



EVOLUTION

Building the Tree of Life, Genome by Genome

Cheaper sequencing has put many more genes into the hands of researchers trying to sort out the degree of relatedness of a menagerie of organisms

Phylogenetic studies have gone 'omic. Whereas researchers used to be satisfied comparing one gene, or a few, to sort out the branching of the tree of life, the push now among those building phylogenies is to consider whole genomes—at the very least, dozens of genes and thousands of DNA bases—in establishing kinships among flora and fauna. In this way, evolutionary biology is joining the bandwagon of data-intensive studies pioneered by genomics.

Thanks to one such phylogenomic analysis reported on page 1763, bird guides may never be the same. According to this new avian family tree, grebes will share a section with flamingos, not loons. Dull brown night jays and iridescent hummingbirds would now go together. Even parrots and songbirds share a closer kinship than has been appreciated, says Shannon Hackett, an ornithologist at The Field Museum of Natural History in Chicago, Illinois.

She and more than a dozen colleagues constructed the new genealogy after analyzing 32,000 bases from 19 genes in 169 species. More than just rearranging which birds perch on what branches of the tree, the results raise questions about the evolution of flight; some birds that don't fly are unexpectedly grouped with those that do. "It's the most impressive paper in the higher level phylogeny of birds to come along in a long time," says Joel Cracraft, an evolutionary biologist at the American

Museum of Natural History in New York City. "It will be used by avian systematists and non-avian systematists for a very long time."

The bird work follows two phylogenomic studies published over the past 3 months that have shaken up perceived evolutionary relationships among animals and more broadly among eukaryotes. In the former effort, a team led by Casey Dunn, now at Brown University, has rearranged the animal kingdom such that comb jellies, not sponges, are among the earliest fauna. In the latter, a European team now divides eukaryotes into two megagroups, not a half-dozen. Together, the three trees speak to the potential of phylogenomics. "We are just beginning to understand what large sequence data sets have to say about the evolution of life on Earth," says Hackett.

Entering the genome age

The term "phylogenomics" was coined by Jonathan Eisen a decade ago to describe incipient efforts to integrate evolutionary thinking into genomic analyses and vice versa. What this evolutionary biologist at the University of California, Davis, had in mind was using information about the relatedness of newly sequenced organisms to help sort out gene function and identify comparable stretches of DNA in genomes that have been deciphered. But the term has been "kidnapped," says Eisen jokingly, by the likes of Hackett and others to

describe large-scale efforts to build family trees based on lots of molecular data.

Systematists may like the label, but there's no agreement about how many genes it takes to make an evolutionary tree phylogenomic. "We would say our study is phylogenomic because we have sampled many different genes from many different chromosomes across a subset of avian species, but others would say we still sampled a small portion of the genome," Hackett points out. And ornithologist Michael Sorenson of Boston University applies an even tougher standard: "I would reserve the term for what lies ahead, i.e., comparisons of whole genomes."

In traditional molecular phylogeny, researchers pick out a short stretch of one gene, often a mitochondrial gene, count up the sequence differences between species in that stretch, and use sophisticated computer programs to come up with the hierarchy of evolutionary relationships between the species. Most simply, the fewer the differences, the more closely related two species were considered to be.

Gradually, however, researchers realized "that single-gene trees are prone to errors and that many genes are necessary," explains Jose Castresana of the CSIC Institute of Molecular Biology of Barcelona in Spain. Because genes can evolve at different rates, it's not always possible to pinpoint the true time a species under consideration diverged from a common ancestor by looking just at the changes in one gene from that species. In some cases, there are too few changes to provide statistically reliable results. Other times, the transfer of a gene from one species to another causes phylogenetic chaos.

When Hackett, her postdoc Sushma Reddy, Rebecca Kimball of the University of Florida, Gainesville, and colleagues started their avian project in 2003, collaborators first

Treed. An in-depth comparison of DNA showed that Western tanagers, parrots, and falcons (*left to right*) are closer kin than expected.

did a computer simulation to determine how much and what kind of DNA sequence would enable them to figure out the early history of birds. The simulation directed the team to collect at least 20,000 bases from introns and intergenic regions, where mutations occur frequently enough for there to be significant differences in the various lineages. At first, the researchers sampled only about 75 species, but after realizing how much more robust results would be with a larger number, they doubled it.

“The ultimate goal is to provide the rest of the ornithological community with the roots and base of the tree that they can leaf out more effectively,” Hackett says. Traditionally, avian systematists have had trouble sorting out those early days of bird evolution, notes Harvard University ornithologist Scott Edwards. The new results are “bold in setting an agenda for future research,” he says.

In agreement with previous avian phylogenies, Hackett, Kimball, and Reddy found that the South American bird family tinamous, along with ratites—kiwis, ostriches, and the like—split off close to the base of the bird tree. Slightly later, chickens, ducks, and their kin branched away from the main group of birds. The subsequent history of birds has been enigmatic, but the new work offers some clarity. Songbirds, for example, are a sister group to parrots, and the two groups encompass all the descendants from their most recent common ancestor. Hummingbirds descended from night jays, evolving bright colors and a diurnal lifestyle along the way.

One of the more controversial results is that tinamous, all capable of flight, belong in the same group as the flightless ratites. This “can change the way people look at the evolution of flight,” Hackett says. Grouping the birds together suggests either that flightlessness evolved multiple times, not once in the ancestor to this group, or that flight evolved more than once in birds, showing up independently in the tinamous and in other flying birds. “This result flies in the face of many other kinds of data,” says Edwards.

Shaky branches

The phylogenomics study of eukaryotes, conducted by Fabien Burki and Jan Pawlowski, both at the University of Geneva, Switzerland,

and Kamran Shalchian-Tabrizi of the University of Oslo, Norway, also upsets old assumptions. Interested in deciphering the deep roots of eukaryotes, which include protists, plants, animals, and fungi, they combed the public databases, coming up with 135 genes from 65 species to compare. Based on the pattern of differences in the sequences, they and their colleagues came up with three early branches, two containing almost all eukaryotes and one tentatively placed branch representing excavates, protists that include *Euglena* and *Giardia*.

Unlike past analyses based on just a few eukaryotic genes, or just one, this phylogenomic effort, published online 3 June in *Biology Letters*, brought all photosynthetic organisms—save *Euglena* and its relatives—into one group. The researchers suggest that the cyanobacterium that gave rise to the modern chloroplasts seen in plants and in green and red algae was acquired much earlier in eukaryotic evolution than had been thought, though more data is needed to confirm this idea, says Burki.

That plants now group with dinoflagellates, diatoms, or freshwater flagellates—all previously considered independent “super-

little-studied creatures, including water bears, comb jellies, sea spiders, and a variety of worms. These data, combined with existing information, enabled them to evaluate 150 genes from 71 animals.

In some cases, the major branches of the new animal family tree confirmed researchers’ suspicions. For example, based on a suite of similar traits seen in the animals, morphologists have long thought that mollusks all stem from a common ancestor. Yet there is no single unifying trait among the phylum, which includes scallops, squid, chitons, and snails. Many, but not all, have a toothlike structure called a radula, and a subset have no shell, even though mollusk means “thin-shelled.” Moreover, the molecular data did not back up the premise that all traditional mollusks belong together. Dunn’s new tree shows that the mollusks are one big family, however. “It’s nice to have tied [this relatedness] down,” says Dunn.

But the conclusion that comb jellies are the oldest animals is a surprise, says Dunn, who adds that the reaction has ranged from “‘That is so cool!’ to ‘There is no way.’” Dunn himself calls that result provisional and sees his 10 April *Nature* paper as just the beginning.

Thanks to new sequencing technologies, “within a year or two, we’ll be seeing studies that have 10 times as many genes from 10 times as many taxa,” he predicts.

And he’s not the only one to soon be awash in data. Burki is generating more sequences for his work with eukaryotes, and Hackett and colleagues are expanding their data set as well. “Phylogenomics is becoming the rule,” says Hervé Philippe, who develops new phylogenetic techniques at the University of Montreal,

Canada. Philippe looks forward to more phylogenomics studies that use gene order, even gene content or intron positions, to infer relationships—approaches that will become “more natural when complete genomes are available,” he says.

Philippe and others caution, however, that more data don’t always guarantee better family trees. “It will be important to reanalyze [data sets] with many different and emerging methods to see if the results change at all,” says Edwards. And a few scientists question whether, even then, the full tree of life can really be resolved. But, Edwards argues, “phylogenomics is our best shot.”

—ELIZABETH PENNISI



Rooting animals. After sequencing DNA from 29 animals, researchers concluded that comb jellies (*above*) are likely the most primitive known animals and that nudibranchs (*left*) and other mollusks are really true kin.

groups”—has raised some eyebrows. “I think this is untenable,” says Patrick Keeling at the University of British Columbia in Vancouver, Canada. Nonetheless, he adds, “this paper represents one of the right ways we should be going to resolve the tree of eukaryotes.” The challenge is to include more organisms in future studies. In doing so, “it’s entirely possible that strong support for many relationships will evaporate,” he notes.

When Dunn and his colleagues wanted to tackle the animal kingdom, they couldn’t find enough publicly available DNA sequence for the many species they needed to examine. So they sequenced 39.9 million bases from 29 of nature’s more peculiar and