

sufficient to generate the cranial changes found in mouse models of Down syndrome with larger numbers of triplicated genes. Furthermore, reducing trisomy of these 33 genes to disomy in the Ts65Dn mouse did not eliminate the phenotype.

Breeding mice with the deleted chromosome segment with trisomy mouse models is a particularly elegant approach to testing the role of the DSCR segment in Down syndrome. It would appear that triplication of the 33 DSCR genes is not necessary at least for the craniofacial alterations characteristic of the disease, thus reducing the likelihood of a contiguous critical region for this aspect of Down syndrome. The authors favor a model in which individual or small numbers of genes can make a “critical” contribution to Down syndrome, but where the effect is highly contextual, depending on the combined effects of altering the dosage of other genes.

The generation of these new mouse

models will allow additional study of the association of these 33 genes with other Down syndrome abnormalities associated with behavior, electrophysiology, and loss of cerebellar granule cells (9). It will also be interesting to increase the size of the duplicated chromosome segments to more closely mimic the human disorder, although if Olson *et al.* are correct that combinations of genes of small (or no) individual effect can contribute to the overall phenotype, the numbers of permutations are daunting.

Mice remain our most useful genetic relative for modeling human disorders, despite numerous differences that complicate analysis. For diseases involving mental retardation, this is a particularly acute problem, as alterations in behavior and learned tasks must suffice to flag differences in mental acuity between mutant mice and their normal counterparts (10). Another challenge is the difference in colinearity of the human and mouse

genomes, and the lack of conservation of gene order. Even though we now have complete genome sequences for both species, there are still many sequences not currently recognized as genes that could prove to be of great importance when designing mouse models of human disorders. Mouse models such as those described here may offer one of the best ways to understand whether such sequences contribute to phenotype.

References

1. L. E. Olson, J. T. Richtsmeier, J. Leszl, R. H. Reeves, *Science* **306**, 687 (2004).
2. M. K. McCormick *et al.*, *Genomics* **5**, 325 (1989).
3. J. R. Korenberg *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 4997 (1994).
4. R. H. Reeves *et al.*, *Nature Genet.* **11**, 177 (1995).
5. H. Sago *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6256 (1998).
6. R. Ramirez-Solis *et al.*, *Nature* **378**, 720 (1995).
7. E. A. Lindsay *et al.*, *Nature* **401**, 379 (1999).
8. D. J. Adams *et al.*, *Nature Genet.* **36**, 867 (2004).
9. L. L. Baxter *et al.*, *Hum. Mol. Genet.* **9**, 195 (2000).
10. M. E. Coussons-Read, L. S. Cnric, *Behav. Genet.* **26**, 7 (1996).

CLIMATE

The Real Color of Climate Change?

Timothy J. Osborn and Keith R. Briffa

How sensitive is the climate to changes in solar irradiance, atmospheric aerosols, greenhouse gases, and other climate forcings? To answer this question, we first need to know the true extent of past climate fluctuations. The changing temperatures over past centuries and millennia have been reconstructed by regressing annually resolved climate proxy records—for example, from ice cores and tree rings—against recent thermometer measurements. On page 679 of this issue, von Storch *et al.* (1) investigate whether climate changes over decades and longer are likely to have been captured realistically in such reconstructions of Northern Hemisphere (NH) mean temperature.

The likelihood that reconstructions of this kind represent accurate “hindcasts” of past climate is usually assessed by verification against a short period of independent thermometer data. Such verification is only possible for short-term (annual to decadal) climate variability, because the instrumental climate record is too short to sample longer (decadal to centennial) time scales adequately.

To overcome this limitation, von Storch *et al.* use a 1000-year simulation from a

coupled ocean-atmosphere model as a test-bed in which the (simulated) NH temperature is known. They then generate pseudo-proxy records by sampling a small selection of the model’s simulated grid-box temperatures (replicating the spatial distribution of existing proxy records) and degrading them with statistical noise.

The authors show that most of their attempts to reconstruct the model’s NH temperature with the pseudo-proxies result in significant underestimates of the amplitude of fluctuations over the last millennium. Published temperature reconstructions for the real world, based on similar calibration methods, may suffer from the same limitation.

Although von Storch *et al.* focus their discussion on the reconstruction method of Mann *et al.* (2), their conclusions are relevant to other attempts to reconstruct NH temperature history. They demonstrate even greater loss of long-term variations with a simple regression-and-averaging method [this observation was also made in (3)]. The results may apply to all regression-based methods. Accepting von Storch *et al.*’s results does not mean that we must also accept that their simulated temperature history is close to reality—merely that it is a reasonable representation of climate behavior for which any valid reconstruction method should perform adequately.

The underestimated long-term variability obtained by von Storch *et al.* is not a result of problems with proxy data or the ability of the proxies to retain information on these time scales (4), because the pseudo-proxies were generated free from such biases. Neither is it simply due to the usual loss of variance associated with any regression-based prediction (this loss already forms the basis for published estimates of reconstruction error). This usual loss of variance is often modeled as a random error, and although a reconstruction may not be perfect, it cannot be scaled by a simple multiplier to achieve a better fit (that is, the reconstruction and its error are uncorrelated) during the calibration period. It is clear from figure 1 of (1) that the underestimation of long-term temperature variability is systematic rather than random: At these time scales, a better fit to the actual NH temperature can be achieved by scaling a reconstruction by a simple multiplier (>1), because the reconstruction and its error are correlated. Such error is not incorporated in the uncertainties associated with any published NH temperature reconstruction.

The source of this systematic error can be traced to differing shapes of the variance spectra of the NH temperature and of the pseudo-proxy data. The authors constructed pseudo-proxies by adding white noise to the simulated temperatures. Doing so alters the variance spectra and leads to a deficiency in variance at longer time scales, even after calibration (see the figure). Hence, for climate reconstructions to be optimal on all time scales, proxy data must have variance spectra that are similar to those of the climate data that they are presumed to represent. It is not only through the noise inherent in proxy records that this requirement may be violated. Using

The authors are with the Climatic Research Unit, University of East Anglia, Norwich NR4 7TJ, UK. E-mail: t.osborn@uea.ac.uk; k.briffa@uea.ac.uk