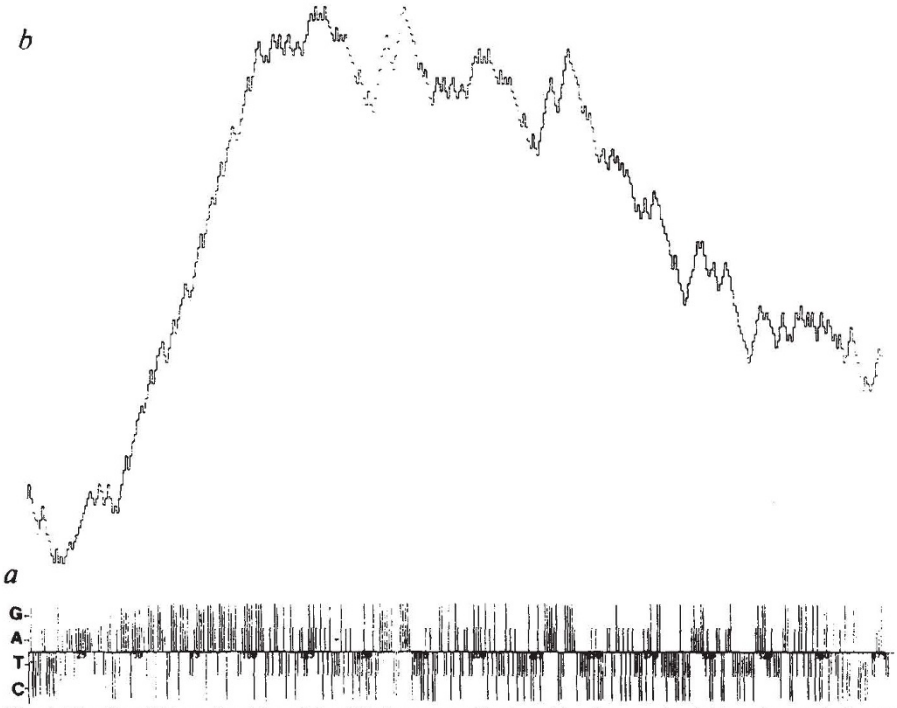

Fig. 1. The first 376 nucleotides of the SV40 genome displayed by the standard (*a*) and cumulative (*b*) line-extension formats. In both diagrams the four nucleotides G, A, T and C are given *y*-axis values of +2, +1, −1 and −2, although the scale of *b* has been reduced in this dimension. The direct 72-base pair repeat is clearly visible (region 100–250). Restriction recognition sequences form (usually) rotationally-symmetrical patterns in *a* and symmetrical peaks or troughs in *b*.

resolving oligonucleotides for nucleic acid sequencing — wandering spot[3] (Hamori) and thin-layer acrylamide gel electrophoresis[4] (our method). This superficial coincidence may reflect a limited number of possible formats.

Our first concern has been to devise a method of transferring printed sequences, using a scanning light-pen, from the printed page to an information retrieval system. Simple variants of bar-thickness identification codes were ruled out since deciphering is particularly taxing. Angle-vector methods of display such as H-curves are not inherently machine-readable although they reveal, as do plots of G/C content, features of the overall sequence profile. The base-composition curve is, to a close approximation, the first derivative of the (projected) H-curve, and as such embodies the same information. We have also examined formats which reflect the chemical structures of the four bases and which might, by comparison of non-identical recognition sequences, reveal nucleic acid–protein contact points. However, such formats were particularly difficult to interpret by eye.

We have subsequently examined a display which combines features of both the line-extension and angle-vector formats. This cumulative line extension format (which respects the extension ratios for the four bases, Fig.1) gives similar information to that yielded by angle-vector diagrams but is less ambiguous than projected H-curves. Restriction recognition sequences appear here as symmetrical peaks. However, the format is not machine-readable and we have not proceeded further with this approach.

The two methods under discussion [1,2] meet different applications. The method we have proposed reveals local features of a DNA sequence and is limited to some 10,000 nucleotides per standard page. The display is machine-readable. In contrast, H-curves accentuate (as do G/C content plots) global features of a large-scale sequence while concealing the detail. They permit 100,000 or more nucleotides to be displayed per page. To this extent the two methods are non-overlapping and provide a visual representation of a particular nucleic acid sequence.

R. LATHE
R. FINDLAY
ARFC Animal Breeding Research Organisation,
West Mains Road,
Edinburgh, EH9 3JQ, UK

1. Hamori, E. *Gene analyt. Techq.* **1**, 69, (1984)
2. Lathe, R. & Findlay, R *Nature* **311**, 610 (1984)
3. Sanger, F., Brownlee, G.G. & Barrell, B.G. *J. molec. Biol.* **13**, 373 (1965)
4. Sanger, F., Nicklen, S & Coulson, A.R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463 (1977)

Powers of ten wrongly expressed

SIR — It is surprising how often authors improperly use powers of 10, in the expression of radioactivity measurements particularly. I often read manuscripts in which the ordinates are labelled as follows: 'Radioactivity (c.p.m. × 10^{-3})' or simply 'c.p.m. × 10^{-3}' with figures ranging from 0 to 10 or so.

Since $a \times b = b \times a$, 'c.p.m. × 10^{-3}' actually means '10^{-3} c.p.m' or 'milli-counts per min'. This is obviously not what the authors mean. The proper way to label such ordinates is either '10^{-3} × radioactivity (c.p.m.)' or 'Radioactivity (c.p.m. × 10^{3})'.

A physical quantity (see page 5 in ref.1) is the product of a pure number and a unit, that is, a particular physical quantity used as a standard, for example:

Mass = 3 × grams (or 3 g) (1)
Radioactivity = 5,000 × c.p.m. (or 5,000 c.p.m.) (2)

If one divides 5,000 by 10^3 in equation (2), then 'radioactivity' should also be divided by 10^3 (or multiplied by 10^{-3}), or 'c.p.m.' should be multiplied by 10^3.

Many authors divide both the number and the unit by 10^3, thus dividing only the right-hand side of equation (2) by 10^3. This is incorrect.

CLAUDE LIÉBECQ
European Journal of Biochemistry,
Boulevard de la Constitution 69/054,
B-4020 Liège, Belgium

1. International Union of Pure and Applied Chemistry. *Manual of Symbols and Terminology for Physicochemical Quantities and Units. 1979 Edition* (Pergamon Press, Oxford 1979).

Mass extinction times and correlations

SIR — Rampino and Stothers[1] make the interesting argument that the correlation they found[2] between ordered sequences of nine mass extinction times and nine galactic plane crossings ($r = 0.996$) is statistically significant, while the correlation that I found[3] between the same ordered extinction times and the first nine prime numbers ($r = 0.986$) is not. Of course, this leaves open the question of the significance of the correlation of the ordered extinction times and the first nine odd numbers ($r = 0.995$).

The question of how to test statistical relations between serial data is a serious and difficult one of long standing, going back at least a century to W. Stanley Jevons' attempts to show a connection between sunspots and the business cycle. I believe its resolution in the present case will probably require more data, and certainly a different analysis, than that given by Rampino and Stothers. All monotone series appear highly correlated, all regular monotone series appear even more so. This artefact aside, and even under the authors' most optimistic assumptions, the statistical evidence is simply too weak to reach any conclusion, due to the large variance (relative to the small sample size) of the times between mass extinction dates. The intriguing connection Rampino and Stothers suggest between these two phenomena may well be a real one; their analysis has not yet proved it to be real.

S. M. STIGLER
Department of Statistics,
University of Chicago,
Chicago, Illinois 60637, USA

1. Rampino, M. R. & Stothers, R.B. *Nature* **313**, 159-160 (1985).
2. Rampino, M.R. & Stothers, R.B. *Nature* **308**, 709-712 (1984).
3. Stigler, S.M. *Nature* **313**, 159 (1985).