samples there are no such signals, in Holocene samples the peaks are very small, but in Pleistocene samples they are quite pronounced.

Although the procedure of ESR dating of samples is tedious and involves many steps, including determination of the ESR spectra, screening requires only this single step. Measuring an ESR spectrum takes only 5–10 min. If another 5–10 min is allowed for sample preparation time (cleaning and grinding the sample of at least 20 mg) hundreds of research dollars can be saved by 10–20 min work. Of course an ESR spectrometer is necessary, but nearly every university has one.

Badly preserved or contaminated samples are not suitable for the ESR screening method (and are probably not suitable for ^{14}C analysis either) but in general, for reasonably well preserved samples, it should be an extremely effective technique. Screening could significantly reduce the number of questionable ^{14}C dates of carbonates. If screening indicates an age that is beyond the limit of ^{14}C dating, then another dating method can be substituted.

ULRICH RADTKE
Geographisches Institut,
Universität Düsseldorf,
D-4000 Düsseldorf 1, FRG

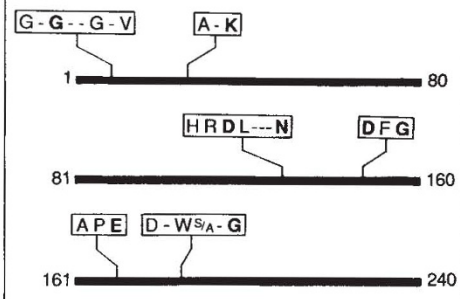
1. Hennig, G.J. & Grün, R. *Quat. Sci. Rev.* **2**, 157–238 (1983).
2. Radtke, U. *et al. Nucl. Tracks* **10**, 879–884 (1985).
3. Radtke, U *et al. Quat. Res.* (in the press).

Identification of protein kinases by computer

SIR—I would like to add a cautionary note to the recent letter of Bairoch and Claverie[1] about recognition by computer of eukaryotic protein kinases by sequence patterns. These authors claim that two patterns can be used to identify protein kinases and discriminate them from other proteins.

Here, I make three points. First, both these patterns are absent from the protein kinases casein kinase II (ref. 2) and *nim* 1 (ref. 3). Second, as the authors kindly informed me when I requested its complete sequence, the TMS1 protein that they predicted to be a protein kinase on the basis of one of these patterns[1] is, in fact, tryptophan 2-monooxygenase[4]. Third, eukaryotic protein kinases can easily and accurately be identified by other means.

Such identification may, if necessary, involve initial sequence comparison with the protein databases (for example, using a program based on the algorithm of Wilbur and Lipman[5]), or direct comparison with protein kinases (for example, using a program based on the algorithm of Needleman and Wunsch[6]). Initial 'matches' of unknown or candidate proteins to known eukaryotic protein kinases must then be examined ('manually' or 'by eye') for the presence of the motifs shown in the figure, which are derived from well-known similarities[7,8]. The amino-acid residues shown in bold have, so far, been found to be totally conserved and, although the other residues can vary, in known protein kinases (including casein kinase II and *nim* 1) no more than a single deviation per motif has been observed. In practice, the presence of other semi-conserved residues, such as the hydrophobics in Bairoch and Claverie's patterns, in the regions of these motifs further assists unambiguous identification.



Highly conserved motifs in eukaryotic protein kinases. The average positions of these in the approximately 240-amino-acid protein kinase domain is indicated (see refs 7 and 8).

It may seem self-evident that all these motifs should be taken into account, but there has been a tendency in some quarters to focus on the G–G– –G of the most amino-terminal motif (equivalent to Bairoch and Claverie's pattern 2) because of its likely location at the nucleotide-binding site of protein kinases (but also of certain other enzymes[9]). Although the functions of the other motifs (except A–K) are unknown, their importance can easily be imagined if one considers the locations of conserved aspartate and asparagine residues[10] in the three-dimensional structure of the GTP-binding protein EF-Tu[11,12]. The extreme conservation in diverse prokaryotic and eukaryotic phosphotransferases of the Asp residues represented in the protein kinase motifs HRDL– – –N and DFG has already been pointed out by Brenner[13].

Computers are indispensible in modern biology. A molecular biological perspective is also helpful.

DAVID P. LEADER
Department of Biochemistry,
University of Glasgow,
Glasgow G12 8QQ, UK

1. Bairoch, A. & Claverie, J.-M. *Nature* **331**, 22 (1988).
2. Saxena, A. *et al. Molec.cell.Biol.* **7**, 3409–3417 (1987).
3. Russell, P. & Nurse, P. *Cell* **49**, 569–576 (1987).
4. Gielen, J. *et al. EMBO J.* **3**, 835–846 (1984).
5. Wilbur, W.J. & Lipman, D.J. *Proc.natn.Acad. Sci U.S.A.* **80**, 726–730 (1983).
6. Needleman, S.B. & Wunsch, C.D. *J.molec.Biol.* **48**, 443–453 (1970).
7. Takio, K. *et al. Biochemistry* **23**, 4207–4218 (1984).
8. Hunter, T. & Cooper, J.A. *A.Rev.Biochem.* **54**, 897 (1985).
9. Möller, W. & Amons, R. *FEBS Lett.* **186**, 1–7 (1985).
10. Dever, T.E., Glynias, M.J. & Merrick, W.C. *Proc.natn. Acad.Sci U.S.A.* **84**, 1814–1818 (1987).
11. Jurnak, F. *Science* **230**, 32–36 (1985).
12. La Cour, T.F.M. *et al. EMBO J.* **4**, 2385–2388 (1985).
13. Brenner, S. *Nature* **329**, 21 (1987).

Lod score or log-likelihood?

SIR—The publication of another paper[1] in *Nature* which uses 'Lod scores' to describe the analysis of human linkage data prompts me to wonder whether it might not be an opportune moment for linkage workers to abandon their idiosyncratic statistical terminology so that others can better follow their analyses.

Their Lod score is simply the log-likelihood to the base 10, standardized as is customary by the addition of a constant, in this case such as to make the log-likelihood zero at a recombination fraction of one-half. The phrase is confusing because in standard likelihood terminology it is neither a lod nor a score: lod ('log-odds' was first defined by Barnard[2] in terms of natural logarithms, and score is now reserved for the first derivative of the log-likelihood[3,4]. Indeed, the word odds itself is nowhere else used in this 'backward' sense introduced by Barnard, the common name being the likelihood ratio. The paper also uses the term 1-Lod confidence interval although it is not a confidence interval in Neyman's sense, as has often been pointed out[5].

I imply no criticism of linkage workers by suggesting that a change to conventional terminology might be timely. On the contrary, linkage analysis in man has been a valuable proving-ground for the use of likelihood as a primary statistical tool, and the terminology introduced by Smith[6] and Morton[7] probably helped to make likelihood methods acceptable (even though that was not the intention of either author at the time[5]). But just as physicists eventually abandoned probable errors in favour of standard deviations, so human geneticists might with advantage abandon Lod scores in favour of the customary log-likelihoods to the base *e*. And rather than using the phrase 'confidence interval' in a non-Neyman sense, they might consider my phrase "support interval"[4] which was specifically coined for likelihood use.

There is, of course, no difficulty in adjusting the standard computer programs for linkage so as to work with natural rather than common logarithms (some already offer this option), and by changing the terminology at the same time as changing the logarithmic base there will be no risk of confusion.

A.W.F. EDWARDS
Department of Community Medicine,
University of Cambridge,
Fenner's, Gresham Road,
Cambridge CB1 2ES, UK

1. Wasmuth, J.J. *et al. Nature* **332**, 734–736 (1988).
2. Barnard, G.A. *J. R. statist. Soc.* **B11**, 115–149 (1949).
3. Bailey, N.T.J. *Introduction to the Mathematical Theory of Genetic Linkage* (Clarendon, Oxford, 1961).
4. Edwards, A.W.F. *Likelihood* (Cambridge University Press, 1972; 1984).
5. Smith, C.A.B. *Ann. hum. Genet.* **50**, 293–311 (1986).
6. Smith, C.A.B. *J. r. statist. Soc.* **B15**, 153–192 (1953).
7. Morton, N.E. *A. J. hum. Genet.* **7**, 277–318 (1955).