Capturing introns: Targeting rapidly evolving regions of the genome for phylogenetics

September 9 2015

Understanding the evolutionary history of organisms is important for myriad reasons. To name a few, information about relationships between species can be used to guide the classification of biodiversity, inform conservation policies aimed at protecting threatened species, aid in tracking the spread of pathogens, and can even play a role in the discovery of new medicines.

Scientists depict the relationships between species with evolutionary trees, also called phylogenies. A phylogeny shows the accumulation of species through time and the relationships between these species, much like a family tree shows how you are related to the other members of your family.

By analyzing the similarities and differences between DNA sequences of different organisms, biologists can infer relatedness between them. But what if the DNA of the species-in-question is so similar that you can't tease the relationships apart?

Luckily, different regions of DNA evolve at different rates. Some regions, such as protein-coding genes, evolve at a relatively slow rate. Introns - the non-coding part of genes - evolve at a much higher rate and are, therefore, potentially useful at resolving evolutionary relationships between closely related species.

"There was a major gap for researchers using genomic DNA sequences to understand the evolution of species complexes," says Ryan Folk, lead author of a study in a recent issue *Applications in Plant Sciences*. "In these groups, the boundaries between species are often quite blurry, making it difficult to accurately infer relationships."

Folk and colleagues at The Ohio State University and University of Memphis have developed a technique to capture the rapidly evolving intronic regions of the genome in hopes of increasing our knowledge of evolution in these difficult groups.

The new study describes a method for targeting introns by utilizing publicly available genomic data to find these rapidly evolving regions. Biotinylated probes are constructed based on these sequences to chemically 'capture' the intron-containing genes from the study species' genome. The rest of the DNA from the sample is then simply washed away, leaving only those targeted sections.

"Previous studies utilizing similar techniques have focused primarily on slowly evolving regions, such as exons, due to concerns with capture success," explains Folk. "The problem with this is wasted DNA sequencing effort. The data are perfectly good, but with a much slower mutation rate it's necessary to increase the amount of data you target to obtain an equivalent amount of evidence for evolutionary relationships among species."

The protocol developed by Folk and colleagues turned out to be more successful than anticipated.

"When we designed this project, we really had no idea how well targeting introns would work. Targeting a DNA region depends on sequence recognition between the unknown sample and a known 'target' sequence from a close relative. Of course, for a rapidly evolving region,

the target will not resemble the sequence of a new sample as closely, resulting in a greater risk of failure for the experiment. But our results show that introns were not particularly difficult to sequence as compared to the more commonly targeted exons."

The method was tested on a recent, rapid radiation of plants in the *Heuchera* group. Over 270 DNA markers were targeted and sequenced using next-generation sequencing. The team found introns to have a fourfold increase in mutation rate compared to exons, allowing them to confidently resolve evolutionary relationships between very closely related taxa, something previously not possible in this group. On top of this, complete organellar (plastid and mitochondrial) genomes were assembled for one species.

"Although often difficult to answer, questions relating to species complexes are sometimes the most interesting because they represent the recent or ongoing formation of new species. We are optimistic that our new method for targeting intronic regions will provide scientists with a valuable tool to answer questions in these difficult groups."

More information: Ryan A. Folk, Jennifer R. Mandel and John V. Freudenstein. 2015. A protocol for targeted enrichment of intron-containing sequence markers for recent radiations: A phylogenomic example from Heuchera (Saxifragaceae). *Applications in Plant Sciences* 3(8): 1500039. DOI: 10.3732/apps.1500039

Provided by Botanical Society of America