The IBEST Computational Resources Core (CRC) provides world class computing facilities for storing and analyzing biological data, and for sophisticated mathematical simulations.

The state of the art computational and data storage systems rival that of larger institutions such as Virginia Tech and UC Davis, and have computational power equivalent to over 800 standard desktop computers. A single computer would require over 250 years to perform the analysis that the CRC system performs in just one year.

The CRC has the capacity to store vast amounts of data, the equivalent of over 1 trillion large documents or over 320 million high quality photos. It also provides computing and educational support to all Idaho universities and colleges.

The CRC facility has been funded by the National Science Foundation and the National Institutes of Health. It is an essential part of projects that have brought tens of millions of dollars into Idaho, and it is recognized nationally for its excellence.

Did you know.....

The IBEST CRC has .5 PETABYTES of storage.

If each star represents a byte, that is 7,500 Milky Way Galaxies.

512 TERABYTES

1,024 laptops with average storage size of 512 GB.

19,968 ft

29,029 ft

768,000 CD roms full of data put in cases and stacked equal nearly the height of Everest.

All servers are named after characters from “The Hitchhiker’s Guide to the Galaxy.”
1999 Computational Biology Seminar becomes the Initiative for Bioinformatics and Evolutionary Studies. Original proposal for establishing a new "Initiative for Bioinformatics and Evolutionary Studies" (IBEST) at the University of Idaho submitted to Vice President for Research Jeanne Sharepey by Drs. Wichman and Foster requesting $15,000 in funding support.

2002 COBRE II grant for $10,823,589 awarded to Larry Forney (PI) by NHR/DoD and establishes a "Center for Research on Processes in Evolution." This grant funds four research projects and establishes a Molecular Biology Core, a Bioinformatics Core and a Structural Biology Core. Larry Forney becomes the Director of IBEST.

2004 IBEST begins a collaboration with the Idaho Idea Network of Biomedical Research Excellence (INBRE) program.

2010 IBEST becomes part of The BEACON Center for the Study of Evolution in Action, an NSF Science and Technology Center led by James Foster (PI). BEACON funds 29 IBEST projects with a five-year grant over $2,700,000.

2013 COBRE II Award continues NIH DoD funding of $5,090,846 for the "Center for Research on Processes in Evolution" led by Larry Forney (PI).

2011 IBEST achieves official "Institute" status at the University of Idaho. Larry Forney named as Director.

1990 Meeting of UI faculty with President Zinter on the topic of research and teaching. An impassioned statement on how research enriches teaching by new Computer Science Assistant Professor James Foster leads to a conversation with Biological Sciences Assistant Professor Holly Wichman.

1993 Emeritus Professor of Physics Larry Johnson joins CBS Lunch. Intensive Hot Tub session facilitated by Olson 14 year old Scot led to the first joint research project focused on multidisciplinary studies in Bioinformatics and Evolution – Foster, Wichman and Bennett.

1991 First weekly lunchtime meeting of the Computational Biology Seminar (CBS) to explore common research interests attended by James Foster, Holly Wichman, John Dickerson and Mike Bennett. Book discussion of Jared Diamond's "The Third Chimpanzee" started the tradition of CBS Lunch.


2000 NSF EPS-048-184 Grant "Multidisciplinary Studies in Bioinformatics and Evolution" for $499,994 awarded to James Foster (PI) to establish and maintain an enduring research community for bioinformatics and evolutionary studies at UI. Collaborators included Robert Heckendorn, Scott Harrison, and Axel Klings from Computer Science; Paul Joyce and Steve Koome from Mathematics; Jack Sullivan and Holly Wichman from Biological Sciences. Larry Forney joins the UI as Chair of the Biological Sciences Department.
Cutthroat trout play a fundamental role in their habitat and are used by many scientists as an indicator species for the health of larger ecosystems. The growing recreational fishing industry has been a primary motivator of introducing a popular sportman fish, the rainbow trout, to many lakes, streams, and reservoirs it is not native to. This introduction has placed pressure on the native cutthroat species of trout, causing a noticeable decline in population trends and resulting in its classification as a threatened species. The replacement of cutthroat trout with non-native trout species is not only about the possibility of endangerment and extinction of a species, it also affects whole ecosystems by changing the energy and nutrient flow, causing a potentially negative ripple effect of influence on the health and balance of the larger system the cutthroat trout is a part of.

Dr. Paul Hohenlohe studies the fitness of cutthroat trout in the waterways of Montana, where rainbow trout have been introduced. Over the years, rainbow and cutthroat trout have been crossbreeding, producing a hybrid called a cutbow trout. Past studies have shown these hybrids have a lower fitness level. The fitness of a species is measured by the number of surviving offspring it has and how well it contributes its genes to the next generation, or in other words, a species survivability. Past studies infer that cutbow hybrids are not a real threat to the cutthroat population. However, it has been discovered that despite hybrids lower fitness, there are some parts of the rainbow trout's genetic information that is being favored in the selection process and continues to spread throughout the native cutthroat population, resulting in a decline in native cutthroat numbers.

“What are these genes and why are they behaving like that?”

“What are these genes and why are they behaving like that?”

Dr. Hohenlohe says, “It’s an interesting question,” Dr. Hohenlohe says, “what are these genes and why are they behaving like that?”

These are the questions that form the foundation of his research. By searching for answers, he hopes to learn what it means for the population fitness and dynamics of the native cutthroat trout and their increasingly threatened status.

How does one go about tracking specific genomic material? Dr. Hohenlohe relies on RAD Sequencing and massive computational power. RAD stands for “Restriction site Association DNA”. RAD sequencing allows researchers to sample the genomes of multiple individuals in a population. Massive amounts of data are produced, which
Dr. Hohenlohe then analyzes. Since the data sets can be large (easily hundreds of gigabytes), high-powered computing is necessary to navigate the information.

“When I was in graduate school I ran a bunch of computer simulations and it was just a single, big computer,” Dr. Hohenlohe says, “I would have to start it running and come back a week later, then start a new one running.” Repeating this process over and over again, Dr. Hohenlohe would process his data, often waiting months for it all to be completed before he could even begin the analysis portion.

Today, with the availability of the IBEST Computational Resources Core, he is able to run many simulations at once, greatly diminishing the time it takes to get his data. After running all his simulations, Dr. Hohenlohe then sorts through the data produced, analyzing and summarizing it for the publication of papers in scientific journals, presentations and advancement of research.

“With genomic technology, people are actually uncovering the structure of real genetic networks in actual organisms, in humans, and other creatures,” he says. “They are understanding how different genes affect each other and how they are interconnected.”

With this information, Dr. Hohenlohe seeks to better understand genetic variation in natural populations, specifically those of the cutthroat trout. In evolutionary terms, this means he can gain a better understanding of how the trout populations adapt and evolve and how genetic variation is structured in these populations and as a result, aid in preserving the populations of native cutthroat trout that are threatened.

© Dreamstime Inc.

DID YOU KNOW?

Cutthroat are the state fish of:

WY ID MT

There are 13 species of Cutthroat trout. Most are named for the body of water they live in.

Most species of Cutthroat trout are either threatened or endangered.

Unlike Salmon, Cutthroat trout can spawn multiple times before they die.

It is a common misconception that Cutbow are sterile.

© Dreamstime Inc.
Tyler Hether scrolls rapidly through the various items open on his computer screen, stopping at the document he wants as he describes his work running quantitative genetics simulations.

“If you know something about the genetic make up of the population and something about the strength of selection,” he says, “then you can predict the next generation.”

On his screen, pairs of red and blue dots rest above graphs, different types of lines run between them, representing their associations. With each graph a complex mathematical equation explains it all. I sigh. This is going to be a tough conversation to follow.

Tyler is a graduate student studying to be an evolutionary biologist under Dr. Paul Hohenlohe at the University of Idaho. The word biologist conjures mental images of Charles Darwin’s bird studies or high school frog dissections, but Tyler spends much of his time in front of a computer screen dealing with complex mathematical equations and analyzing data in a multi-planar, multi-axis theoretical universe. Sounds complicated because it is.

Being exposed to the work of biologists on a day-to-day basis has made me realize that often the center of their research world is technology.

“We are interested in how genetic networks in general can effect phenotypic variation,” Tyler says. I give him a blank look.

“Let’s back up a bit here. Genetic networks? Phenotypes? I was digging deep for any sort of information gleaned from freshman biology years ago, coming up mostly empty handed and grateful for Tyler’s patience in explaining “the basics” to me. A genetic network, or gene regulatory network (GRN), is a collection of DNA (our genetic information) segments in a cell that interact with each other or other substances in the cell. These interactions distribute genetic information and determine how and to what level this information is expressed. The outcomes of the interactions and expressions are called phenotypes or more broadly, traits (such as blue eyes or red hair). Some genes in the network have more of an effect on certain outcomes then others.

Along with the variety of genes contributing, you also have environmental factors affecting the selection and expression of these traits. All these things come together to create a complicated and multi-faceted realm in which to navigate.

“There’s this idea,” says Tyler, “that you can predict the response to some selection, whether it be natural or artificial. For example, if a breeder wants to get higher percent milk fat in a cow they might do something similar to this.”

Tyler’s simulations model interactions between genes, allowing him to see different results that occur depending on how he manipulates the network. He explores how wiring the networks differently causes changes in the gene interactions. In a simulation, populations are formed, they mate and produce offspring, and they die, all in the computer. In order to get good information from these types of simulations, he has to run lots of replicates where he applies different parameters to the networks. This makes it computationally intense because Tyler ends up doing iteration upon iteration of these simulations, each producing different variations – such as migration between populations, rate of evolution and other factors – to be studied and analyzed.

“So the result is essentially that genetic regulatory networks, even really simple ones,” says Dr. Hohenlohe, who oversees Tyler’s research, “can have interesting effects on patterns of variation in populations and you can see those effects on how populations adapt through natural selection.”

Of course, when Dr. Hohenlohe was in school, the computational resources Tyler now accesses did not exist. The computing resources available to biologists today are much more powerful than when Dr. Hohenlohe was a student, allowing for more complex simulations and data analysis.

““If you had knowledge of the networks, then you could form decisions on management for an endangered species.”

TYLER HETHER
A Ph.D. student in the Hohenlohe Lab, Tyler’s research focuses on merging concepts from ecology, physiology, mathematics, computer science and molecular biology in an attempt to elucidate how diversity arises.
Tyler is able to explore and test common assumptions about quantitative genetics that have been generally accepted over the last decade. He must have noticed my pained expression while trying to connect the dots, because he explains that he is also contributing to the growing need for biologists to understand genetic networks and their interactions. Despite being theoretical in nature, Tyler contributes to work that is used often in important applications, especially in the agriculture realm, people wanting to increase the yield in their wheat or produce cattle to a certain level of meat quality.

"If you had knowledge of the networks then you could form decisions on management for an endangered species," Tyler says, "or populations changing from different land use, such as larger, regional and global climate change."

Even within these simplified networks, Tyler Hether’s work demonstrates the important role of genetic networks when it comes to adaptation. The wiring of certain networks makes it difficult for a population to adapt and change to a new environment, while others aid in the change. As more and more people pay attention to and comprehend how genes are networked together, profound impacts on how populations adapt and our role in that adaptation will continue to be understood.
Hundreds of years ago people looked to their genealogical lines to see if they were related to nobility. Today it has become an over 1.6 billion dollar industry of curiosity and history. Now scientist Dr. James Rosindell has taken genealogy to a whole new level, asking not whom you are related to, but what. While working for Dr. Harmon as a postdoctoral student, Dr. Rosindell decided to apply his love of fractals to create One Zoom, an online, interactive map allowing anyone to navigate the complex and extensive phylogenetic tree of life. Phylogenetics refers to the study of branching diagrams called trees that represent the relationships among various biological species or entities. One Zoom compiles all these trees into one giant tree, the tree of life.

“One Zoom took some serious programming,” Says Dr. Harmon, “When James would come in and tell me what he was doing, I thought whoa, keep doing that.”

The goal of One Zoom is to raise awareness about the diversity of life, its history and threats of extinction. Every living organism on planet earth, from bacteria to humans, chimps, fish, reptiles, and amphibians – over 2 million known species are mapped on the One Zoom tree. Most of the tree represents things we cannot see with the naked eye. Visible creatures and organisms only make up a small twig.

“It is important to understand where we came from, our history, the history of species on the earth, how we are all related to each other;” says Dr. Harmon, “It’s the story of life on earth, which I think is important to know.”

There are many practical implications to understanding relationships among species. For example, Dr. Harmon uses the tree to figure out how often new species form. This is important because if we are wiping species out at a rapid rate, knowing how often new species form will tell us just how dire the consequences our human caused extinctions might be. If it takes a million years for an extinct species to be replaced by others, then we can gain a sense of how devastating our actions will be for future life on the tree, such as fish known as cichlids in east African lakes. These cichlids are showing some of the fastest vertebrate speciation rates in the whole tree of life. This pattern may have gone unnoticed without viewing the cichlids in the broader context of the tree. Now researchers are investigating why there is such rapid forming of new species in cichlids and what that might mean in regards to other parts of the tree.

“All of these things require really heavy-duty computational tools that we wouldn’t be able to do without the cluster here at Idaho,” says Dr. Harmon, “These trees are actually too big to print out, they are too big for the human brain to comprehend, so we use the computational cluster to do the processing and statistical analysis that tells us what parts of the tree are interesting and unusual, to characterize their general properties and to measure them.”

“It is important to understand where we came from.”

**LUKE HARMON**

Research in the Harmon Lab investigates ecological and evolutionary aspects of adaptive radiations. They explore the causes and effects of both speciation and trait change and how species interactions shape macroevolution.

Department of Biological Sciences
University of Idaho
208.885.0346
lukeh@uidaho.edu
webpages.uidaho.edu/~lukeh/people.html
Each species’ information is obtained primarily through genetic sequencing. If you have a tree with 50,000 species on it, each with a large amount of genetic data, the files for these trees become massive. They would easily crash your laptop or desktop computer, a problem Drs. Rosindell and Harmon had to wrestle with before utilizing the IBEST Computational Resources Core.

In the Inland Northwest there is a species called the tailed frog. It has its own branch on the tree of life. Only males of the species have tails, which they use to internally fertilize the females. This is really unique in amphibians, which typically fertilize eggs externally, after females have expelled them. When researchers started looking at the tailed frog in the context of the whole tree of life, they discovered there are only two species of this frog in the whole world, one in the Inland Northwest and the other on the west coast around Seattle. These two species are connected to the rest of the frogs on a branch that spans 200 million years or so. Not having any other relatives anywhere near them on the tree of life makes them absolutely and completely unique and relatively unchanged over millions of years.

“I think we’ve learned over the last 20 years in biology that it matters species are related to one another,” says Dr. Harmon.

As a rough guide, the One Zoom comprehensive tree of life helps Dr. Harmon and other researchers like him navigate the complicated realm of how species form and evolve, how we are all interconnected and what that might mean for future generations.
Then and Now

The IBEST Computational Resources Core came from humble beginnings. Drs. James Foster and Robert Heckendorn built the first computers with a class of Computer Science students in the basement of the Janssen Engineering Building for a total of $35,000. Funding was provided by NSF-EPSCoR. They jokingly called the computer cluster LOBOS, “Lots of boxes on shelves” and later, The Cube. (In the words of the Borg, “You will be assimilated!”)

After burning out two window AC units and suffering several system failures due to overheating while housed in the Janssen Engineering Buildings, Beowulf was moved to the basement of the McClure building. The University paid for a comprehensive renovation to accommodate the needs of such a large computer system, creating a state of the art facility. Here it began its transition into the large computing cluster it is today supported by funding from Idaho INBRE and the IBEST COBRE.

Across
3. Dr. Paul Hohenlohe studies the fitness of __________ trout.
6. Along with gene contribution, __________ factors affect the selection and expression of traits.
9. One __________ is an interactive map allowing anyone to navigate the phylogenetic tree of life.
10. The hybrid of a rainbow trout and cutthroat trout is called a __________.
11. The IBEST Genomics Resources Core produces twenty _________ reads per full sequencing cycle.

Down
1. __________ trees represent the relationships among various biological species or entities.
2. RAD Sequencing allows researchers to sample the __________ of multiple individuals in a population.
4. The IBEST Computational Resources Core has 512 __________ of storage.
5. Birds are a lineage of the __________.
7. A genetic __________, is a collection of DNA segments in a cell which interact with each other indirectly and with other substances in the cell.
8. __________, or “lots of boxes on shelves,” was the nickname of the NSF EPSCOR computer cluster that grew into the now IBEST Computational Resources Core.
Institute for Bioinformatics and Evolutionary Studies
875 Perimeter Drive MS 3051
Moscow, ID 83844
Phone 208.885.9076
Fax 208.885.5003
ibest@uidaho.edu
ibest.uidaho.edu