CONFLICT OF INTEREST

The authors declare no conflict of interest.

Gunnar Houge[1,2], Camilla Tøndel[3,4], Øyvind Kaarbøe[5],
Asle Hirth[3,4], Leif Bostad[6,7] and Einar Svarstad[4,8]

[1]Center for Medical Genetics and Molecular Medicine,
Haukeland University Hospital, Bergen, Norway;
[2]Institute of Clinical Medicine, University of Bergen,
Bergen, Norway;
[3]Department of Pediatrics, Haukeland University Hospital,
Bergen, Norway;
[4]Institute of Medicine, University of Bergen, Bergen, Norway;
[5]Department of Medicine, Ålesund Hospital, Ålesund, Norway;
[6]Department of Pathology, Haukeland
University Hospital, Bergen, Norway;
[7]The Gade Institute, University of Bergen, Bergen, Norway;
[8]Department of Medicine, Haukeland
University Hospital, Bergen, Norway
E-mail: gunnar.houge@helse-bergen.no

1 Schiffmann R: Fabry disease. *Pharmacol Ther* 2009; **122**: 65–77.
2 Houge G, Skarbovik AJ: Fabry disease – a diagnostic and therapeutic challenge. *Tidsskr Nor Laegeforen* 2005; **125**: 1004–1006.
3 Linthorst GE, Bouwman MG, Wijburg FA, Aerts JM, Poorthuis BJ, Hollak CE: Screening for Fabry disease in high-risk populations: a systematic review. *J Med Genet* 2010; **47**: 217–222.
4 Aasebø W, Strøm EH, Hovig T, Undset LH, Heiberg A, Jenssen T: Fabry disease in donor kidneys with 3- and 12-year follow-up after transplantation. *NDT Plus* 2010; **3**: 303–305.
5 Dobrovolny R, Dvorakova L, Ledvinova J *et al*: Recurrence of Fabry disease as a result of paternal germline mosaicism for alpha-galactosidase a gene mutation. *Am J Med Genet A* 2005; **134A**: 84–87.
6 Tondel C, Bostad L, Laegreid LM, Houge G, Svarstad E: Prominence of glomerular and vascular changes in renal biopsies in children and adolescents with Fabry disease and microalbuminuria. *Clin Ther* 2008; **30**(Suppl B): S42.

The causality of *de novo* copy number variants is overestimated

The application of array CGH or chromosomal microarrays is causing a revolutionary change in clinical genetics and especially cytogenetics, as it enables the genome wide identification of submicroscopic copy number variations (CNVs).[1] Given the significant increase in diagnostic yield compared with conventional karyotyping in patients with intellectual disability (ID) and the technical ease of use, the technique is now recommended as a first tier diagnostic test for patients with ID and/or multiple congenital anomalies (MCA).[2,3] Arrays not only enable detection of disease-causing CNVs in patients with ID/MCA, but also in patients with isolated heart defects, neurological diseases and psychiatric disorders.

Therefore, besides pediatricians and clinical geneticists, more and more other medical specialists request array analysis arrays.[4–7] In addition, there is a rapid implementation of array CGH in prenatal diagnosis.[8–11]

For a small subset of CNVs, the association with an ID/MCA phenotype is beyond doubt. However, many CNVs detected using high-resolution arrays remain private, or the functional relationship with the phenotype is, at best, vague. To enable genotype–phenotype correlations, databases collecting phenotypes and genotypes have been established.[3,12,13] Although those databases have been successful in establishing a functional relationship for recurrent CNVs, a genotype–phenotype relationship has yet to be established for a majority of rare CNVs. Variable expressivity and reduced penetrance often confound significant associations to be made and is especially challenging for rare variants.[1] Due to the difficulty of associating CNVs with a phenotype, several reports provide guidelines on how clinical laboratories can interpret array results. If a clear association between the phenotype under investigation and the CNV is lacking, a series of steps are guiding the interpretation of the clinical significance. In all guidelines, a rule of thumb is that *de novo* CNVs, not occurring in normal individuals, are considered causal for the abnormal phenotype.[1,3,14–16]

During our screen of patients with mental retardation and developmental disorders, we identified several private *de novo* CNVs. Previously, we reported a 250 kb *de novo* deletion in C20orf133, nowadays known as *MACROD2*, in a patient with Kabuki syndrome.[17] A highly conserved region of C20orf133, likely to have a role in chromatin or chromosome biology, was deleted, and the gene was shown to be expressed in mice in the tissues affected in Kabuki syndrome. Although screening of 62 Kabuki syndrome patients failed to identify mutations in C20orf133, the disorder was hypothesized to be genetically heterogeneous.[17,18] Recently, mutations were identified in *MLL2* in 33 out of 50 Kabuki syndrome patients.[19] Sequencing of this gene in the patient with the C20orf133 deletion, shows the presence of a *de novo* mutation in *MLL2*.[20]

Most recently, we had a similar experience with another *de novo* CNV. During a screen of patients with ID and eye disorders, we identified a *de novo* 86.5 kb deletion in a patient referred because of an eye malformation associated with mental retardation. The deletion harbored only a single gene, *AMBRA1*. As the gene is expressed in the neural retina and brain, and mice knock-outs result in exencephaly,[21] the deletion was considered a likely cause for the observed phenotype. To further establish the clinical relationship between the gene and the phenotype, morpholino knockdowns were performed in zebrafish, which resulted in eye coloboma as well as equilibrium defects. On the basis of these observations and during the preparation of a manuscript, a more detailed phenotypic description of the patient was requested. The patient was characterized by bilateral coloboma, anosmia, disturbance of the equilibrium and ID, a CHARGE like phenotype. Once the diagnosis of CHARGE was uttered, a mutation analysis of *CHD7* was instigated.[22] A *de novo* mutation was identified causing a splice site mutation, thus disrupting the *CHD7* gene.

Although smaller modifier effects of the deletions on the phenotype cannot be excluded, the overall phenotype can in both cases be explained by *de novo* point mutations rather than to *de novo* CNVs. Note that CNV detection in both cases was performed on DNA extracted from blood and not on cell cultures. The latter are known to accumulate chromosomal rearrangements and hence, its use for CNV detection should be avoided for clinical diagnosis. We believe that these two case reports are representative of a larger number of misinterpretations that are currently made in diagnostic laboratories

offering array CGH testing. Two important messages for the human (cyto) genetics community can now be drawn.

First, despite the flurry of schemes suggesting that *de novo* CNVs can be interpreted and counseled to patients as causal for the investigated phenotype,[1,3,14–16] this is precocious. These guidelines are based on the (unwritten) assumption that the co-occurrence of both are a rare phenotype with a rare mutational event is statistically unlikely. For large, microscopically visible CNVs harboring multiple genes, this assumption is true. This is the type of rearrangement that (cyto) geneticists have come acquainted with over the last 50 years. However, for smaller CNVs, the chance may be significantly higher. Direct estimates of the genome-wide CNV mutation rates from family studies have estimated the CNV mutation rate to be in the range of 1.2×10^{-2} CNVs per haploid genome per transmission at a median resolution of 150 kb, amounting to about 2.5 CNVs/100 live births.[23–25] One study also demonstrated that larger imbalances (> 500 kb) are significantly more enriched than smaller imbalances in the autism spectrum population as compared with a control cohort of normal individuals, which is not surprising, as larger imbalances harbor more genes.[26] On the basis of the fourfold increased incidence of *de novo* CNVs in this autism population, as compared with the unaffected siblings, the number of CNVs in the size range of 60–500 kb wrongly classified as causative can be estimated at about one in five. At present, a direct estimate of the mutation rate for imbalances < 60 kb is lacking, but it can be conceived that the frequency of *de novo* mutation events not leading to developmental anomalies exceeds the number of events that do hit genes causing such disorders. The indirect estimate of *de novo* CNV < 500 bp rate is estimated at a minimum of $6 \times 10^{-2}/$ diploid genome per generation.[24,26] In an extrapolation of the frequency, CNVs cause Duchenne muscular dystrophy, leading to an estimate of one deletion every eight generations and a duplication of 1/50 generations.[27] As a consequence, smaller *de novo* imbalances cannot automatically be classified as likely causal for the investigated phenotype in the absence of strong evidence from other data sources. Recently, novel bioinformatic approaches have been developed to aid the clinical interpretation. These include the incorporation of structural and functional genomic features to distinguish pathogenic from benign CNVs.[28,29] Although those approaches will greatly improve the interpretation, they remain indirect. The extensive collection of CNVs and associated phenotypes in common databases will be a prerequisite for proper clinical interpretation of CNVs.

A second observation is that a molecular diagnosis can only be as good as the clinical diagnosis. In the second patient, before having a thorough clinical work-up, arrays were requested and an apparently causal CNV was detected. However, completeness of the clinical information directed the genetic testing, which subsequently enabled a proper molecular diagnosis to be made. It can be anticipated that in the near future, full exome or genome sequencing will be offered as a clinical diagnostic test.[30,31] A genomic sequence will offer the apparent security that a full genome is analyzed. However, a genomic analysis of all variants will only enable the causative variants to be identified in relation to a well-defined clinical question.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

Joris R Vermeesch[1], Irina Balikova[1], Connie Schrander-Stumpel[2],
Jean-Pierre Fryns[1] and Koenraad Devriendt[1]
[1]Centre for Human Genetics, University Hospital,
K.U. Leuven, Leuven, Belgium;
[2]Departments of Clinical Genetics, Maastricht UMC+,
Maastricht, The Netherlands
E-mail: Joris.Vermeesch@med.kuleuven.be

1 Lee C, Iafrate AJ, Brothman AR: Copy number variations and clinical cytogenetic diagnosis of constitutional disorders. *Nat Genet* 2007; **39**: S48–S54.

2 Vermeesch JR, Fiegler H, de Leeuw N *et al*: Guidelines for molecular karyotyping in constitutional genetic diagnosis. *Eur J Hum Genet* 2007; **15**: 1105–1114.

3 Miller DT, Adam MP, Aradhya S *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.

4 Thienpont B, Breckpot J, Vermeesch JR, Gewillig M, Devriendt K: A complex sub-microscopic chromosomal imbalance in 19p13.11 with one microduplication and two microtriplications. *Eur J Med Genet* 2008; **51**: 219–225.

5 Balikova I, de Ravel T, Ayuso C *et al*: High frequency of submicroscopic chromosomal deletions in patients with idiopathic congenital eye malformations. *Am J Ophthalmol* 2011; e-pub ahead of print 23 February 2011.

6 Breckpot J, Thienpont B, Peeters H *et al*: Array comparative genomic hybridization as a diagnostic tool for syndromic heart defects. *J Pediatr* 2010; **156**: 810–817.

7 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.

8 Beaudet AL: Ethical issues raised by common copy number variants and single nucleotide polymorphisms of certain and uncertain significance in general medical practice. *Genome Med* 2010; **2**: 42.

9 Faas BH, van der Burgt I, Kooper AJ *et al*: Identification of clinically significant, submicroscopic chromosome alterations and UPD in fetuses with ultrasound anomalies using genome-wide 250k SNP array analysis. *J Med Genet* 2010; **47**: 586–594.

10 Tyreman M, Abbott KM, Willatt LR *et al*: High resolution array analysis: diagnosing pregnancies with abnormal ultrasound findings. *J Med Genet* 2009; **46**: 531–541.

11 Van den Veyver I, Patel A, Shaw CA *et al*: Clinical use of array comparative genomic hybridization (aCGH) for prenatal diagnosis in 300 cases. *Prenat Diagn* 2009; **29**: 29–39.

12 Feenstra I, Fang J, Koolen DA *et al*: European Cytogeneticists Association Register of Unbalanced Chromosome Aberrations (ECARUCA); an online database for rare chromosome abnormalities. *Eur J Med Genet* 2006; **49**: 279–291.

13 Firth HV, Richards SM, Bevan AP *et al*: DECIPHER: database of chromosomal imbalance and phenotype in humans using ensembl resources. *Am J Hum Genet* 2009; **84**: 524–533.

14 Buysse K, Delle CB, Van Coster R *et al*: Challenges for CNV interpretation in clinical molecular karyotyping: lessons learned from a 1001 sample experience. *Eur J Med Genet* 2009; **52**: 398–403.

15 Gijsbers AC, Lew JY, Bosch CA *et al*: A new diagnostic workflow for patients with mental retardation and/or multiple congenital abnormalities: test arrays first. *Eur J Hum Genet* 2009; **17**: 1394–1402.

16 Koolen DA, Pfundt R, de Leeuw N *et al*: Genomic microarrays in mental retardation: a practical workflow for diagnostic applications. *Hum Mutat* 2009; **30**: 283–292.

17 Maas NM, Van de PT, Melotte C *et al*: The C20orf133 gene is disrupted in a patient with Kabuki syndrome. *J Med Genet* 2007; **44**: 562–569.

18 Kuniba H, Tsuda M, Nakashima M *et al*: Lack of C20orf133 and FLRT3 mutations in 43 patients with Kabuki syndrome in Japan. *J Med Genet* 2008; **45**: 479–480.

19 Ng SB, Bigham AW, Buckingham KJ *et al*: Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nat Genet* 2010; **42**: 790–793.

20 Paulussen AD, Stegmann KM, Blok MJ *et al*: MLL2 mutation spectrum in 45 patients with Kabuki syndrome. *Hum Mutat* 2011; **32**: E2018–E2025.

21 Fimia GM, Stoykova A, Romagnoli A *et al*: Ambra1 regulates autophagy and development of the nervous system. *Nature* 2007; **447**: 1121–1125.

22 Vissers LE, van Ravenswaaij CM, Admiraal R *et al*: Mutations in a new member of the chromodomain gene family cause CHARGE syndrome. *Nat Genet* 2004; **36**: 955–957.

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

24 Conrad DF, Pinto D, Redon R *et al*: Origins and functional impact of copy number variation in the human genome. *Nature* 2010; **464**: 704–712.

25 Xu B, Roos JL, Levy S, van Rensburg EJ, Gogos JA, Karayiorgou M: Strong association of *de novo* copy number mutations with sporadic schizophrenia. *Nat Genet* 2008; **40**: 880–885.

26 Itsara A, Wu H, Smith JD *et al*: De novo rates and selection of large copy number variation. *Genome Res* 2010; **20**: 1469–1481.

27 van Ommen GJ: Frequency of new copy number variation in humans. *Nat Genet* 2005; **37**: 333–334.

28 Hehir-Kwa JY, Wieskamp N, Webber C *et al*: Accurate distinction of pathogenic from benign CNVs in mental retardation. *PLoS Comput Biol* 2010; **6**: e1000752.

29 Huang N, Lee I, Marcotte EM, Hurles ME: Characterising and predicting haploinsufficiency in the human genome. *PLoS Genet* 2010; **6**: e1001154.

30 Vissers LE, de Ligt J, Gilissen C *et al*: A *de novo* paradigm for mental retardation. *Nat Genet* 2010; **42**: 1109–1112.

31 Katsnelson A. Human genome: genomes by the thousand. *Nature* 2010; **467**: 1026–1027.