

When Does Homology Mean Something Else?

Molecular biologists routinely compare sequences between different genes, searching for degrees of "homology" between the two; "foul" cry evolutionary biologists, who say the word is misused, thereby causing confusion

“WHEN I use a word,” said Humpty Dumpty in a rather scornful tone, “it means just what I choose it to mean—neither more nor less.” But, in the scientific realm at least, such an Alice Through the Looking Glass approach to language can lead to misunderstandings: a prime example is the word homology.

“‘Homology’ has the precise meaning in biology of ‘having a common evolutionary origin,’ but it also carries the loose meaning of ‘possessing similarity or being matched.’ Its rampant use in the loose sense is clogging the literature on protein and nucleic acid sequence comparisons with muddy writing and, in some cases, muddy thinking.”

This verbal volley, launched in the pages of *Cell*, is the most recent attack on an old problem. “Yes, this battle has been fought for more than a decade, but it has usually been fought by individuals,” says Gerald Reeck of the Kansas State University. “I thought it was time for a more concerted effort.” That effort is in the form of an appeal signed by Reeck and ten prominent evolutionarily-oriented molecular biologists, including, Richard Dickerson, Thomas Dukes, Emanuel Margoliash, and Emile Zuckerkandl. “With a collective effort to mend our ways, proper usage [of homology] would soon become fashionable and therefore easy,” they write. “We believe that we and our scientific heirs would benefit significantly.”

The problem arises in the comparison of sequences, either of proteins or genes, in which, say, a 20% identity of sequence is typically described as 20% homology. “Molecular biologists know what they mean by such a statement,” says Walter Fitch, a molecular biologist at the University of Southern California and a cosigner of the letter. “But in fact they are mixing together two different, but related, properties. To classical biologists, homology means not just similarity of structure, it means common descent. It may be true in many cases that similarity of sequences between, say, two genes is the result of common descent, of homology. But I believe it is important to distinguish

the observation from the conclusion.”

Russell Doolittle, a molecular biologist at the University of California at San Diego, and a cosigner of the letter, traces the abuse of homology back through almost two decades. “Most people in protein chemistry in the late 1960s were not classically trained biologists,” he explains, “and to them the word homology simply meant similarity.

“When I use a word,” said Humpty Dumpty in a rather scornful tone, “it means just what I choose it to mean—neither more nor less.”

This word mutation became fixed in this group, and continues through to today. A second group led by Walter Fitch and others—people with a much greater awareness of evolutionary biology—saw the need for correct usage, and invented all kinds of other terms that simply made the whole thing very complicated and esoteric.”

In recent years the sequencing fraternity has far outweighed Fitch and his ilk, and homology meaning similarity has become common usage, as any quick glimpse of the literature reveals. “We’ve been swamped,” says Reeck. At the same time, many molecular biologists have become interested in evolutionary questions. As the Whitehead Institute’s David Baltimore once remarked, “everything we look at is evolution.” Once this happened, the potential for confusion became acute.

The first skirmish over the use of homology was between Fitch and others in the pages of *Science* some 15 years ago. Fitful exchanges continued in various vehicles, with a flurry of letters breaking out in *Nature* 4 years ago. Reeck’s current multi-authored appeal is an attempt to raise the issue beyond the level of individual sparring.

“I’ve been spending summers collaborating with Christoph de Haen and David Teller,” says Reeck, “and we would discuss this from time to time. Eventually we decided that a group effort was the only way to achieve anything. The result was the letter to *Cell*.” De Haen and Teller, both at the University of Washington, became cosigners.

“People tell us that things have gone too far,” says Fitch, “even people who were sympathetic with what we are trying to do. Maybe that’s true. I’m not interested in fighting for lost causes. I just think things should be clear.”

The clarity Fitch seeks is this. “In its precise biological meaning, ‘homology’ is a concept of quality. The word asserts a type of relationship between two or more things,” write Reeck and his colleagues. That relationship is common descent, and therefore homology cannot be partial—10%, 20%, and so on. “Things can’t be partially homologous any more than a woman can be partially pregnant,” quips Fitch.

“If using ‘homology’ loosely did not interfere with our thinking about evolutionary relationships,” write Reeck and his colleagues, “the way in which we use the term would be a rather unimportant semantic issue. The fact is, however, that loose usage in sequence comparison papers often makes it difficult to know the author’s intent and can lead to confusion for the reader (and even for the author).”

One key source of confusion is that a degree of structural similarity is an irrefutable, quantified fact, supposing the sequencing has been done correctly. By contrast, the suggestion of common descent must always be an hypothesis, however strongly supported by the evidence. Structural similarity and homology are clearly very closely tied together, but they are not necessarily the same thing.

Keeping the two things apart requires using different words, urge Reeck and his colleagues. “Sequence similarity” should be used to describe what is observed between two structures. “Homology” is the inference of common evolutionary origin. Period. The cosigners say that many people argue that attempting to enforce such terminology is anachronistic, that the word homology itself is evolving and taking on new meanings. “If that evolution is toward vagueness and if it results in making our scientific discourse unclear, surely we should intervene.” ■

ROGER LEWIN

ADDITIONAL READING

G. R. Reeck *et al.*, “Homology in proteins and nucleic acids: A terminology muddle and a way out of it,” *Cell* 50, 667 (1987).