microlesions in the *KAT6B* gene in genitopatellar syndrome (MIM 606170).[16 17] Genitopatellar syndrome is a rare disorder in which patellar aplasia or hypoplasia is associated with external genital anomalies and severe intellectual disability. In line with the assumption that the underlying causal mutations would act in a dominant manner and would have arisen de novo, *KAT6B* was identified as the only candidate gene harbouring previously unidentified (not previously reported in a public database) heterozygous variants in five of the six individuals whose exomes were sequenced. The mutations included a single nonsense variant and three frameshift indels, one of which (a 4 bp microdeletion) was observed in two unrelated cases. The de novo status was subsequently discovered by Sanger sequencing the parents who had not been subjected to WES, as none of the mutations was present in genomic DNA from the unaffected parents.[16] De novo heterozygous truncating mutations in the *KAT6B* gene were independently identified in three subjects with genitopatellar syndrome by another study (also through WES), and subsequent Sanger sequencing detected similar *KAT6B* gene mutations in three additional subjects.[17]

The consistency of these findings provided strong support for the causative role of deleterious de novo mutations of *KAT6B* as a cause of genitopatellar syndrome. Intriguingly, de novo protein truncating mutations in *KAT6B* were also implicated in another rare disorder, Say–Barber–Biesecker–Young–Simpson syndrome (SBBYSS or Ohdo syndrome) (MIM 603736).[14] This is a multiple anomaly syndrome which is also characterised by severe intellectual disability in addition to blepharophimosis and a mask-like facial appearance. *KAT6B* is a gene encoding a highly conserved histone acetyltransferase involved in chromatin modification. This finding, together with the identification of the genes for Coffin–Siris syndrome,[22 23] has provided evidence for a key role for chromatin modifying genes and epigenetic abnormalities in human developmental and congenital disorders characterised by different developmental anomalies. The findings so far reported are likely to be merely the tip of the iceberg, since many more mutations affecting genes involved in chromatin remodelling or modification are expected to be identified in other rare congenital disorders.

In addition to the application of WES to unrelated cases (with subsequent further validation of the putative disease variants in parental DNA to confirm their de novo status), other studies have applied WES directly to trios. Thus, Lin *et al*[24] applied WES to an individual with Olmsted syndrome (MIM 607066) and her parents; to detect de novo mutations, heterozygous variants from her parents were filtered out. In total, 45 putative de novo variants were found in the affected individual which were predicted to be damaging by SIFT (Sorting Intolerant From Tolerant). It is very likely that among these putative de novo variants, there are quite a few false positives as the number appears too large by comparison with other studies in apparently normal healthy individuals.[2] The false positives could be due to variants which were detected in offspring but overlooked in parents. Indeed, further validation of these candidate variants by Sanger sequencing yielded a solitary de novo heterozygous point mutation in *TRPV3*. Support for the causative role of this mutation was then garnered by the discovery of additional heterozygous missense mutations in *TRPV3* in five additional individuals; all five mutations were found to have occurred de novo. Further, these mutations were absent in 216 ethnically matched normal controls.[24] This is debatable, as for de novo mutations, there are chances that they will be absent in normal controls; therefore this does not constitute additional independent proof.

The major advantage of applying WES to parent–case trios is that it narrows down the list of potential candidate disease genes very significantly because of the limited number of de novo events to be expected in protein coding sequences. This is of course especially useful in the case of disorders in which de novo mutations are strongly suspected from the outset, for example, Baraitser–Winter syndrome (MIM 243310), a rare but well defined developmental disorder.[25] No familial recurrence or consanguinity has ever been observed in families affected with this syndrome, and hence the genetic basis in the known cases of Baraitser–Winter syndrome was always likely to be due to the occurrence of de novo microlesions (particularly as no obviously pathogenic CNVs had been detected using microarrays). Thus, WES was applied in the case of three probands and their unaffected parents. This allowed the direct investigation of de novo mutation and resulted in the identification of de novo missense changes in the *ACTG1* gene in two probands and in the *ACTB* gene in the third proband.[25] However, the potential drawback of this study design is the additional cost of WES incurred by sequencing the parents as well as the child, in contrast to testing a limited panel of candidate disease mutations or genes in the corresponding parents in order to determine their de novo status.

Although these different approaches come with their own particular advantages and disadvantages, both have been successful in identifying genes harbouring de novo mutations causing rare disorders. WES would be sufficient to identify the mutations or genes underlying most rare Mendelian disorders as long as the causative mutations reside within the coding regions. In contrast to WGS, WES is more cost effective and analytically less challenging. However, some studies have also applied the WGS approach, for example, to a family quartette of a subject affected by a sporadic case of severe epileptic encephalopathy and her unaffected parents and sibling.[79] WGS revealed a de novo missense mutation in *SCN8A*. The advantage of WGS is that it does not (like WES) exclude the possibility that the sought after pathological mutations might occur in functional regulatory or splicing elements, some of which might be remote from the genes whose expression they help to regulate.[80] The disadvantage of WGS is that there are potentially many more de novo mutations that need to be assessed before pathogenicity can be reliably attributed. For example, from variants detected in the quartette, 34 violated the Mendelian inheritance rules, suggesting the occurrence of multiple de novo mutations within the proband. However, 10 of these variants were removed because they were found during the course of the 1000 Genomes Project, and excluded by other criteria such as in error-prone regions and known segmental duplications. Finally, Sanger sequencing demonstrated that 23 of the remaining 24 candidates were false positives, leaving a single, true de novo variant in the proband.

Although numerous de novo mutations have now been identified as being responsible for rare Mendelian disorders (see online supplementary table S1), it remains unclear what constitutes evidence for causality. It is difficult and challenging to establish the causative role of newly identified mutations even when these events have occurred de novo.[81 82] Cosegregation of the putative causal mutation with the disease phenotype in large multigenerational pedigrees can provide strong genetic evidence of causality, but this is obviously not feasible with de novo mutations that severely impair reproductive fitness and are therefore not transmitted down the generations. Therefore, for any newly identified putative pathological lesion which appears to have occurred de novo, further screening of additional cases is invariably required. Detection of recurrent deleterious mutation or different mutations in the same gene in additional cases constitutes strong evidence of causality.[83] However, it can be difficult, in the context of extremely rare disorders, to find additional cases to validate the newly identified de novo mutation. Further, this search might be unsuccessful in disorders characterised by extensive locus heterogeneity. Moreover, de novo mutations tend to be individually extremely rare. Although sometimes very challenging, further validation in additional patient samples has nevertheless been achieved in recent WES studies of rare disorders. Thus, after the discovery of de novo mutations in the *ACTG1* and *ACTB* genes by WES of three parent–case trios with Baraitser–Winter syndrome, Sanger sequencing was employed to screen the coding sequences of both genes in 15 additional affected individuals, with the result that pathogenic mutations in one or other of these genes were detected in all subjects under study.[25] These mutations were then shown to have occurred de novo in all 11 subjects for whom parental DNA was available. Moreover, further validation in normal controls demonstrated that none of the mutations identified in Baraitser–Winter syndrome was present in several large control datasets.[25]

This approach was also successfully employed by Rosewich *et al*[33] who aimed to identify the genetic cause of alternating hemiplegia of childhood type 2 (AHC2; MIM 614820). A total of 24 patients with a clinical diagnosis of AHC2, together with their healthy parents, were recruited for this study. Even though WES was initially performed in only three proband–parent trios, the analysis identified *ATP1A3* as the disease causing gene for AHC2 on the basis that it harboured three quite distinct heterozygous de novo missense mutations in the different patients. However, to further strengthen the evidence to support the contention that de novo mutations in ATP1A3 were causative, this gene was sequenced in the remaining 21 AHC2 patients and their healthy parents; in line with expectation, mutations in the *ATP1A3* gene were identified in all patients and all the heterozygous mutations were shown to have occurred de novo.[33] Finally, confirmation of *ATP1A3* as the causal gene for AHC2 was provided by another WES based study. This study also applied WES on seven patients with AHC2 and their unaffected parents, which identified de novo non-synonymous mutations in *ATP1A3* in all patients.[32] This suggested that *ATP1A3* might be the only disease gene for this disorder. However, subsequent analysis of *ATP1A3* in an additional 98 patients with AHC2 found that *ATP1A3* mutations were likely to be responsible for at least 74% of the cases. Most AHC2 cases appear to be caused by one of seven recurrent *ATP1A3* mutations found by Heinzen *et al*.[32] The identification of *ATP1A3* as the disease gene for AHC2 is critical because AHC2 is currently diagnosed only on the basis of clinical criteria, and the variable phenotypic manifestations of this condition could easily give rise to diagnostic confusion. Delineation of the underlying molecular basis of AHC2 should allow the development of a genetic test to identify unequivocally children with the disease at the same time as ascertaining its clinical/phenotypic spectrum.

In addition to genetic evidence as described above, molecular functional studies can be performed to confirm the functional significance of newly identified de novo mutations. Evidence from such studies can be considered confirmatory only when the mutated gene can be shown to have a clear and well defined role in the molecular pathology of the disease. For example, *ATP1A3*, identified as the underlying disease gene in AHC2, is an α subunit of the Na^+/K^+ ATPase pump that is partly responsible for establishing and maintaining electrochemical gradients of sodium and potassium ions across the plasma membrane of neurons.[32] To better understand how *ATP1A3* mutations cause two clinically distinct disorders, as mutations in *ATP1A3* have also been shown to cause rapid onset dystonia–parkinsonism,[84] Heinzen *et al*[32] investigated the in vitro functional consequences of the *ATP1A3* mutations underlying AHC2 and rapid onset dystonia–parkinsonism, respectively. It was found that unlike the *ATP1A3* mutations associated with rapid onset dystonia–parkinsonism, the AHC2-causing mutations in this gene caused consistent reductions in ATPase activity without affecting the level of protein expression.

Nonsense and splice site mutations are very likely to be deleterious. However, most of the de novo SNVs identified to date are missense and it is often difficult to establish an unequivocal causative link between a specific missense mutation and a disease phenotype.[85] Amino acid substitutions in evolutionarily conserved residues can provide evidence for pathogenicity.[86] If the function of the protein is known, assessment of the biological effect of the missense mutation can be performed by in vitro mutagenesis and functional assay.[87–92] Without in depth analytical studies, however, missense mutations may often be difficult to distinguish from polymorphisms with little or no clinical significance, either in the context of candidate gene sequencing studies[93] or in the context of WES studies.[94] Evidence for pathological authenticity usually comes from one or more different lines of evidence,[93 95] including: (1) independent occurrence in additional patients; (2) absence in normal controls (only applicable to inherited variants as it does not constitute an additional proof for de novo mutations); (3) cosegregation of the lesion and disease phenotype through the family pedigree (not applicable in the case of de novo mutations); (4) non-conservative substitutions being more likely to disrupt protein function; (5) location in a protein region of structural or functional importance; (6) location in an evolutionarily conserved nucleotide sequence and/or amino acid residue; and albeit rarely (7) reversal of the pathological phenotype in patient/cultured cells by gene replacement. These lines of evidence should not of course be given equal weight; in most cases, the independent occurrence of de novo mutations in additional patients provides by far the most compelling support for pathogenicity.

De novo mutations located outside coding regions, for example, in promoter, intronic, intergenic regions or untranslated regions, can also be of pathological significance. However, in the context of a whole genome screen, most such pathogenic mutations could only be identified using WGS rather than WES. This notwithstanding, the ability of WES to identify de novo mutations residing within flanking non-coding regions was exemplified by the identification of a heterozygous de novo mutation in the 5′-untranslated region of *IFITM5* (it was located 14 bp upstream of the annotated translation initiation codon) underlying osteogenesis imperfecta type V (MIM 610967).[29] Subsequently, the study identified an identical heterozygous de novo mutation in an additional patient with osteogenesis imperfecta type V by Sanger sequencing, thereby strongly suggesting that this was the causal mutation for the disease. This finding concurred with another study, which also found the same heterozygous mutation in the 5′-untranslated region of *IFITM5*. This mutation occurred de novo in five unrelated cases, and cosegregated completely with the disease in three families—providing yet another line of evidence to support the pathological significance of this mutation (irrespective of whether it was acquired de novo or was inherited from the parents) in osteogenesis imperfecta type V.[30]

The identification of the causal genes for rare Mendelian disorders is important for the development of future molecular diagnostic tests to confirm the clinical diagnosis, especially for those disorders characterised by diverse clinical manifestations lacking specific phenotypes/phenotypic characteristics, or for those disorders which share common clinical features, and where ambiguity in diagnosis is common. For example, Weaver syndrome (MIM 277590) is a rare congenital anomaly syndrome characterised by generalised overgrowth, advanced bone age, pronounced macrocephaly, hypertelorism, and characteristic facial features. In addition, intellectual disability is common. Some patients with Weaver syndrome harbour mutations in *NSD1*, the gene that is mutated (or deleted) in most patients with classic Sotos syndrome (MIM 117550).[96 97] This molecular finding has led to uncertainty as to whether the Sotos and Weaver syndromes represent variable expressivity of a single locus with allelic heterogeneity or whether they represent distinct disorders caused by mutations in different genes. To compound the problem, these disorders have some shared and some distinguishing clinical features.[96 97] The application of WES to two trios affected by Weaver syndrome successfully identified two different de novo mutations in the *EZH2* gene.[19] Sanger sequencing of *EZH2* in a third classically affected proband identified a third de novo mutation in this gene. In addition, the study ruled out rare variants in *NSD1* by means of Sanger sequencing in all three probands.[19] The identification of this new disease gene has extended our definition of the genetic basis of Weaver syndrome. It may well be that other genes still remain to be identified, since Weaver syndrome is genetically and clinically quite a heterogeneous condition.

SUMMARY

In the context of the entirety of the human genome, de novo point mutations remained largely refractory to analysis until the arrival of NGS and its twin fruits, WES and WGS. Microarray technologies have been used widely and successfully to identify de novo CNVs in a number of common neurodevelopmental conditions such as schizophrenia and ASD. By contrast, since SNVs and other microlesions occurring de novo were not amenable to analysis by these microarray based methods, little was known about either their frequency or their impact upon neurodevelopmental disease until the advent of WES. However, by means of WES of case–parent trios, de novo SNVs have recently been implicated in schizophrenia, ASD, and intellectual disability. Taken together, these findings strengthen the hypothesis that the occurrence of de novo mutations could account for the high prevalence of those diseases which are associated with a notable reduction in reproductive fitness.

Many rare Mendelian disorders that are associated with multiple congenital malformations or anomalies, developmental delay, and intellectual disability occur sporadically because the severity and/or early onset of the disorders tend to preclude the transmission of the casual mutations to subsequent generations. As a result, these causal mutations are under strong negative selection that ensures they are quickly eliminated from the population. Since de novo mutations are by definition refractory to traditional linkage analysis, many genes causing sporadic cases of rare diseases still remain to be identified. Although the importance of de novo mutations in rare disorders without family transmission is well recognised, the genome wide screening methodology to identify novel disease de novo mutations within the coding regions was not widely available until 3 years ago. Recent studies have demonstrated the effectiveness of WES in the elucidation of the de novo genetic basis of an ever increasing number of rare Mendelian diseases. We are only at the beginning of the process of elucidation of the molecular basis of a large number of individually rare diseases whose genetic aetiologies have previously been elusive. It is really just a matter of time until WES is applied to all those rare diseases which were not previously amenable to study by pre-genomic era analytical methods. In addition to improving our knowledge of the pathogenesis of human genetic disease and the biological roles of the newly identified genes and the proteins they encode, the identification of disease-causing genes (whether the mutation has been parentally transmitted or has instead occurred de novo) will form the basis of future molecular diagnostic tests for these conditions. While technological advances have driven studies of de novo mutations, the collection of DNA samples from parent–offspring trios affected by various diseases is an important prerequisite to further enhance the rate of discovery of novel disease genes characterised by de novo mutations.

Confirming the pathological authenticity of missense mutations occurring de novo which have been identified in unique individuals/families is extremely challenging. This is because, in such cases, there is often little or no possibility of confirming the involvement of the candidate mutation or gene in the disease through analysis of a second individual or family. In addition to the in silico (bioinformatics) prediction of a protein damaging effect or the assessment of the evolutionary conservation of the amino acid residues affected by these 'unique' de novo missense mutations, the identification of these or similar mutations in other closely related conditions could be held to be suggestive of their pathogenicity. A further line of evidence might be the biological plausibility of the gene harbouring these mutations. However, the biological functions of the candidate gene(s) might not be fully known, and the biology underlying the disease might not be well characterised. Individualised functional studies of the 'unique' de novo missense mutations will provide additional evidence for their pathogenicity, and this functional characterisation will be greatly facilitated if the gene encodes a protein that plays a role in a pathophysiological pathway known to be involved in the disease. Currently, there are no well established criteria with which to distinguish de novo pathological lesions from other missense mutations that may have occurred de novo in a particular individual. It is even more challenging to show that a de novo missense mutation has no pathological significance in unique cases, as such mutations cannot be screened for in a general population of apparently healthy individuals to confirm their non-involvement in pathology.

Contributors CSK and DNC contributed to the conceptualisation of this manuscript. KCS contributed to the writing of this manuscript. DNC contributed to the editing of this manuscript. All authors contributed to proofreading. CSK approved the final version of this manuscript.

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Roach JC, Glusman G, Smit AF, Huff CD, Hubley R, Shannon PT, Rowen L, Pant KP, Goodman N, Bamshad M, Shendure J, Drmanac R, Jorde LB, Hood L, Galas DJ. Analysis of genetic inheritance in a family quartet by whole-genome sequencing. *Science* 2010;328:636–9.

2 Conrad DF, Keebler JE, DePristo MA, Lindsay SJ, Zhang Y, Casals F, Idaghdour Y, Hartl CL, Torroja C, Garimella KV, Zilversmit M, Cartwright R, Rouleau GA, Daly M, Stone EA, Hurles ME, Awadalla P, Genomes P. Variation in genome-wide mutation rates within and between human families. *Nat Genet* 2011;43:712–14.

3 Kong A, Frigge ML, Masson G, Besenbacher S, Sulem P, Magnusson G, Gudjonsson SA, Sigurdsson A, Jonasdottir A, Jonasdottir A, Wong WS, Sigurdsson G, Walters GB, Steinberg S, Helgason H, Thorleifsson G, Gudbjartsson DF, Helgason A, Magnusson OT, Thorsteinsdottir U, Stefansson K. Rate of *de novo* mutations and the importance of father's age to disease risk. *Nature* 2012;488:471–5.

4 Sebat J, Lakshmi B, Malhotra D, Troge J, Lese-Martin C, Walsh T, Yamrom B, Yoon S, Krasnitz A, Kendall J, Leotta A, Pai D, Zhang R, Lee YH, Hicks J, Spence SJ, Lee AT, Puura K, Lehtimaki T, Ledbetter D, Gregersen PK, Bregman J, Sutcliffe JS, Jobanputra V, Chung W, Warburton D, King MC, Skuse D, Geschwind DH, Gilliam TC, Ye K, Wigler M. Strong association of *de novo* copy number mutations with autism. *Science* 2007;316:445–9.

5 Xu B, Roos JL, Levy S, van Rensburg EJ, Gogos JA, Karayiorgou M. Strong association of *de novo* copy number mutations with sporadic schizophrenia. *Nat Genet* 2008;40:880–5.

6 Rees E, Moskvina V, Owen MJ, O'Donovan MC, Kirov G. *De novo* rates and selection of schizophrenia-associated copy number variants. *Biol Psychiatry* 2011;70:1109–14.

7 Kirov G, Pocklington AJ, Holmans P, Ivanov D, Ikeda M, Ruderfer D, Moran J, Chambert K, Toncheva D, Georgieva L, Grozeva D, Fjodorova M, Wollerton R, Rees E, Nikolov I, van de Lagemaat LN, Bayes A, Fernandez E, Olason PI, Bottcher Y, Komiyama NH, Collins MO, Choudhary J, Stefansson K, Stefansson H, Grant SG, Purcell S, Sklar P, O'Donovan MC, Owen MJ. *De novo* CNV analysis implicates specific abnormalities of postsynaptic signalling complexes in the pathogenesis of schizophrenia. *Mol Psychiatry* 2012;17:142–53.

8 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 NM, Fernandez TV, Fielding-Singh V, Fishman DO, Frahm S, Garagaloyan R, Goh GS, Kammela S, Klei L, Lowe JK, Lund SC, McGrew AD, Meyer KA, Moffat WJ, Murdoch JD, O'Roak BJ, Ober GT, Pottenger RS, Raubeson MJ, Song Y, Wang Q, Yaspan BL, Yu TW, Yurkiewicz IR, Beaudet AL, Cantor RM, Curland M, Grice DE, Gunel M, Lifton RP, Mane SM, Martin DM, Shaw CA, Sheldon M, Tischfield JA, Walsh CA, Morrow EM, Ledbetter DH, Fombonne E, Lord C, Martin CL, Brooks AI, Sutcliffe JS, Cook EH Jr, Geschwind D, Roeder K, Devlin B, State MW. Multiple recurrent *de novo* CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. *Neuron* 2011;70:863–85.

9 Devlin B, Scherer SW. Genetic architecture in autism spectrum disorder. *Curr Opin Genet Dev* 2012;22:229–37.

10 Mardis ER. A decade's perspective on DNA sequencing technology. *Nature* 2011;470:198–203.

11 Ng SB, Bigham AW, Buckingham KJ, Hannibal MC, McMillin MJ, Gildersleeve HI, Beck AE, Tabor HK, Cooper GM, Mefford HC, Lee C, Turner EH, Smith JD, Rieder MJ, Yoshiura K, Matsumoto N, Ohta T, Niikawa N, Nickerson DA, Bamshad MJ, Shendure J. Exome sequencing identifies *MLL2* mutations as a cause of Kabuki syndrome. *Nat Genet* 2010;42:790–3.

12 Hoischen A, van Bon BW, Gilissen C, Arts P, van Lier B, Steehouwer M, de Vries P, de Reuver R, Wieskamp N, Mortier G, Devriendt K, Amorim MZ, Revencu N, Kidd A, Barbosa M, Turner A, Smith J, Oley C, Henderson A, Hayes IM, Thompson EM, Brunner HG, de Vries BB, Veltman JA. *De novo* mutations of *SETBP1* cause Schinzel-Giedion syndrome. *Nat Genet* 2010;42:483–5.

13 Hoischen A, van Bon BW, Rodriguez-Santiago B, Gilissen C, Vissers LE, de Vries P, Janssen I, van Lier B, Hastings R, Smithson SF, Newbury-Ecob R, Kjaergaard S, Goodship J, McGowan R, Bartholdi D, Rauch A, Peippo M, Cobben JM, Wieczorek D, Gillessen-Kaesbach G, Veltman JA, Brunner HG, de Vries BB. *De novo* nonsense mutations in *ASXL1* cause Bohring-Opitz syndrome. *Nat Genet* 2011;43:729–31.

14 Clayton-Smith J, O'Sullivan J, Daly S, Bhaskar S, Day R, Anderson B, Voss AK, Thomas T, Biesecker LG, Smith P, Fryer A, Chandler KE, Kerr B, Tassabehji M, Lynch SA, Krajewska-Walasek M, McKee S, Smith J, Sweeney E, Mansour S, Mohammed S, Donnai D, Black G. Whole-exome-sequencing identifies mutations in histone acetyltransferase gene *KAT6B* in individuals with the Say-Barber-Biesecker variant of Ohdo syndrome. *Am J Hum Genet* 2011;89:675–81.

15 Bochukova E, Schoenmakers N, Agostini M, Schoenmakers E, Rajanayagam O, Keogh JM, Henning E, Reinemund J, Gevers E, Sarri M, Downes K, Offiah A, Albanese A, Halsall D, Schwabe JW, Bain M, Lindley K, Muntoni F, Khadem FV, Dattani M, Farooqi IS, Gurnell M, Chatterjee K. A mutation in the thyroid hormone receptor alpha gene. *N Engl J Med* 2012;366:243–9.

16 Simpson MA, Deshpande C, Dafou D, Vissers LE, Woollard WJ, Holder SE, Gillessen-Kaesbach G, Derks R, White SM, Cohen-Snuijf R, Kant SG, Hoefsloot LH, Reardon W, Brunner HG, Bongers EM, Trembath RC. *De novo* mutations of the gene encoding the histone acetyltransferase *KAT6B* cause Genitopatellar syndrome. *Am J Hum Genet* 2012;90:290–4.

17 Campeau PM, Kim JC, Lu JT, Schwartzentruber JA, Abdul-Rahman OA, Schlaubitz S, Murdock DM, Jiang MM, Lammer EJ, Enns GM, Rhead WJ, Rowland J, Robertson SP, Cormier-Daire V, Bainbridge MN, Yang XJ, Gingras MC, Gibbs RA, Rosenblatt DS, Majewski J, Lee BH. Mutations in *KAT6B*, encoding a histone acetyltransferase, cause Genitopatellar syndrome. *Am J Hum Genet* 2012;90:282–9.

18 Sirmaci A, Spiliopoulos M, Brancati F, Powell E, Duman D, Abrams A, Bademci G, Agolini E, Guo S, Konuk B, Kavaz A, Blanton S, Digilio MC, Dallapiccola B, Young J, Zuchner S, Tekin M. Mutations in *ANKRD11* cause KBG syndrome, characterized by intellectual disability, skeletal malformations, and macrodontia. *Am J Hum Genet* 2011;89:289–94.

19 Gibson WT, Hood RL, Zhan SH, Bulman DE, Fejes AP, Moore R, Mungall AJ, Eydoux P, Babul-Hirji R, An J, Marra MA, Consortium FC, Chitayat D, Boycott KM, Weaver DD, Jones SJ. Mutations in *EZH2* cause Weaver syndrome. *Am J Hum Genet* 2012;90:110–18.

20 Hood RL, Lines MA, Nikkel SM, Schwartzentruber J, Beaulieu C, Nowaczyk MJ, Allanson J, Kim CA, Wieczorek D, Moilanen JS, Lacombe D, Gillessen-Kaesbach G, Whiteford ML, Quaio CR, Gomy I, Bertola DR, Albrecht B, Platzer K, McGillivray G, Zou R, McLeod DR, Chudley AE, Chodirker BN, Marcadier J, Consortium FC, Majewski J, Bulman DE, White SM, Boycott KM. Mutations in *SRCAP*, encoding

SNF2-related CREBBP activator protein, cause Floating-Harbor syndrome. *Am J Hum Genet* 2012;90:308–13.

21 Rademakers R, Baker M, Nicholson AM, Rutherford NJ, Finch N, Soto-Ortolaza A, Lash J, Wider C, Wojtas A, DeJesus-Hernandez M, Adamson J, Kouri N, Sundal C, Shuster EA, Aasly J, MacKenzie J, Roeber S, Kretzschmar HA, Boeve BF, Knopman DS, Petersen RC, Cairns NJ, Ghetti B, Spina S, Garbern J, Tselis AC, Uitti R, Das P, Van Gerpen JA, Meschia JF, Levy S, Broderick DF, Graff-Radford N, Ross OA, Miller BB, Swerdlow RH, Dickson DW, Wszolek ZK. Mutations in the colony stimulating factor 1 receptor (*CSF1R*) gene cause hereditary diffuse leukoencephalopathy with spheroids. *Nat Genet* 2012;44:200–5.

22 Santen GW, Aten E, Sun Y, Almomani R, Gilissen C, Nielsen M, Kant SG, Snoeck IN, Peeters EA, Hilhorst-Hofstee Y, Wessels MW, den Hollander NS, Ruivenkamp CA, van Ommen GJ, Breuning MH, den Dunnen JT, van Haeringen A, Kriek M. Mutations in SWI/SNF chromatin remodeling complex gene *ARID1B* cause Coffin-Siris syndrome. *Nat Genet* 2012;44:379–80.

23 Tsurusaki Y, Okamoto N, Ohashi H, Kosho T, Imai Y, Hibi-Ko Y, Kaname T, Naritomi K, Kawame H, Wakui K, Fukushima Y, Homma T, Kato M, Hiraki Y, Yamagata T, Yano S, Mizuno S, Sakazume S, Ishii T, Nagai T, Shiina M, Ogata K, Ohta T, Niikawa N, Miyatake S, Okada I, Mizuguchi T, Doi H, Saitsu H, Miyake N, Matsumoto N. Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. *Nat Genet* 2012;44:376–8.

24 Lin Z, Chen Q, Lee M, Cao X, Zhang J, Ma D, Chen L, Hu X, Wang H, Wang X, Zhang P, Liu X, Guan L, Tang Y, Yang H, Tu P, Bu D, Zhu X, Wang K, Li R, Yang Y. Exome sequencing reveals mutations in *TRPV3* as a cause of Olmsted Syndrome. *Am J Hum Genet* 2012;90:558–64.

25 Riviere JB, van Bon BW, Hoischen A, Kholmanskikh SS, O'Roak BJ, Gilissen C, Gijsen S, Sullivan CT, Christian SL, Abdul-Rahman OA, Atkin JF, Chassaing N, Drouin-Garraud V, Fry AE, Fryns JP, Gripp KW, Kempers M, Kleefstra T, Mancini GM, Nowaczyk MJ, van Ravenswaaij-Arts CM, Roscioli T, Marble M, Rosenfeld JA, Siu VM, de Vries BB, Shendure J, Verloes A, Veltman JA, Brunner HG, Ross ME, Pilz DT, Dobyns WB. De novo mutations in the actin genes ACTB and ACTG1 cause Baraitser-Winter syndrome. *Nat Genet* 2012;44:440–4, S1–2.

26 Winkelmann J, Lin L, Schormair B, Kornum BR, Faraco J, Plazzi G, Melberg A, Cornelio F, Urban AE, Pizza F, Poli F, Grubert F, Wieland T, Graf E, Hallmayer J, Strom TM, Mignot E. Mutations in DNMT1 cause autosomal dominant cerebellar ataxia, deafness and narcolepsy. *Hum Mol Genet* 2012;21:2205–10.

27 Lines MA, Huang L, Schwartzentruber J, Douglas SL, Lynch DC, Beaulieu C, Guion-Almeida ML, Zechi-Ceide RM, Gener B, Gillessen-Kaesbach G, Nava C, Baujat G, Horn D, Kini U, Caliebe A, Alanay Y, Utine GE, Lev D, Kohlhase J, Grix AW, Lohmann DR, Hehr U, Bohm D, Consortium FC, Majewski J, Bulman DE, Wieczorek D, Boycott KM. Haploinsufficiency of a spliceosomal GTPase encoded by *EFTUD2* causes mandibulofacial dysostosis with microcephaly. *Am J Hum Genet* 2012;90:369–77.

28 Caputo V, Cianetti L, Niceta M, Carta C, Ciolfi A, Bocchinfuso G, Carrani E, Dentici ML, Biamino E, Belligni E, Garavelli L, Boccone L, Melis D, Andria G, Gelb BD, Stella L, Silengo M, Dallapiccola B, Tartaglia M. A restricted spectrum of mutations in the *SMAD4* tumor-suppressor gene underlies Myhre syndrome. *Am J Hum Genet* 2012;90:161–9.

29 Semler O, Garbes L, Keupp K, Swan D, Zimmermann K, Becker J, Iden S, Wirth B, Eysel P, Koerber F, Schoenau E, Bohlander SK, Wollnik B, Netzer C. A mutation in the 5′-UTR of *IFITM5* creates an in-frame start codon and causes autosomal-dominant osteogenesis imperfecta type V with hyperplastic callus. *Am J Hum Genet* 2012;91:349–57.

30 Cho TJ, Lee KE, Lee SK, Song SJ, Kim KJ, Jeon D, Lee G, Kim HN, Lee HR, Eom HH, Lee ZH, Kim OH, Park WY, Park SS, Ikegawa S, Yoo WJ, Choi IH, Kim JW. A single recurrent mutation in the 5′-UTR of *IFITM5* causes osteogenesis imperfecta type V. *Am J Hum Genet* 2012;91:343–8.

31 Barcia G, Fleming MR, Deligniere A, Gazula VR, Brown MR, Langouet M, Chen H, Kronengold J, Abhyankar A, Cilio R, Nitschke P, Kaminska A, Boddaert N, Casanova JL, Desguerre I, Munnich A, Dulac O, Kaczmarek LK, Colleaux L, Nabbout R. *De novo* gain-of-function *KCNT1* channel mutations cause malignant migrating partial seizures of infancy. *Nat Genet* 2012;44:1255–9.

32 Heinzen EL, Swoboda KJ, Hitomi Y, Gurrieri F, Nicole S, de Vries B, Tiziano FD, Fontaine B, Walley NM, Heavin S, Panagiotakaki E, European Alternating Hemiplegia of Childhood Genetics C, Biobanca e Registro Clinico per l'Emiplegia Alternante C, European Network for Research on Alternating Hemiplegia for S, Medium-sized Enterpriese C, Fiori S, Abiusi E, Di Pietro L, Sweney MT, Newcomb TM, Viollet L, Huff C, Jorde LB, Reyna SP, Murphy KJ, Shianna KV, Gumbs CE, Little L, Silver K, Ptacek LJ, Haan J, Ferrari MD, Bye AM, Herkes GK, Whitelaw CM, Webb D, Lynch BJ, Uldall P, King MD, Scheffer IE, Neri G, Arzimanoglou A, van den Maagdenberg AM, Sisodiya SM, Mikati MA, Goldstein DB. *De novo* mutations in *ATP1A3* cause alternating hemiplegia of childhood. *Nat Genet* 2012;44:1030–4.

33 Rosewich H, Thiele H, Ohlenbusch A, Maschke U, Altmuller J, Frommolt P, Zirn B, Ebinger F, Siemes H, Nurnberg P, Brockmann K, Gartner J. Heterozygous *de-novo* mutations in *ATP1A3* in patients with alternating hemiplegia of childhood: a whole-exome sequencing gene-identification study. *Lancet Neurol* 2012;11:764–73.

34 Jones WD, Dafou D, McEntagart M, Woollard WJ, Elmslie FV, Holder-Espinasse M, Irving M, Saggar AK, Smithson S, Trembath RC, Deshpande C, Simpson MA. *De novo* mutations in *MLL* cause Wiedemann-Steiner syndrome. *Am J Hum Genet* 2012;91:358–64.

35 Michot C, Le Goff C, Goldenberg A, Abhyankar A, Klein C, Kinning E, Guerrot AM, Flahaut P, Duncombe A, Baujat G, Lyonnet S, Thalassinos C, Nitschke P, Casanova JL, Le Merrer M, Munnich A, Cormier-Daire V. Exome sequencing identifies *PDE 4D* mutations as another cause of acrodysostosis. *Am J Hum Genet* 2012;90:740–5.

36 Lee H, Graham JM Jr, Rimoin DL, Lachman RS, Krejci P, Tompson SW, Nelson SF, Krakow D, Cohn DH. Exome sequencing identifies *PDE4D* mutations in acrodysostosis. *Am J Hum Genet* 2012;90:746–51.

37 Ku CS, Wu M, Cooper DN, Naidoo N, Pawitan Y, Pang B, Iacopetta B, Soong R. Technological advances in DNA sequence enrichment and sequencing for germline genetic diagnosis. *Expert Rev Mol Diagn* 2012;12:159–73.

38 Mamanova L, Coffey AJ, Scott CE, Kozarewa I, Turner EH, Kumar A, Howard E, Shendure J, Turner DJ. Target-enrichment strategies for next-generation sequencing. *Nat Methods* 2010;7:111–18.

39 Clark MJ, Chen R, Lam HY, Karczewski KJ, Chen R, Euskirchen G, Butte AJ, Snyder M. Performance comparison of exome DNA sequencing technologies. *Nat Biotechnol* 2011;29:908–14.

40 Asan Xu Y, Jiang H, Tyler-Smith C, Xue Y, Jiang T, Wang J, Wu M, Liu X, Tian G, Wang J, Wang J, Yang H, Zhang X. Comprehensive comparison of three commercial human whole-exome capture platforms. *Genome Biol* 2011;12:R95.

41 Parla JS, Iossifov I, Grabill I, Spector MS, Kramer M, McCombie WR. A comparative analysis of exome capture. *Genome Biol* 2011;12:R97.

42 Sulonen AM, Ellonen P, Almusa H, Lepisto M, Eldfors S, Hannula S, Miettinen T, Tyynismaa H, Salo P, Heckman C, Joensuu H, Raivio T, Suomalainen A, Saarela J. Comparison of solution-based exome capture methods for next generation sequencing. *Genome Biol* 2011;12:R94.

43 Ng SB, Turner EH, Robertson PD, Flygare SD, Bigham AW, Lee C, Shaffer T, Wong M, Bhattacharjee A, Eichler EE, Bamshad M, Nickerson DA, Shendure J. Targeted capture and massively parallel sequencing of 12 human exomes. *Nature* 2009;461:272–6.

44 Majewski J, Schwartzentruber J, Lalonde E, Montpetit A, Jabado N. What can exome sequencing do for you? *J Med Genet* 2011;48:580–9.

45 Gilissen C, Hoischen A, Brunner HG, Veltman JA. Unlocking Mendelian disease using exome sequencing. *Genome Biol* 2011;12:228.

46 Bamshad MJ, Ng SB, Bigham AW, Tabor HK, Emond MJ, Nickerson DA, Shendure J. Exome sequencing as a tool for Mendelian disease gene discovery. *Nat Rev Genet* 2011;12:745–55.

47 Singleton AB. Exome sequencing: a transformative technology. *Lancet Neurol* 2011;10:942–6.

48 Ku CS, Naidoo N, Pawitan Y. Revisiting Mendelian disorders through exome sequencing. *Hum Genet* 2011;129:351–70.

49 Veltman JA, Brunner HG. *De novo* mutations in human genetic disease. *Nat Rev Genet* 2012;13:565–75.

50 Vissers LE, de Ligt J, Gilissen C, Janssen I, Steehouwer M, de Vries P, van Lier B, Arts P, Wieskamp N, del Rosario M, van Bon BW, Hoischen A, de Vries BB, Brunner HG, Veltman JA. A *de novo* paradigm for mental retardation. *Nat Genet* 2010;42:1109–12.

51 O'Roak BJ, Deriziotis P, Lee C, Vives L, Schwartz JJ, Girirajan S, Karakoc E, Mackenzie AP, Ng SB, Baker C, Rieder MJ, Nickerson DA, Bernier R, Fisher SE, Shendure J, Eichler EE. Exome sequencing in sporadic autism spectrum disorders identifies severe *de novo* mutations. *Nat Genet* 2011;43585–9.

52 Xu B, Roos JL, Dexheimer P, Boone B, Plummer B, Levy S, Gogos JA, Karayiorgou M. Exome sequencing supports a *de novo* mutational paradigm for schizophrenia. *Nat Genet* 2011;43:864–8.

53 Girard SL, Gauthier J, Noreau A, Xiong L, Zhou S, Jouan L, Dionne-Laporte A, Spiegelman D, Henrion E, Diallo O, Thibodeau P, Bachand I, Bao JY, Tong AH, Lin CH, Millet B, Jaafari N, Joober R, Dion PA, Lok S, Krebs MO, Rouleau GA. Increased exonic *de novo* mutation rate in individuals with schizophrenia. *Nat Genet* 2011;43:860–3.

54 Neale BM, Kou Y, Liu L, Ma'ayan A, Samocha KE, Sabo A, Lin CF, Stevens C, Wang LS, Makarov V, Polak P, Yoon S, Maguire J, Crawford EL, Campbell NG, Geller ET, Valladares O, Schafer C, Liu H, Zhao T, Cai G, Lihm J, Dannenfelser R, Jabado O, Peralta Z, Nagaswamy U, Muzny D, Reid JG, Newsham I, Wu Y, Lewis L, Han Y, Voight BF, Lim E, Rossin E, Kirby A, Flannick J, Fromer M, Shakir K, Fennell T, Garimella K, Banks E, Poplin R, Gabriel S, DePristo M, Wimbish JR, Boone BE, Levy SE, Betancur C, Sunyaev S, Boerwinkle E, Buxbaum JD, Cook EH Jr, Devlin B, Gibbs RA, Roeder K, Schellenberg GD, Sutcliffe JS, Daly MJ. Patterns and rates of exonic *de novo* mutations in autism spectrum disorders. *Nature* 2012;485:242–5.

55 Sanders SJ, Murtha MT, Gupta AR, Murdoch JD, Raubeson MJ, Willsey AJ, Ercan-Sencicek AG, DiLullo NM, Parikshak NN, Stein JL, Walker MF, Ober GT, Teran NA, Song Y, El-Fishawy P, Murtha RC, Choi M, Overton JD, Bjornson RD, Carriero NJ, Meyer KA, Bilguvar K, Mane SM, Sestan N, Lifton RP, Gunel M, Roeder K, Geschwind DH, Devlin B, State MW. *De novo* mutations revealed by

whole-exome sequencing are strongly associated with autism. *Nature* 2012;485:237–41.

56 O'Roak BJ, Vives L, Girirajan S, Karakoc E, Krumm N, Coe BP, Levy R, Ko A, Lee C, Smith JD, Turner EH, Stanaway IB, Vernot B, Malig M, Baker C, Reilly B, Akey JM, Borenstein E, Rieder MJ, Nickerson DA, Bernier R, Shendure J, Eichler EE. Sporadic autism exomes reveal a highly interconnected protein network of *de novo* mutations. *Nature* 2012;485:246–50.

57 Malhotra D, Sebat J. CNVs: harbingers of a rare variant revolution in psychiatric genetics. *Cell* 2012;148:1223–41.

58 Rodriguez-Murillo L, Gogos JA, Karayiorgou M. The genetic architecture of schizophrenia: new mutations and emerging paradigms. *Annu Rev Med* 2012;63:63–80.

59 Endele S, Rosenberger G, Geider K, Popp B, Tamer C, Stefanova I, Milh M, Kortum F, Fritsch A, Pientka FK, Hellenbroich Y, Kalscheuer VM, Kohlhase J, Moog U, Rappold G, Rauch A, Ropers HH, von Spiczak S, Tonnies H, Villeneuve N, Villard L, Zabel B, Zenker M, Laube B, Reis A, Wieczorek D, Van Maldergem L, Kutsche K. Mutations in *GRIN 2A* and *GRIN2B* encoding regulatory subunits of NMDA receptors cause variable neurodevelopmental phenotypes. *Nat Genet* 2010;42:1021–6.

60 Hamdan FF, Daoud H, Rochefort D, Piton A, Gauthier J, Langlois M, Foomani G, Dobrzeniecka S, Krebs MO, Joober R, Lafreniere RG, Lacaille JC, Mottron L, Drapeau P, Beauchamp MH, Phillips MS, Fombonne E, Rouleau GA, Michaud JL. *De novo* mutations in *FOXP1* in cases with intellectual disability, autism, and language impairment. *Am J Hum Genet* 2010;87:671–8.

61 Horn D, Kapeller J, Rivera-Brugues N, Moog U, Lorenz-Depiereux B, Eck S, Hempel M, Wagenstaller J, Gawthrope A, Monaco AP, Bonin M, Riess O, Wohlleber E, Illig T, Bezzina CR, Franke A, Spranger S, Villavicencio-Lorini P, Seifert W, Rosenfeld J, Klopocki E, Rappold GA, Strom TM. Identification of *FOXP1* deletions in three unrelated patients with mental retardation and significant speech and language deficits. *Hum Mutat* 2010;31:E1851–60.

62 Ogiwara I, Ito K, Sawaishi Y, Osaka H, Mazaki E, Inoue I, Montal M, Hashikawa T, Shike T, Fujiwara T, Inoue Y, Kaneda M, Yamakawa K. De novo mutations of voltage-gated sodium channel alphaII gene *SCN2A* in intractable epilepsies. *Neurology* 2009;73:1046–53.

63 Bartnik M, Chun-Hui Tsai A, Xia Z, Cheung SW, Stankiewicz P. Disruption of the *SCN2A* and *SCN3A* genes in a patient with mental retardation, neurobehavioral and psychiatric abnormalities, and a history of infantile seizures. *Clin Genet* 2011;80:191–5.

64 Shi X, Yasumoto S, Kurahashi H, Nakagawa E, Fukasawa T, Uchiya S, Hirose S. Clinical spectrum of *SCN2A* mutations. *Brain Dev* 2012;34:541–5.

65 van Bokhoven H. Genetic and epigenetic networks in intellectual disabilities. *Annu Rev Genet* 2011;45:81–104.

66 de Ligt J, Willemsen MH, van Bon BW, Kleefstra T, Yntema HG, Kroes T, Vulto-van Silfhout AT, Koolen DA, de Vries P, Gilissen C, del Rosario M, Hoischen A, Scheffer H, de Vries BB, Brunner HG, Veltman JA, Vissers LE. Diagnostic exome sequencing in persons with severe intellectual disability. *N Engl J Med* 2012;367:1921–9.

67 Giannandrea M, Bianchi V, Mignogna ML, Sirri A, Carrabino S, D'Elia E, Vecellio M, Russo S, Cogliati F, Larizza L, Ropers HH, Tzschach A, Kalscheuer V, Oehl-Jaschkowitz B, Skinner C, Schwartz CE, Gecz J, Van Esch H, Raynaud M, Chelly J, de Brouwer AP, Toniolo D, D'Adamo P. Mutations in the small GTPase gene *RAB39B* are responsible for X-linked mental retardation associated with autism, epilepsy, and macrocephaly. *Am J Hum Genet* 2010;86:185–95.

68 Hamdan FF, Gauthier J, Spiegelman D, Noreau A, Yang Y, Pellerin S, Dobrzeniecka S, Cote M, Perreau-Linck E, Carmant L, D'Anjou G, Fombonne E, Addington AM, Rapoport JL, Delisi LE, Krebs MO, Mouaffak F, Joober R, Mottron L, Drapeau P, Marineau C, Lafreniere RG, Lacaille JC, Rouleau GA, Michaud JL, Synapse to Disease G. Mutations in *SYNGAP1* in autosomal nonsyndromic mental retardation. *N Engl J Med* 2009;360:599–605.

69 Plagnol V, Curtis J, Epstein M, Mok KY, Stebbings E, Grigoriadou S, Wood NW, Hambleton S, Burns SO, Thrasher AJ, Kumararatne D, Doffinger R, Nejentsev S. A robust model for read count data in exome sequencing experiments and implications for copy number variant calling. *Bioinformatics* 2012;28:2747–54.

70 Krumm N, Sudmant PH, Ko A, O'Roak BJ, Malig M, Coe BP, Project NES, Quinlan AR, Nickerson DA, Eichler EE. Copy number variation detection and genotyping from exome sequence data. *Genome Res* 2012;22:1525–32.

71 Karakoc E, Alkan C, O'Roak BJ, Dennis MY, Vives L, Mark K, Rieder MJ, Nickerson DA, Eichler EE. Detection of structural variants and indels within exome data. *Nat Methods* 2012;9:176–8.

72 Berg JS, Khoury MJ, Evans JP. Deploying whole genome sequencing in clinical practice and public health: meeting the challenge one bin at a time. *Genet Med* 2011;13:499–504.

73 Ku CS, Cooper DN, Polychronakos C, Naidoo N, Wu M, Soong R. Exome sequencing: dual role as a discovery and diagnostic tool. *Ann Neurol* 2012;71:5–14.

74 Jackson L, Goldsmith L, O'Connor A, Skirton H. Incidental findings in genetic research and clinical diagnostic tests: a systematic review. *Am J Med Genet A* 2012;158A:3159–67.

75 Ng SB, Buckingham KJ, Lee C, Bigham AW, Tabor HK, Dent KM, Huff CD, Shannon PT, Jabs EW, Nickerson DA, Shendure J, Bamshad MJ. Exome sequencing identifies the cause of a mendelian disorder. *Nat Genet* 2010;42:30–5.

76 Pierce SB, Walsh T, Chisholm KM, Lee MK, Thornton AM, Fiumara A, Opitz JM, Levy-Lahad E, Klevit RE, King MC. Mutations in the DBP-deficiency protein *HSD17B4* cause ovarian dysgenesis, hearing loss, and ataxia of Perrault Syndrome. *Am J Hum Genet* 2010;87:282–8.

77 Gilissen C, Arts HH, Hoischen A, Spruijt L, Mans DA, Arts P, van Lier B, Steehouwer M, van Reeuwijk J, Kant SG, Roepman R, Knoers NV, Veltman JA, Brunner HG. Exome sequencing identifies *WDR 35* variants involved in Sensenbrenner syndrome. *Am J Hum Genet* 2010;87:418–23.

78 Lalonde E, Albrecht S, Ha KC, Jacob K, Bolduc N, Polychronakos C, Dechelotte P, Majewski J, Jabado N. Unexpected allelic heterogeneity and spectrum of mutations in Fowler syndrome revealed by next-generation exome sequencing. *Hum Mutat* 2010;31:918–23.

79 Veeramah KR, O'Brien JE, Meisler MH, Cheng X, Dib-Hajj SD, Waxman SG, Talwar D, Girirajan S, Eichler EE, Restifo LL, Erickson RP, Hammer MF. *De novo* pathogenic *SCN8A* mutation identified by whole-genome sequencing of a family quartet affected by infantile epileptic encephalopathy and SUDEP. *Am J Hum Genet* 2012;90:502–10.

80 Cooper DN, Chen JM, Ball EV, Howells K, Mort M, Phillips AD, Chuzhanova N, Krawczak M, Kehrer-Sawatzki H, Stenson PD. Genes, mutations, and human inherited disease at the dawn of the age of personalized genomics. *Hum Mutat* 2010;31:631–55.

81 Gimelli G, Gimelli S, Dimasi N, Bocciardi R, Di Battista E, Pramparo T, Zuffardi O. Identification and molecular modelling of a novel familial mutation in the SRY gene implicated in the pure gonadal dysgenesis. *Eur J Hum Genet* 2007;15:76–80.

82 Benson AP, Al-Owais M, Holden AV. Quantitative prediction of the arrhythmogenic effects of de novo hERG mutations in computational models of human ventricular tissues. *Eur Biophys J* 2011;40:627–39.

83 Schuurs-Hoeijmakers JH, Oh EC, Vissers LE, Swinkels ME, Gilissen C, Willemsen MA, Holvoet M, Steehouwer M, Veltman JA, de Vries BB, van Bokhoven H, de Brouwer AP, Katsanis N, Devriendt K, Brunner HG. Recurrent *de novo* mutations in *PACS1* cause defective cranial-neural-crest migration and define a recognizable intellectual-disability syndrome. *Am J Hum Genet* 2012;91:1122–7.

84 de Carvalho Aguiar P, Sweadner KJ, Penniston JT, Zaremba J, Liu L, Caton M, Linazasoro G, Borg M, Tijssen MA, Bressman SB, Dobyns WB, Brashear A, Ozelius LJ. Mutations in the Na+/K+ -ATPase alpha3 gene *ATP1A3* are associated with rapid-onset dystonia parkinsonism. *Neuron* 2004;43:169–75.

85 Cotton RG, Scriver CR. Proof of "disease causing" mutation. *Hum Mutat* 1998;12:1–3.

86 Miller MP, Kumar S. Understanding human disease mutations through the use of interspecific genetic variation. *Hum Mol Genet* 2001;10:2319–28.

87 Drost M, Zonneveld JB, van Hees S, Rasmussen LJ, Hofstra RM, de Wind N. A rapid and cell-free assay to test the activity of lynch syndrome-associated *MSH2* and *MSH6* missense variants. *Hum Mutat* 2012;33:488–94.

88 Kansikas M, Kariola R, Nystrom M. Verification of the three-step model in assessing the pathogenicity of mismatch repair gene variants. *Hum Mutat* 2011;32:107–15.

89 Naruse H, Ikawa N, Yamaguchi K, Nakamura Y, Arai M, Ishioka C, Sugano K, Tamura K, Tomita N, Matsubara N, Yoshida T, Moriya Y, Furukawa Y. Determination of splice-site mutations in Lynch syndrome (hereditary non-polyposis colorectal cancer) patients using functional splicing assay. *Fam Cancer* 2009;8:509–17.

90 Walker LC, Whiley PJ, Couch FJ, Farrugia DJ, Healey S, Eccles DM, Lin F, Butler SA, Goff SA, Thompson BA, Lakhani SR, Da Silva LM, kConFab I, Tavtigian SV, Goldgar DE, Brown MA, Spurdle AB. Detection of splicing aberrations caused by *BRCA1* and *BRCA2* sequence variants encoding missense substitutions: implications for prediction of pathogenicity. *Hum Mutat* 2010;31:E1484–505.

91 Biswas K, Das R, Eggington JM, Qiao H, North SL, Stauffer S, Burkett SS, Martin BK, Southon E, Sizemore SC, Pruss D, Bowles KR, Roa BB, Hunter N, Tessarollo L, Wenstrup RJ, Byrd RA, Sharan SK. Functional evaluation of *BRCA2* variants mapping to the *PALB2*-binding and C-terminal DNA-binding domains using a mouse ES cell-based assay. *Hum Mol Genet* 2012;21:3993–4006.

92 Webb M. Developing functional assays for BRCA1 unclassified variants. *Methods Mol Biol* 2010;653:281–91.

93 Thusberg J, Vihinen M. Pathogenic or not? And if so, then how? Studying the effects of missense mutations using bioinformatics methods. *Hum Mutat* 2009;30:703–14.

94 Cooper DN, Bacolla A, Ferec C, Vasquez KM, Kehrer-Sawatzki H, Chen JM. On the sequence-directed nature of human gene mutation: the role of genomic architecture and the local DNA sequence environment in mediating gene mutations underlying human inherited disease. *Hum Mutat* 2011;32:1075–99.

95 Thusberg J, Olatubosun A, Vihinen M. Performance of mutation pathogenicity prediction methods on missense variants. *Hum Mutat* 2011;32:358–68.

96 Douglas J, Hanks S, Temple IK, Davies S, Murray A, Upadhyaya M, Tomkins S, Hughes HE, Cole TR, Rahman N. *NSD1* mutations are the major cause of Sotos syndrome and occur in some cases of Weaver syndrome but are rare in other overgrowth phenotypes. *Am J Hum Genet* 2003;72:132–43.

97 Rio M, Clech L, Amiel J, Faivre L, Lyonnet S, Le Merrer M, Odent S, Lacombe D, Edery P, Brauner R, Raoul O, Gosset P, Prieur M, Vekemans M, Munnich A, Colleaux L, Cormier-Daire V. Spectrum of *NSD1* mutations in Sotos and Weaver syndromes. *J Med Genet* 2003;40:436–40.

Table 1 - Summary of whole-exome sequencing studies identifying *de novo* mutations underlying rare Mendelian disorders

Study	Disorder	Characteristics	Major findings
Ng et al. (2010) [11]	Kabuki syndrome [MIM# 147920]	A rare, multiple malformation disorder characterized by a distinctive facial appearance, cardiac anomalies, skeletal abnormalities, immunological defects and mild to moderate mental retardation.	Seven probands had newly identified nonsense or frameshift mutations in *MLL2* gene. Follow-up Sanger sequencing detected *MLL2* mutations in two of the three remaining individuals with Kabuki syndrome and in 26 of 43 additional cases. In families where parental DNA was available, the mutation was confirmed to be *de novo* (n = 12) or transmitted (n = 2) in concordance with the phenotype.
Hoischen et al. (2010) [12]	Schinzel-Giedion syndrome [MIM# 269150]	Characterized by severe mental retardation, distinctive facial features, multiple congenital malformations (including skeletal abnormalities, genitourinary and renal malformations, and cardiac defects) and a higher-than-normal prevalence of tumours, notably neuroepithelial neoplasia.	Heterozygous *de novo* variants identified in *SETBP1* gene in four affected individuals. The study also identified *SETBP1* mutations in eight additional cases using Sanger sequencing. For six of the eight follow-up cases, parental DNA was available, and the mutations present in the affected individuals were again shown to have occurred *de novo*.
Hoischen et al. (2011) [13]	Bohring-Opitz syndrome [MIM# 605039]	Characterized by severe intellectual disability, prominent metopic suture, exophthalmus, forehead nevus flammeus, upslanting palpebral fissures, ulnar deviation and flexion of the wrists and metacarpophalangeal joints, and severe feeding problems.	Heterozygous *de novo* nonsense mutations identified in the *ASXL1* gene in three individuals with Bohring-Opitz syndrome. Analysis of additional samples found that in total, 7 out of 13 subjects with a Bohring-Opitz phenotype harboured *de novo ASXL1* mutations.
Clayton-Smith et al. (2011) [14]	Say-Barber-Biesecker-Young-Simpson syndrome (SBBYSS or Ohdo syndrome) [MIM# 603736]	A multiple anomaly syndrome characterized by severe intellectual disability, blepharophimosis, and a mask-like facial appearance. A number of individuals with SBBYSS also have thyroid abnormalities and cleft	Identified *de novo* protein-truncating mutations in the highly conserved histone acetyltransferase gene, *KAT6B,* in three out of four individuals. Sanger sequencing was used to confirm truncating mutations of *KAT6B*, clustering in the final exon of the gene in all four

		palate.	individuals and in a further nine persons with typical SBBYSS. Where parental samples were available, the mutations were shown to have occurred *de novo*.
Bochukova et al. (2012) [15]	Hypothyroidism [MIM# 614450]	A child with classic features of hypothyroidism (growth retardation, developmental retardation, skeletal dysplasia, and severe constipation) but only borderline-abnormal thyroid hormone levels.	*De novo* heterozygous nonsense mutation identified in the *THRA* gene generating a mutant protein that inhibits wild-type receptor action in a dominant negative manner.
Simpson et al. (2012) [16]	Genitopatellar syndrome (GPS) [MIM# 606170]	A rare disorder in which patellar aplasia or hypoplasia is associated with external genital anomalies and severe intellectual disability.	*De novo* mutations of the *KAT6B* gene identified in five individuals with GPS; a single nonsense variant and three frameshift indels, including a 4bp deletion observed in two cases.
Campeau et al. (2012) [17]	Genitopatellar syndrome (GPS) [MIM# 606170]	A rare skeletal dysplasia combining hypoplastic or absent patellae, genital anomalies, craniofacial defects, and intellectual disability among other features.	*De novo* heterozygous truncating mutations identified in the *KAT6B* gene in three subjects; Sanger sequencing of *KAT6B* was used to identify similar mutations in three additional subjects.
Sirmaci et al. (2011) [18]	KBG syndrome [MIM# 148050]	Characterized by intellectual disability associated with macrodontia of the upper central incisors as well as distinct craniofacial findings, short stature, and skeletal anomalies.	Deleterious heterozygous mutations identified in the *ANKRD11* gene. A splice-site mutation, c.7570-1G>C, cosegregated with the disease in a family with three affected members, whereas in a simplex case a *de novo* truncating mutation, c.2305delT, was detected. Sanger sequencing revealed additional *de novo* truncating *ANKRD11* mutations in three other simplex cases.
Gibson et al. (2012) [19]	Weaver syndrome [MIM# 614421]	A rare congenital anomaly syndrome comprising generalized overgrowth, advanced bone age, marked macrocephaly, hypertelorism, and characteristic facial features. Intellectual disability is	Two different *de novo* mutations identified in the *EZH2* gene. Sanger sequencing of *EZH2* in a third classically affected proband identified a third *de novo* mutation in this gene.

		common.	
Hood et al. (2012) [20]	Floating-Harbor syndrome (FHS) [MIM# 136140]	A rare condition characterized by short stature, delayed osseous maturation, expressive-language deficits, and a distinctive facial appearance.	Heterozygous truncating mutations identified in the *SRCAP* gene in five unrelated individuals with sporadic FHS. Sanger sequencing identified mutations in *SRCAP* in eight additional affected persons. Mutations were found to have occurred *de novo* in all six instances in which parental DNA was available.
Rademakers et al. (2012) [21]	Hereditary diffuse leukoencephalopathy with spheroids (HDLS) [MIM# 221820]	An autosomal dominant central nervous system white-matter disease with variable clinical presentations, including personality and behavioural changes, dementia, depression, parkinsonism, seizures and other phenotypes.	14 different mutations affecting the tyrosine kinase domain of the *CSF1R* gene identified in 14 families with HDLS. In one kindred, the study confirmed the *de novo* occurrence of the mutation.
Santen et al. (2012) [22]	Coffin-Siris syndrome [MIM# 135900]	Characterized by developmental delay, severe speech impairment, coarse facial features, hypertrichosis, hypoplastic or absent fifth fingernails or toenails and agenesis of the corpus callosum.	*De novo* truncating mutations identified in the *ARID1B* gene in three individuals with Coffin-Siris syndrome. All variants served to truncate the *ARID1B* reading frame (two nonsense variants and one frameshift). The mutations were validated using Sanger sequencing and shown to occur *de novo* in all three individuals.
Tsurusaki et al. (2012) [23]	Coffin-Siris syndrome [MIM# 135900]	A rare congenital anomaly syndrome characterized by growth deficiency, intellectual disability, microcephaly, coarse facial features and hypoplastic nail of the fifth finger and/or toe.	*De novo SMARCB1* gene mutations identified in two of five individuals with typical Coffin-Siris syndrome. Two *de novo* coding sequence mutations occurring within a specific gene is *a priori* an extremely unlikely event, supporting the idea that *SMARCB1* is the causative gene in Coffin-Siris syndrome.
Lin et al. (2012) [24]	Olmsted syndrome [MIM# 607066]	A rare congenital disorder characterized by bilateral mutilating palmoplantar keratoderma (PPK) and periorificial keratotic	*De novo* missense mutation identified in the *TRPV3* gene yielding a p.Gly573Ser substitution in an individual with Olmsted syndrome. Nucleotide sequencing of

		plaques with severe itching at all lesions.	five additional affected individuals also revealed missense mutations in *TRPV3* (p.Gly573Ser in three cases plus p.Gly573Cys and p.Trp692Gly).
Rivière et al. (2012) [25]	Baraitser-Winter syndrome [MIM #243310]	A rare but well-defined developmental disorder recognized by the combination of congenital ptosis, high-arched eyebrows, hypertelorism, ocular colobomata and a brain malformation consisting of anterior predominant lissencephaly. Other typical features include postnatal short stature and microcephaly, intellectual disability, seizures and hearing loss.	*De novo* missense changes identified in the cytoplasmic actin–encoding genes *ACTB* and *ACTG1* in one and two probands, respectively. Sequencing of both genes in 15 additional affected individuals identified disease-causing mutations in all probands, including two recurrent *de novo* alterations.
Winkelmann et al. (2012) [26]	Autosomal dominant cerebellar ataxia, deafness and narcolepsy (ADCA-DN) [MIM# 604121]	Characterized by late onset (age 30–40 years) narcolepsy–cataplexy, sensory neuronal deafness, cerebellar ataxia, dementia and, more variably, psychosis, optic atrophy and other symptoms.	*DNMT1* identified as the only gene with mutations found in all five affected individuals. Sanger sequencing confirmed a *de novo* mutation p.Ala570Val in one family, and showed co-segregation of p.Val606Phe and p.Ala570Val with the ADCA-DN phenotype in two other kindreds.
Lines et al. (2012) [27]	Mandibulofacial dysostosis with microcephaly (MFDM) [MIM# 610536]	A rare sporadic syndrome comprising craniofacial malformations, microcephaly, developmental delay, and a recognizable dysmorphic appearance. Major sequelae, including choanal atresia, sensorineural hearing loss, and cleft palate, each occur in a significant proportion of affected individuals.	Heterozygous mutations or deletions of the *EFTUD2* gene identified in all four unrelated affected individuals. Validation studies of eight additional individuals with MFDM demonstrated causative *EFTUD2* mutations in all affected individuals tested. A range of *EFTUD2* mutation types, including null alleles and frameshifts, is seen in MFDM, consistent with haploinsufficiency; segregation is *de novo* in all cases assessed to date.
Caputo et al. (2012) [28]	Myhre syndrome [MIM# 139210]	A developmental disorder characterized by reduced growth, generalized	Heterozygous mutations were identified in the *SMAD4* gene, which encodes for a transducer mediating

		muscular hypertrophy, facial dysmorphism, deafness, cognitive deficits, joint stiffness, and skeletal anomalies.	transforming growth factor β and bone morphogenetic protein signaling branches. Two recurrent *de novo SMAD4* mutations were identified in eight unrelated subjects.
Semler et al. (2012) [29]	Osteogenesis imperfecta type V [MIM# 610967]	A clinically and genetically heterogeneous disorder associated with bone fragility and susceptibility to fractures after minimal trauma. Some individuals with osteogenesis imperfecta also have hearing loss, dentinogenesis imperfecta, hypermobility of joints, and/or blue sclera.	An heterozygous de *novo* mutation was found in the 5'-untranslated region of *IFITM5*, 14 bp upstream of the annotated translation initiation codon (c. -14C>T). An identical heterozygous *de novo* mutation was identified in a second individual with osteogenesis imperfecta type V by Sanger sequencing
Cho et al. (2012) [30]	Osteogenesis imperfecta type V [MIM# 610967]	A heterogeneous group of disorders characterized by intrinsic bone fragility, resulting in frequent fractures as well as in deformities of the spine and limbs.	An heterozygous mutation of c.-14C>T was identified in the 5'-untranslated region of *IFITM5*. It cosegregated with the disease in three families and occurred *de novo* in five simplex individuals.
Barcia et al. (2012) [31]	Malignant migrating partial seizures of infancy (MMPSI)	Characterized by polymorphous focal seizures and arrest of psychomotor development in the first 6 months of life.	*De novo* gain-of-function mutations identified which affected the C-terminal domain of the *KCNT1* potassium channel. Nine additional individuals with MMPSI were sequenced and *KCNT1* mutations were identified in in 4 of them, in total identifying mutations in 6 out of 12 unrelated affected individuals.
Heinzen et al. (2012) [32]	Alternating hemiplegia of childhood (AHC) [MIM #614820]	A rare, severe neurodevelopmental syndrome characterized by recurrent hemiplegic episodes and distinct neurological manifestations.	*De novo* nonsynonymous mutations identified in the *ATP1A3* gene in all seven individuals with AHC.
Rosewich et al. (2012) [33]	Alternating hemiplegia of childhood (AHC) [MIM #614820]	Characterized by transient episodes of hemiplegia combined with other paroxysmal symptoms, including dystonia, nystagmus,	Three heterozygous *de novo* missense mutations identified in the *ATP1A3* gene. Sequencing of 21 affected individuals identified disease-associated mutations in the *ATP1A3* gene in all patients,

		autonomic disturbances, and seizures.	including six *de novo* missense mutations and one *de novo* splice-site mutation.
Jones et al. (2012) [34]	Wiedemann-Steiner syndrome [MIM# 605130]	Individuals with hypertrichosis cubiti associated with short stature, intellectual disability, and a distinctive facial appearance.	Identified *de novo* mutations in the *MLL* gene in five of the six individuals studied.
Michot et al. (2012) [35]	Acrodysostosis [MIM# 101800]	A dominantly inherited condition consisting of (1) skeletal dysplasia characterized by facial dysostosis with nasal hypoplasia, severe brachydactyly with short broad metatarsals, metacarpals, and phalanges, cone-shaped epiphyses, advanced bone maturation, spinal stenosis, and short stature; (2) resistance to multiple hormones, including parathyroid hormone and thyrotropin; and (3) possible neurological involvement (moderate to mild intellectual disability).	Detected heterozygous *PDE4D* mutations (c.673C>A and c.677T>C) in two individuals. Screening of *PDE4D* identified heterozygous mutations (c.568T>G and c.1759A>C) in two additional acrodysostosis cases. These mutations occurred *de novo* in all four cases.
Lee et al. (2012) [36]	Acrodysostosis [MIM# 101800]	A pleiotropic disorder characterized by skeletal, endocrine, and neurological abnormalities. Skeletal features include brachycephaly, midface hypoplasia with a small upturned nose, brachydactyly, and lumbar spinal stenosis. Endocrine abnormalities have been reported and include hypothyroidism and hypogonadism in males and irregular menses in females.	Three different missense mutations in the *PDE4D* gene were found to be heterozygous in three of the cases. Two of the mutations were shown to have occurred *de novo.*

NB: We include above all recent studies which applied the whole-exome sequencing approach to identify *de novo* mutations in rare Mendelian disorders. However, the list is by no means complete. This table briefly summarizes those studies which applied whole-exome sequencing to identify *de novo* mutations in rare disorders together with their major findings. Readers are encouraged to refer to the full papers for more detailed information.