

# Researchers use supercomputers to study snake evolution, unique traits

June 22 2017, by Aaron Dubrow

---



A Burmese python superimposed on an analysis of gene expression that uncovers how the species changes in its organs upon feeding. Credit: Todd Castoe

Evolution takes eons, but it leaves marks on the genomes of organisms that can be detected with DNA sequencing and analysis.

As methods for studying and comparing genetic data improve, scientists are beginning to decode these marks to reconstruct the evolutionary history of species, as well as how variants of [genes](#) give rise to unique traits.

A research team at the University of Texas at Arlington led by assistant professor of biology Todd Castoe has been exploring the genomes of snakes and lizards to answer critical questions about these creatures' [evolutionary history](#). For instance, how did they develop venom? How do they regenerate their organs? And how do evolutionarily-derived variations in genes lead to variations in how organisms look and function?

"Some of the most basic questions drive our research. Yet trying to understand the genetic explanations of such questions is surprisingly difficult considering most vertebrate genomes, including our own, are made up of literally billions of DNA bases that can determine how an organism looks and functions," says Castoe. "Understanding these links between differences in DNA and differences in form and function is central to understanding biology and disease, and investigating these critical links requires massive computing power."

To uncover new insights that link variation in DNA with variation in vertebrate form and function, Castoe's group uses supercomputing and data analysis resources at the Texas Advanced Computing Center or TACC, one of the world's leading centers for computational discovery.

Recently, they used TACC's supercomputers to understand the mechanisms by which Burmese pythons regenerate their organs—including their heart, liver, kidney, and small intestines—after feeding.

Burmese pythons (as well as other snakes) massively downregulate their

metabolic and physiological functions during extended periods of fasting. During this time their organs atrophy, saving energy. However, upon feeding, the size and function of these organs, along with their ability to generate energy, dramatically increase to accommodate digestion.

Within 48 hours of feeding, Burmese pythons can undergo up to a 44-fold increase in metabolic rate and the mass of their major organs can increase by 40 to 100 percent.

Writing in *BMC Genomics* in May 2017, the researchers described their efforts to compare gene expression in pythons that were fasting, one day post-feeding and four days post-feeding. They sequenced pythons in these three states and identified 1,700 genes that were significantly different pre- and post-feeding. They then performed statistical analyses to identify the key drivers of organ regeneration across different types of tissues.

What they found was that a few sets of genes were influencing the wholesale change of pythons' internal organ structure. Key proteins, produced and regulated by these important genes, activated a cascade of diverse, tissue-specific signals that led to regenerative organ growth.

Intriguingly, even mammalian cells have been shown to respond to serum produced by post-feeding pythons, suggesting that the signaling function is conserved across species and could one day be used to improve human health.

"We're interested in understanding the molecular basis of this phenomenon to see what genes are regulated related to the feeding response," says Daren Card, a doctoral student in Castoe's lab and one of the authors of the study. "Our hope is that we can leverage our understanding of how snakes accomplish organ regeneration to one day

help treat human diseases."

## Making Evolutionary Sense of Secondary Contact

Castoe and his team used a similar genomic approach to understand gene flow in two closely related species of western rattlesnakes with an intertwined genetic history.

The two species live on opposite sides of the Continental Divide in Mexico and the U.S. They were separated for thousands of years and evolved in response to different climates and habitat. However, over time their geographic ranges came back together to the point that the rattlesnakes began to crossbreed, leading to hybrids, some of which live in a region between the two distinct climates.

The work was motivated by a desire to understand what forces generate and maintain distinct species, and how shifts in the ranges of species (for example, due to global change) may impact species and speciation.



This is a Burmese python consuming a rodent meal. Within 48 hours of feeding, Burmese pythons can undergo up to a 44-fold increase in metabolic rate and the mass of their major organs can increase by 40 to 100 percent. Credit: Stephen Secor

The researchers compared thousands of genes in the rattlesnakes' nuclear DNA to study genomic differentiation between the two lineages. Their comparisons revealed a relationship between genetic traits that are most important in evolution during isolation and those that are most important during secondary contact, with greater-than-expected overlap between genes in these two scenarios.

However, they also found regions of the rattlesnake genome that are important in only one of these two scenarios. For example, genes functioning in venom composition and in reproductive differences—distinct traits that are important for adaptation to the local habitat—likely diverged under selection when these species were isolated. They also found other sets of genes that were not originally important for diversification of form and function, that later became important in reducing the viability of hybrids. Overall, their results provide a genome-scale perspective on how speciation might work that can be tested and refined in studies of other species.

The team published their results in the April 2017 issue of [\*Ecology and Evolution\*](#).

## **The Role of Supercomputing in Genomics Research**

The studies performed by members of the Castoe lab rely on advanced computing for several aspects of the research. First, they use advanced computing to create genome assemblies—putting millions of small chunks of DNA in the correct order.

"Vertebrate genomes are typically on the larger side, so it takes a lot of computational power to assemble them," says Card. "We use TACC a lot for that."

Next, the researchers use advanced computing to compare the results

among many different samples, from multiple lineages, to identify subtle differences and patterns that would not be distinguishable otherwise.

Castoe's lab has their own in-house computers, but they fall short of what is needed to perform all of the studies the group is interested in working on.

"In terms of genome assemblies and the very intensive analyses we do, accessing larger resources from TACC is advantageous," Card says.

"Certain things benefit substantially from the general output from TACC machines, but they also allow us to run 500 jobs at the same time, which speeds up the research process considerably."

A third computer-driven approach lets the team simulate the process of genetic evolution over millions of generations using synthetic biological data to deduce the rules of evolution, and to identify genes that may be important for adaptation.

For one such project, the team developed a new software tool called GppFst that allows researchers to differentiate genetic drift - a neutral process whereby genes and gene sequences naturally change due to random mating within a population - from genetic variations that are indicative of evolutionary changes caused by natural selection.

The tool uses simulations to statistically determine which changes are meaningful and can help biologists better understand the processes that underlie genetic variation. They described the tool in the May 2017 issue of [\*Bioinformatics\*](#).

Lab members are able to access TACC resources through a unique initiative, called the University of Texas Research Cyberinfrastructure, which gives researchers from the state's 14 public universities and health centers access to TACC's systems and staff expertise.

"It's been integral to our research," said Richard Adams, another doctoral student in Castoe's group and the developer of GppFst. "We simulate large numbers of different evolutionary scenarios. For each, we want to have hundreds of replicates, which are required to fully vet our conclusions. There's no way to do that on our in-house systems. It would take 10 to 15 years to finish what we would need to do with our own machines—frankly, it would be impossible without the use of TACC systems."

Though the roots of evolutionary biology can be found in field work and close observation, today, the field is deeply tied to computing, since the scale of genetic material—tiny but voluminous—cannot be viewed with the naked eye or put in order by an individual.

"The massive scale of genomes, together with rapid advances in gathering genome sequence information, has shifted the paradigm for many aspects of life science research," says Castoe.

"The bottleneck for discovery is no longer the generation of data, but instead is the analysis of such massive datasets. Data that takes less than a few weeks to generate can easily take years to analyze, and flexible shared supercomputing resources like TACC have become more critical than ever for advancing discovery in our field, and broadly for the life sciences."

**More information:** Audra L. Andrew et al, Growth and stress response mechanisms underlying post-feeding regenerative organ growth in the Burmese python, *BMC Genomics* (2017). [DOI: 10.1186/s12864-017-3743-1](https://doi.org/10.1186/s12864-017-3743-1)

Provided by University of Texas at Austin

Citation: Researchers use supercomputers to study snake evolution, unique traits (2017, June 22)  
retrieved 20 April 2024 from  
<https://phys.org/news/2017-06-supercomputers-snake-evolution-unique-traits.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.