

Genotype to phenotype—discovery and characterization of novel genomic disorders in a “genotype-first” era

Heather C. Mefford, MD, PhD

Abstract: Recent advances in technology for detecting copy number changes have enabled genome-wide detection of submicroscopic deletions and duplications. As these technologies are applied to large cohorts of patients with diverse phenotypes, novel genomic disorders are defined by a common genotype (deletion or duplication) rather than a common phenotype. The discovery of new genomic disorders using this “genotype-first” approach has increased dramatically, and several recently described recurrent rearrangements are associated with a surprisingly wide range of phenotypes. This review will discuss the importance of genomic architecture for generating recurrent rearrangements, implications of the genotype-first approach for medical genetics, and features of several new genomic disorders with highly variable phenotypes. *Genet Med* 2009;11(12):836–842.

Key Words: Genomic disorder, array comparative genomic hybridization, phenotypic diversity, microdeletion syndrome

Chromosomal imbalances and rearrangements have long been implicated in the etiology of mental retardation (MR) and congenital anomalies, beginning with the recognition in 1959 that an extra copy of Chromosome 21 is the cause of Down syndrome.¹ Standard karyotype analysis can detect microscopically visible chromosome rearrangements, which include trisomies, monosomies, supernumerary marker chromosomes, unbalanced translocations, and large (roughly >5 Mb) deletions and duplications; such large events are responsible for 10% to 15% of severe MR,² a common diagnosis for which karyotype studies are performed. In the early 1980s, improvements in cytogenetic techniques including high-resolution karyotyping and fluorescence in situ hybridization (FISH) resulted in the identification of smaller cytogenetic rearrangements. Examples include the common deletions responsible for clinically recognizable syndromes such as Prader-Willi and Angelman (15q11-q13 deletions)³ and Smith-Magenis (17p11.2 deletions)⁴ syndromes as well as the subtelomeric rearrangements that underlie another 2.5% to 5% of MR.^{5–7}

Along with improved cytogenetics and the recognition that deletions of specific genomic regions result in common phenotypic features came molecular diagnosis of several syndromes including Prader-Willi,⁸ Angelman,⁹ Smith-Magenis,¹⁰ velocardiofacial,¹¹ and Williams-Beuren¹² syndromes. However, the molecular tests available (primarily FISH) were targeted and largely used for molecular confirmation of a specific suspected clinical diagnosis—a “phenotype first” approach. Efforts to fine map the breakpoints of deletions responsible for these disorders

revealed that each breakpoint lies within large blocks of duplicated sequence.^{8–12} The same flanking structure had been noted for reciprocal duplication and deletion events on Chromosome 17p12 that cause Charcot-Marie-Tooth type IA¹³ and hereditary neuropathy with liability to pressure palsies,¹⁴ two neurologic disorders. Together, these discoveries suggested nonallelic homologous recombination (NAHR) as a possible mechanism for rearrangement, and the term genomic disorder was coined to describe conditions due to recurrent rearrangements that occur because of regional genomic architecture.¹⁵

More recent advances in technology for detecting copy number changes—most notably array comparative genomic hybridization (CGH) and single nucleotide polymorphism (SNP) microarrays—enable genome-wide detection of submicroscopic deletions and duplications (referred to here as microdeletions and microduplications). Since the introduction of these technologies, the rate of discovery of submicroscopic rearrangements in both affected and unaffected individuals has increased dramatically. Genome-wide studies of copy number variation were first performed in control cohorts^{16–24} to catalog the copy number changes that the human genome can tolerate without apparently deleterious consequences. Almost simultaneously, discovery of microdeletions and duplications underlying various disorders including MR and developmental delay,^{25–30} autism,^{31–36} and congenital anomalies^{37–40} were underway. Several large studies reported apparently pathogenic deletions or duplications detected by array CGH in ~5% to 15% of individuals with MR,^{25,26,29,41–43} and results from large diagnostic laboratories performing clinical array CGH are similar.^{28,44} Although initial reports of novel genomic disorders primarily described MR syndromes, studies of copy number variation have now expanded beyond phenotypes traditionally thought to have a chromosomal basis (i.e., MR, congenital anomalies, and autism) to include schizophrenia,^{45–49} epilepsy,^{50–52} amyotrophic lateral sclerosis,⁵³ autoimmune diseases,⁵⁴ craniosynostosis,⁵⁵ and many other disorders.

Of course, the microdeletions and microduplications that are identified by these methods include both nonrecurrent and recurrent events. Although both types are clearly important to disease etiology, this review will focus on recurrent events, defined here as those that are found in multiple individuals, are the same size, and have the same breakpoints each time they occur. As described earlier, recurrent events occur (and recur) because of a specific underlying genomic architecture in which a unique stretch of DNA is flanked by large, highly homologous segmental duplications that facilitate NAHR.^{56,57} Later, the importance of genomic architecture, the implications of a genotype-first approach and the burden of proving pathogenicity are discussed. Finally, I will review the findings associated with four of the genomic disorders that have been described primarily since the widespread introduction of whole-genome technologies and that exhibit a wide range of phenotypes.

From the Department of Pediatrics, Division of Genetic Medicine, University of Washington, Seattle, Washington.

Heather C. Mefford, MD, PhD, 1705 NE Pacific Street, Box 355065, Seattle, WA 98195. E-mail: hmefford@u.washington.edu.

Disclosure: The author declares no conflict of interest.

Submitted for publication September 9, 2009.

Accepted for publication September 14, 2009.

DOI: 10.1097/GIM.0b013e3181c175d2

Importance of genomic architecture for recurrent rearrangements

All genomic disorders described to date share a common genomic landscape: flanking duplications tend to be >10 kb in size, >95% identical at the sequence level, and separated by 50 kb to 10 Mb of DNA.⁵⁷ Importantly, in order for NAHR to result in deletion or duplication, the flanking duplications must be in direct orientation; if they lie in inverted orientation, the result of NAHR is inversion of the intervening sequence. Given the large size and high degree of sequence identity between flanking duplications, it is reasonable to hypothesize that recombination events might take place anywhere within the flanking duplications. However, detailed studies in some disorders have revealed recombination hotspots ranging from 500 bp to 12 kb within the larger segmental duplication blocks.^{58–64} Therefore, although the large blocks of duplicated sequence may facilitate misalignment at meiosis, there are likely additional local sequence requirements for NAHR to occur.

Using the criteria described earlier (segmental duplications >10 kb, >95% identical, and separated by 50 kb–10 Mb), Bailey et al.⁵⁶ mined the sequenced human genome (NCBI Build 34) and identified 130 regions of the genome with this architecture that were predicted to be susceptible to NAHR. Several of these genomic “hotspot” regions were already known to be associated with disease but the majority was not. Using a bacterial artificial chromosome array targeted to the hotspots, they investigated copy number variation in controls²² as well as individuals with MR and/or congenital anomalies.^{29,39} Interestingly, in a study of 290 individuals with idiopathic MR and 360 controls, six hotspots were deleted or duplicated only in affected individuals and all have since been confirmed as pathogenic rearrangements. These include deletions of 17q21.31,^{26,29,30,64} 15q13.3,^{51,52,65–71} 15q24,^{33,72,73} and 17q12³⁹; and reciprocal deletions and duplications of 16p13.11^{74–76} and 1q21.1,^{20,68,69,77–80} some of which are discussed in further detail later. Another structural feature that seems to contribute to the susceptibility of some regions to NAHR is polymorphic inversion. In Williams-Beuren syndrome, for example, 25% to 30% of parents who transmit a de novo deletion to their child have a 1.5-Mb inversion encompassing the commonly deleted region; the same inversion is present in only 6% of the general population.^{60,81} Similarly, mothers who transmit a de novo deletion causing Angelman syndrome are more likely to carry a polymorphic inversion of 15q11-q13 than those in the general population.⁸² Two more striking examples are found in Sotos syndrome⁶³ and the 17q21.31 microdeletion syndrome.^{26,29,30,64,83} In each of these disorders, every parent studied to date in which a de novo microdeletion arises carries an inversion of the same region. Because there are balanced rearrangements, inversions are more difficult to detect than copy number variations. Sequence-based studies have been used to detect inversions, however, and it is interesting to note that polymorphic inversions are also present at several additional sites of recurrent rearrangement including 1q21.1, 3q29, 15q13.3, 15q24, and 17q12.^{84–86} Further studies will be required to determine whether any (or all) of these inversions are predisposing factors for disease-causing rearrangements. Whether they increase risk of rearrangement, it is clear that inversions can be markers of vulnerable genomic regions.

Implications of a “genotype-first” approach

The use of genome-wide assays—or even technologies that are targeted but evaluate multiple genomic regions—has several important implications. In contrast to targeted FISH tests,

which require clinical suspicion of a specific disorder, genome-wide studies can be (and often are) performed without a suspected diagnosis. In the clinical setting, this can be advantageous in that an individual may be diagnosed before he has developed all of the features of a given disorder, perhaps allowing earlier intervention as well as recurrence risk counseling for the family. In another case, a given individual may not have the classic features associated with a known disorder but the molecular diagnosis will clarify the clinical diagnosis. Importantly, the ability to perform a single test in a patient with nonspecific findings and end the “diagnostic odyssey” can be very beneficial to both families and physicians. Finally, it can be argued that such an approach, being available to the nonspecialist, can expedite diagnosis by primary care providers if such technologies are more widely—but carefully and appropriately—applied and interpreted.

Conversely, casting such a wide net inevitably leads to unexpected findings. These can include detecting an unexpected pathogenic change unrelated to the patient’s phenotype (e.g., a 17p12 duplication predicting Charcot-Marie-Tooth disease in a patient being evaluated for congenital heart disease [CHD] and developmental delay). Another troubling finding is the identification of a deletion or duplication of “uncertain significance”—one that has not previously been associated with disease but has also not been reported as a being copy number variant. Such a finding can prolong the diagnostic odyssey because of evaluation of parents and other family members, which may or may not help clarify the significance of the aberration, or create new diagnostic dilemmas with attendant problems.

In both the research and clinical settings, the unbiased whole-genome approach has led to unprecedented discovery of novel genomic disorders. Since the introduction of array CGH and SNP microarrays circa 2005, 18 new genomic disorders involving 12 regions of the genome have been described, more than doubling the number of disorders described in the previous 20 years (see Refs. 88–90 for more thorough discussion of these). New syndromes identified by this genotype-first approach are often defined by a genomic location and rearrangement type (e.g., 15q13.3 microdeletion syndrome), and clinical features are compared among patients after a common rearrangement is identified. Because more diverse phenotypes are being evaluated with similar platforms, it is now clear that for some deletions and duplications, the associated phenotypes are so diverse that a phenotype-first approach to identify affected individuals would never be successful. In addition to phenotypic diversity, there is also incomplete penetrance associated with many of the recently identified genomic disorders. As described in the next section, these features complicate the process of proving pathogenicity, not to mention genetic counseling and clinical care.

Determining pathogenicity

The criteria for determining whether a given deletion or duplication is pathogenic have traditionally been the following: (i) the event is observed in affected individuals; (ii) the same event is not found in unaffected parents or control populations; and (iii) individuals who have the same event have the same or similar phenotypic features. Examples of microdeletion syndromes that generally follow these “rules” include Smith-Magenis syndrome, Prader-Willi and Angelman syndromes, and Williams syndrome. Many of the recently identified microdeletion or duplication “syndromes,” however, tend to bend these rules. Deletions and duplications can no longer be dismissed because they are inherited from an apparently unaffected parent.⁸⁷ One must consider other factors including size, gene

content, overlap with known benign copy number changes, and frequency of the same event in controls, although none of these are necessary or sufficient to unambiguously determine pathogenicity. For the novel genomic disorders discussed in the next section, deletions or duplications can be inherited or de novo; most exhibit incomplete penetrance; and affected individuals may have any one or more of a wide range of neuropsychiatric conditions. As a result, new rules are in play for determining the pathogenic significance of any given event. In most cases, statistical arguments have been used to determine whether a given event is pathogenic. That is, the deletion or duplication is deemed pathogenic if it is clearly enriched in affected individuals compared with unaffected individuals. This task is easier for recurrent rearrangements because of the simple fact that they recur; as a result, far fewer cases and controls need to be evaluated to identify multiple individuals with the same event.

There are several excellent reviews describing new genomic disorders identified during the past few years,^{88–91} some of which cause clinically recognizable syndromes. Later, I will review the findings associated with four recently described genomic disorders that exhibit extreme phenotypic diversity and elude syndromic classification. These genomic disorders cannot be recognized based on shared clinical features, but the rearrangements (and genes within) underlying each disorder clearly play an important role in development.

Recurrent rearrangements of 15q13.3

Microdeletions of 15q13.3 were first described in a series of six probands and three affected family members with intellectual disability and minor dysmorphic features,^{29,67} and the deletion was clearly enriched in affected individuals compared with controls. Seven of the nine affected individuals also had seizures or abnormal electroencephalography findings. Since that first report, there have been several additional series of patients with MR (with and without seizures) or autism.^{65,66,70,71} The severity of MR phenotypes in all of these studies ranges from mild to severe; the frequency of seizures and of autistic features differs among series; and there are no consistent dysmorphic features (if any) among affected individuals. Although some of the deletions reported in these studies were de novo, in many cases the deletion is inherited, suggesting incomplete penetrance in addition to variable expressivity.

Studies that have focused on patients with neuropsychiatric conditions other than MR or autism reveal that the phenotypes associated with deletions of 15q13.3 are even more variable. Two large studies independently reported enrichment of the 15q13.3 deletion in patients with schizophrenia compared with controls (0.2% vs. 0.02% combined).^{68,69} Investigating yet another phenotype, Helbig et al.⁵¹ followed up on the finding that 7 of 9 patients reported by Sharp et al.⁶⁷ suffered from seizures. In their study, 12 of 1223 (1%) patients with idiopathic generalized epilepsy (IGE) were found to carry deletions of 15q13.3 compared with 0 of 3699 controls. Unlike patients from previous studies, the majority of individuals in the IGE series had normal intellect. Two studies have confirmed the association with epilepsy^{50,52} in independent IGE cohorts, and one study estimated an odds ratio of 68 for carriers of the deletion to develop epilepsy.⁵² The frequency of the deletion in epilepsy probands (~1%) is notably higher than in patients who present with MR or autism (0.3%)^{65,67} or schizophrenia (0.2%).^{68,69} Remarkably, the 15q13.3 deletion is now the most prevalent known genetic risk factor for epilepsy.

Candidate genes

There are seven genes within the commonly deleted region on 15q13.3, one of which is *CHRNA7*, encoding a subunit of the nicotinic acetylcholine receptor. Through linkage studies, *CHRNA7* has been identified as a candidate gene for both epilepsy^{92,93} and schizophrenia.⁹⁴ In addition, mutations in several other acetylcholine receptor subunit genes have been shown to cause epilepsy. Although emphasis has been placed on *CHRNA7*, one of more of the other genes in the region is likely to play a role as well.

Reciprocal rearrangements of 16p13.11

A similar story is emerging for rearrangements of the 16p13.11 region. Deletions of 16p13.11 have now been reported in patients with a wide range of neuropsychiatric conditions. In studies of patients presenting with MR or autism,^{74–76} affected individuals with deletions had a range of phenotypes that include mild to severe MR, autistic features, brain anomalies, and epilepsy. Deletions of 16p13.11 may also be a risk factor for schizophrenia. Need et al.⁴⁹ identified deletions of this region in 3 of 1013 probands with schizophrenia and 0 of 1084 controls. Although not statistically significant, this finding is intriguing and evaluation of additional cohorts will be required to confirm this finding. Like deletions of 15q13.3, deletions of 16p13.11 are also significantly enriched in probands with IGE⁵⁰ and may be as important as 15q13.3 deletions in determining genetic risk for IGE. Although occasionally inherited from mildly affected or unaffected parents, the deletion is often de novo, is rarely reported in control cohorts,^{68,74,95} is significantly enriched in affected individuals, and appears to be highly (though not fully) penetrant.

The reciprocal duplication on 16p13.11 is also a risk factor for neurocognitive disease, albeit with reduced penetrance. Ulmann et al.⁷⁶ first reported four male patients with severe autistic features from three families who all carried duplications of this region. In this study, several mildly affected or unaffected family members had the same duplication and the frequency of the duplication in controls was not reported. A subsequent study found duplications of 16p13.11 in both affected and unaffected individuals, suggesting that the duplication was either incompletely penetrant or benign variant.⁷⁴ In yet another study of patients with unexplained MR, we found the duplication in ~1% of affected individuals, representing a significant enrichment compared with controls.⁷⁵ Interestingly, duplications of 16p13.11 have also been reported in patients with schizophrenia.^{48,68} Although not statistically significant, the data suggest that duplications of this region may be a risk factor for schizophrenia, too. Taken together, these data suggest that duplications of 16p13.11 are likely to be pathogenic and contribute to the genetic etiology of neuropsychiatric illness but exhibit reduced penetrance.

Candidate genes

The 16p13.11 deletion region contains 15 *RefSeq* genes, at least two of which are plausible candidate genes for the some of the phenotypes associated with imbalances of this region. The *NDE1* gene is predominantly brain expressed and interacts with both *LIS1*,⁹⁶ mutations in which cause lissencephaly, and *DISC1*, an important gene in schizophrenia.^{97,98} *NTAN1* is also brain expressed, and mice lacking the gene exhibit abnormal social behavior and learning phenotypes.⁹⁹

Recurrent reciprocal rearrangements of 1q21.1

Deletions of 1q21.1 are also associated with a broad range of phenotypes. Early and relatively small studies reported dele-

tions of this region in a few patients with disparate phenotypes including CHD,⁷⁸ unexplained MR,²⁹ schizophrenia,⁴⁷ autism,³⁴ and müllerian aplasia.¹⁰⁰ Much larger studies in patients with various phenotypes have confirmed that all of these initial associations are likely valid. Two large studies of probands with schizophrenia independently reported a significant enrichment of 1q21.1 deletions in affected individuals (0.26% of affected, combined).^{68,69} In a large study of ~5000 patients ascertained for MR, autism, or congenital anomalies and ~5000 controls, we confirmed that deletions of 1q21.1 are indeed pathogenic and result in a range of phenotypes,⁸⁰ a finding that was replicated in a second large study⁷⁷ with similar patient ascertainment. Cognitive phenotypes of affected individuals in these studies range from normal intellect to severe MR; dysmorphic features are reported in some but not in all patients; and other anomalies in some patients include CHD, microcephaly, cataracts, and seizures. Deletions have also been reported in patients with isolated IGE without cognitive delays⁵⁰; this finding was not a significant enrichment compared with controls and further studies of copy number variation in epilepsy will be required to confirm an association. Reciprocal duplications of the same region on 1q21.1 have also been reported in patients with MR and developmental delay,^{77,80} autism,^{34,80} and isolated CHD.^{77,79} In a recent study of 512 patients with isolated tetralogy of Fallot, one patient carried a 1q21.1 deletion, consistent with previous reports of CHD with 1q21.1 deletions.⁷⁸ Interestingly, four additional patients in this study had the reciprocal duplication,⁷⁹ which represented a statistically significant enrichment compared with controls and a higher frequency of 1q21.1 duplications than that reported in MR or autism. In most of the studies cited earlier, both de novo and inherited rearrangements of 1q21.1 have been reported, again suggesting incomplete penetrance.

Candidate genes

The 1q21.1 critical region contains eight *RefSeq* genes with several good candidates for some of the phenotypes reported to date. *GJA8* has previously been implicated in schizophrenia¹⁰¹ and cataracts^{102,103} and may very well influence these phenotypes. *GJA5* is highly expressed in the heart, and null mice have a high incidence of cardiac malformations.¹⁰⁴ Although these genes are plausible candidates, given the very diverse phenotypes reported so far, other genes in the region and elsewhere likely play a role.

Rearrangements of 16p11.2

The genetics of autism is clearly complex, and large efforts have been made to identify genetic risk factors contributing to this condition. An exciting advance came with the discovery by several groups that recurrent microdeletion of 16p11.2 is present in 0.5% to 1% of affected individuals.^{32–34,36} Early studies hinted at phenotypic diversity. The deletion was also present in individuals with language or psychiatric disorders, although at a lower frequency (0.1%), as well as in the general population (0.01%).³⁶ Subsequent studies have confirmed that the deletion is not specific to autism but is also enriched in patients with MR, developmental delay, or schizophrenia without autistic features.^{75,95,105} Reciprocal duplications of 16p11.2 have also been reported, but the pathogenic significance is less clear. Duplications are often inherited from an unaffected parent and are found at a higher frequency in the population than deletions (0.03%). Even larger studies may be required to determine whether the duplication is a low penetrance risk factor or a benign variant.

Candidate genes

The ~500-kb critical region on 16p11.2 is gene rich, with >25 genes. Pathway analysis has shown that nearly half of the genes map to a single genetic network involving cell–cell interactions.³² There has been one follow-up study to look for association or mutation of individual genes within the region in patients with autism.¹⁰⁶ This study failed to find strong association, but additional studies are sure to follow.

What are the modifiers and how do we find them?

These studies all suggest a common genetic and developmental etiology for MR, autism, schizophrenia, and epilepsy. Given the extreme variability in phenotypes among individuals with the same rearrangements, there must be other factors—so-called modifiers—that contribute to phenotypic outcome. Several possibilities are discussed briefly later. One explanation for variable phenotypes could be that the deletions (or duplications) are not really identical and that differences in size reflect differences in phenotype. We have performed very high-density oligonucleotide array CGH to evaluate the breakpoints of 1q21.1 and 15q13.11 deletions in patients with different phenotypes and cannot detect appreciable differences in the breakpoints.^{51,80,88} Given the complexity of the segmental duplications in many of the breakpoint regions, it is also possible that there are sequences not represented in the current assembly that contribute to phenotypic diversity but have not yet been evaluated. Genetic modifiers could include differences in the sequence of one or more genes on the nondeleted allele (in the case of deletions). For example, sequence variants in the *COMT* gene within the 22q11 deletion region have been shown to influence cognitive and psychiatric phenotypes in individuals with 22q11 deletions.^{107,108} Comprehensive studies of sequence variation on the nondeleted allele have not been published for any of the disorders described earlier. Of course, sequence variation influencing phenotypic outcome need not be restricted to the region of genomic rearrangement. SNPs or mutations in genes that interact with deleted or duplicated genes—or even in the same developmental pathway—could potentially influence phenotype. With improvements in sequence capture and sequencing technologies,^{109–112} both targeted (e.g., the nondeleted allele or all genome-wide exonic sequence) and whole-genome studies of sequence variants in affected individuals will soon be feasible. Other copy number variants could also contribute to phenotype diversity. Because most copy number studies are now carried out on whole-genome platforms, analysis of copy number changes that are common among affected individuals should be possible. Finally, epigenetic changes may play a role. An extreme but well-known example of this is imprinting in Prader-Willi (paternally derived deletions) and Angelman (maternally derived deletions) syndromes, which accounts for very different outcomes associated with the same deletion. In the genomic disorders discussed here, there has been no clear difference in phenotype associated with parent of origin, whether de novo or inherited. However, as with sequence variation, no large studies have been performed to evaluate epigenetic changes at the site of rearrangement of genome wide. Finally, it must be remembered that nongenetic and nonepigenetic factors such as the environment may be critically important in influencing the penetrance of an underlying genomic rearrangement. Clearly, there are many studies that can and should be performed to answer the question of why there is such a range of phenotypes associated with each of the rearrangements described earlier. One of the challenges will be to collect a large cohort of well-phenotyped patients for each deletion and

duplication to detect significant differences among subphenotypes. This will require large collaborative studies with standardized phenotyping evaluations. These studies are essential and will undoubtedly increase our understanding of genes and pathways critical for normal development. Furthermore, understanding the modifying factors will allow better recurrence risk counseling and hopefully facilitate interventions to prevent or ameliorate the adverse outcomes associated with the deletions and duplications described.

Implications for medical genetics and conclusions

With the introduction of whole-genome technologies to identify copy number changes, the rate of discovery of pathogenic rearrangements, including novel recurrent genomic disorders, has rapidly increased. In many ways, discovery of rearrangements in diverse cohorts of patients has outpaced our ability to interpret those changes in the clinical setting. Although statistical arguments have been helpful at the level of case-control studies for determining pathogenic significance, they are not as comforting in the setting of trying to counsel an individual family. Although we have learned a great deal in the past few years, careful genetic counseling of patients and families will remain critical as this emerging field has a greater and greater impact on medical care. The translation of such information in the clinic will remain difficult until we have a much better understanding of the other factors—genetic, epigenetic, or environmental—that influence phenotypic outcomes.

ACKNOWLEDGMENT

The author would like to thank her mentor, Dr. Evan Eichler, for his invaluable support and guidance. She also thanks the William K. Bowes, Jr. Award committee for their generosity and support of medical geneticists. This work was supported in part by NIH Grant K12 HD043376.

REFERENCES

- Lejeune J, Gautier M, Turpin R. Etudes des chromosomes somatique de neuf enfants mongoliens. *CR Acad Sci Paris* 1959;248:1721–1722.
- Ropers HH. Genetics of intellectual disability. *Curr Opin Genet Dev* 2008;18:241–250.
- Butler MG, Meaney FJ, Palmer CG. Clinical and cytogenetic survey of 39 individuals with Prader-Labhart-Willi syndrome. *Am J Med Genet* 1986;23:793–809.
- Smith AC, McGavran L, Robinson J, et al. Interstitial deletion of (17)(p11.2p11.2) in nine patients. *Am J Med Genet* 1986;24:393–414.
- Ballif BC, Sulpizio SG, Lloyd RM, et al. The clinical utility of enhanced subtelomeric coverage in array CGH. *Am J Med Genet A* 2007;143A:1850–1857.
- Knight SJ, Regan R, Nicod A, et al. Subtle chromosomal rearrangements in children with unexplained mental retardation. *Lancet* 1999;354:1676–1681.
- Ravnan JB, Tepperberg JH, Papenhausen P, et al. Subtelomere FISH analysis of 11 688 cases: an evaluation of the frequency and pattern of subtelomere rearrangements in individuals with developmental disabilities. *J Med Genet* 2006;43:478–489.
- Carrozzo R, Rossi E, Christian SL, et al. Inter- and intrachromosomal rearrangements are both involved in the origin of 15q11–q13 deletions in Prader-Willi syndrome. *Am J Hum Genet* 1997;61:228–231.
- Amos-Landgraf JM, Ji Y, Gottlieb W, et al. Chromosome breakage in the Prader-Willi and Angelman syndromes involves recombination between large, transcribed repeats at proximal and distal breakpoints. *Am J Hum Genet* 1999;65:370–386.
- Chen KS, Manian P, Koeuth T, et al. Homologous recombination of a flanking repeat gene cluster is a mechanism for a common contiguous gene deletion syndrome. *Nat Genet* 1997;17:154–163.
- Edelmann L, Pandita RK, Spiteri E, et al. A common molecular basis for rearrangement disorders on chromosome 22q11. *Hum Mol Genet* 1999;8:1157–1167.
- Perez Jurado LA, Peoples R, Kaplan P, Hamel BC, Francke U. Molecular definition of the chromosome 7 deletion in Williams syndrome and parent-of-origin effects on growth. *Am J Hum Genet* 1996;59:781–792.
- Lupski JR, de Oca-Luna RM, Slaugenhaupt S, et al. DNA duplication associated with Charcot-Marie-Tooth disease type 1A. *Cell* 1991;66:219–232.
- Chance PF, Alderson MK, Leppig KA, et al. DNA deletion associated with hereditary neuropathy with liability to pressure palsies. *Cell* 1993;72:143–151.
- Lupski JR. Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits. *Trends Genet* 1998;14:417–422.
- de Stahl TD, Sandgren J, Piotrowski A, et al. Profiling of copy number variations (CNVs) in healthy individuals from three ethnic groups using a human genome 32 K BAC-clone-based array. *Hum Mutat* 2008;29:398–408.
- Iafrate AJ, Feuk L, Rivera MN, et al. Detection of large-scale variation in the human genome. *Nat Genet* 2004;36:949–951.
- Locke DP, Sharp AJ, McCarroll SA, et al. Linkage disequilibrium and heritability of copy-number polymorphisms within duplicated regions of the human genome. *Am J Hum Genet* 2006;79:275–290.
- Pinto D, Marshall C, Feuk L, Scherer SW. Copy-number variation in control population cohorts. *Hum Mol Genet* 2007;16 Spec No. 2:R168–R173.
- Redon R, Ishikawa S, Fitch KR, et al. Global variation in copy number in the human genome. *Nature* 2006;444:444–454.
- Sebat J, Lakshmi B, Troge J, et al. Large-scale copy number polymorphism in the human genome. *Science* 2004;305:525–528.
- Sharp AJ, Locke DP, McGrath SD, et al. Segmental duplications and copy number variation in the human genome. *Am J Hum Genet* 2005;77:78–88.
- Simon-Sanchez J, Scholz S, Fung HC, et al. Genome-wide SNP assay reveals structural genomic variation, extended homozygosity and cell-line induced alterations in normal individuals. *Hum Mol Genet* 2007;16:1–14.
- Zogopoulos G, Ha KC, Naqib F, et al. Germ-line DNA copy number variation frequencies in a large North American population. *Hum Genet* 2007;122:345–353.
- de Vries BB, Pfundt R, Leisink M, et al. Diagnostic genome profiling in mental retardation. *Am J Hum Genet* 2005;77:606–616.
- Koolen DA, Vissers LE, Pfundt R, et al. A new chromosome 17q21.31 microdeletion syndrome associated with a common inversion polymorphism. *Nat Genet* 2006;38:999–1001.
- Sagoo G, Butterworth A, Sanderson S, Shaw-Smith C, Higgins J, Burton H. Array CGH in patients with learning disability (mental retardation) and congenital anomalies: updated systematic review and meta-analysis of 19 studies and 13,926 subjects. *Genet Med* 2009;11:139–146.
- Shaffer LG, Kashork CD, Saleki R, et al. Targeted genomic microarray analysis for identification of chromosome abnormalities in 1500 consecutive clinical cases. *J Pediatr* 2006;149:98–102.
- Sharp AJ, Hansen S, Selzer RR, et al. Discovery of previously unidentified genomic disorders from the duplication architecture of the human genome. *Nat Genet* 2006;38:1038–1042.
- Shaw-Smith C, Pittman AM, Willatt L, et al. Microdeletion encompassing MAPT at chromosome 17q21.3 is associated with developmental delay and learning disability. *Nat Genet* 2006;38:1032–1037.
- Christian SL, Brune CW, Sudi J, et al. Novel submicroscopic chromosomal abnormalities detected in autism spectrum disorder. *Biol Psychiatry* 2008;63:1111–1117.
- Kumar RA, KaraMohamed S, Sudi J, et al. Recurrent 16p11.2 microdeletions in autism. *Hum Mol Genet* 2008;17:628–638.
- Marshall CR, Noor A, Vincent JB, et al. Structural variation of chromosomes in autism spectrum disorder. *Am J Hum Genet* 2008;82:477–488.
- Sebat J, Lakshmi B, Malhotra D, et al. Strong association of de novo copy number mutations with autism. *Science* 2007;316:445–449.
- Szatmari P, Paterson AD, Zwaigenbaum L, et al. Mapping autism risk loci using genetic linkage and chromosomal rearrangements. *Nat Genet* 2007;39:319–328.
- Weiss LA, Shen Y, Korn JM, et al. Association between microdeletion and microduplication at 16p11.2 and autism. *N Engl J Med* 2008;358:667–675.
- Erdogan F, Larsen LA, Zhang L, et al. High frequency of submicroscopic genomic aberrations detected by tiling path array comparative genome hybridisation in patients with isolated congenital heart disease. *J Med Genet* 2008;45:704–709.
- Klopocki E, Schulze J, Strauss G, et al. Complex inheritance pattern resembling autosomal recessive inheritance involving a microdeletion in thrombocytopenia-absent radius syndrome. *Am J Hum Genet* 2007;80:232–240.
- Mefford HC, Clauin S, Sharp AJ, et al. Recurrent reciprocal genomic rearrangements of 17q12 are associated with renal disease, diabetes, and epilepsy. *Am J Hum Genet* 2007;81:1057–1069.
- Richards AA, Santos LJ, Nichols HA, et al. Cryptic chromosomal abnormalities identified in children with congenital heart disease. *Pediatr Res* 2008;64:358–363.
- Menten B, Maas N, Thienpont B, et al. Emerging patterns of cryptic chromosomal imbalance in patients with idiopathic mental retardation and

- multiple congenital anomalies: a new series of 140 patients and review of published reports. *J Med Genet* 2006;43:625–633.
42. Rosenberg C, Knijnenburg J, Bakker E, et al. Array-CGH detection of micro rearrangements in mentally retarded individuals: clinical significance of imbalances present both in affected children and normal parents. *J Med Genet* 2006;43:180–186.
 43. Schoumans J, Ruivenkamp C, Holmberg E, Kyllerman M, Anderlid BM, Nordenskjöld M. Detection of chromosomal imbalances in children with idiopathic mental retardation by array based comparative genomic hybridisation (array-CGH). *J Med Genet* 2005;42:699–705.
 44. Lu X, Shaw CA, Patel A, et al. Clinical implementation of chromosomal microarray analysis: summary of 2513 postnatal cases. *PLoS One* 2007;2:e327.
 45. Stefansson H, Rujescu D, Cichon S, et al. Large recurrent microdeletions associated with schizophrenia. *Nature* 2008;455:232–236.
 46. Stone JL, O'Donovan MC, Gurling H, et al. Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature* 2008;455:237–241.
 47. Walsh T, McClellan JM, McCarthy SE, et al. Rare structural variants disrupt multiple genes in neurodevelopmental pathways in schizophrenia. *Science* 2008;320:539–543.
 48. Kirov G, Grozeva D, Norton N, et al. Support for the involvement of large CNVs in the pathogenesis of schizophrenia. *Hum Mol Genet* 2009;18:1497–1503.
 49. Need AC, Ge D, Weale ME, et al. A genome-wide investigation of SNPs and CNVs in schizophrenia. *PLoS Genet* 2009;5:e1000373.
 50. de Kovel C, Trucks H, Helbig I, et al. Recurrent microdeletions at 15q11.2 and 16p13.11 predispose to idiopathic generalised epilepsies. *Brain*. In press.
 51. Helbig I, Mefford HC, Sharp AJ, et al. 15q13.3 microdeletions increase risk of idiopathic generalized epilepsy. *Nat Genet* 2009;41:160–162.
 52. Dibbens LM, Mullen S, Helbig I, et al. Familial and sporadic 15q13.3 microdeletions in idiopathic generalized epilepsy: precedent for disorders with complex inheritance. *Hum Mol Genet* 2009;18:3626–3631.
 53. Shoichet SA, Waibel S, Endruhn S, et al. Identification of candidate genes for sporadic amyotrophic lateral sclerosis by array comparative genomic hybridization. *Amyotroph Lateral Scler* 2009;10:162–169.
 54. de Cid R, Riveira-Munoz E, Zeeuwen PL, et al. Deletion of the late cornified envelope LCE3B and LCE3C genes as a susceptibility factor for psoriasis. *Nat Genet* 2009;41:211–215.
 55. Jehue FS, Krepisch-Santos AC, Rocha KM, et al. High frequency of submicroscopic chromosomal imbalances in patients with syndromic craniosynostosis detected by a combined approach of microsatellite segregation analysis, multiplex ligation-dependent probe amplification and array-based comparative genome hybridisation. *J Med Genet* 2008;45:447–450.
 56. Bailey JA, Gu Z, Clark RA, et al. Recent segmental duplications in the human genome. *Science* 2002;297:1003–1007.
 57. Stankiewicz P, Lupski JR. Genome architecture, rearrangements and genomic disorders. *Trends Genet* 2002;18:74–82.
 58. Reiter LT, Hastings PJ, Nelis E, De Jonghe P, Van Broeckhoven C, Lupski JR. Human meiotic recombination products revealed by sequencing a hotspot for homologous strand exchange in multiple HNPP deletion patients. *Am J Hum Genet* 1998;62:1023–1033.
 59. Lopez-Corra C, Dorschner M, Brems H, et al. Recombination hotspot in NF1 microdeletion patients. *Hum Mol Genet* 2001;10:1387–1392.
 60. Bayes M, Magano LF, Rivera N, Flores R, Perez Jurado LA. Mutational mechanisms of Williams-Beuren syndrome deletions. *Am J Hum Genet* 2003;73:131–151.
 61. Bi W, Park SS, Shaw CJ, Withers MA, Patel PI, Lupski JR. Reciprocal crossovers and a positional preference for strand exchange in recombination events resulting in deletion or duplication of chromosome 17p11.2. *Am J Hum Genet* 2003;73:1302–1315.
 62. Kurotaki N, Stankiewicz P, Wakui K, Niikawa N, Lupski JR. Sotos syndrome common deletion is mediated by directly oriented subunits within inverted Sos-REP low-copy repeats. *Hum Mol Genet* 2005;14:535–542.
 63. Visser R, Shimokawa O, Harada N, et al. Identification of a 3.0-kb major recombination hotspot in patients with Sotos syndrome who carry a common 1.9-Mb microdeletion. *Am J Hum Genet* 2005;76:52–67.
 64. Koolen DA, Sharp AJ, Hurst JA, et al. Clinical and molecular delineation of the 17q21.31 microdeletion syndrome. *J Med Genet* 2008;45:710–720.
 65. Miller DT, Shen Y, Weiss LA, et al. Microdeletion/duplication at 15q13.2q13.3 among individuals with features of autism and other neuropsychiatric disorders. *J Med Genet* 2008;46:242–248.
 66. Pagnamenta AT, Wing K, Akha ES, et al. A 15q13.3 microdeletion segregating with autism. *Eur J Hum Genet* 2009;17:687–692.
 67. Sharp AJ, Mefford HC, Li K, et al. A recurrent 15q13.3 microdeletion syndrome associated with mental retardation and seizures. *Nat Genet* 2008;40:322–328.
 68. International Schizophrenia Consortium. Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature* 2008;455:237–241.
 69. Stefansson H, Rujescu D, Cichon S, et al. Large recurrent microdeletions associated with schizophrenia. *Nature* 2008;455:232–236.
 70. van Bon BW, Mefford HC, Menten B, et al. Further delineation of the 15q13 microdeletion and duplication syndromes: a clinical spectrum varying from nonpathogenic to a severe outcome. *J Med Genet* 2009;46:511–523.
 71. Ben-Shachar S, Lanpher B, German JR, et al. Microdeletion 15q13.3: a locus with incomplete penetrance for autism, mental retardation, and psychiatric disorders. *J Med Genet* 2009;46:382–388.
 72. Klopocki E, Graul-Neumann LM, Grieben U, et al. A further case of the recurrent 15q24 microdeletion syndrome, detected by array CGH. *Eur J Pediatr* 2008;167:903–908.
 73. Sharp AJ, Selzer RR, Veltman JA, et al. Characterization of a recurrent 15q24 microdeletion syndrome. *Hum Mol Genet* 2007;16:567–572.
 74. Hannes FD, Sharp AJ, Mefford HC, et al. 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* 2008;46:223–232.
 75. Mefford HC, Cooper GM, Zerr T, et al. A method for rapid, targeted CNV genotyping identifies rare variants associated with neurocognitive disease. *Genome Res* 2009;19:1579–1585.
 76. Ullmann R, Turner G, Kirchhoff M, et al. Array CGH identifies reciprocal 16p13.1 duplications and deletions that predispose to autism and/or mental retardation. *Hum Mutat* 2007;28:674–682.
 77. Brunetti-Pierri N, Berg JS, Scaglia F, et al. Recurrent reciprocal 1q21.1 deletions and duplications associated with microcephaly or macrocephaly and developmental and behavioral abnormalities. *Nat Genet* 2008;40:1466–1471.
 78. Christiansen J, Dyck JD, Elyas BG, et al. Chromosome 1q21.1 contiguous gene deletion is associated with congenital heart disease. *Circ Res* 2004;94:1429–1435.
 79. Greenway SC, Pereira AC, Lin JC, et al. De novo copy number variants identify new genes and loci in isolated sporadic tetralogy of Fallot. *Nat Genet* 2009;41:931–935.
 80. Mefford HC, Sharp AJ, Baker C, et al. Recurrent rearrangements of chromosome 1q21.1 and variable pediatric phenotypes. *N Engl J Med* 2008;359:1685–1699.
 81. Osborne LR, Li M, Pober B, et al. A 1.5 million-base pair inversion polymorphism in families with Williams-Beuren syndrome. *Nat Genet* 2001;29:321–325.
 82. Gimelli G, Pujana MA, Patricelli MG, et al. Genomic inversions of human chromosome 15q11-q13 in mothers of Angelman syndrome patients with class II (BP2/3) deletions. *Hum Mol Genet* 2003;12:849–858.
 83. Zody MC, Jiang Z, Fung HC, et al. Evolutionary toggling of the MAPT 17q21.31 inversion region. *Nat Genet* 2008;40:1076–1083.
 84. Kidd JM, Cooper GM, Donahue WF, et al. Mapping and sequencing of structural variation from eight human genomes. *Nature* 2008;453:56–64.
 85. Tuzun E, Sharp AJ, Bailey JA, et al. Fine-scale structural variation of the human genome. *Nat Genet* 2005;37:727–732.
 86. Antonacci F, Kidd JM, Marques-Bonet T, et al. Characterization of six human disease-associated inversion polymorphisms. *Hum Mol Genet* 2009;18:2555–2566.
 87. Mencarelli MA, Katzaki E, Papa FT, et al. Private inherited microdeletion/microduplications: implications in clinical practice. *Eur J Med Genet* 2008;51:409–416.
 88. Mefford H, Eichler E. Duplication hotspots, rare genomic disorders, and common disease. *Curr Opin Genet Dev* 2009;19:196–204.
 89. Slavotinek AM. Novel microdeletion syndromes detected by chromosome microarrays. *Hum Genet* 2008;124:1–17.
 90. Lupski JR. Genomic disorders ten years on. *Genome Med* 2009;1:42.
 91. Gu W, Lupski JR. CNV and nervous system diseases—what's new? *Cytogenet Genome Res* 2008;123:54–64.
 92. Elmslie FV, Rees M, Williamson MP, et al. Genetic mapping of a major susceptibility locus for juvenile myoclonic epilepsy on chromosome 15q. *Hum Mol Genet* 1997;6:1329–1334.
 93. Neubauer BA, Fiedler B, Himmelein B, et al. Centrotemporal spikes in families with rolandic epilepsy: linkage to chromosome 15q14. *Neurology* 1998;51:1608–1612.
 94. Freedman R, Coon H, Myles-Worsley M, et al. Linkage of a neurophysiological deficit in schizophrenia to a chromosome 15 locus. *Proc Natl Acad Sci USA* 1997;94:587–592.
 95. Itsara A, Cooper GM, Baker C, et al. Population analysis of large copy number variants and hotspots of human genetic disease. *Am J Hum Genet* 2009;84:148–161.
 96. Reiner O, Sapir T. Abnormal cortical development; towards elucidation of the LIS1 gene product function (review). *Int J Mol Med* 1998;1:849–853.
 97. Bradshaw NJ, Christie S, Soares DC, Carlyle BC, Porteous DJ, Millar JK. NDE1 and NDEL1: multimerisation, alternate splicing and DISC1 interaction. *Neurosci Lett* 2009;449:228–233.
 98. Mackie S, Millar JK, Porteous DJ. Role of DISC1 in neural development and schizophrenia. *Curr Opin Neurobiol* 2007;17:95–102.

99. Kwon YT, Balogh SA, Davydov IV, et al. Altered activity, social behavior, and spatial memory in mice lacking the NTAN1p amidase and the asparagine branch of the N-end rule pathway. *Mol Cell Biol* 2000;20:4135–4148.
100. Cheroki C, Krepischi-Santos AC, Szuhai K, et al. Genomic imbalances associated with mullerian aplasia. *J Med Genet* 2008;45:228–232.
101. Ni X, Valente J, Azevedo MH, Pato MT, Pato CN, Kennedy JL. Connexin 50 gene on human chromosome 1q21 is associated with schizophrenia in matched case control and family-based studies. *J Med Genet* 2007;44:532–536.
102. Shiels A, Mackay D, Ionides A, Berry V, Moore A, Bhattacharya S. A missense mutation in the human connexin50 gene (GJA8) underlies autosomal dominant “zonular pulverulent” cataract, on chromosome 1q. *Am J Hum Genet* 1998;62:526–532.
103. Vanita V, Hennies HC, Singh D, Nurnberg P, Sperling K, Singh JR. A novel mutation in GJA8 associated with autosomal dominant congenital cataract in a family of Indian origin. *Mol Vis* 2006;12:1217–1222.
104. Gu H, Smith FC, Taffet SM, Delmar M. High incidence of cardiac malformations in connexin40-deficient mice. *Circ Res* 2003;93:201–206.
105. Bijlsma EK, Gijsbers AC, Schuurs-Hoeijmakers JH, et al. Extending the phenotype of recurrent rearrangements of 16p11.2: deletions in mentally retarded patients without autism and in normal individuals. *Eur J Med Genet* 2009;52:77–87.
106. Kumar RA, Marshall CR, Badner JA, et al. Association and mutation analyses of 16p11.2 autism candidate genes. *PLoS One* 2009;4:e4582.
107. Bearden CE, Jawad AF, Lynch DR, et al. Effects of a functional COMT polymorphism on prefrontal cognitive function in patients with 22q11.2 deletion syndrome. *Am J Psychiatry* 2004;161:1700–1702.
108. Gothelf D, Eliez S, Thompson T, et al. COMT genotype predicts longitudinal cognitive decline and psychosis in 22q11.2 deletion syndrome. *Nat Neurosci* 2005;8:1500–1502.
109. Ng SB, Turner EH, Robertson PD, et al. Targeted capture and massively parallel sequencing of twelve human exomes. *Nature* 2009;461:272–276.
110. Shendure J, Ji H. Next-generation DNA sequencing. *Nat Biotechnol* 2008;26:1135–1145.
111. Shendure J, Mitra RD, Varma C, Church GM. Advanced sequencing technologies: methods and goals. *Nat Rev Genet* 2004;5:335–344.
112. Turner EH, Lee C, Ng SB, Nickerson DA, Shendure J. Massively parallel exon capture and library-free resequencing across 16 genomes. *Nat Methods* 2009;6:315–316.