FEBS 16745

FEBS Letters 380 (1996) 301

Letter to the Editor

Primary structure and homology

Barry Whyte, Bengt Persson, Hans Jörnvall*

Department of Medical Biochemistry and Biophysics, Karolinska Institutet, S-171 77 Stockholm, Sweden

The number of reports dealing with sequences and structural analysis is constantly increasing. Presently, 'amino acid sequence' is mentioned in 6% of the abstracts of scientific papers listed in Medline. This number has increased linearly since about 1988 (Fig. 1). In contrast to this growth, the total number of scientific reports listed in Medline has stayed surprisingly constant for several years, at almost 375,000 for each year during 1990–1993.

As a consequence of the increase of structural reports, a rise in the use of sequence-related expressions that are unclear or even incorrect is observed. One misusage that occurs frequently is 'extent of homology', i.e. high homology, low homology, percent homology, or other numerical expressions. These phrases contradict the correct usage of homology as a qualitative term indicating the existence of evolutionary relationships [1,2]. A search of Medline abstracts indicates that this misusage started in 1975 and increased in frequency, with some fluctuations, until 1986. After a decrease until the 1990s, the homology misuse has persisted at a substantial level of about 3% of all Medline-listed abstracts which deal with the topic 'amino acid sequence'. Although this level is approximately constant in relative terms, its absolute values are in fact increasing because of the rise in the number of structural reports.

A misnomer of even older origin is 'primary sequence'. This hybrid expression of 'primary structure' and 'amino acid or nucleotide sequence' was first found in Medline-listed abstracts in 1968. It was for a time more common than the homology misnomers, but was overtaken by the latter in 1982. The incidence of primary sequence has since fluctuated and this misnomer is presently used in about 1% of all scientific reports containing the phrase 'amino acid sequence' (Fig. 1).

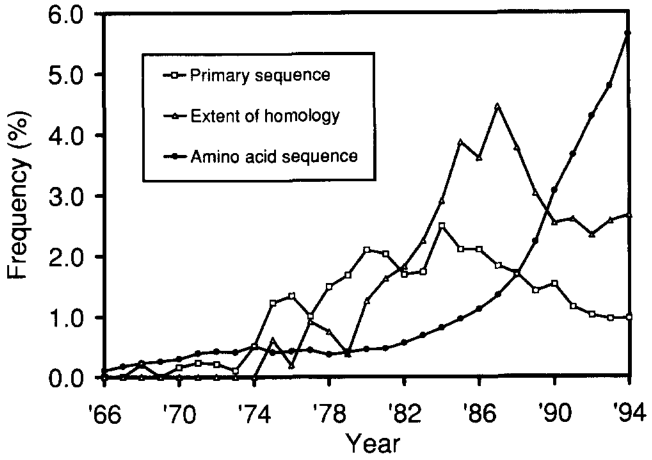

Fig. 1. Frequencies per year of Medline-listed reports containing the misnomers 'primary sequence' (□) and 'extent of homology' (△) in abstracts including the keyword 'amino acid sequence'. The increase of papers (○) containing 'amino acid sequence' is also shown.

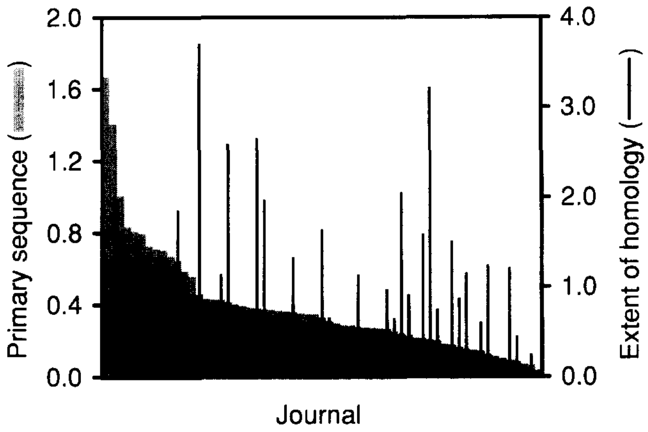

Fig. 2. Proportion of Medline-listed abstracts (%) containing the misnomers 'primary sequence' (gray bars) and 'extent of homology' (black bars) in journals with above average values. The lack of covariance between the two misnomers in different journals is evident.

The homology and primary sequence misusages are well represented in most journals (Fig. 2), including those devoted to protein and nucleic acid structures. The main reason for high misusage must be non-observance by authors, editors, referees, and copy-editors. Perhaps it is now time for a reminder to provoke a second substantial decrease similar to the one observed in the 1980s. The observation that some journals misusing primary sequence score well in not using extent of homology, and vice versa (Fig. 2), indicates that the task of avoiding the incorrect terms is not hopeless. Authors and journals should adopt the correct terminology in the interest of clear definitions in the expanding area of structural research.

References

[1] Reeck, G.R., de Haen, C., Teller, D.C., Doolittle, R.F., Fitch, W.M., Dickerson, R.E., Chambon, P., McLachlan, A.D., Margoliash, E., Jukes, T.H. and Zuckerkandl, E. (1987) 'Homology' in Proteins and Nucleic Acids: A Terminology Muddle and a Way out of It. Cell 50, 667.

[2] Lewin, R. (1987) When Does Homology Mean Something Else? Science 237, 1570.

---

*Corresponding author. Fax: (46) (8) 337462.
E-mail: Hans.Jornvall@mbb.ki.se