Incomplete lineage sorting and other 'rogue' data fell the tree of life

Jeffrey Tomkins and Jerry Bergman

The 'tree of life' (TOL) popularized by Darwin and used as the inferred pattern of life's history is the centrepiece of evolutionary biology. The molecular genetics revolution has presented many contradictions for the TOL and the modern Darwinian synthesis. Incomplete lineage sorting (ILS) is a discordant and pervasive outcome produced when constructing phylogenetic trees using homologous biological sequence data across all types of life studied. The ILS paradigm is characterized by segments of DNA that produce phylogenetic trees with different topologies compared to hypothetical inferred evolutionary trees. While ILS within closely related taxonomic groups can largely be explained by horizontal genetic variation and limitations on accurately sampling large populations, ILS across clearly different and unrelated kinds of organisms represents a mosaic of DNA sequence patterns that cannot be explained by common ancestry. Other 'rogue' genetic data that defy the TOL are microRNA genes and taxonomically restricted genes. MicroRNAs produce completely different trees compared to other gene sequences and appear unexpectedly in taxa. Taxonomically restricted genes also appear abruptly without evolutionary precursors, lack homology to other genes, and uniquely define taxon. Genetics research consistently reveals patterns of DNA mosaics that defy evolution and vindicate biblical creation 'after their kinds'.

The dominant metaphor of evolutionary biology is the overall concept of a branching tree described by Darwin in 1859, in his book titled *On the Origin of Species by Means of Natural Selection, or The Preservation of Favored Races in the Struggle for Life.*[1] Using a single illustration of a tree diagram with branching patterns and calculations, Darwin illustrated the gradualistic divergence of species over time (figure 1). However, from the time of Darwin to the early molecular protein work of Zuckerkandl and Pauling in the 1960s, these trees were largely based on closely-related species and groups of organisms.[2] For an example of a phylogenetic tree, see figure 2.

According to evolutionary theorists, the simple assumption of phylogenetics and the development of evolutionary trees from biological sequence implies that "as the time increases since two sequences diverged from their last common ancestor, so does the number of differences between them, tree estimation seems to be a relatively simple exercise: count the number of differences between sequences and group those that are most similar".[3] Nevertheless, evolutionary biologists also recognize that "The simplicity of such an algorithm underestimates the complexity of the phylogenetic-inference problem".[3] In fact, the main problem with phylogenetic inference is that of discordant data. This rogue data provides no support for gradualistic Darwinian assumptions and the inferred common ancestry across the spectrum of life. In the case of phylogenetics, where certain homologous sequences across taxa exist and make possible the use of comparative techniques, the discordant data is typically referred to as incomplete lineage sorting or ILS. For a simple graphical example of ILS as displayed in phylogenetic trees, see figure 2.

Prior to the recent advent in DNA sequencing, a 1965 report by Throckmorton using morphological characters in the genus *Drosophila* (fruit fly) described how similarity in individual phenotypic traits did not consistently predict assumed evolutionary relationships when evaluated independently.[4] Later, in 1978, Farris made one of the first attempts at using one of the early tools of molecular genetics (chromosome inversion data) to infer evolutionary phylogenies and ran into the same enigmatic issue of ILS.[5] It should be noted, however, that Drosophila is an animal with large populations and short generation times. In interrelated and interfertile populations that may be largely separate, chromosome inversions that are tolerated will not completely inhibit gene flow.[6] Thus, ILS among closely related taxa can largely be explained as a common feature of horizontal genetic variation within kinds as recently demonstrated among a tribe of cichlid fishes.[7] The presence of ILS among closely related taxa is also affected by the fact that accurately sampling and characterizing large populations, such as cichlid fish in multiple lakes and rivers, can be very difficult.[8]

While the findings and reports of ILS within single taxa are noteworthy, they do not provide an adequate evolutionary explanation for many recent studies in which ILS is observed across completely unrelated kinds of organisms that are obviously not interfertile and have no flow of genetic information between them. Evolutionists like to extrapolate the observed variation within kinds and