

[Home](#) > [ISGTW 16 December 2009](#) > Feature - GRAPPAling with evolutionary history

## Feature - GRAPPAling with evolutionary history

We've known for several years now that chimpanzees share 96 percent of our DNA. Our technology tells us how closely humans and chimps are related. But it doesn't tell us how we're related. We need new technology for that.

Enter [GRAPPA](#) – or Genome Rearrangements Analysis under Parsimony and other Phylogenetic Algorithms if you want a mouthful. GRAPPA has already been used to analyze the evolution of organelles such as chloroplasts and mitochondria, running on cluster computers with upwards of 500 processors. To analyze more complex organisms, however, the team that develops GRAPPA will have to take the code to an entirely new level – the petascale level.

With the help of a \$1 million grant from the National Science Foundation, computational science researchers [David Bader](#) of Georgia Tech and [Jijun Tang](#) of the University of Southern Carolina have joined forces with genetics researcher [Stephen Schaeffer](#), from Pennsylvania State University, to do just that.

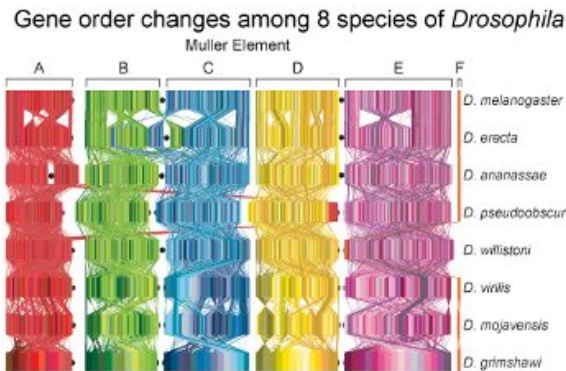
"One of the first open science petascale systems will be the IBM BlueWaters resource supercomputer," Bader said. "Our goal is to scale GRAPPA up to use that magnificent computer."

The trick is in the algorithms, according to Bader. "As an example, our first version of GRAPPA eight years ago took an hour and a half to run a problem involving the chloroplasts of a dozen species of bluebell flowers on a 512 processor linux cluster," Bader said.



A "Tree of Life" diagram. Courtesy of David Bader.

schools. Not only will they be creating invaluable tools for advancing science, but they'll also be learning very specialized skills in the process.



This figure illustrates how gene order changes among the eight species. Each thin line represents a single gene and its position in the different species. Most genes are conserved on the same chromosomal arm or Muller element, but gene order is shuffled between species. *This figure appeared in the July 2008 issue of Genetics. Image courtesy of Stephen Schaeffer.*

[ISGTW 16 December 09](#)

[Feature - GRAPPAling with evolutionary history](#)

[Feature - An interview with the "Particle-Zoo Keeper"](#)

[Feature - Arie Shoshani talks about Scientific Data Management](#)

[Link of the week - GridCast at e-Science Oxford 2009](#)

[Special Announcement - ISGTW on winter holidays](#)

**Announcements**

[Call for papers: PRACE Award for young scientists](#)

[Call for papers: First ACM Symposium on Cloud Computing](#)

[ParaPLOP 2010 accepting papers](#)

[Jobs in grid](#)

**Subscribe**

Enter your email address to subscribe to *ISGTW*.

[Unsubscribe](#)

**ISGTW Blog Watch**

[Keep up with the grid's blogosphere](#)

**Mark your calendar**

[December 2009](#)

9-10, [SEE-GRID-SCI User Forum](#), Istanbul, Turkey

9-11, [IEEE e-Science](#), Oxford, UK

9-11, [ICPADS](#), Shenzhen, China

"I think we're the pioneers in large-scale evolutionary or reconstruction of evolutionary trees," Bader said. "We're breaking new ground with each new algorithm and implementation, where no one has studied before."

It isn't clear how far these funds will take the team. Over the course of the grants lifetime, they hope to continue to refine the code, doing test runs on prototypical petascale systems. Perhaps they will be able to study fish. Or mammals. Or even humans.

Either way, a new and improved GRAPPA will be valuable in innumerable ways. Already, it has been used to develop biochemical products, identify target drug receptors, create safer pesticides, and study how viruses evolve in order to make better vaccines.

"For instance, if you know that there's a specific plant that has a property but that plant is rare or difficult to find, you may be interested in what plants are phylogenetically close or closely related in the family tree," Bader said. "They may be more abundant or easier to use to produce the right biochemical substance."

It is clear that Bader believes this is only the beginning of a new era. "If I look at other communities in scientific computing, they've matured their methods and techniques over the course of decades to centuries," Bader said. "We've only known about the structure of DNA for fifty years, and we've only had the ability to sequence full genomes in the last several years. So we're really at the infancy of what we see in the area of understanding biological sciences through computational methods."

—Miriam Boon, iSGTW

Tags: [Americas](#) [Biology and genomics](#) [Project Profile](#)

Share this page:



[Email this page](#)



[digg.com](#)



[reddit](#)



[del.icio.us](#)



[StumbleUponMa.gnolia.com](#)



**Disclaimer:**

These are external Web sites and iSGTW cannot guarantee their security nor endorse their content.

17, [EGI Industry Day](#),  
Vilnius, Lithuania

[More calendar items...](#)