Late Breaking Chromosomes

Beware of Hemizygous Deletions That May Unmask Deleterious Variants

With array-based genome-wide analyses of copy number variations (CNVs), in up to 19% of patients with idiopathic mental retardation or developmental delay and other complex disorders, pathogenic segmental aneuploidies have been detected [Hochstenbach et al., 2011]. Deletions, being either terminal or interstitial and including recurrent microdeletion syndromes, account for roughly half of the genomic imbalances in such patients. De novo deletions in sporadic patients are generally believed to be pathogenic [Gijsbers et al., 2011]. Genes within an inherited deletion, however, cannot provoke a phenotypic effect by mere haploinsufficiency, since these genes were also found in a single copy in a healthy parent. Such inherited deletions may, however, still contribute to disease when combined with rare de novo or inherited deleterious variants on the other transmitted chromosome [Coman and Gardner, 2007]. Thus, by screening for variants in regions corresponding to inherited hemizygous deletions, we may be able to identify genes that exert a phenotypic effect by the independent loss of both alleles. This mechanism, termed ‘unmasking heterozygosity’, may often have escaped detection and is therefore considered an infrequent cause of genetic abnormalities [Coman and Gardner, 2007].

Two groups of investigators took 2 distinct approaches to delineate possible unmasking heterozygosity. Hochstenbach et al. [2012] identified 20 patients with mental retardation or developmental delay and inherited deletions. Subsequently, they performed multiplexed genomic enrichment and next-generation sequencing of the entire coding sequence of all genes in these deletion regions. After mapping the sequence reads to the human reference

genome, they evaluated 703 single nucleotide variants (SNVs) with a predicted effect at the protein level. Thus, they identified 8 SNVs that were located exclusively within the corresponding deletions of their patients. Upon evaluation for their phenotypic impact using the Grantham matrix score [Grantham, 1974], Genomic Evolutionary Rate Profile (GERP) [Cooper et al., 2005; Davydov et al., 2010], PANTHER [Thomas et al., 2003], and PolyPhen2 [Adzhubei et al., 2010], the authors retained no plausibly pathogenic SNV [Hochstenbach et al., 2012].

In one of the patients, the inherited deletion unmasked another small inherited deletion on the second allele, which comprised the entire *HSBP1* gene [Hochstenbach et al., 2012]. Population analyses of large CNVs has revealed a frequency of 5–10% in healthy individuals for CNVs of 500 kb and larger, which is in the same range as in patients with developmental delay [Cooper et al., 2011]. Since size and gene content of CNVs are negatively correlated with their population frequency [Itsara et al., 2009], smaller structural variants may occur at even higher frequencies in the healthy population. Two independent studies indicate that CNVs of 500 bp and larger occur at a median frequency of more than 1,000 CNVs per healthy individual [Itsara et al., 2009; Conrad et al., 2010]. Thus, it is conceivable that compound inheritance of 2 deletions, albeit of different size, may have caused the disease of the patient who has retained no intact *HSBP1* alleles [Hochstenbach et al., 2012].

Albers et al. [2012], in contrast, selected 55 patients with the thrombocytopenia absent radius syndrome (TAR; OMIM 274000), which is a contiguous gene deletion syndrome with apparent autosomal recessive inheritance [Klopocki et al., 2007] and performed exome sequencing of 5 TAR patients. These authors too, did not find any coding mutations in the retained alleles of the

genes in the deletion region. However, in 4 out of 5 patients they found the same variant in the 5' UTR of the *RBM8A* gene (position 1:145,507,646, G/A), which is listed in dbSNP (v. 135) as rs139428292. Analysis of *RBM8A* by standard capillary sequencing in all other patients uncovered unmasking of this and 3 other variants in 46 out of 48 patients with a transmitted hemizygous deletion of the ATR-associated region. With elegant functional assays, the authors show that these variants indeed provoke less transcription of *RBM8A* and affect platelet function, thus corroborating that these variants are indeed pathogenic. These data convincingly show that the TAR syndrome may result from compound inheritance of a hemizygous deletion and either a low-frequency regulatory SNP or a null mutation, each affecting expression of the *RBM8A* gene.

While the paper by Albers et al. [2012] conclusively demonstrated the pathogenic impact of the combination of mutated alleles of *RBM8A*, the frequency of such combinations needs some consideration. Based on their population screening data, a compound heterozygote of a 5' UTR and an intronic SNV would occur in 1.28 per 10,000 people. A combination of the 200 kbp hemizygous deletion with the SNP in the 5' UTR would occur in less than 5 per million. While there are no data on the frequency of TAR in the human population, the reported

frequency of *RBM8A* variants suggest that some patients may have gone unnoticed or may show a phenotype that is not recognized as TAR. It is conceivable that combinations of mutations in *RBM8A* may provoke a whole spectrum of clinical phenotypes. In addition, the estimated contribution of inherited hemizygous deletions to clinical TAR, 5 per million versus 1.28 per 10,000, suggests that hemizygous deletions involving *RBM8A* may be involved in roughly 4% of cases. The latter compares well with mutation data on the autosomal recessive Werner syndrome [Friedrich et al., 2011].

Taking together these 2 papers demonstrate that unmasking heterozygosity may indeed be involved in some patients with an inherited deletion in which the disease seems to follow an autosomal recessive pattern of inheritance. By limiting genetic analyses to the coding regions of genes, as is done by most exome sequencing studies, one stands the risk to actually miss contributing variants, such as the variant in the 5' UTR of *RBM8A* or a deletion of the second allele being transmitted by a healthy parent. These 2 papers have widened the scope of pathogenic variants that may underlie even 'simple Mendelian diseases'. With this in mind, a reappraisal of the myriad genetic mechanisms that may be involved in 'complex disorders' is certainly needed [Poot et al., 2011].

M. Poot

References

- Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, et al: A method and server for predicting damaging missense mutations. *Nat Methods* 7:248–249 (2010).
- Albers CA, Paul DS, Schulze H, Freson K, Stephens JC, et al: Compound inheritance of a low-frequency regulatory SNP and a rare null mutation in exon-junction complex subunit *RBM8A* causes TAR syndrome. *Nat Genet* 44:435–439 (2012).
- Coman DJ, Gardner RJ: Deletions that reveal recessive genes. *Eur J Hum Genet* 15:1103–1104 (2007).
- Conrad DF, Pinto D, Redon R, Feuk L, Gokcumen O, et al: Origins and functional impact of copy number variation in the human genome. *Nature* 464:704–712 (2010).
- Cooper GM, Stone EA, Asimenos G; NISC Comparative Sequencing Program, Green ED, et al: Distribution and intensity of constraint in mammalian genomic sequence. *Genome Res* 15:901–913 (2005).
- Cooper GM, Coe BP, Girirajan S, Rosenfeld JA, Vu TH, et al: A copy number variation morbidity map of developmental delay. *Nat Genet* 43:838–846 (2011).
- Davydov EV, Goode DL, Sirota M, Cooper GM, Sidow A, Batzoglou S: Identifying a high fraction of the human genome to be under selective constraint using GERP++. *PLoS Comput Biol* 6:e1001025 (2010).
- Friedrich K, Lee L, Leistritz DF, Nürnberg G, Saha B, et al: *WRN* mutations in Werner syndrome patients: genomic rearrangements, unusual intronic mutations and ethnic-specific alterations. *Hum Genet* 128:103–111 (2010).
- Gijsbers AC, Schoumans J, Ruivenkamp CA: Interpretation of array comparative genome hybridization data: a major challenge. *Cytogenet Genome Res* 135: 222–227 (2011).
- Grantham R: Amino acid difference formula to help explain protein evolution. *Science* 185: 862–864 (1974).
- Hochstenbach R, Buizer-Voskamp JE, Vorstman JA, Ophoff RA: Genome arrays for the detection of copy number variations in idiopathic mental retardation, idiopathic generalized epilepsy and neuropsychiatric disorders: lessons for diagnostic workflow and research. *Cytogenet Genome Res* 135: 174–202 (2011).
- Hochstenbach R, Poot M, Nijman IJ, Renkens I, Duran KJ, et al: Discovery of variants unmasked by hemizygous deletions. *Eur J Hum Genet* (2012), E-pub ahead of print.
- Itsara A, Cooper GM, Baker C, Girirajan S, Li J, et al: Population analysis of large copy number variants and hotspots of human genetic disease. *Am J Hum Genet* 84:148–161 (2009).
- Klopocki E, Schulze H, Strauss G, Ott CE, Hall J, et al: Complex inheritance pattern resembling autosomal recessive inheritance involving a microdeletion in thrombocytopenia-absent radius syndrome. *Am J Hum Genet* 80:232–240 (2007).
- Poot M, van der Smagt JJ, Brilstra EH, Bourgeron T: Disentangling the myriad genomics of complex disorders, specifically focusing on autism, epilepsy, and schizophrenia. *Cytogenet Genome Res* 135:228–240 (2011).
- Thomas PD, Campbell MJ, Kejariwal A, Mi H, Karlak B, et al: PANTHER: a library of protein families and subfamilies indexed by function. *Genome Res* 13:2129–2141 (2003).