Reply to Dr. Robert A. Hegele

To the Editor: Dr. Hegele called attention to the disparities between the observed frequencies of copy-number variations (CNVs) and the reported frequencies of some OMIM diseases.[1(in this issue)] We share Dr. Hegele's concern and emphasize the importance of interpreting CNV data with caution, with respect to correlations with genes or phenotypes.

A number of issues must be considered when reported CNV data are used. One is that the exact boundaries of CNVs are rarely known; thus, direct correlations with genes can be challenging. The *BSCL2* (MIM 606158) gene, noted by Dr. Hegele, overlaps the boundary of a BAC clone that we reported as variable in copy number.[2] In this case, approximately one-third of the *BSCL2* gene overlaps the end of a CNV clone (this can be viewed using our CNV custom track, now publicly available at the UCSC Genome Browser). However, the gene may not be part of the CNV. It is therefore necessary to confirm, using alternative validation technologies, whether a gene is actually affected by a particular CNV. Depending on the detection sensitivities and resolutions of the array platforms, the boundaries of the reported CNVs will be within tens to hundreds of kilobases from the actual boundaries. When a combined data set such as the Database of Genomic Variants[3] is used, it is very important to be mindful of the strengths and weaknesses of the platforms used to derive the data, such as resolution, detection sensitivity, and false-positive and false-negative rates.

In our study, we reported genes as CNV associated if any part of the gene overlapped a BAC clone that we measured to be variable in copy number.[2] There is a need for standardization in the reporting of which genes are potentially associated with CNVs. Furthermore, it is biologically difficult to know when a CNV will influence the expression of a gene, given that position effects are known to influence the expression of genes across hundreds of kilobases.[4]

An additional confounding issue when interpreting CNV data obtained from array comparative genomic hybridization is the comparative nature of the technique. Gains and losses are called in relation to a reference DNA, which will vary by study and are often simplistically interpreted as a single-copy change from diploid. In fact, the exact copy number is often not known for either the sample or the reference. In the case of the CNV associated with the *BSCL2* gene, we detected two samples that showed a gain relative to the reference and one sample that showed a loss in copy number relative to the reference, demonstrating the complexity of changes occurring in the genome. A further consideration is that, in some cases, the baseline copy number could be greater than two. Genes in such regions may be particularly resistant to disease-causing mutations because of functional redundancy.

In summary, CNV data provide valuable information for studies involving human genetics, and the abundance of CNVs means that they are likely to include or influence many genes; however, the data need to be used with caution.

Kendy K. Wong, Ronald J. deLeeuw, Carolyn J. Brown, and Wan L. Lam

Web Resources

The URLs for data presented herein are as follows:

Database of Genomic Variants, http://projects.tcag.ca/variation/
Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for *BSCL2*)
UCSC Genome Browser, http://genome.ucsc.edu/goldenPath/customTracks/custTracks.html

References

1. Hegele RA (2007) Copy-number variations and human disease. Am J Hum Genet 81:414–415 (in this issue)
2. Wong KK, deLeeuw RJ, Dosanjh NS, Kimm LR, Cheng Z, Horsman DE, MacAulay C, Ng RT, Brown CJ, Eichler EE, et al (2007) A comprehensive analysis of common copy-number variations in the human genome. Am J Hum Genet 80:91–104
3. Zhang J, Feuk L, Duggan GE, Khaja R, Scherer SW (2006) Development of bioinformatics resources for display and analysis of copy number and other structural variants in the human genome. Cytogenet Genome Res 115:205–214
4. Kleinjan DJ, van Heyningen V (1998) Position effect in human genetic disease. Hum Mol Genet 7:1611–1618

From the Department of Cancer Genetics and Developmental Biology, British Columbia Cancer Research Centre (K.K.W.; R.J.d.; W.L.L.), and the Department of Medical Genetics, University of British Columbia (C.J.B.), Vancouver

Address for correspondence and reprints: Dr. Kendy K. Wong, 675 W. 10th Avenue, Vancouver, BC, V5Z 1L3 Canada. E-mail: kwong@bccrc.ca

Am. J. Hum. Genet. 2007;81:415.

0002-9297/2007/8102-0023$15.00
DOI: 10.1086/519223