

**This Week's Citation Classic<sup>®</sup>****Fitch W M & Margoliash E.** Construction of phylogenetic trees.*Science* 155:279-84, 1967.

[University of Wisconsin Medical School, Madison, WI and Protein Section, Department of Molecular Biology, Abbott Laboratories, North Chicago, IL]

This paper provided a way to construct an evolutionary tree using a measure of genetic distance obtained from amino acid sequence differences interpreted in terms of the genetic code. [The SC1<sup>®</sup> indicates that this paper has been cited in over 670 publications.]

Walter M. Fitch

Ahmanson Center for Biological Research  
University of Southern California  
Los Angeles, CA 90089-1481

April 27, 1988

In 1964 H.G. Khorana told me how the genetic code was about to be broken. I could see how that event would make it possible to fulfill the vision of L. Pauling and E. Zuckerkandl<sup>1</sup> of a molecular paleontological record in proteins and nucleic acids. If I could see it, so would everyone else. I figured the only way to beat out everyone else was to develop a computer program that, to work, would need only the minimum base differences for every possible pair of amino acids. That way, if I were lucky, after the code was solved I'd be analyzing data while everyone else was still writing their programs.

In the spring of 1966, Emanuel Margoliash, whom I had never met, came to give a seminar in Madison on cytochrome *c*. Robert Bock, the new dean of our graduate school, knew about my effort and that I was using cytochrome *c* for this study because there were 10 published references, more than for any other protein. He arranged for me to have lunch with Margoliash, during which I unfolded my scheme. Margoliash was very interested and

revealed that there was some slowness in the rate of sequence publication from his laboratory and that he had another 10 completed sequences. He offered to provide them to me. What a windfall! At the stroke of a conversational noon-hour clock, the sample size was doubled to 20, permitting us to present a tree spanning the largest part of the eukaryotic kingdom, and I got a coauthor whose prose was pellucid.

The genetic code was solved (except we then thought that the methionine and tryptophan codons were twofold degenerate), the minimum base differences were plugged into the program, and the analysis was performed. We were lucky that cytochrome *c* was so slowly evolving, or our first tree would have been garbage. But, as cytochrome *c* was slowly evolving, the resulting tree, spanning fungi, plants, and metazoans, looked quite respectable despite a few imperfections.

The importance of this work lay in its demonstrating two future potentials. One lay in the data source. With only one small macromolecular sequence, an excellent view of evolutionary relations in nearly all of the eukaryotic kingdom was obtained. The data were minimal estimates of the number of nucleotide substitutions separating every pair of sequences, data that—unlike morphological attributes—could be examined across all of life, data that had intrinsic units of change that could be added, subtracted, multiplied, and divided, and data that were comparable across different genes. The other potential lay in the methods. The principle method used improved on the unweighted pair-group method<sup>2</sup> by allowing unequal rates of evolution since a common ancestor. A subsidiary consideration introduced parsimony, a method that was not fully developed until later<sup>3</sup> but that was destined to be a major method in future evolutionary efforts.<sup>4</sup>

1. Pauling L & Zuckerkandl E. Chemical paleogenetics: molecular "restoration studies" of extinct forms of life. *Acta Chem. Scand.* 17:S9-16, 1963.
2. Sokal R R & Sneath P H A. *Principles of numerical taxonomy*. San Francisco, CA: Freeman, 1963. 359 p. (Cited 1,425 times.) [See also: Sokal R R. Citation Classic. (Barrett J T, comp.) *Contemporary classics in plant, animal, and environmental sciences*. Philadelphia: ISI Press, 1986. p. 327.]
3. Fitch W M. Toward defining the course of evolution: minimum change for a specific tree topology. *Syst. Zool.* 20:406-16, 1971. (Cited 145 times.) [See also: Fitch W M. Citation Classic. *Current Contents/Agriculture, Biology & Environmental Sciences* 18(27):14, 6 July 1987.]
4. Pace N R, Olsen G J & Woese C R. Ribosomal RNA phylogeny and the primary lines of evolutionary descent. *Cell* 45:325-6, 1986.