elements have comparable cell sizes to those of warm, shallow waters of equatorial Laurentia and Gondwana (Australia). Thereafter, conodonts appear restricted to within 40° of the paleoequator (*33*), and no link between cell size and latitudinal position is evident.

At present, therefore, variations in inferred genome size lack an obvious adaptive explanation. Suggestions that variation in genome size is linked to paedomorphosis (*5, 9, 18*) indicate one avenue of inquiry, but little is now known of heterochronous processes in conodonts (*34*). Investigations of cell size in the fossil record of conodonts and other groups where comparable data are potentially available (for example, brachiopods, arthropods, vertebrates, and vascular plants) can be used to test further adaptive explanations of changes in inferred genome size.

REFERENCES AND NOTES

1. B. J. Shuter *et al.*, *Am. Nat.* **122**, 26 (1983).
2. T. Cavalier-Smith, *Annu. Rev. Biophys. Bioeng.* **11**, 273 (1982).
3. ———, in *Evolution of Genome Size*, T. Cavalier-Smith, Ed. (Wiley, Chichester, 1985), pp. 105–184.
4. M. D. Bennett, *New Phytol.* **106** (suppl.), 177 (1987).
5. O. B. Goin *et al.*, *Copeia* **1968**, 532 (1968).
6. K. Bachmann, O. B. Goin, C. J. Goin, *Brookhaven Symp. Biol.* **23**, 419 (1972).
7. M. D. Bennett, *Proc. R. Soc. London* **B181**, 109 (1972).
8. H. A. Horner and H. C. Macgregor, *J. Cell Sci.* **63**, 135 (1983).
9. S. K. Sessions and A. Larson, *Evolution* **41**, 1239 (1987).
10. M. D. Bennett, *Envir. Exp. Bot.* **16**, 93 (1976).
11. L. E. Orgel *et al.*, *Nature* **288**, 645 (1980).
12. W. F. Doolittle, in *Genome Evolution*, G. A. Dover and R. B. Flavell, Eds. (Academic Press, London, 1982), pp. 3–28.
13. S. Ohno, *Brookhaven Symp. Biol.* **23**, 366 (1972).
14. G. Dover, *Nature* **285**, 618 (1980).
15. K. Bachmann and H. J. Price, *Chromosoma* **61**, 267 (1977).
16. S. W. Sherwood and J. L. Patton, *ibid.* **85**, 163 (1982).
17. R. Hinegardner, in *Molecular Evolution*, F. J. Ayala, Ed. (Sinauer, Sunderland, MA, 1976), pp. 179–199.
18. E. Olmo and A. Morescalchi, *Experientia* **31**, 804 (1975).
19. M. Vialli and G. Sacchi Vialli, *Rendiconti Istituto Lombardo Sci. Lett.* **B103**, 234 (1969).
20. B. Runnegar, *Palaeontology* **29**, 1 (1986).
21. K. S. Thomson, *J. Exp. Zool.* **180**, 363 (1972).
22. ——— and K. Muraszko, *ibid.* **205**, 315 (1978).
23. R. J. Aldridge *et al.*, *Lethaia* **19**, 279 (1986).
24. S. Bengtson, *ibid.* **9**, 185 (1976).
25. R. W. Pierce and R. L. Langenheim, *Geol. Soc. Am. Bull.* **81**, 3225 (1970).
26. For many species several elements were available for measurement. Despite the wide variety of sources, measurements of the same species illustrated by different authors usually showed good agreement. In low magnification illustrations the number of epithelial polygons per unit length was measured. Illustrations at higher magnifications allowed more reliable estimates of size frequency and variation to be obtained. Because polygons tend to be somewhat elongate, here the maximum dimension was used.
27. More than 110 separate articles and papers were consulted; full details of this data source will be provided elsewhere. In addition, H. Armstrong, R. L. Austin, R. D. Burnett, J. Long, A. D. McCracken, G. S. Nowlan, and J. E. Repetski provided photographs of conodonts from which we derived cell measurements.
28. E. Olmo, *Basic Appl. Histochem.* **27**, 227 (1983).
29. H. J. Price, *Bot. Rev.* **42**, 27 (1976).
30. D. L. Clark, *J. Paleontol.* **57**, 652 (1983).
31. D. L. Clark *et al.*, *Treatise on Invertebrate Paleontology*, part W. Miscellanea. Supplement 2, Conodonta. (Geological Society of America, Boulder, CO, and University of Kansas, Lawrence, 1981).
32. A. W. Ebeling *et al.*, *Am. Nat.* **105**, 549 (1971).
33. W. C. Sweet, *J. Paleontol.* **59**, 485 (1985).
34. T. W. Broadhead and R. McComb, *Fossils Strata* **15** 149 (1983).
35. R. J. Aldridge, *Syst. Assoc. Spec.* **34**, 231 (1988).

36. R. J. Aldridge, M. D. Bennett, and G. Dover are thanked for reviewing the manuscript and helpful discussions, while a number of investigators (*27*) generously made available unpublished photographs and stratigraphic data. We thank M. Orchard for photographs for Fig. 1, S. Ripper for drafting Fig. 2, and S. Last for typing the manuscript. We acknowledge financial support from Gonville and Caius College for E. H. Cambridge Earth Sciences Publication 1263.

14 March 1988; accepted 1 July 1988

Technical Comments

Do Short-Term Tests Predict Rodent Carcinogenicity?

Long-term rodent studies are expensive ($1 to $2 million) and time-consuming (3 to 4 years), so there is great interest in being able to predict their results with the use of short-term tests (STTs). This interest is particularly keen since a battery of such tests costs approximately $10,000. Tennant *et al.* (*1*) examined the results of the rodent test and genetic toxicity tests for 73 compounds recently tested by the National Cancer Institute and concluded that, of the four STTs examined, a single test, the Ames *Salmonella*-microsome test, is 60% concordant with the rodent test and that any one of the other three tests do not help in predicting rodent carcinogenicity. One might conclude that the other three tests are superfluous; either they lack additional information or they are not cost-effective.

The conclusion of Tennant *et al.* appears to be based largely on looking at pairs of STTs. A rearrangement of their results, Table 1, shows a good correlation between the number of STT positives and the probability of a positive rodent result. When all four STTs were positive, the rodent test was positive about 80% of the time. In only three instances were all STTs positive and the rodent test negative.

Even with no STT positives, the rodent test was positive about 40% of the time. What are the possible explanations? Nongenetic mechanisms of carcinogenicity might be operable. These are likely because the rodent tests are conducted at maximal doses, which might be expected to upset hormonal or other physiological balances. It is also likely that some of the positive rodent results are statistical false-positives. There are many types of spontaneous tumors; for rats and mice, males and females, there can be 40 to 100 statistical tests for each compound. In multiple testing situations, Gill (*2*) has argued that one should expect positive results by chance alone. Haseman *et al.* (*3*), in examining paired control groups from 18 studies, found one or more statistically significant differences between the control groups 44% of the time. In six of the 18 studies, there were two or more statistically significant results. Also, Brown and Fears (*4*) estimated that false-positive results could be expected 30% of the time in long-term rodent tests.

Figure 1 summarizes the above points: one might expect 30 to 44% of the rodent results to be false-positives; the rodent-STT regression line predicts about a 40% nongenetic or statistical false-positive rate. Even with four STT positives, one should expect about 20% of the rodent tests to be negative.

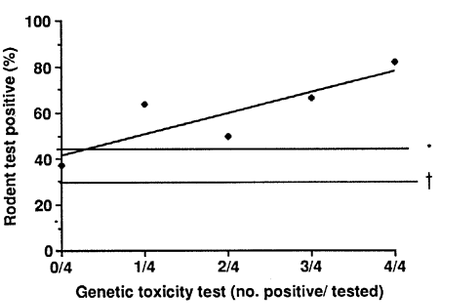

Fig. 1. Positive results in rodent tests versus positive results in genetic toxicity tests. *Estimated false-positive rate (*3*). †Estimated false-positive rate (*4*) ($y = 41.5 + 3.7x$, $P < 0.007$).

Table 1. Summary of genetic toxicity STTs and rodent tests, positive and tested. Cochran-Armitage linear trend test, $P < 0.007$.

STTs positive/ tested	Rodent	
	Positive/ tested	Positive (%)
4/4	14/17	82.3
3/4	10/15	66.7
2/4	7/14	50.0
1/4	7/11	63.6
0/4	6/16	37.5
Total	44/73	60.3

Since a linear response fits the data better than a step function comparing "no STT positive" to "any STT positive," the data indicate that additional STTs provide additional information (as would be expected biologically). Since individual STTs are inexpensive relative to a long-term rodent test, the additional information is cost-effective (*5*).

All of this begs the question, "Are rodent tests predictive of humans?" Readers interested in that question might study (*6*) and (*7*) and reach their own conclusions.

S. STANLEY YOUNG
Glaxo Inc.,
Research Triangle Park, NC 27709

REFERENCES

1. R. W. Tennant *et al.*, *Science* **236**, 933 (1987).
2. J. L. Gill, *J. Anim. Sci.* **60**, 867 (1985).
3. J. K. Haseman, J. S. Winbush, M. W. O'Connel, *Fundam. Appl. Toxicol.* **7**, 573 (1986).
4. C. C. Brown and T. R. Fears, *Biometrics* **37**, 763 (1981).
5. L. B. Lave and G. S. Omenn, *Nature* **324**, 29 (1986).
6. E. J. Calabrese, *Principles of Animal Extrapolation* (Wiley, New York, 1983), table 14.1.
7. F. J. C. Roe, in *Human Risk Assessment,* M. V. Roloff, Ed. (Taylor and Francis, New York, 1987), pp. 31–44.

18 September 1987; revised 22 April 1988; accepted 28 April 1988

Response: Young (*1*) raises two issues: (i) the comparative performance of a battery of short-term tests (STTs) versus the *Salmonella* mutagenesis assay (SAL) for predicting rodent carcinogenicity and (ii) the false-positive rate associated with rodent carcinogenicity studies.

Regarding the first issue, we concluded that, for a set of 73 chemicals evaluated by the National Toxicology Program (NTP), a battery of four STTs was not significantly more predictive of the results of rodent carcinogenicity studies than was SAL alone (*2*). Young apparently questions this conclusion, asserting that, for the battery of four STTs (including SAL), a trend test shows that there is "a good correlation between the number of STT positives and the probability of a positive rodent result." This is misleading because the "good correlation" of the battery reflects primarily the high predictivity of SAL. When SAL is excluded from the battery and separate comparisons are made for SAL positive and SAL negative chemicals, Young's trend test analysis shows no significant association between the number of STT positives and rodent carcinogenicity, as indicated in Table 1.

To put this matter into perspective, one could consider a comparison of the predictivity and concordance of the two approaches. Young states that "when all four STTs were positive, the rodent test was positive about 80% of the time." However, the predictivity of a positive SAL is even greater (83%; see Table 1). Young further states that SAL "is 60% concordant [62% actually]" with the rodent test. However, the corresponding concordance of the battery of four STTs is essentially the same, that is, 55 to 66%, depending upon the decision rule employed (*2*). Thus, for the 73 NTP chemicals the predictivity and concordance of the battery of four STTs is similar to that of SAL alone.

Regarding the second issue, Young asserts that rodent carcinogenicity studies have a high statistical false-positive rate. His conclusion is based on what appears to be a misinterpretation of the results of Brown and Fears (*3*) and of Haseman *et al.* (*4*), who emphasized that such high false-positive rates (30 to 44%) would occur only if every statistically significant ($P < 0.05$) increase in tumor incidence were regarded as a biologically meaningful effect. This does not occur in practice because biological as well as statistical factors are taken into consideration in the overall evaluation of the data. Most investigators in this area are familiar with the multiple comparisons issue, and thus it is generally recognized that the actual false-positive rate is much lower than 30 to 44%.

What is the actual false-positive rate? The International Agency for Research on Cancer concludes that "rules which attempt to model the actual decision process indicate that false-positive rates are close to the nominal level" (*5*). The Office of Science and Technology Policy (*6*) reaches a similar conclusion. Moreover, one of us (J.K.H.) (*7*) has estimated that the false-positive rate associated with NTP carcinogenicity studies (such as those used by Tennant *et al.*) is no greater than 7 to 8%. Many NTP "nongenotoxic carcinogens" showed markedly increased tumor incidences at multiple sites and for multiple doses or in three to four sex-species groups, or both. It is extremely unlikely that these striking effects are statistical false-positives, as suggested by Young.

In summary, for the particular 73 chemicals considered by Tennant *et al.* (*2*), the evidence is compelling that the other three STTs did not improve significantly the performance of SAL for predicting rodent carcinogenicity. Additional studies are now in progress to determine whether these results also hold for a second set of chemicals recently evaluated by the NTP.

J. K. HASEMAN
B. H. MARGOLIN
M. D. SHELBY
E. ZEIGER
R. W. TENNANT
National Toxicology Program,
Post Office Box 12233,
Research Triangle Park, NC 27709

Table 1. Performance of a battery of three STTs (excluding SAL) for predicting for carcinogenicity of 73 NTP chemicals.

Proportion of STTs positive	Chemicals positive in SAL [proportion of carcinogens (%)]	Chemicals negative in SAL [proportion of carcinogens (%)]
3/3	14/17 (82)	5/9 (56)
2/3	5/6 (83)	7/14 (50)
1/3	–	6/10 (60)
0/3	1/1 (100)	6/16 (38)
Total	20/24 (83)	24/49 (49)
Cochran-Armitage linear trend test	$P > 0.50$	$P > 0.20$

REFERENCES

1. S. S. Young, *Science* **241**, 1232 (1988).
2. R. W. Tennant *et al.*, *ibid.* **236**, 933 (1987).
3. C. C. Brown and T. R. Fears, *Biometrics* **37**, 763 (1981).
4. J. K. Haseman, J. S. Winbush, W. M. O'Donnell, Jr., *Fundam. Appl. Toxicol.* **7**, 573 (1986).
5. J. J. Gart, D. Krewski, P. N. Lee, R. E. Tarone, J. Wahrendorf, *Statistical Methods in Cancer Research,* vol. **3**, *The Design and Analysis of Long-term Animal Experiments* (IARC Scientific Publication 79, International Agency for Research on Cancer, Lyons, France, 1986).
6. Office of Science and Technology Policy, *Fed. Reg.* **50**, 10371 (14 March 1985).
7. J. K. Haseman, *Fundam. Appl. Toxicol.* **3**, 334 (1983).

29 July 1988; accepted 4 August 1988