3. Lockhart, D. *et al.*. Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat. Biotechnol.* **14**, 1675–1680 (1996).
4. Affymetrix. *Affymetrix Microarray Suite User Guide. Version 5* (Affymetrix, Santa Clara, CA, USA, 2001).
5. Durbin, B.P., Hardin, J.S., Hawkins, D.M. & Rocke, D.M. *Bioinformatics* **18**, S105–S110 (2002).
6. Wu, Z. & Irizarry, R.A. *Proceedings of RECOMB 2004*, 98–106 (2004).
7. Irizarry, R.A. *et al. Nucleic Acids Research* **31**, e15 (2003).
8. Hekstra, D., Taussig, A.R., Magnasco, M. & Naef, F. *Nucleic Acids Res.* **31**, 1962–1968 (2003).
9. Wu, Z., Irizarry, R.A., Gentleman, R., Murillo, F.M. & Spencer, F. *A Model Based Background Adjustment for Oligonucleotide Expression Arrays. Technical Report* (John Hopkins University, Department of Biostatistics Working Papers, Baltimore, MD, 2003).
10. Li, C. & Wong, W. *Proc. Natl. Acad. Sci. USA*, **98**, 31–36 (2001).

Zhijin Wu & Rafael A Irizarry

Department of Biostatistics, Bloomberg School of Public Health, Johns Hopkins University, 615 North Wolfe Street, Baltimore, Maryland 21205, USA. e-mail: ririzarr@jhsph.edu

Zhang et al. respond:

Our study showed that the log-transformed gene expression level estimated by the PerfectMatch algorithm is linearly related to the log-transformed nominal concentration. This log-linear relationship has a slope <1, which leads to an underestimation of fold-change in expression, as noted in our paper. The bias is easily correctable by rescaling the log-transformed gene expression level by a fixed factor, according to the slope. Our analysis of the spike-in data s*et al*lowed us to calibrate this factor to be around 2. Because the slope bias appears to be consistent for all spike-in genes, this factor is expected to be generally applicable so that it is unnecessary to recalibrate it in routine use of the technology.

As long as the log-linear relationship holds, the slope bias *per se* should have no effect on our power to identify differentially expressed genes, nor should it change the shape of gene expression profiles. Therefore, the slope bias doesn't seem to us the most important issue in practical studies of gene expression. We think it is more relevant to evaluate algorithms in terms of sensitivity and specificity in the context of identifying differential gene expression. A commonly used tool for this purpose is AUC. In a previous publication by Irizarry's group, Cope *et al.*[1] compute AUC defined as area under the Receiver operating characteristic (ROC) curve up to 100 false positives for GCRMA, RMA, MAS 5.0, dChip and PerfectMatch algorithms and find AUC values of 0.82, 0.82, 0.36, 0.67, and 0.84, respectively[2]. These results indicate that GCRMA, RMA and PerfectMatch gave comparable performances, whereas MAS 5.0 and dChip performed poorly.

However, these results are inconsistent with **Supplementary Figure 2** of Irizarry's correspondence. The origin of the inconsistency is not clear because of insufficient information given in the correspondence. For a more detailed discussion of the related issues, please see our own **Supplementary Note** online together with **Supplementary Figure 1**.

Note: Supplementary information is available on the Nature Biotechnology website.

1. Cope, L.M., Irizarry, R.A., Jaffee, H.A., Wu, Z. & Speed TP. *Bioinformatics* **20**, 323–331 (2004).
2. http://affycomp.biostat.jhsph.edu/AFFY2/TABLES.hgu/0.html

Li Zhang, Chunlei Wu, Roberto Carta, Keith Baggerly & Kevin R Coombes

Department of Biostatistics and Applied Mathematics, the University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd, Box 447, Houston, Texas 77030, USA. email: lzhang@mdanderson.org