LETTER TO THE EDITOR

Reply to Lynch et al.

To the Editor: The letter by Lynch et al.[1(in this issue)] described the application of more-robust statistical modeling for the determination of false-negative and false-positive rates in our copy-number variation (CNV) study. Their conclusion is that the inclusion of dependence in the model increases the false-negative rate while leaving the false-positive rate unaltered.

These findings raised key questions as to what methodology should be employed to quantify false-positive and false-negative rates in CNV data. To determine false-detection rates, are single experiments repeated multiple times preferable to single replication of many experiments or, alternatively, use of self- versus self-hybridization experiments? Within currently published CNV studies,[2–4] which were derived from different array platforms, these methods have been employed individually or in combinations in some studies, whereas others employed completely different methods of quality assessment.[5] Clearly, there is a need for standardization of methods for determining these rates.

We acknowledge that our analysis of false-positive and false-negative rates did not account for the dependence between repeated experiments, although Lynch et. al.[1] determined that the false-positive rate (denoted as "*q*") was not "dramatically altered."[1(p419)] In fact, on the basis of their criteria, we have gained confidence in a greater number of CNV calls than the 800 reported as "high-frequency CNVs" in our original publication[4(p99)]—that is, an additional 736 CNVs seen in only 2 of the 95 individuals (see data set 2 in the online version of our article[4]). The increase in the false-negative rate (i.e., decrease in *p*) would have broad implications. If the false-negative rate is as high as Lynch et al. proposed (~60%), the benefit of repeating every experiment with the fluorochromes reversed and eliminating the CNVs not seen in both experiments (also know as "flip-fluor experiments") would be offset by the erroneous elimination of a major portion of real data. Specifically, by achieving a relatively small false-positive rate, flip-fluor repeat experiments (with a false-negative rate of 60%) will capture only 16% of the true CNVs in a given experiment. This raises the question of whether such a practice would be unacceptable if we wish to identify all CNVs in the human population.

Currently, there are >6,000 CNVs noted in the Database of Genomic Variants that affect >3,500 loci.[6] The meta-analysis of the various CNV studies is a major challenge. With the diverse array of platforms employed, it is important to consider the advantages and limitations of each study, since array resolution, DNA reference, genome coverage, and cohort composition vary greatly. Given the limited overlap between individual studies and the indication by Lynch et al.[1] that we are vastly underestimating their prevalence, there are likely tens of thousands of CNVs to be discovered.

Ronald J. deLeeuw, Kendy K. Wong,
Raymond T. Ng, and Wan L. Lam

Web Resource

The URL for data presented herein is as follows:

Database of Genomic Variants,http://projects.tcag.ca/variation/

References

1. Lynch AG, Marioni JC, Tavaré S (2007) Numbers of copy-number variations and false-negative rates will be underestimated if we do not account for the dependence between repeated experiments. Am J Hum Genet 81:418–420 (in this issue)
2. Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, et al (2006) Global variation in copy number in the human genome. Nature 444: 444–454
3. Sharp AJ, Locke DP, McGrath SD, Cheng Z, Bailey JA, Vallente RU, Pertz LM, Clark RA, Schwartz S, Segraves R, et al (2005) Segmental duplications and copy-number variation in the human genome. Am J Hum Genet 77:78–88
4. 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
5. Hinds DA, Kloek AP, Jen M, Chen X, Frazer KA (2006) Common deletions and SNPs are in linkage disequilibrium in the human genome. Nat Genet 38:82–85
6. 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

From the Department of Cancer Genetics and Developmental Biology, British Columbia Cancer Research Centre (R.J.d.; K.K.W.; W.L.L.), and the Department of Computer Science, University of British Columbia (R.T.N.), Vancouver

Address for correspondence and reprints: Dr. Ronald J. deLeeuw, 675 W. 10th Avenue, Vancouver, BC V5Z 1L3, Canada. E-mail: rdeleeuw@bccrc.ca

Am. J. Hum. Genet. 2007;81:420.

0002-9297/2007/8102-0027$15.00
DOI: 10.1086/519395