



Reconstructing/Deconstructing the Tree of Life

W. Ford Doolittle

The Haida Indians of British Columbia believed that after Raven (a mischievous, mythological creature) created this earth *ex nihilo* (out of nothing), he coaxed the first humans from a clam shell and made them the gift of fire.

The Judaeo-Christian tradition, on the other hand, endorses a sequential creation of the non-living and living elements of the cosmos—plants and animals first, then humans.

Creation narratives like these play a vital moral and inspirational role in all our cultures. They ease the burden of individual mortality and stake our species' claim to at least one tiny corner of an immense, indifferent universe. Although at different ends of the spectrum, each narrative is an essential component of a revealed spiritual belief system. There is a parallel narrative about the history of life in the universe that is, at least in principle, discoverable through scientific inquiry. But can this narrative truly provide spiritual solace or guidance for those seeking it? Where do these individuals turn for answers to the ancient questions about the origins of our species?

Narrative, spirituality, and culture aside, perhaps these answers lie in the tiniest organisms that were a major feature on the landscape billions of years before man and woman ever appeared on the scene.

Looking for answers in the tiniest of places

Large organisms with recognizable hard parts have left a fossil record for modern man to use as a resource and guide in the quest for answers. Because of this record, we can formulate, test, and sometimes even prove detailed evolutionary theories. For instance, using this approach we can learn about the origins of mammals and birds from reptiles between 150 million and 300 million years ago, or about the more recent separation of humans, chimps, and other primates. In science, however, it's not always enough to travel back a mere 300 million years. Sometimes a longer trip is necessary. In fact, my colleagues and I are interested in events that happened much longer ago—about two billion to four billion years—and that involved the tiniest microscopic cells with few if any distinguishable morphological features. In particular, to assist us in our efforts to piece together history, we are interested in the origins of the very first cells, and the evolution of complexity at the cellular level.

So how can we hope to reconstruct this history? My field, usually called “molecular evolution,” is based on the premise that a comparative analysis of sequences of molecules, in particular of DNA, can replace the study of fossils. The underlying logic is simple. As two species diverge in evolution from a common ancestor, their genes will accumulate mutations. Some of these mutations will affect the function of the protein or RNA that the gene encodes, and some will be simply neutral. The longer the divergence, the more the accumulated mutations. And the more the accumulated

Reconstructing/Deconstructing the Tree of Life

mutations, the greater the difference between the genes. It is this difference between genes (more precisely, gene sequences) that can be used as a measure of evolutionary divergence. Typically, the pattern of differences is represented as a phylogenetic tree, or genealogy—the Tree of Life. As a result, we can reconstruct past events that gave rise to the current panoply of similar and dissimilar creatures by isolating and sequencing specific genes from their DNA.

Reconstructing the Tree of Life

Ever since our field was invented in the mid-1960s, reconstructing phylogenetic trees from gene sequences in this way has been the major preoccupation of molecular evolutionists. Since the beginning of the 1970s, most of us interested in recovering life's most ancient history have used sequences of a single kind of gene, called SSU rRNA. The gene is known to be present in all organisms and is thought to exhibit a regular, clock-like rate of accumulation of neutral mutations. In fact, the publicly operated databases now contain almost 10,000 versions of this gene's sequence, representing organisms across the entire spectrum from *E. coli* to elephants.

What does the analysis of all this data show us? It shows us an enormous twiggy tree, with three main branches. By studying this "Tree of Life," it becomes clear that all living things are located on one of its three branches; they are either eukaryotes (*you-carry-oats*), or one of two kinds of prokaryotes (*pro-carry-oats*)—bacteria or archaea.

Eukaryotic cells are the kind with which we are most familiar. That's because almost all life forms that are big enough to be seen are eukaryotes. The distinguishing characteristic is the nucleus, a membrane-bounded compartment that surrounds the DNA. The DNA is the genetic material or blueprint of the cell. In prokaryotes, however, the DNA is in free contact with the rest of the cell contents. Undoubtedly, the first eukaryotes (single, amoeba-like cells) arose somehow from the simpler prokaryotes. Geochemical evidence suggests that this happened between two billion and three billion years ago.

Of the prokaryotes, we are already well acquainted with bacteria. Many of our worst diseases (plague, tuberculosis, necrotizing fasciitis) are caused by bacteria. But many essential environmental services (decomposition, recycling of carbon, nitrogen, and oxygen) also depend on them. Archaea, however, are the unexpected group and have shown themselves to have unexpected biological idiosyncrasies. Some can grow at temperatures as high as 115 degrees Celsius (under the sea); some are killed by the smallest traces of oxygen; while others thrive only in saturated brines.

The last type of archaea, the halophiles, have long been a Canadian specialty. Al Matheson and Mak Yaguchi at the National Research Council, and Donn Kushner of the University of Ottawa, pioneered research on halophile ecology and physiology even before their evolutionary position was revealed. My lab at Dalhousie University and that of Pat Dennis at the University of British Columbia, were the first to apply modern genetic tools to understanding their biochemistry.

Reconstructing the phylogenetic tree with SSU rRNA data has shown us how complex eukaryotic cells could have arisen from simpler prokaryotic antecedents. The current view is that eukaryotes are evolutionary chimeras (named after the mythological

Reconstructing/Deconstructing the Tree of Life

monster with a lion's head, goat's body, and serpent's tail). That means that the thousands of genes that make up the eukaryotic genome ultimately derive from different prokaryotic sources. Those genes responsible for key hereditary processes and the expression of hereditary information (the “hardware” of the cell) are derived from the archaea. Those genes responsible for at least two key energy-generating processes—respiration and photosynthesis (more akin to “software”)—are of bacterial origin.

For molecular evolutionists, the biggest surprise in recent years has been the discovery that bacteria and archaea are themselves chimeric (derived from different prokaryotic sources). This knowledge came after the completion of more than three dozen bacterial and archaeal genome sequences, each with its own complement of 500 to 5,000 genes. By no means do all these genes produce the same phylogenetic tree as SSU rRNA.

The driving force behind evolution

It is clear that the borrowing of genes across lineages (from one kind of bacterium to another, from archaea to bacteria, or from either to eukaryotes) has played a major role in the evolutionary process. In some groups, it seems that it is gene borrowing not mutation that has been the major source of functional novelty. As a result, it has been the real driving force behind evolution. Some examples of consequences of immediate relevance to us? The rise of antibiotic resistance in known pathogenic bacteria (due almost entirely to transfer), and the appearance of new pathogens (such as *E. coli* O157:H7)—often the result of the transfer of clusters of virulence-promoting genes called “pathogenicity islands.”

A new century with new ideas

The renowned British naturalist and evolutionist Charles Darwin (1809 to 1892) believed that a single phylogenetic tree would be the best representation of life's history. It would also be the best way to account for the patterns of similarity and difference that we see among contemporary organisms.

Today, at the beginning of the 21st Century, is there room for a different viewpoint? Is our model of the Tree of Life still as relevant and fruitful as we once believed? Perhaps not. If different genes have followed different evolutionary paths, then there can be no such single phylogenetic tree that can be reconstructed from gene sequences. For bacteria and archaea, in order to visualize the pattern of history, a web or net may be more appropriate than a tree. In a sense, bacteria and archaea evolve as if they were a single global species, albeit one within which rates of genetic exchange are highly variable and generally slow. Sorin Sonea, a microbiologist at the University of Montreal, actually said this about 30 years ago. However, at that time, the implications of such a view seemed far too radical and the support in evidence far too scant.

What does the future hold?

My own lab at Dalhousie University is now running full tilt to elaborate the web or net model. We are testing its general validity and exploring its implications through the acquisition and analysis of genomic (whole genome) data. The general parameters that affect gene transfer—the relatedness and physical proximity of the partners in exchange; the genetic mechanisms involved; the classes of genes that can or cannot be

Reconstructing/Deconstructing the Tree of Life

exchanged; selective pressures that drive the integration of new genes; the consequences for the functions of resident genes—all need to be worked out.

What do we hope to achieve? At the practical level, we would hope to understand the spread of antibiotic resistance and pathogenicity in an experimental/theoretical framework that would also encompass the micro-organisms that do good things for us (in both polluted and natural environments). At the philosophical level, we want to ask how gene transfer threatens the very concept of species. At all levels, we want to get a good feeling for the prevalence and power of gene transfer. At the moment, too much of the data is still anecdotal or ambiguous.

Research like this is so basic some might consider it arcane. We prefer to see it as foundational. I consider myself blessed to have been able to spend most of my career building this foundation. In this effort, I have had as colleagues some of the world's brightest and most deeply motivated biologists, chemists, geneticists, and mathematicians—no small number of them Canadians, and gratifyingly, many of them my own students and postdoctoral fellows.

All that we understand and do in clinical and environmental biology depends on a structure of prior knowledge, which we impart to our students and apply in a variety of practical contexts. This structure is not static. Although science does build incrementally, so that much of what we learn today is a further detailed elaboration of established generalities, sometimes those generalities are themselves overthrown. This is where I think and hope we are in our understanding of the first few billion years of the history of life, and of the genetic structure of the microbial world—then and still.