

Voss Replies: In [1] I demonstrated that ensemble averages over more than 25000 DNA sequences in the GenBank databank show systematic changes in long-range equal-base correlations with evolutionary category. Those measurements used all of the available long sequence data and made no distinction between, and no specific claims about, coding (exon) vs noncoding (intron) portions. Other researchers [2-4], using different methods on a few (≈ 20) samples, reported correlations only in noncoding sequences. Although a detailed criticism and comparison [5] has already been prepared, the major points are relevant to [6].

As shown in Fig. 1, DNA sequences exhibit remarkably large variations in their equal-base correlations. It is possible to select individual sequences or portions [4,6] to display almost any desired behavior [7]. Moreover, $F(l)$ vs l curves mask the typical change from uncorrelated to long-range correlations in the range 10-100 base posi-

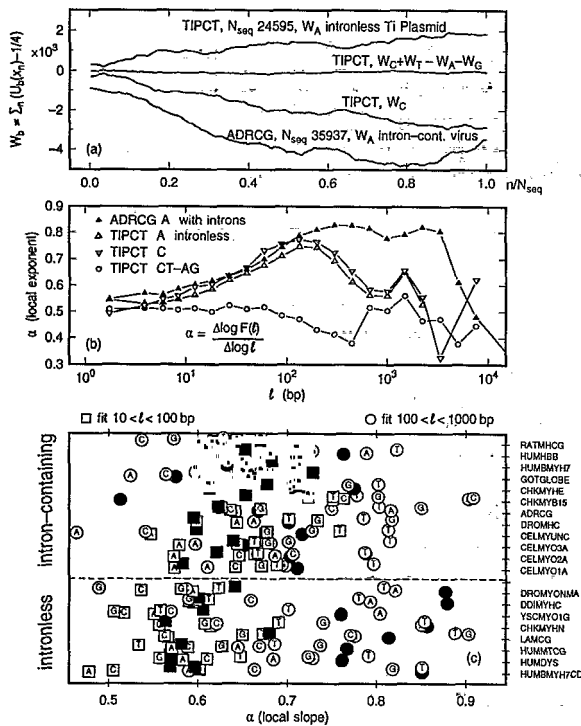


FIG. 1. Wide sample variations of DNA sequences demonstrate the impossibility of drawing reliable conclusions from small samples. (a) Sample walks for the intronless Ti plasmid [6]. The pyrimidine-purine walk, $W_P = W_C + W_T - W_A - W_G$, has reduced long-ranged variations due to the anticorrelation of CT vs AG. The individual base walks, W_A and W_C , however, are similar to the intron-containing virus from [4]. (b) Local slope α vs l from $F(l)$ analysis. (c) Local slope α fit from different l regions for A, C, G, T, and P walks with $N_{seq} > 6000$ from the ensembles of [4].

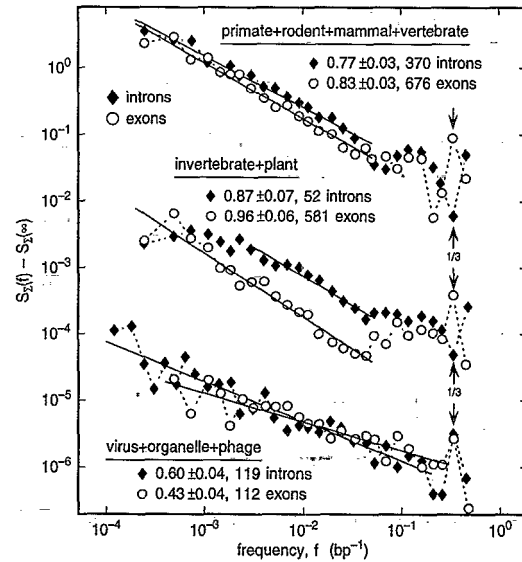


FIG. 2. Category ensemble average $S_Z(f) - S_Z(\infty)$ for all annotated intron and exon features from GenBank release 73 with length $N \geq 512$ bp.

tions (bp) that is easily visible in $S(f)$ or local a plots.

Thus, statistical differences between coding and noncoding portions can best be examined by using all of the available data for all bases without *a priori* assumptions about the relevance of *only* certain combinations (CT-AG). Figure 2 demonstrates that both exons and introns have similar long-range fractal correlations with exponents consistent with previous [1] category averages.

Richard F. Voss
 IBM Thomas J. Watson Research Center
 Yorktown Heights, New York 10598

Received 2 July 1993
 PACS numbers: 87.10.+e, 05.40.+j, 06.50.-x, 72.70.+m

[1] R. F. Voss, Phys. Rev. Lett. **68**, 3805 (1992).
 [2] W. Li, Int. J. Bifurcation Chaos **2**, 137 (1992).
 [3] W. Li and K. Kaneko, Europhys. Lett. **17**, 655 (1992).
 [4] C. K. Peng, S. V. Buldyrev, A. L. Goldberger, S. Havlin, F. Sciortino, M. Simons, and H. E. Stanley, Nature (London) **356**, 168 (1992).
 [5] R. F. Voss, Fractals (to be published).
 [6] S. V. Buldyrev *et al.*, preceding Comment, Phys. Rev. Lett. **71**, 1776 (1993).
 [7] Although the sequences in Fig. 1 of [4] maximized the difference between intronless and intron containing, other choices from the same ensembles are indistinguishable [5]. Different $F(l)$ analysis procedures for the two ensembles [4] contributed to the apparently clear distinction.