

**Table 1.** Numbers and mean lengths for proteins and pseudogenes in four eukaryotes

Organism	Number of proteins	Number of pseudogenes <sup>a</sup>	Number of processed pseudogenes	Mean length of protein <sup>b</sup>	Mean length of matching protein for pseudogenes
Human <sup>c</sup>	927	384 (2.4)	189	317 ( $\pm 43$ )	342
Nematode worm	20732	1100 (18.9)	104	435 ( $\pm 15$ )	450
Budding yeast	6340	221 (28.7)	0	467 ( $\pm 29$ )	424 <sup>d</sup>
Fruit fly	14332	110 (130.3)	34 <sup>e</sup>	500 ( $\pm 50$ )	808 <sup>f</sup>

<sup>a</sup>The proportion of proteins to pseudogenes is given in brackets.

<sup>b</sup>Standard deviation of the sample mean is given in brackets.

<sup>c</sup>These data are for chromosomes 21 and 22 only (7).

<sup>d</sup>The difference between mean lengths of yeast proteins in general and those that are closest matches to pseudogenes is marginally significant ( $P < 0.06$ ) using normal statistics.

<sup>e</sup>This value is for pseudogenes (i) that are of substantial length (>70% the length of the closest matching organismal protein) and have no introns (where a matching protein does have introns) or (ii) that have some evidence of polyadenylation. See Materials and Methods for more detail. These procedures are described in (7).

<sup>f</sup>The difference between mean lengths of fruit fly proteins in general and those that are closest matches to pseudogenes is very significant ( $P < 0.0001$ ) using normal statistics. Removing the outlying matchers of seven fragments whose lengths exceed 3000 amino acids, reduces the mean to 610 residues ( $P < 0.02$ ).