Mousterian and Solutrian levels at Le Placard, and between the Mousterian and Aurignacian levels at Abri Caminade[2,11]. Positive indications of this hiatus can be seen in recently published sections of the site[14]. The suggestion that the Neanderthal burials at La Ferrassie represent later instrusions into the Mousterian levels is intriguing, but difficult to evaluate, 80 years after the original excavations.

Finally, Cook and Ashton make no allowance for the wealth of purely archaeological evidence bearing on the wider problems of Mousterian chronology and intercorrelations between sites. Any attempt to support the stratigraphic correlations of Laville[2,12] would need to explain why at least 12 sites in south-west France show a consistent sequence of Mousterian of Acheulian tradition above Quina/Ferrassie industries, why Mousterian of Acheulian tradition industries invariably occur at the top of Mousterian sequences in cave and rock shelter sites, and why Mousterian of Acheulian tradition industries are entirely lacking from all except the uppermost 4 of the 55 levels of Mousterian occupation at Combe Grenal[13]. No explanation has so far been offered as to how these observations can be reconciled with the hypothesis of a close synchronism between the archaeological sequences at Le Moustier and Combe Grenal.

PAUL MELLARS

Department of Archaeology,
Downing Street,
Cambridge CB2 3DZ, UK

1. Bowman, S.G. & Sieveking, G. de G. *Pact* **9**, 253 (1983).
2. Laville, H., Rigaud, J-P. & Sackett, J.R. *Rock Shelters of the Perigord: Geological Stratigraphy and Archaeological Succession* (Academic, New York, 1980).
3. Bordes, F., Laville, H. & Paquereau, M-M. *Actes de la Société Linnéene de Bordeaux* **103B**, 1–19 (1966).
4. Bowman, S.G.E. *et al. Pact* **6**, 362 (1982).
5. Laville, H. *et al. Bull. Inst. Géol. Bassin Aquitaine* **34**, 219 (1983).
6. Dennell, R. *Nature* **301**, 199 (1983).
7. Sancetta, C. *et al. Quat. Res.* **3**, 110 (1973).
8. Woillard, G. & Mook, W.G. *Science* **215**, 159 (1982).
9. Turon, J-L *Nature* **309**, 673 (1984).
10. Bordes, F. & Prat, F. *L'Anthropologie* **69**, 31 (1965).
11. de Lumley, H. (ed.) *La Préhistoire Française* (CNRS, Paris, 1976).
12. Laville, H. *World Archaeol.* **4**, 323 (1973).
13. Mellars, P.A. *Proc. Prehist. Soc.* **35**, 134 (1969).
14. Delporte, H. (ed.) *Le Grand Abri de la Ferrassie* (Institut de Paléontologie Humaine, Paris, 1984).

What can AIDS virus codon usage tell us?

SIR—Macromolecular sequence comparison is extraordinarily useful in elucidating evolutionary relationships, largely because divergence rates are approximately constant, and the chance of convergence is very small. The close evolutionary relationship of the human immunodeficiency virus (HIV) and visna virus has been well characterized by this approach[1,2]. Grantham, largely responsible for demonstrating that patterns of codon choice are species are species specific[3], and Perrin now suggest that comparison of such patterns in HIV and visna may provide a useful approximate evolutionary distance[4].

Decomposition of a DNA sequence into a codon usage table loses a great deal of information — comparisons of the intact sequences of homologous genes and proteins must be more informative than comparisons of codon choice. In fact, the distance measure used by Grantham and Perrin (how many of 18 amino acids have the same most-preferred codon in two codon usage tables) can yield some strange results. For example, among the viruses in Table 1, it is thought that Moloney murine leukaemia virus (MMLV, another retrovirus) is the most closely related to HIV, and yet HIV codon usage is most similar to that in the influenza virus, and even cauliflower mosaic virus (CAMV) is more similar to HIV than is MMLV. Also, a compilation of 66 mouse genes[5] yields the same preferred codons as the human (total) for each of the 18 amino acids, and yet differs from 60 rat genes[5] for two amino acids. Thus, unlike divergence in protein and DNA sequences, differentiation at the level of codon usage is clearly not linear with time. It is also highly susceptible to convergence.

Furthermore, it is clear that there is considerable heterogeneity in codon choice among genes within species. In *Escherichia coli* and yeast (the two best studied organisms) groups of genes with high and low codon bias can be identified[3,6,7], that differ somewhat in their preferred codons (Table 1). In mammals there are also large differences among genes in the direction of codon preference[8]. These differences seem to be related to local genomic G+C content[8], and can yield very different sets of preferred codons.

Thus, when Grantham and Perrin use their comparison to estimate the extent to which viruses imitate the codon preference of their hosts[4], and point out that codon usage in HIV genes is different to that in a compilation of all human (excluding immune system) genes, this is somewhat misleading. In fact, HIV genes show considerable similarity in codon choice to a group of human genes with low G+C content. Comparing species totals; HIV codon usage shows more similarity to that in *Bacillus subtilis* and yeast, than to that in the compilation of human genes.

On the positive side, it is possible that the differences among mammalian genes in pattern of codon usage may partly reflect the diverse tissues in which those genes are expressed. One investigation of this possibility[9] gave negative results, but examined very few genes. It may be that codon usage in HIV genes reflects adaption to expression in particular tissues, in which case the report[4] that there is some similarity in this respect between HIV genes and T-cell receptor genes would be rather interesting.

PAUL M. SHARP

Department of Genetics,
Trinity College.
Dublin 2, Ireland

Table 1 Similarities of codon choice between organisms

	(*N*)	HIV	E LO	E HI	Baci	Y LO	Y HI	H AT	H CG
Human (total)	(135)	2	8	8	1	2	7	3	14
Human G+C rich	(8)	0	8	10	1	0	7	1	
Human A+T rich	(8)	12	8	5	13	14	7		
Yeast HI	(38)	5	2	11	6	11			
Yeast LO	(66)	11	6	6	13				
Bacillus	(21)	11	9	7					
E.coli HI	(42)	2	8				5	(3)	MMLV
E.coli LO	(58)	6				10	1	(28)	AD2
					3	6	12	(6)	SV40
				11	1	5	13	(8)	Flu
			10	7	5	8	8	(2)	CAMV
			Flu	SV40	AD2	MMLV	HIV		

Values are the number of amino acids (out of 18) where the two data sets show the same preferred codon. *N* is the number of genes in each data set. Yeast, *Saccharomyces cerevisiae*; *Bacillus*, *B. subtilis*. HI and LO (for *E.coli* and yeast) indicate groups of genes identified as having high and low codon bias. Human GC-rich : α-actin, apolipoprotein A1, β-chorionic gonadotropin, α-globin, ζ-globin, metallothionein II, insulin, α-tubulin. Human AT-rich : albumin, α-amylase (pancreatic and salivary), factor IX, γ-fibrinogen, γ-interferon, HPRT, parathyroid hormone. Viruses: HIV, 5 genes (*gag, pol, sor, env, P27*) of ARV-2; Moloney murine leukaemia virus; Flu, influenza virus isolate B/Lee/40; SV40, simian polyoma; CAMV, cauliflower mosaic virus; AD2, adenovirus 2. All codon usage data was taken from ref.5, except: yeast[6], *Bacillus*[10], *E.coli*[7] and HIV (calculated from GenBank).

1. .Sonigo, P. *et al. Cell* **42**, 369–382 (1985).
2. Yasunaga, T., Sagata, N. & Ikawa, Y. *FEBS Lett.* **199**, 145–150 (1986).
3. Grantham, R. *et al. Nucleic Acids Res.* **8**, r49–r62 (1980); **9**, r43–r47 (1981).
4. Grantham, R. & Perrin, P. *Nature* **319**, 727–728 (1986).
5. Maruyama, T., Gojobori, T., Aota, S. & Ikemura, T. *Nucleic Acids Res.* **14**, r151–r197 (1986).
6. Sharp, P.M., Tuohy, T.M.F. & Mosurski, K.R. *Nucleic Acids Res.* **14**, 5125–5143 (1986).
7. Sharp, P.M. & Li, W-H. *Nucleic Acids Res.* **14**, (in the press)
8. Ikemura, T. *Molec. Biol. Evol.* **2**, 13–14 (1985).
9. Hastings, K.E.M. & Emerson, C.P. *J. molec. Evol.* **19**, 214–218 (1983).
10. Ogasawara, N. *Gene* **40**, 145–150 (1985).