MOLECULAR
CYTOGENETICS

Deciphering the pathogenic consequences of chromosomal aberrations in human genetic disease

Wigard P Kloosterman[1*] and Ron Hochstenbach[2]

Abstract

Chromosomal aberrations include translocations, deletions, duplications, inversions, aneuploidies and complex rearrangements. They underlie genetic disease in roughly 15% of patients with multiple congenital abnormalities and/or mental retardation (MCA/MR). In genetic diagnostics, the pathogenicity of chromosomal aberrations in these patients is typically assessed based on criteria such as phenotypic similarity to other patients with the same or overlapping aberration, absence in healthy individuals, *de novo* occurrence, and protein coding gene content. However, a thorough understanding of the molecular mechanisms that lead to MCA/MR as a result of chromosome aberrations is often lacking. Chromosome aberrations can affect one or more genes in a complex manner, such as by changing the regulation of gene expression, by disrupting exons, and by creating fusion genes. The precise delineation of breakpoints by whole-genome sequencing enables the construction of local genomic architecture and facilitates the prediction of the molecular determinants of the patient's phenotype. Here, we review current methods for breakpoint identification and their impact on the interpretation of chromosome aberrations in patients with MCA/MR. In addition, we discuss opportunities to dissect disease mechanisms based on large-scale genomic technologies and studies in model organisms.

Keywords: Chromosomal aberration, Paired-end sequencing, Karyotyping, arrayCGH, Gene dosage, Gene fusion, Expression regulation, Expression profiling, Genome engineering

Introduction

Structural genomic variations (SVs) are generally regarded as genetic changes with a size larger than 50 bp [1]. SVs form a major source of common genetic variation in the human population and they primarily comprise deletions, (mobile element) insertions and tandem duplications [2]. Smaller structural variations occur much more frequently than larger ones. Furthermore, there is significant purifying selection against large and gene-disruptive copy number variations (CNVs), indicating their considerable phenotypic impact. The rates at which large CNVs (>100 kb) arise *de novo* in the general population have been conservatively estimated to be around 1.2×10^{-2} CNVs per germ line transmission [3]. This is relatively low when compared to the rates of base substitutions

[4-6]. However, large CNVs and SVs have a high chance of affecting important genomic elements, which could lead to congenital disease. In line with this, large copy number changes are strongly enriched among patients with idiopathic multiple congenital abnormalities and/or mental retardation (MCA/MR) phenotypes compared to normal individuals [7]. In addition, *de novo* CNVs are found at higher frequency in MCA/MR patients (3.6×10^{-2} per germ line transmission) compared to control populations [8].

In spite of the clear association between large (*de novo*) CNVs and other SVs (collectively known as chromosomal aberrations) with MCA/MR phenotypes [9], precise characterization of molecular mechanisms that cause disease in individual patients is often unknown. This is largely a result of the complex effects of breakpoints on gene structure, function and expression. In the following sections we first provide an overview of current methods for diagnostic detection and interpretation of

* Correspondence: w.kloosterman@umcutrecht.nl
[1]Department of Medical Genetics, Center for Molecular Medicine, University Medical Center Utrecht, P.O. Box 85060, 3508 AB Utrecht, The Netherlands
Full list of author information is available at the end of the article

chromosomal aberrations in MCA/MR patients. Subsequently, the introduction and relevance of new methods for high-resolution dissection of breakpoints of chromosomal aberrations is discussed. Finally, we summarize the possible molecular consequences of chromosome aberrations for gene expression and function, and we discuss approaches for identification and unraveling the molecular determinants of congenital disease phenotypes.

Review

Diagnostics of chromosome aberrations in patients with MCA/MR

Current methods for detection of chromosome aberrations in genetic diagnosis of idiopathic MCA/MR

(Mosaic) aneuploidies and chromosomal rearrangements are a frequent cause of idiopathic MCA/MR. Starting in 1959 with the identification of trisomy-21 as the genetic basis of Down syndrome [10], microscopic observation of metaphase chromosomes has for several decades been the method of choice for detecting chromosome abnormalities in MCA/MR patients. This includes both karyotyping of banded chromosomes and, since the nineties of the last century, also fluorescence in situ hybridization (FISH). In consecutive, unselected MCA/MR patients karyotyping enables the detection of pathogenic chromosomal abnormalities in about 4% of cases (excluding Down syndrome), despite its limited resolution to about 5–10 Mb. The aberrations mainly include a single loss or gain of a chromosomal segment (>90% of cases) [11].

The application of targeted FISH to detect rearrangements that are beyond the resolution of chromosome banding, adds another ~4% of diagnosed MCA/MR patients [11]. Targeted FISH on metaphase chromosomes has been particularly successful in the diagnosis of recurrent microdeletions that are mediated by nonhomologous allelic recombination (NAHR) between flanking low copy repeats (LCRs), such as Velo-Cardio-Facial (VCF)/DiGeorge syndrome in 22q11.2, Williams-Beuren syndrome in 7q11.23 and Smith-Magenis syndrome in 17p11.2 (and many others, that were more recently discovered using microarrays [12], or computationally predicted [13]). These recurrent aberrations can be efficiently identified by targeted FISH if the patient's phenotype has been recognized by a clinican ("phenotype-first" approach).

In contrast, most non-recurrent, sporadic chromosome rearrangements are characterized by breakpoints that are more or less randomly located, as they are typically driven by mechanisms of non-homologous repair such as nonhomologous end joining (NHEJ) and microhomology mediated break induced replication (MMBIR) [14-16]. Many of these private rearrangements may be associated with unique clinical phenotypes, requiring a "genotype-first" approach. The introduction of microarray-based aneuploidy detection has enabled such a "genotype-first",

discovery-based approach. With high-resolution microarray platforms, on which all protein-coding genes are addressed with multiple probes, the diagnostic yield in the population with MCA/MR referral is about 15% [11,17]. Therefore, microarray-based aneuploidy detection has become the initial test in the laboratory workup of patients with idiopathic MCA/MR [17]. Limitations of arrays include the inability to detect low-level chromosomal mosaicism (<7-10%) and balanced rearrangements. However, in diagnostics of idiopathic MCA/MR, karyotyping would add less than 1% of pathogenic cases to those detected by microarray, as shown by a survey of 36,325 consecutive MCA/MR cases [11].

Current methods for the assessment of the clinical relevance of chromosomal imbalances in idiopathic MCA/MR

Before CNVs could be precisely delineated using microarrays, traditional approaches for assessing whether an imbalance is causal for a patient's clinical symptoms have been based on (i) the phenotypic resemblance to other patients with an identical or largely overlapping imbalance, (ii) the absence of the imbalance in a large number of healthy individuals, and (iii) the segregation of the CNV in the family, for example, *de novo* imbalances are considered to be likely pathogenic whereas those inherited from an apparently healthy parent are not. These criteria still hold true nowadays for the interpretation of CNVs [18]. However, the advent of the human genome sequence (and all its associated databases for genome annotation) has - together with the ability to precisely delineate imbalances using arrays - provided additional, sequence content-dependent features of CNVs. Most important, the presence of a gene in the imbalance with proven dosage effect or association with a known clinical disorder is considered as robust evidence for a pathogenic effect. Public databases such as ECARUCA, DECIPHER, ISCA, and PubMed (for comparison to other patients), and the Database of Genomic Variants (DGV) (for checking whether the CNV occurs in healthy individuals) are indispensable for the routine diagnostic workflow in the clinic [17-20].

Additional methods are being developed based on functional enrichment analysis of the genetic content of CNVs [21]. For example, CNVs are more likely to contribute to the patient's phenotype if they contain genes with temporal and spatial patterns of expression that are in line with the phenotype seen in the patient, or if they contain genes with mouse phenotypes similar to symptoms in the patient. Furthermore, it has been shown that CNVs identified in patients with idiopathic MR are more likely to be pathogenic if they contain genes with mouse orthologues that cause abnormal neuron morphology or neurodegenerative disease when disrupted [22,23]. More recently, it was shown that CNVs in patients with MR or

schizophrenia are enriched in miRNA genes with brain related functions compared to common CNVs [24,25].

Sequencing of breakpoints from chromosomal aberrations to decipher rearrangement structure and identify affected genes

Detection of structural variation using next-generation sequencing

The introduction of next-generation DNA sequencing has accelerated the discovery of SVs in the human genome. In a landmark paper by Korbel *et al.*, paired-end sequencing was introduced to identify more than thousand SVs in the human genome at unprecedented resolution [26]. Paired-end sequencing involves the mapping of pairs of sequence reads to the human reference genome [1]. These pairs are derived from the two ends of a single genomic DNA segment. A larger distance between the two reads that form a pair can be achieved by mate-pair sequencing [26,27].

Genomic breakpoints are detected following mapping of the read pairs to the reference genome. The mapping locations and orientations of the two reads will be concordant relative to each other if no breakpoint is present in between the two reads with respect to the reference genome. On the contrary, discordant mapping locations and orientations indicate a possible structural genomic change. Discordant read pair signatures that denote different types of structural variation are shown in Figure 1. Because the analysis of discordant reads is dependent on mapping of sequence reads in the vicinity of breakpoints, this approach is less feasible for recurrent chromosomal rearrangements.

Besides discordant read pair analysis, two other approaches are often used to detect structural genomic variation based on next-generation sequencing [1]. The analysis of the depth of sequence read coverage (DOC) is of great help to determine copy number status, while split read mapping gives insight into the precise breakpoint junction sequence.

The use of paired-end sequencing to fine-map rearrangement breakpoints in patients with MCA/MR

The first efforts to characterize genomic breakpoints of chromosomal aberrations in MCA/MR patients were performed by Chen *et al.* [28]. To limit sequencing costs, these authors used paired-end sequencing of derivative chromosomes isolated by flow sorting. Another strategy to limit sequencing costs and efforts concerns the sequence enrichment of breakpoint regions which have been previously established based on karyotyping or FISH studies [29,30]. Such methods involve the design of capture probes within breakpoint regions and subsequent enrichment of these regions using on-array or in-solution enrichment of genomic libraries.

With decreasing costs of next-generation sequencing, various studies have described the use of whole-genome sequencing to identify genomic breakpoints of balanced and unbalanced chromosomal rearrangements at nucleotide resolution [31-37]. Differences in approach in these studies mostly comprise the generation of genomic libraries with large (mate-pair) or short (paired-end) insert sizes. This has implications for the amount of sequencing reads that are needed to capture breakpoints: the smaller the insert-size of the library, the more sequencing reads are needed. The major conclusion from these studies is that underlying gene defects can be directly identified. In addition, the orientation of junction fragments provides a precise view of the rearrangement structure, which is impossible to reach by microarray or cytogenetic investigation only. This is particularly important for complex genomic rearrangements where multiple genomic segments have been rearranged, such as those caused by chromothripsis (Figure 2) [35-37]. In fact, sequencing has often revealed an unanticipated complexity of chromosomal aberrations [29,32,38], such as a high frequency of inversions associated with balanced chromosomal rearrangements [37]. The high resolution at which rearrangement breakpoints are mapped using next-generation sequencing requires new nomenclature for their clinical reporting [39].

Mapping disease genes using paired-end sequencing

Next generation sequencing of chromosomal breakpoints provides direct insight into the genomic elements – e.g. genes or regulatory DNA sequences - which are affected by the breakpoints. The precise definition of the genomic breakpoints and the local architecture of the chromosomal rearrangement is an essential first step towards the understanding of genetic disease mechanisms. For example, sequencing the breakpoints of a *de novo* translocation or inversion provides a direct view on the disrupted genes that may be relevant candidates genes for disease [29,31,33,34,40-45]. What is clearly hampering in most of these studies is the sporadic nature of the breakpoints in patients with a balanced rearrangement. Thus, it is difficult to establish a causal relationship between a gene and a clinical phenotype. Only in exceptional cases multiple translocation carriers were found, leading to firm associations between a gene and a phenotype [46-48]. An additional strategy to strengthen genotype-phenotype relationships involves the assessment of patients with copy number changes encompassing the same gene(s) as disrupted by the balanced rearrangement breakpoints [49-52]. The information derived from the sequencing of breakpoints from balanced translocations can also be of benefit to identify disease genes for microdeletion syndromes [52].
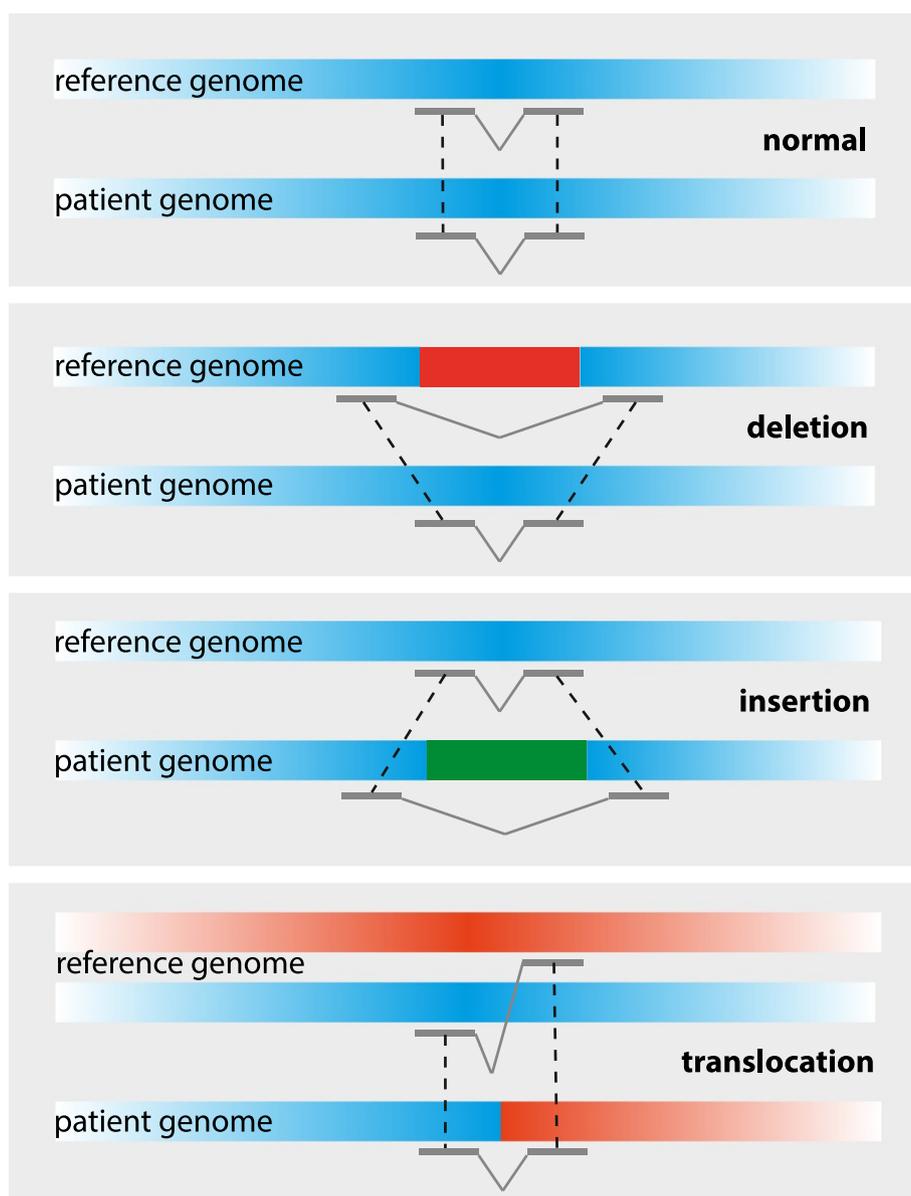
Figure 1 Examples of read pair signatures that indicate the presence of different types of chromosomal aberrations. Read pairs are indicated in grey. Read pair analysis uses information about the mapping position and orientation of sequence reads within a pair relative to each other.
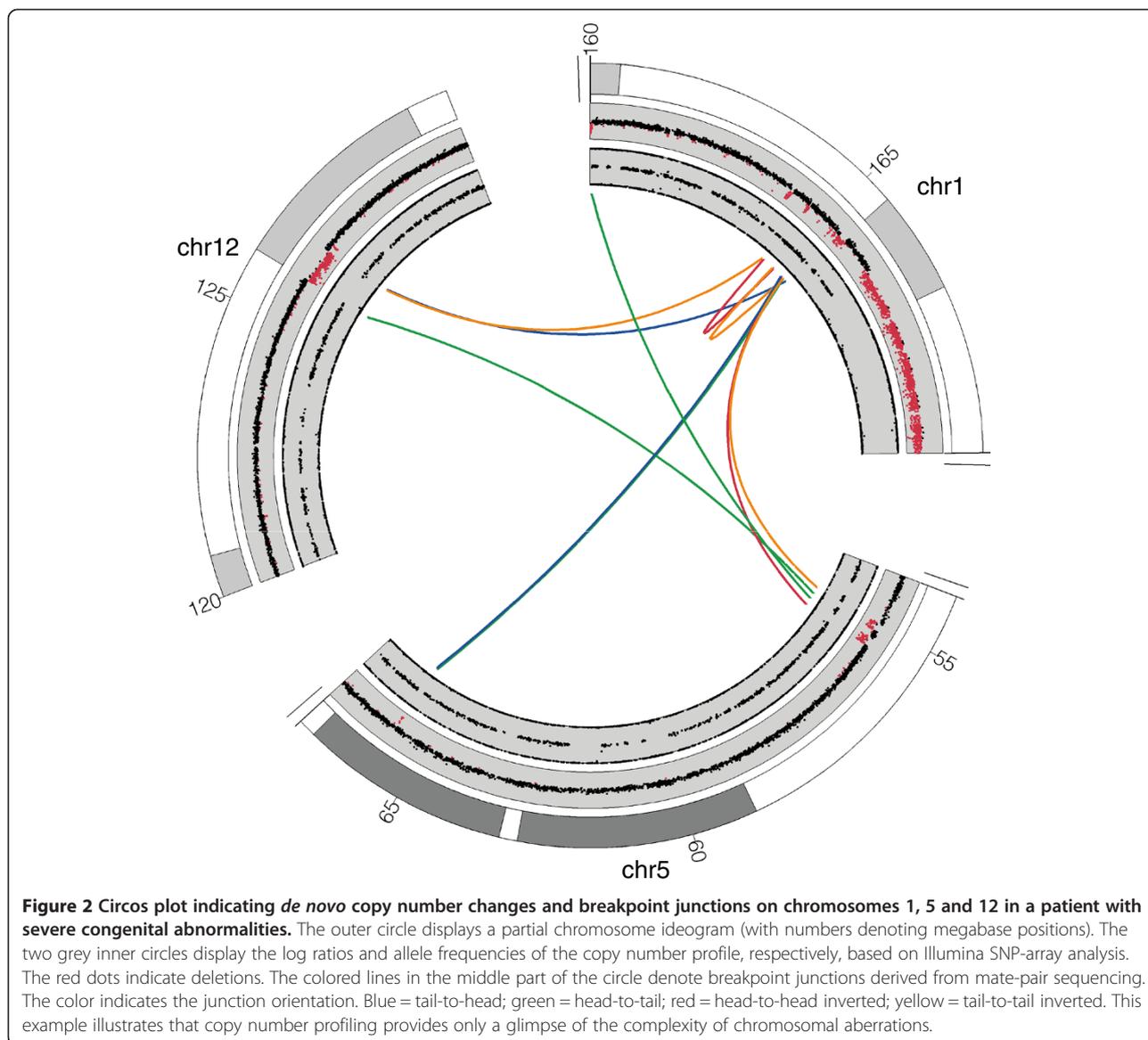
The molecular consequences of chromosome aberrations

The effects of breakpoints or copy number changes can be versatile and multiple effects may result from a single chromosomal aberration. The following paragraphs highlight different molecular consequences of chromosomal aberrations observed in patients with congenital phenotypes.

Dosage effects

Genes that are completely encompassed by a CNV undergo a dosage change, which is expected to coincide with a change in mRNA expression level. This may have severe effects, because for a substantial amount of genes a single functional copy is insufficient for normal gene function [53]. The impact of CNVs on mRNA expression differences has been observed at the genome-wide level in human cell lines [54]. Copy number status correlates with mRNA expression but also genes outside of CNV intervals show altered expression [55,56]. Combined analysis of CNV and transcriptome data from patients with autism spectrum disorder (ASD) revealed significant enrichment of dysregulated mRNA expression for genes within *de novo* CNVs when compared to common CNVs [57]. At the level of individual CNVs, studies have

Figure 2 Circos plot indicating *de novo* copy number changes and breakpoint junctions on chromosomes 1, 5 and 12 in a patient with severe congenital abnormalities. The outer circle displays a partial chromosome ideogram (with numbers denoting megabase positions). The two grey inner circles display the log ratios and allele frequencies of the copy number profile, respectively, based on Illumina SNP-array analysis. The red dots indicate deletions. The colored lines in the middle part of the circle denote breakpoint junctions derived from mate-pair sequencing. The color indicates the junction orientation. Blue = tail-to-head; green = head-to-tail; red = head-to-head inverted; yellow = tail-to-tail inverted. This example illustrates that copy number profiling provides only a glimpse of the complexity of chromosomal aberrations.

reported the effects of gene dosage on mRNA levels for e.g. VCF syndrome deletion, Williams Beuren syndrome deletion [58] and 16p11.2 deletion and duplication [59]. In the latter case, for virtually all genes spanned by the CNV interval, gene dosage correlated well with mRNA expression level.

To counteract the loss or gain of a chromosome (or chromosomal segment), a phenomenon known as dosage compensation can occur [60]. This effect may occur both at the level of gene transcription and translation [61]. Dosage compensation could explain the lack of correlation between copy number loss and mRNA expression level for some common CNVs in the human population [55]. Also, only 29% of genes on chromosome 21 in patients with Down syndrome show higher transcript levels, whereas the other 71% of genes show dosage compensation or

variable expression among different patients [62]. A major hurdle to precisely define effects of CNVs on gene expression is the general variability in mRNA expression between different individuals. Therefore, a matched control experiment is often not available. A recent and very elegant study has overcome this limitation and reported a unique case: a monozygotic twin pair discordant for trisomy 21. Genome-wide differential expression patterns between the twins uncovered the presence of gene expression dysregulation domains (GEDDs), which contain either up- or downregulated genes [63].

Gene fusions

Gene fusions (also known as chimeric genes) can occur when two genomic segments – each containing (part of) a gene - join together and form a novel protein or a new

promoter-gene fusion. Gene fusions are well studied in cancer, where they form active subjects of positive selection during tumor development [64]. Yet, in patients with MCA/MR the role of gene fusion is largely unclear. A few reports have characterized gene fusions associated with congenital disease phenotypes, often resulting from a balanced translocation. The gene fusions can be classified into several different classes. First, bona-fide in-frame gene fusions leading to novel, chimeric proteins can arise [65,66]. In a patient with mental retardation, ataxia and atrophy of the brain, a *PAFAH1B3-CLK2* fusion was found, but a gain-of-function effect could not be observed [65]. The most likely explanation for the phenotype in this patient was the truncation (functional hemizygosity) of *PAFAH1B3*. Functional hemizygosity was also suggested for an in-frame *TNS3-FGFR1* gene fusion, which led to impaired FGFR1 function [67]. In other patients novel fusion proteins were observed, which consisted of part of a known protein coupled to a novel open reading frame that lacks homology to known proteins [68-70]. A striking negative effect on mitochondrial function was observed for novel *DISC1* fusions, which involve part of the DISC1 protein coupled to a novel peptide sequence, possibly accounting for psychiatric disease in a large family [71]. A second possibility entails the fusion of a promoter to the coding regions from another gene. However, this has not yet been described for MCA/MR patients.

Two large-scale studies have used CNV datasets from patients with autism spectrum disorders (ASD) and schizophrenia, respectively [72,73]. An increased frequency of brain-expressed in-frame gene fusions was found in schizophrenia cases, but not in patients with ASD. Two fusion genes in the schizophrenia cohort displayed different subcellular locations compared to their parental genes, suggesting a role in disease etiology.

In all cases described above, the observed gene fusions were unique to a single patient, precluding the establishment of a causal relationship to the disease phenotype. Clearly, determining the precise role of gene fusion in congenital disease requires detection of breakpoints at nucleotide resolution, identification of recurrent fusions and in-depth functional characterization of individual cases.

Deregulation of gene expression through dislocation and disruption of non-coding elements

In addition to direct disruption of coding sequences through a CNV or SV breakpoint, chromosomal rearrangements may also affect regulatory elements, which in turn alter gene transcription. In fact, the altered regulation caused by disruption of regulatory elements may result in subtler gene expression changes - such as during specific developmental stages or in specific tissues –

and not be a full loss of function (for review see: [74]). This leads to different phenotypes depending on the type of regulatory element that is affected or the severity of the effect. For example, Pierre Robin sequence is caused by chromosomal aberrations affecting noncoding enhancer elements upstream of the *SOX9* gene [75,76]. The deregulation of *SOX9* is determining the final phenotypic outcome: translocation breakpoints close to *SOX9* typically result in campomelic dysplasia, whereas those further upstream cause acampomelic campomelic dysplasia [77]. Overall, the skeletal phenotypes tend to be less severe with increased distance between breakpoints and *SOX9* [77]. Besides distance to a gene, a second factor of importance for the molecular consequences of the disruption of non-coding elements by chromosomal aberrations is the type of element that is disrupted [74]. Loss of enhancer function in the case of Pierre Robin sequence, causes reduced *SOX9* expression. In contrast, deletion of a small genomic silencer element just upstream of the delta-globin gene, provokes fetal hemoglobin expression [78]. Another mechanism concerns the relocation of enhancer elements relative to coding regions, causing them to control activity of new genes. Such a position effect has been observed in a patient with an inversion that dislocates the *SHH* gene and places it under the control of a highly active limb bud enhancer element [79]. Systematic survey of deletions from the DECIPHER database has revealed that 11.8% of patient phenotypes could best be explained by enhancer adoption rather than mere dosage effects of genes within the deletion interval [80]. These and other effects of chromosomal aberrations on noncoding DNA elements have only recently been recognized, but are essential for diagnostic interpretation [74].

Unmasking of recessive mutations by chromosomal deletions

Whenever a deletion occurs, any recessive mutation on the non-deleted homologous chromosome would become unmasked, and could contribute to the phenotype. Indeed, there is anecdotal evidence for the unmasking of a mutated, pathogenic allele on the homologous, non-deleted chromosome [81]. This phenomenon may be one of the explanations for the phenotypic variability between unrelated patients with identical deletions. The best studied case is the 3 Mb 22q11.21 deletion mediated by NAHR between LCRs A and D that occurs *de novo* in about 90% of patients with DiGeorge/VCF syndrome [82]. For example, mutations affecting the *GP1BB* gene have been identified in several patients with features of both DiGeorge/VCF syndrome and Bernard-Soulier syndrome, a rare autosomal recessive bleeding disorder [83,84]. Other examples include the unmasking of a *SCARF2* splice site mutation in a patient with DiGeorge/VCF syndrome and additional features of Van den Ende-Gupta syndrome,

such as arachnocamptodactily and typical facial anomalies such as blepharophimosis, beaked nose and malar hypoplasia [85], and the unmasking of *SNAP29* frameshift mutations in two patients who also had features of CEDNIK syndrome (cerebral dysgenesis, neuropathy, ichthyosis and keratoderma) [86]. It should be noted that these patients were investigated in-depth because they displayed features of more than one syndrome or condition. This suggests that such patients are rare, exceptional cases. This is supported by a the lack of evidence for the unmasking of selected candidate genes in other recurrent deletion syndromes, such as the 1q21.1 deletion associated with thrombocytopenia-absent radius syndrome [87], the 16p13.11 deletion associated with MR and autism [88], and the 16p11.2 deletion associated with autism [89].

There are also some examples of inherited deletions that unmask a recessive pathogenic allele [81]. In theory, this mechanism could explain the phenotypic difference when there are affected and unaffected carriers of the deletion in a family. This situation occurs, for example, when there is an affected child with the deletion and an unaffected, apparently healthy carrier parent. This theory was tested in a cohort of 20 familial deletion cases [81]. Only in one case a 16.0-31.7 kb deletion containing only the *HSBP1* gene in 16q23.3 inherited from one parent was unmasked by a 2.2 Mb deletion inherited from the other parent.

In summary, the unmasking of pathogenic mutations may be a rare phenomenon, both in recurrent, *de novo* deletions and in familial deletions. It is possible, however, that this mechanism operates in more subtle ways, i.e. by unmasking sequence variants that occur in the general population. For example, the risk of developing schizophrenia and other neuropsychiatric conditions in 22q11.2 deletion patients is dependent on DNA-sequence polymorphisms on the intact chromosome that affect the function or expression of genes such as *COMT*, *PIK4CA*, and *GNB1L* [90]. There is a clear need for systematic studies of genetic variation in the non-deleted genes, including miRNA genes [91] in large cohorts to assess the contribution of unmasking to our understanding of the phenotypic variation between patients with identical deletions.

Strategies for functional characterization of chromosome aberrations

Extensive functional characterization of chromosomal aberrations is needed to identify the critical molecular determinants of a patient's phenotype. It may well be that some of the effects of breakpoints on gene function are not harmful and can be regarded as passenger events, while other effects are predominantly driving disease. To discriminate between passenger and driver effects a variety of approaches can be applied, ranging from large-scale genomics to functional studies in model organisms.

Large-scale genomics technologies

The molecular effects of structural genomic variations may expand far beyond the genomic region that is directly affected by breakpoints or copy number change. Genome-wide molecular profiling is helpful to pinpoint consequences of SVs in an unbiased fashion. For example, mRNA expression analysis has been performed in families including patients with ASD to identify autism-susceptibility genes associated with CNVs [57]. One important limitation of mRNA expression studies is the tissue type chosen for investigation. In most cases blood is used as a proxy for expression in the brain. A second limitation is the lack of a proper control dataset. Healthy family members can be used as a control, but differences in genetic background contribute already substantially to gene-expression differences. Indeed, the study by Luo *et al*. [57] showed that significantly misexpressed genes are not restricted to probands, but are equally prevalent in unaffected siblings. However, gene ontology analysis revealed an enrichment of misexpressed genes in neural-related pathways and misexpressed genes are localized within pathogenic CNV intervals. Finally, the effects of CNVs on gene expression may not be extending throughout lifetime, as studies in mouse revealed temporary compensatory loops of brain-expressed genes at specific developmental time-points [92].

Application of mRNA sequencing to a cohort of patients with reciprocal 16p11.2 duplication or deletion revealed most prominent changes within CNV intervals, but also several effects in *cis* and *trans* [59]. These effects were associated with changes in long-range physical interactions identified by Hi-C chromosome conformation capture technology. Chromosome conformation capture methods provide insight into nuclear organization and identify chromosomal regions that physically interact with each other [93]. SVs can change the organization of chromatin and thereby affect transcription of nearby genes, as was recently shown for deletion of the Williams-Beuren Syndrome region [94].

Finally, the impact of chromosomal rearrangements and SVs on regulatory regions in the genome can be identified by the analysis of histone modifications that mark active promoters and enhancers, such as H3K4Me3 and H3K27Ac, respectively [95]. Effects of SVs on histone modifications have hardly been studied, but may provide important information on global changes in gene regulation [94].

Engineering rearrangements using genome editing technology

Genome engineering with programmable nucleases provides an extremely powerful approach for introducing specific mutations into human cells or model organisms [96]. These systems are often used to enable gene disruption by

generating a nuclease-mediated double-stranded break, leading to frame-shifts or other disruptive mutations. However, expression of two site-specific nucleases allows engineering of chromosomal rearrangements, such as deletions, inversions and translocations [97-100]. For example, cancer–related translocations and inversions leading to fusion genes were mimicked in human cells [99]. These technological breakthroughs provide unique opportunities to engineer specific chromosomal rearrangements that occur in patients with MCA/MR. Such an approach circumvents measurement noise resulting from differences in genetic background, which has, for example, hampered the detection of gene-expression changes in case–control studies for various microdeletion or microduplication syndromes [57,59,101,102]. Large-scale genomics analyses or specific functional assays on cells with engineered rearrangements and controls will reveal detailed insight into genome-wide effects. Furthermore, introduction of chromosomal aberrations in induced pluripotent stem (iPS) cell lines [103] or primary neuronal cell cultures [104] allows modeling facets of human disease, such as neuronal differentiation. Finally, the effects of chromosomal rearrangements on non-coding elements in the genome can best be studied by deleting an entire locus by cutting with two site-specific nucleases.

Understanding functional consequences of chromosomal aberrations in model organisms

Functional characterization of chromosomal aberrations in model organisms can be of high value to understand disease mechanisms or strengthen genotype-phenotype associations. Several mouse models have been generated to investigate recurrent microdeletion and microduplication syndromes. An elegant study described mouse models with CNVs that affect regions syntenic to human chromosomal band 17p11.2 [105]. Deletions and duplications in 17p11.2 lead to Smith-Magenis and Potocki-Lupski syndrome, respectively. These mouse models recapitulate much of the phenotypes observed in human subjects. The dosage of the *RAI1* gene appears to be most critical to these phenotypes [106,107]. The effects of the CNVs were further studied in several mouse tissues revealing that many expression changes map to the engineered CNV intervals. Strikingly, restoring copy number by generating a mouse strain with a deletion and a duplication chromosome did not completely rescue some neurobehavioral phenotypes. Thus, a disturbance of local genomic architecture also plays a role in disease, in addition to gene dosage effects. Similar studies in mouse models of the 16p11.2 deletion and duplication were instrumental to map the changes in expression networks that account for the ASD phenotypes that are associated with these CNVs [59].

Zebrafish (*Danio rerio*) form an alternative to mouse models and are exceptionally suited for studying early embryonic development. Furthermore, zebrafish embryos can be easily manipulated allowing quick evaluation of many genes and mutations [108]. Overexpression in zebrafish embryos of genes within the 16p11.2 CNV interval revealed *KCTD13* as the major driver of the microcephaly phenotype seen in patients with the 16p11.2 duplications [109]. On the contrary, knockdown of *KCTD13* resulted in macrocephaly, which is typical for 16p11.2 deletion carriers. These results indicate that a single gene may form the primary driver of disease even though the entire 16p11.2 CNV region encompasses 29 genes. Thus, screening of large numbers of genes that are affected by dosage changes or breakpoints of chromosomal aberrations is extremely informative to understand disease mechanisms. A similar approach of functional testing in zebrafish was used to dissect the phenotypic determinants of the 8q24.3 copy number variant [110]. This revealed that two genes, *SCRIB* and *PUF60*, are primary drivers of disease. These selected examples demonstrate that zebrafish appears a very powerful model system to systematically dissect the functional consequences of chromosomal aberrations.

Conclusions
The genetics clinic of the future

The resolution, speed and breadth at which a patient's genome can be scanned for pathogenic changes have increased dramatically over the last decade. The use of microarrays as the initial genetic test is likely to change in the coming years. By the application of whole exome sequencing, pathogenic gains and losses can be identified on a genome-wide scale, in parallel to pathogenic mutations in (protein coding) genes [111-113]. Although whole exome sequencing has lead to rapid discovery of disease genes, trends are shifting towards whole genome sequencing, as this is now an affordable method to identify all classes of genetic changes across the entire genome. In a recent study, whole genome sequencing was used to map *de novo* genetic changes in a group of patients with severe intellectual disability, which had not received a diagnosis based on previous CNV profiling and exome sequencing [114]. In this cohort additional mutations and CNVs were identified in coding regions, enabling a genetic diagnosis in 20 out of 50 patients.

In the genetics clinic of the future, whole genome sequencing to detect pathogenic SVs and CNVs will become a routine analysis. The application of microscopy will then be limited to the determination of the structure of chromosomal rearrangements, such as, for example, the discrimination between trisomy-21 and Robertsonian translocations involving chromosome 21 in Down syndrome, as in 1959. Also the detection of low-level mosaicism for aneuploidy of entire or rearranged chromosomes

may remain dependent on microscopy in the near future. A suspicion of mosaicism should be raised for patients with normal NGS findings who display asymmetric body features, and "deep karyotyping" of multiple tissues may be required to arrive at a clinical diagnosis [115].

The major challenge will be to address the functional consequences of pathogenic variants by large-scale genomic approaches and systematic studies in model systems. These efforts are needed to fully understand the molecular effects of chromosomal aberrations and their role in disease etiology. Detailed insight in disease mechanisms will improve diagnostics, allow for pre-symptomatic screening for complications, support short and long-term prognosis and design of targeted and personalized medical treatment strategies.

Abbreviations
ASD: Autism spectrum disorder; CEDNIK: Cerebral dysgenesis, neuropathy, ichthyosis and keratoderma; CGH: Comparative genomic hybridization; CNV: Copy number variation; DECIPHER: DatabasE of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources; DGV: Database of Genomic Variants; DOC: Depth of coverage; ECARUCA: European Cytogeneticists Association Register of Unbalanced Chromosome Aberrations; FISH: Fluorescence in situ hybridization; GEDD: Gene expression dysregulation domains; IPS: Induced pluripotent stem; ISCA: The International Standards for Cytogenomic Arrays Consortium; LCR: Low copy repeat; MCA/MR: Multiple congenital abnormalities and/or mental retardation; MMBIR: Microhomology-mediated break-induced replication; MR: Mental retardation; NAHR: Non-allelic homologous recombination; NHEJ: Nonhomologous end-joining; SNP: Single nucleotide polymorphism; SV: Structural variation; VCF: Velo cardio facial.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
WPK and RH jointly wrote this review. WPK designed and generated Figures 1 and 2. All authors read and approved the final manuscript.

Author details
[1]Department of Medical Genetics, Center for Molecular Medicine, University Medical Center Utrecht, P.O. Box 85060, 3508 AB Utrecht, The Netherlands. [2]Department of Medical Genetics, Genome Diagnostics, P.O. Box 850903508 AB Utrecht, The Netherlands.

References
1. Alkan C, Coe BP, Eichler EE: **Genome structural variation discovery and genotyping.** *Nat Rev Genet* 2011, **12**:363–376.
2. Mills RE, Walter K, Stewart C, Handsaker RE, Chen K, Alkan C, Abyzov A, Yoon SC, Ye K, Cheetham RK, Chinwalla A, Conrad DF, Fu Y, Grubert F, Hajirasouliha I, Hormozdiari F, Iakoucheva LM, Iqbal Z, Kang S, Kidd JM, Konkel MK, Korn J, Khurana E, Kural D, Lam HYK, Leng J, Li R, Li Y, Lin C-Y, Luo R, et al: **Mapping copy number variation by population-scale genome sequencing.** *Nature* 2011, **470**:59–65.
3. Itsara A, Wu H, Smith JD, Nickerson DA, Romieu I, London SJ, Eichler EE: **De novo rates and selection of large copy number variation.** *Genome Res* 2010, **20**:1469–1481.
4. Campbell CD, Eichler EE: **Properties and rates of germline mutations in humans.** *Trends Genet* 2013, **29**:575–584.
5. Kong A, Frigge ML, Masson G, Besenbacher S, Sulem P, Magnusson G, Gudjonsson SA, Sigurdsson A, Jonasdottir A, Jonasdottir A, Wong WSW, 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**:471–475.
6. Michaelson JJ, Shi Y, Gujral M, Zheng H, Malhotra D, Jin X, Jian M, Liu G, Greer D, Bhandari A, Wu W, Corominas R, Peoples Á, Koren A, Gore A, Kang S, Lin GN, Estabillo J, Gadomski T, Singh B, Zhang K, Akshoomoff N, Corsello C, McCarroll S, Iakoucheva LM, Li Y, Wang J, Sebat J: **Whole-genome sequencing in autism identifies hot spots for de novo germline mutation.** *Cell* 2012, **151**:1431–1442.
7. Cooper GM, Coe BP, Girirajan S, Rosenfeld JA, Vu 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, Johnson K, Rehder C, Ballif BC, Shaffer LG, Eichler EE: **A copy number variation morbidity map of developmental delay.** *Nat Genet* 2011, **43**:838–846.
8. Hehir-Kwa JY, Rodriguez-Santiago B, Vissers LE, De Leeuw N, Pfundt R, Buitelaar JK, Perez-Jurado LA, Veltman JA: **De novo copy number variants associated with intellectual disability have a paternal origin and age bias.** *J Med Genet* 2011, **48**:776–778.
9. Stankiewicz P, Lupski JR: **Structural variation in the human genome and its role in disease.** *Annu Rev Med* 2010, **61**:437–455.
10. Lejeune J, Gautier M, Turpin R: **A study of somatic chromosomes in nine infants with mongolism.** *CR Acad Sci* 1959, **248**:1721–1722.
11. Hochstenbach R, van Binsbergen E, Engelen J, Nieuwint A, Polstra A, Poddighe P, Ruivenkamp C, Sikkema-Raddatz B, Smeets D, Poot M: **Array analysis and karyotyping: Workflow consequences based on a retrospective study of 36,325 patients with idiopathic developmental delay in the Netherlands.** *Eur J Med Genet* 2009, **52**:161–169.
12. Shaffer LG, Bejjani BA, Torchia B, Kirkpatrick S, Coppinger J, Ballif BC: **The identification of microdeletion syndromes and other chromosome abnormalities: Cytogenetic methods of the past, new technologies for the future.** *Am J Med Genet Part C Semin Med Genet* 2007, **145**:335–345.
13. Liu P, Carvalho CMB, Hastings PJ, Lupski JR: **Mechanisms for recurrent and complex human genomic rearrangements.** *Curr Opin Genet Dev* 2012, **22**:211–220.
14. Hastings PJ, Lupski JR, Rosenberg SM, Ira G: **Mechanisms of change in gene copy number.** *Nat Rev Genet* 2009, **10**:551–564.
15. Zhang F, Gu W, Hurles ME, Lupski JR: **Copy number variation in human health, disease, and evolution.** *Annu Rev Genomics Hum Genet* 2009, **10**:451–481.
16. Hastings PJ, Ira G, Lupski JR: **A microhomology-mediated break-induced replication model for the origin of human copy number variation.** *PLoS Genet* 2009, **5**:e1000327.
17. Miller DT, Adam MP, Aradhya 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, Kaminsky EB, Kok K, Krantz ID, Kuhn RM, Lee C, Ostell JM, Rosenberg C, Scherer SW, Spinner NB, Stavropoulos DJ, Tepperberg JH, Thorland EC, Vermeesch JR, Waggoner 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.
18. Riggs ER, Church DM, Hanson K, Horner VL, Kaminsky EB, Kuhn RM, Wain KE, Williams ES, Aradhya S, Kearney HM, Ledbetter DH, South ST, Thorland EC, Martin CL: **Towards an evidence-based process for the clinical interpretation of copy number variation.** *Clin Genet* 2012, **81**:403–412.
19. Kearney HM, Thorland EC, Brown KK, Quintero-Rivera F, South ST: **American College of Medical Genetics standards and guidelines for interpretation and reporting of postnatal constitutional copy number variants.** *Genet Med* 2011, **13**:680–685.
20. De Leeuw N, Dijkhuizen T, Hehir-Kwa JY, Carter NP, Feuk L, Firth HV, Kuhn RM, Ledbetter DH, Martin CL, van Ravenswaaij CMA, Scherer SW, Shams S, Van Vooren S, Sijmons R, Swertz M, Hastings R: **Diagnostic interpretation of array data using public databases and internet sources.** *Hum Mutat* 2012, **33**:930–940.
21. Webber C: **Functional enrichment analysis with structural variants: pitfalls and strategies.** *Cytogenet Genome Res* 2011, **135**:277–285.
22. Webber C, Hehir-Kwa JY, Nguyen DQ, De Vries BBA, Veltman JA, Ponting CP: **Forging links between human mental retardation-associated CNVs and mouse gene knockout models.** *PLoS Genet* 2009, **5**:e1000531.

23. Hehir-Kwa JY, Wieskamp N, Webber C, Pfundt R, Brunner HG, Gilissen C, de Vries BBA, Ponting CP, Veltman JA: Accurate distinction of pathogenic from benign CNVs in mental retardation. *PLoS Comput Biol* 2010, **6**:e1000752.

24. Qiao Y, Badduke C, Mercier E, Lewis SME, Pavlidis P, Rajcan-Separovic E: miRNA and miRNA target genes in copy number variations occurring in individuals with intellectual disability. *BMC Genomics* 2013, **14**:544.

25. Warnica W, Merico D, Costain G, Alfred SE, Wei J, Marshall CR, Scherer SW, Bassett AS: Copy number variable micrornas in schizophrenia and their neurodevelopmental gene targets. *Biol Psychiatry* 2014. doi:10.1016/j.biopsych.2014.05.011

26. Korbel JO, Urban AE, Affourtit JP, Godwin B, Grubert F, Simons JF, Kim PM, Palejev D, Carriero NJ, Du L, Taillon BE, Chen Z, Tanzer A, Saunders a CE, Chi J, Yang F, Carter NP, Hurles ME, Weissman SM, Harkins TT, Gerstein MB, Egholm M, Snyder M: Paired-End Mapping Reveals Extensive Structural Variation in the Human Genome. *Science* 2007, **318**:420–426.

27. Hanscom C, Talkowski M: Design of large-insert jumping libraries for structural variant detection using Illumina sequencing. *Curr Protoc Hum Genet* 2014, **80**:1–9.

28. Chen W, Kalscheuer V, Tzschach A, Menzel C, Ullmann R, Schulz MH, Erdogan F, Li N, Kijas Z, Arkesteijn G, Pajares IL, Goetz-Sothmann M, Heinrich U, Rost I, Dufke A, Grasshoff U, Glaeser B, Vingron M, Ropers HH: Mapping translocation breakpoints by next-generation sequencing. *Genome Res* 2008, **18**:1143–1149.

29. Sobreira NLM, Gnanakkan V, Walsh M, Marosy B, Wohler E, Thomas G, Hoover-Fong JE, Hamosh A, Wheelan SJ, Valle D: Characterization of complex chromosomal rearrangements by targeted capture and next-generation sequencing. *Genome Res* 2011, **21**:1720–1727.

30. Talkowski ME, Ernst C, Heilbut A, Chiang C, Hanscom C, Lindgren A, Kirby A, Liu S, Muddukrishna B, Ohsumi TK, Shen Y, Borowsky M, Daly MJ, Morton CC, Gusella JF: Next-generation sequencing strategies enable routine detection of balanced chromosome rearrangements for clinical diagnostics and genetic research. *Am J Hum Genet* 2011, **88**:469–481.

31. Schluth-Bolard C, Labalme A, Cordier M-P, Till M, Nadeau G, Tevissen H, Lesca G, Boutry-Kryza N, Rossignol S, Rocas D, Dubruc E, Edery P, Sanlaville D: Breakpoint mapping by next generation sequencing reveals causative gene disruption in patients carrying apparently balanced chromosome rearrangements with intellectual deficiency and/or congenital malformations. *J Med Genet* 2013, **50**:144–150.

32. Vergult S, Van Binsbergen E, Sante T, Nowak S, Vanakker O, Claes K, Poppe B, Van der Aa N, van Roosmalen MJ, Duran K, Tavakoli-Yaraki M, Swinkels M, van den Boogaard M-J, van Haelst M, Roelens F, Speleman F, Cuppen E, Mortier G, Kloosterman WP, Menten B: Mate pair sequencing for the detection of chromosomal aberrations in patients with intellectual disability and congenital malformations. *Eur J Hum Genet* 2014, **22**:652–659.

33. Utami KH, Hillmer AM, Aksoy I, Chew EGY, Teo ASM, Zhang Z, Lee CWH, Chen PJ, Seng CC, Ariyaratne PN, Rouam SL, Soo LS, Yousoof S, Prokudin I, Peters G, Collins F, Wilson M, Kakakios A, Haddad G, Menuet A, Perche O, Tay SKH, Sung KWK, Ruan X, Ruan Y, Liu ET, Briault S, Jamieson RV, Davila S, Cacheux V: Detection of chromosomal breakpoints in patients with developmental delay and speech disorders. *PLoS One* 2014, **9**:e90852.

34. Rasmussen MB, Nielsen JV, Lourenço CM, Melo JB, Halgren C, Geraldi CVL, Marques W, Rodrigues GR, Thomassen M, Bak M, Hansen C, Ferreira SI, Venâncio M, Henriksen KF, Lind-Thomsen A, Carreira IM, Jensen NA, Tommerup N: Neurodevelopmental disorders associated with dosage imbalance of ZBTB20 correlate with the morbidity spectrum of ZBTB20 candidate target genes. *J Med Genet* 2014, **51**:605–613.

35. Kloosterman WP, Guryev V, van Roosmalen M, Duran KJ, de Bruijn E, Bakker SCM, Letteboer T, van Nesselrooij H, Hochstenbach R, Poot M, Cuppen E: Chromothripsis as a mechanism driving complex de novo structural rearrangements in the germline. *Hum Mol Genet* 2011, **20**:1916–1924.

36. Nazaryan L, Stefanou EG, Hansen C, Kosyakova N, Bak M, Sharkey FH, Mantziou T, Papanastasiou AD, Velissariou V, Liehr T, Syrrou M, Tommerup N: The strength of combined cytogenetic and mate-pair sequencing techniques illustrated by a germline chromothripsis rearrangement involving FOXP2. *Eur J Hum Genet* 2013, **22**:338–343.

37. Chiang C, Jacobsen JC, Ernst C, Hanscom C, Heilbut A, Blumenthal I, Mills RE, Kirby A, Lindgren AM, Rudiger SR, McLaughlan CJ, Bawden CS, Reid SJ, Faull RLM, Snell RG, Hall IM, Shen Y, Ohsumi TK, Borowsky ML, Daly MJ, Lee C, Morton CC, MacDonald ME, Gusella JF, Talkowski ME: Complex reorganization

38. Kloosterman WP, Tavakoli-Yaraki M, Van Roosmalen MJ, Van Binsbergen E, Renkens I, Duran K, Ballarati L, Vergult S, Giardino D, Hansson K, Ruivenkamp CAL, Jager M, Van Haeringen A, Ippel EF, Haaf T, Passarge E, Hochstenbach R, Menten B, Larizza L, Guryev V, Poot M, Cuppen E: Constitutional chromothripsis rearrangements involve clustered double-stranded DNA breaks and nonhomologous repair mechanisms. *Cell Rep* 2012, **1**:648–655.

39. Ordulu Z, Wong KE, Currall BB, Ivanov AR, Pereira S, Althari S, Gusella JF, Talkowski ME, Morton CC: Describing sequencing results of structural chromosome rearrangements with a suggested next-generation cytogenetic nomenclature. *Am J Hum Genet* 2014, **94**:695–709.

40. Blake J, Riddell A, Theiss S, Gonzalez AP, Haase B, Jauch A, Janssen JWG, Ibberson D, Pavlinic D, Moog U, Benes V, Runz H: Sequencing of a patient with balanced chromosome abnormalities and neurodevelopmental disease identifies disruption of multiple high risk loci by structural variation. *PLoS One* 2014, **9**:e90894.

41. Lango Allen H, Caswell R, Xie W, Xu X, Wragg C, Turnpenny PD, Turner CLS, Weedon MN, Ellard S: Next generation sequencing of chromosomal rearrangements in patients with split-hand/split-foot malformation provides evidence for DYNC1I1 exonic enhancers of DLX5/6 expression in humans. *J Med Genet* 2014, **51**:264–267.

42. Hooper SD, Johansson ACV, Tellgren-Roth C, Stattin E-L, Dahl N, Cavelier L, Feuk L: Genome-wide sequencing for the identification of rearrangements associated with Tourette syndrome and obsessive-compulsive disorder. *BMC Med Genet* 2012, **13**:123.

43. Slade I, Stephens P, Douglas J, Barker K, Stebbings L, Abbaszadeh F, Pritchard-Jones K, Cole R, Pizer B, Stiller C, Vujanic G, Scott RH, Stratton MR, Rahman N: Constitutional Translocation Breakpoint Mapping by Genome-Wide Paired-End Sequencing Identifies HACE1 as a Putative Wilms Tumour Susceptibility Gene. *J Med Genet* 2010, **47**:342–347.

44. Utami KH, Winata CL, Hillmer AM, Aksoy I, Long HT, Liany H, Yan ECG, Mathavan S, Hong STK, Korzh V, Sarda P, Davila S, Cacheux V: Impaired development of neural-crest cell derived organs and intellectual disability caused by mED13L haploinsufficiency. *Hum Mutat* 2014.

45. Talkowski ME, Ordulu Z, Pillalamarri V, Benson CB, Blumenthal I, Connolly S, Hanscom C, Hussain N, Pereira S, Picker J, Rosenfeld JA, Shaffer LG, Wilkins-Haug LE, Gusella JF, Morton CC: Clinical diagnosis by whole-genome sequencing of a prenatal sample. *N Engl J Med* 2012, **367**:2226–2232.

46. Kalscheuer VM, FitzPatrick D, Tommerup N, Bugge M, Niebuhr E, Neumann LM, Tzschach A, Shoichet SA, Menzel C, Erdogan F, Arkesteijn G, Ropers HH, Ullmann R: Mutations in autism susceptibility candidate 2 (AUTS2) in patients with mental retardation. *Hum Genet* 2007, **121**:501–509.

47. Kim HG, Kishikawa S, Higgins AW, Seong IS, Donovan DJ, Shen Y, Lally E, Weiss LA, Najm J, Kutsche K, Descartes M, Holt L, Braddock S, Troxell R, Kaplan L, Volkmar F, Klin A, Tsatsanis K, Harris DJ, Noens I, Pauls DL, Daly MJ, MacDonald ME, Morton CC, Quade BJ, Gusella JF: Disruption of neurexin 1 associated with autism spectrum disorder. *Am J Hum Genet* 2008, **82**:199–207.

48. Kim HG, Kim HT, Leach NT, Lan F, Ullmann R, Silahtaroglu A, Kurth I, Nowka A, Seong IS, Shen Y, Talkowski ME, Ruderfer D, Lee JH, Glotzbach C, Ha K, Kjaergaard S, Levin AV, Romeike BF, Kleefstra T, Bartsch O, Elsea SH, Jabs EW, MacDonald ME, Harris DJ, Quade BJ, Ropers HH, Shaffer LG, Kutsche K, Layman LC, Tommerup N, *et al*: Translocations disrupting PHF21A in the Potocki-Shaffer-syndrome region are associated with intellectual disability and craniofacial anomalies. *Am J Hum Genet* 2012, **91**:56–72.

49. Lindgren AM, Hoyos T, Talkowski ME, Hanscom C, Blumenthal I, Chiang C, Ernst C, Pereira S, Ordulu Z, Clericuzio C, Drautz JM, Rosenfeld JA, Shaffer LG, Velsher L, Pynn T, Vermeesch J, Harris DJ, Gusella JF, Liao EC, Morton CC: Haploinsufficiency of KDM6A is associated with severe psychomotor retardation, global growth restriction, seizures and cleft palate. *Hum Genet* 2013, **132**:537–552.

50. Talkowski ME, Maussion G, Crapper L, Rosenfeld JA, Blumenthal I, Hanscom C, Chiang C, Lindgren A, Pereira S, Ruderfer D, Diallo AB, Lopez JP, Turecki G, Chen ES, Gigek C, Harris DJ, Lip V, An Y, Biagioli M, MacDonald ME, Lin M, Haggarty SJ, Sklar P, Purcell S, Kellis M, Schwartz S, Shaffer LG, Natowicz MR, Shen Y, Morton CC, *et al*: Disruption of a large intergenic noncoding RNA in subjects with neurodevelopmental disabilities. *Am J Hum Genet* 2012, **91**:1128–1134.

51. Hodge JC, Mitchell E, Pillalamarri V, Toler TL, Bartel F, Kearney HM, Zou YS, Tan WH, Hanscom C, Kirmani S, Hanson RR, Skinner SA, Rogers RC, Everman

DB, Boyd E, Tapp C, Mullegama SV, Keelean-Fuller D, Powell CM, Elsea SH, Morton CC, Gusella JF, DuPont B, Chaubey A, Lin AE, Talkowski ME: **Disruption of MBD5 contributes to a spectrum of psychopathology and neurodevelopmental abnormalities.** *Mol Psychiatry* 2014, 19:368–379.

52. Talkowski ME, Rosenfeld JA, Blumenthal I, Pillalamarri V, Chiang C, Heilbut A, Ernst C, Hanscom C, Rossin E, Lindgren AM, Pereira S, Ruderfer D, Kirby A, Ripke S, Harris DJ, Lee JH, Ha K, Kim HG, Solomon BD, Gropman AL, Lucente D, Sims K, Ohsumi TK, Borowsky ML, Loranger S, Quade B, Lage K, Miles J, Wu BL, Shen Y, et al: **Sequencing chromosomal abnormalities reveals neurodevelopmental loci that confer risk across diagnostic boundaries.** *Cell* 2012, 149:525–537.

53. Huang N, Lee I, Marcotte EM, Hurles ME: **Characterising and predicting haploinsufficiency in the human genome.** *PLoS Genet* 2010, 6:1–11.

54. Stranger BE, Forrest MS, Dunning M, Ingle CE, Beazley C, Thorne N, Redon R, Bird CP, de Grassi A, Lee C, Tyler-Smith C, Carter N, Scherer SW, Tavaré S, Deloukas P, Hurles ME, Dermitzakis ET: **Relative impact of nucleotide and copy number variation on gene expression phenotypes.** *Science* 2007, 315:848–853.

55. Schlattl A, Anders S, Waszak SM, Huber W, Korbel JO: **Relating CNVs to transcriptome data at fine resolution: assessment of the effect of variant size, type, and overlap with functional regions.** *Genome Res* 2011, 21:2004–2013.

56. Henrichsen CN, Vinckenbosch N, Zöllner S, Chaignat E, Pradervand S, Schütz F, Ruedi M, Kaessmann H, Reymond A: **Segmental copy number variation shapes tissue transcriptomes.** *Nat Genet* 2009, 41:424–429.

57. Luo R, Sanders SJ, Tian Y, Voineagu I, Huang N, Chu SH, Klei L, Cai C, Ou J, Lowe JK, Hurles ME, Devlin B, State MW, Geschwind DH: **Genome-wide transcriptome profiling reveals the functional impact of rare de novo and recurrent CNVs in autism spectrum disorders.** *Am J Hum Genet* 2012, 91:38–55.

58. Merla G, Howald C, Henrichsen CN, Lyle R, Wyss C, Zabot M-T, Antonarakis SE, Reymond A: **Submicroscopic deletion in patients with Williams-Beuren syndrome influences expression levels of the nonhemizygous flanking genes.** *Am J Hum Genet* 2006, 79:332–341.

59. Blumenthal I, Ragavendran A, Erdin S, Klei L, Sugathan A, Guide JR, Manavalan P, Zhou JQ, Wheeler VC, Levin JZ, Ernst C, Roeder K, Devlin B, Gusella JF, Talkowski ME: **Transcriptional consequences of 16p11.2 deletion and duplication in mouse cortex and multiplex autism families.** *Am J Hum Genet* 2014, 94:870–883.

60. Disteche CM: **Dosage compensation of the sex chromosomes.** *Annu Rev Genet* 2012, 46:537–560.

61. Veitia RA, Bottani S, Birchler JA: **Cellular reactions to gene dosage imbalance: genomic, transcriptomic and proteomic effects.** *Trends Genet* 2008, 24:390–397.

62. Aït Yahya-Graison E, Aubert J, Dauphinot L, Rivals I, Prieur M, Golfier G, Rossier J, Personnaz L, Creau N, Bléhaut H, Robin S, Delabar JM, Potier M-C: **Classification of human chromosome 21 gene-expression variations in Down syndrome: impact on disease phenotypes.** *Am J Hum Genet* 2007, 81:475–491.

63. Letourneau A, Santoni FA, Bonilla X, Sailani MR, Gonzalez D, Kind J, Chevalier C, Thurman R, Sandstrom RS, Hibaoui Y, Garieri M, Popadin K, Falconnet E, Gagnebin M, Gehrig C, Vannier A, Guipponi M, Farinelli L, Robyr D, Migliavacca E, Borel C, Deutsch S, Feki A, Stamatoyannopoulos JA, Herault Y, van Steensel B, Guigo R, Antonarakis SE: **Domains of genome-wide gene expression dysregulation in Down's syndrome.** *Nature* 2014, 508:345–350.

64. Mitelman F, Johansson B, Mertens F: **The impact of translocations and gene fusions on cancer causation.** *Nat Rev Cancer* 2007, 7:233–245.

65. Nothwang HG, Kim HG, Aoki J, Geisterfer M, Kübart S, Wegner RD, van Moers A, Ashworth LK, Haaf T, Bell J, Arai H, Tommerup N, Ropers HH, Wirth J: **Functional Hemizygosity of PAFAH1B3 due to a PAFAH1B3-CLK2 Fusion Gene in a Female with Mental Retardation.** *Ataxia and Atrophy of the Brain. Hum Mol Genet* 2001, 10:797–806.

66. Backx L, Seuntjens E, Devriendt K, Vermeesch J, Van Esch H: **A balanced translocation t(6;14)(q25.3;q13.2) leading to reciprocal fusion transcripts in a patient with intellectual disability and agenesis of corpus callosum.** *Cytogenet Genome Res* 2011, 132:135–143.

67. Kim HG, Herrick SR, Lemyre E, Kishikawa S, Salisz JA, Seminara S, MacDonald ME, Bruns GAP, Morton CC, Quade BJ, Gusella JF: **Hypogonadotropic hypogonadism and cleft lip and palate caused by a balanced translocation producing haploinsufficiency for FGFR1.** *J Med Genet* 2005, 42:666–672.

68. Mansouri MR, Carlsson B, Davey E, Nordenskjöld A, Wester T, Annerén G, Läckgren G, Dahl N: **Molecular genetic analysis of a de novo balanced translocation t(6;17)(p21.31;q11.2) associated with hypospadias and anorectal malformation.** *Hum Genet* 2006, 119:162–168.

69. Borsani G, Piovani G, Zoppi N, Bertini V, Bini R, Notarangelo L, Barlati S: **Cytogenetic and molecular characterization of a de-novo t(2p;7p) translocation involving TNS3 and EXOC6B genes in a boy with a complex syndromic phenotype.** *Eur J Med Genet* 2008, 51:292–302.

70. Yue Y, Grossmann B, Holder SE, Haaf T: **De novo t(7;10)(q33;q23) translocation and closely juxtaposed microdeletion in a patient with macrocephaly and developmental delay.** *Hum Genet* 2005, 117:1–8.

71. Eykelenboom JE, Briggs GJ, Bradshaw NJ, Soares DC, Ogawa F, Christie S, Malavasi ELV, Makedonopoulou P, Mackie S, Malloy MP, Wear MA, Blackburn EA, Bramham J, Mcintosh AM, Blackwood DH, Muir WJ, Porteous DJ, Millar JK: **A t(1;11) translocation linked to schizophrenia and affective disorders gives rise to aberrant chimeric DISC1 transcripts that encode structurally altered, deleterious mitochondrial proteins.** *Hum Mol Genet* 2012, 21:3374–3386.

72. Holt R, Sykes NH, Conceição IC, Cazier J-B, Anney RJL, Oliveira G, Gallagher L, Vicente A, Monaco AP, Pagnamenta AT: **CNVs leading to fusion transcripts in individuals with autism spectrum disorder.** *Eur J Hum Genet* 2012, 20:1141–1147.

73. Rippey C, Walsh T, Gulsuner S, Brodsky M, Nord AS, Gasperini M, Pierce S, Spurrell C, Coe BP, Krumm N, Lee MK, Sebat J, McClellan JM, King MC: **Formation of chimeric genes by copy-number variation as a mutational mechanism in schizophrenia.** *Am J Hum Genet* 2013, 93:697–710.

74. Spielmann M, Mundlos S: **Structural variations, the regulatory landscape of the genome and their alteration in human disease.** *BioEssays* 2013, 35:533–543.

75. Gordon CT, Attanasio C, Bhatia S, Benko S, Ansari M, Tan TY, Munnich A, Pennacchio LA, Abadie V, Temple IK, Goldenberg A, van Heyningen V, Amiel J, Fitzpatrick D, Kleinjan DA, Visel A, Lyonnet S: **Identification of novel craniofacial regulatory domains located far upstream of SOX9 and Disrupted in pierre robin sequence.** *Hum Mutat* 2014, 35:1011–1020.

76. Benko S, Fantes JA, Amiel J, Kleinjan D-J, Thomas S, Ramsay J, Jamshidi N, Essafi A, Heaney S, Gordon CT, McBride D, Golzio C, Fisher M, Perry P, Abadie V, Ayuso C, Holder-Espinasse M, Kilpatrick N, Lees MM, Picard A, Temple IK, Thomas P, Vazquez M-P, Vekemans M, Roest Crollius H, Hastie ND, Munnich A, Etchevers HC, Pelet A, Farlie PG, et al: **Highly Conserved Non-Coding Elements on Either Side of SOX9 Associated with Pierre Robin Sequence.** *Nat Genet* 2009, 41:359–364.

77. Gordon CT, Tan TY, Benko S, Fitzpatrick D, Lyonnet S, Farlie PG: **Long-range regulation at the SOX9 locus in development and disease.** *J Med Genet* 2009, 46:649–656.

78. Sankaran VG, Xu J, Byron R, Greisman HA, Fisher C, Weatherall DJ, Sabath DE, Groudine M, Orkin SH, Premawardhena A, Bender MA: **A functional element necessary for fetal hemoglobin silencing.** *N Engl J Med* 2011, 365:807–814.

79. Lettice LA, Daniels S, Sweeney E, Venkataraman S, Devenney PS, Gautier P, Morrison H, Fantes J, Hill RE, Fitzpatrick DR: **Enhancer-adoption as a mechanism of human developmental disease.** *Hum Mutat* 2011, 32:1492–1499.

80. Ibn-Salem J, Köhler S, Love MI, Chung H-R, Huang N, Hurles ME, Haendel M, Washington NL, Smedley D, Mungall CJ, Lewis SE, Ott C-E, Bauer S, Schofield PN, Mundlos S, Spielmann M, Robinson PN: **Deletions of chromosomal regulatory boundaries are associated with congenital disease.** *Genome Biol* 2014, 15:423.

81. Hochstenbach R, Poot M, Nijman IJ, Renkens I, Duran KJ, van'T Slot R, van Binsbergen E, van der Zwaag B, Vogel MJ, Terhal PA, van Amstel HK P, Kloosterman WP, Cuppen E: **Discovery of variants unmasked by hemizygous deletions.** *Eur J Hum Genet* 2012, 20:748–753.

82. Shaikh TH, Kurahashi H, Saitta SC, O'Hare AM, Hu P, Roe BA, Driscoll DA, McDonald-McGinn DM, Zackai EH, Budarf ML, Emanuel BS: **Chromosome 22-specific low copy repeats and the 22q11.2 deletion syndrome: genomic organization and deletion endpoint analysis.** *Hum Mol Genet* 2000, 9:489–501.

83. Ludlow LB, Schick BP, Budarf ML, Driscoll DA, Zackai EH, Cohen A, Konkle BA: **Identification of a mutation in a GATA binding site of the platelet glycoprotein Ibβ promoter resulting in the Bernard-Soulier Syndrome.** *J Biol Chem* 1996, 271:22076–22080.

84. Kunishima S, Imai T, Kobayashi R, Kato M, Ogawa S, Saito H: **Bernard-Soulier syndrome caused by a hemizygous GPIb(beta) mutation and 22q11.2 deletion.** *Pediatr Int* 2013, **55**:434–437.

85. Bedeschi MF, Colombo L, Mari F, Hofmann K, Rauch A, Gentilin B, Renieri A, Clerici D: **Unmasking of a recessive SCARF2 mutation by a 22q11.12 de novo deletion in a patient with Van den Ende-Gupta syndrome.** *Mol Syndromol* 2011, **1**:239–245.

86. McDonald-McGinn DM, Fahiminiya S, Revil T, Nowakowska BA, Suhl J, Bailey A, Mlynarski E, Lynch DR, Yan AC, Bilaniuk LT, Sullivan KE, Warren ST, Emanuel BS, Vermeesch JR, Zackai EH, Jerome-Majewska LA: **Hemizygous mutations in SNAP29 unmask autosomal recessive conditions and contribute to atypical findings in patients with 22q11.2DS.** *J Med Genet* 2013, **50**:80–90.

87. Mefford HC, Sharp AJ, Baker C, Itsara A, Jiang Z, Buysse K, Huang S, Maloney VK, Crolla JA, Baralle D, Collins A, Mercer C, Norga K, de Ravel T, Devriendt K, Bongers EMHF, de Leeuw N, Reardon W, Gimelli S, Bena F, Hennekam RC, Male A, Gaunt L, Clayton-Smith J, Simonic I, Park SM, Mehta SG, Nik-Zainal S, Woods CG, Firth HV, *et al*: **Recurrent rearrangements of chromosome 1q21.1 and variable pediatric phenotypes.** *N Engl J Med* 2008, **359**:1685–1699.

88. Hannes FD, Sharp AJ, Mefford HC, de Ravel T, Ruivenkamp CA, Breuning MH, Fryns J-P, Devriendt K, Van Buggenhout G, Vogels A, Stewart H, Hennekam RC, Cooper GM, Regan R, Knight SJL, Eichler EE, Vermeesch JR: **Recurrent reciprocal deletions and duplications of 16p13.11: the deletion is a risk factor for MR/MCA while the duplication may be a rare benign variant.** *J Med Genet* 2009, **46**:223–232.

89. Kumar RA, Marshall CR, Badner JA, Babatz TD, Mukamel Z, Aldinger KA, Sudi J, Brune CW, Goh G, Karamohamed S, Sutcliffe JS, Cook EH, Geschwind DH, Dobyns WB, Scherer SW, Christian SL: **Association and mutation analyses of 16p11.2 autism candidate genes.** *PLoS One* 2009, **4**:e4582.

90. Williams NM: **Molecular mechanisms in 22q11 deletion syndrome.** *Schizophr Bull* 2011, **37**:882–889.

91. Forstner AJ, Degenhardt F, Schratt G, Nöthen MM: **MicroRNAs as the cause of schizophrenia in 22q11.2 deletion carriers, and possible implications for idiopathic disease: a mini-review.** *Front Mol Neurosci* 2013, **6**:47.

92. Chaignat E, Yahya-Graison EA, Henrichsen CN, Chrast J, Schütz F, Pradervand S, Reymond A: **Copy number variation modifies expression time courses.** *Genome Res* 2011, **21**:106–113.

93. De Wit E, de Laat W: **A decade of 3C technologies: insights into nuclear organization.** *Genes Dev* 2012, **26**:11–24.

94. Gheldof N, Witwicki RM, Migliavacca E, Leleu M, Didelot G, Harewood L, Rougemont J, Reymond A: **Structural variation-associated expression changes are paralleled by chromatin architecture modifications.** *PLoS One* 2013, **8**:e79973.

95. Zentner GE, Henikoff S: **Regulation of nucleosome dynamics by histone modifications.** *Nat Struct Mol Biol* 2013, **20**:259–266.

96. Kim H, Kim J-S: **A guide to genome engineering with programmable nucleases.** *Nat Rev Genet* 2014, **15**:321–334.

97. Öllü C, Pars K, Cornu TI, Thibodeau-Beganny S, Maeder ML, Joung JK, Heilbronn R, Cathomen T: **Autonomous zinc-finger nuclease pairs for targeted chromosomal deletion.** *Nucleic Acids Res* 2010, **38**:8269–8276.

98. Piganeau M, Ghezraoui H, De Cian A, Guittat L, Tomishima M, Perrouault L, René O, Katibah GE, Zhang L, Holmes MC, Doyon Y, Concordet JP, Giovannangeli C, Jasin M, Brunet E: **Cancer translocations in human cells induced by zinc finger and TALE nucleases.** *Genome Res* 2013, **23**:1182–1193.

99. Choi PS, Meyerson M: **Targeted genomic rearrangements using CRISPR/Cas technology.** *Nat Commun* 2014, **5**:3728.

100. Xiao A, Wang Z, Hu Y, Wu Y, Luo Z, Yang Z, Zu Y, Li W, Huang P, Tong X, Zhu Z, Lin S, Zhang B: **Chromosomal deletions and inversions mediated by TALENs and CRISPR/Cas in zebrafish.** *Nucleic Acids Res* 2013, **41**:e141.

101. Henrichsen CN, Csárdi G, Zabot MT, Fusco C, Bergmann S, Merla G, Reymond A: **Using transcription modules to identify expression clusters perturbed in Williams-Beuren Syndrome.** *PLoS Comput Biol* 2011, **7**:e1001054.

102. Antonell A, Vilardell M, Jurado LAP: **Transcriptome profile in Williams-Beuren syndrome lymphoblast cells reveals gene pathways implicated in glucose intolerance and visuospatial construction deficits.** *Hum Genet* 2010, **128**:27–37.

103. Kim H, Bernitz J: **Genomic editing tools to model human diseases with isogenic pluripotent stem cells.** *Stem cells* 2014, **23**:2673–2686.

104. Incontro S, Asensio CS, Edwards RH, Nicoll RA: **Efficient, Complete Deletion of Synaptic Proteins using CRISPR.** *Neuron* 2014, **83**:1051–7.

105. Ricard G, Molina J, Chrast J, Gu W, Gheldof N, Pradervand S, Schütz F, Young JI, Lupski JR, Reymond A, Walz K: **Phenotypic consequences of copy number variation: insights from smith-magenis and Potocki-Lupski syndrome mouse models.** *PLoS Biol* 2010, **8**:e1000543.

106. Bi W, Ohyama T, Nakamura H, Yan J, Visvanathan J, Justice MJ, Lupski JR: **Inactivation of Rai1 in mice recapitulates phenotypes observed in chromosome engineered mouse models for Smith-Magenis syndrome.** *Hum Mol Genet* 2005, **14**:983–995.

107. Slager RE, Newton TL, Vlangos CN, Finucane B, Elsea SH: **Mutations in RAI1 associated with Smith-Magenis syndrome.** *Nat Genet* 2003, **33**:466–468.

108. Davis EE, Frangakis S, Katsanis N: **Interpreting human genetic variation with in vivo zebrafish assays.** *Biochim Biophys Acta - Mol Basis Dis* 1842, **2014**:1311–1320.

109. Golzio C, Willer J, Talkowski ME, Oh EC, Taniguchi Y, Jacquemont S, Reymond A, Sun M, Sawa A, Gusella JF, Kamiya A, Beckmann JS, Katsanis N: **KCTD13 is a major driver of mirrored neuroanatomical phenotypes of the 16p11.2 copy number variant.** *Nature* 2012, **485**:363–367.

110. Dauber A, Golzio C, Guenot C, Jodelka FM, Kibaek M, Kjaergaard S, Leheup B, Martinet D, Nowaczyk MJM, Rosenfeld JA, Zeesman S, Zunich J, Beckmann JS, Hirschhorn JN, Hastings ML, Jacquemont S, Katsanis N: **Scrib and puf60 are primary drivers of the multisystemic phenotypes of the 8q24.3 copy-Number variant.** *Am J Hum Genet* 2013, **93**:798–811.

111. Vissers LELM, de Ligt J, Gilissen C, Janssen I, Steehouwer M, de Vries P, van Lier B, Arts P, Wieskamp N, del Rosario M, van Bon BWM, Hoischen A, de Vries BBA, Brunner HG, Veltman JA: **A de novo paradigm for mental retardation.** *Nat Genet* 2010, **42**:1109–1112.

112. De Ligt J, Willemsen MH, van Bon BWM, 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 BBA, Brunner HG, Veltman JA, Vissers LELM: **Diagnostic Exome Sequencing in Persons with Severe Intellectual Disability.** *N Engl J Med* 2012, **367**:1921–1929.

113. Rauch A, Wieczorek D, Graf E, Wieland T, Endele S, Schwarzmayr T, Albrecht B, Bartholdi D, Beygo J, Di Donato N, Dufke A, Cremer K, Hempel M, Horn D, Hoyer J, Joset P, Röpke A, Moog U, Riess A, Thiel CT, Tzschach A, Wiesener A, Wohlleber E, Zweier C, Ekici AB, Zink AM, Rump A, Meisinger C, Grallert H, Sticht H, *et al*: **Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: An exome sequencing study.** *Lancet* 2012, **380**:1674–1682.

114. Gilissen C, Hehir-Kwa JY, Thung DT, van de Vorst M, van Bon BWM, Willemsen MH, Kwint M, Janssen IM, Hoischen A, Schenck A, Leach R, Klein R, Tearle R, Bo T, Pfundt R, Yntema HG, de Vries BBA, Kleefstra T, Brunner HG, Vissers LELM, Veltman JA: **Genome sequencing identifies major causes of severe intellectual disability.** *Nature* 2014, **511**:344–347.

115. Biesecker LG, Spinner NB: **A genomic view of mosaicism and human disease.** *Nat Rev Genet* 2013, **14**:307–320.