international BARCODE OF LIFE

BARCODE BULLETIN

Vol. 1, No. 2 - June 2010

The newsletter of the International Barcode of Life project

Big funding boost for iBOL

New grants raise total investment to $80m

The International Barcode of Life Secretariat has announced major new funding for the world's largest biodiversity genomics project. iBOL Scientific Director Paul Hebert said that four Canadian agencies have made new commitments to iBOL totaling $35 million, raising total investments by these funders to $80 million (all figures in Canadian dollars).

Building on an earlier $5 million award, the Ontario Ministry of Research and Innovation has committed another $8.1 million over the next five years to support BOLD, the informatics platform for DNA barcode data, and further expand the barcode reference library.

Genome Canada is also extending its support for the iBOL project for another year with a second funding installment of $4.6 million. This follows the $2 million provided by Genome Canada in 2009-10 to initiate the project through its International Consortium Initiative Program.

Continued on page 10

iBOL releases 150K barcode records to BOLD, GenBank

The first large blocks of iBOL barcode sequence data have been released to NCBI GenBank and are publicly accessible on the Internet from the Barcode of Life Data (BOLD) Systems web site (www.boldsystems.org/views/datarelease.php).

Continued on page 10

BOLD developer wins Ebbe Nielsen Prize

Sujeevan Ratnasingham, the bioinformatics expert who developed the Barcode of Life Data Systems (BOLD) platform for the global DNA barcoding community, has won this year's Ebbe Nielsen Prize from the Global Biodiversity Information Facility (GBIF).

Continued on page 11

Index:

Making every species count

Section of DNA barcode for *Lasiurus borealis* - story on page 8

New iBOL Executive Director to focus on developing international partnerships

The interim chair of the International Barcode of Life Consortium (iBOL) Board of Directors, Dr. Christian Burks, has announced the appointment, beginning July 2010, of Dr. Peter Freeman as the new Executive Director of iBOL.

Dr. Freeman will be responsible for the overall operations of the iBOL Secretariat, including keeping stakeholders apprised of progress and building the participation of key international funders and research institutions. He will work closely with and report to iBOL's Scientific Director, Dr. Paul Hebert.

Originally from Ireland, Dr. Freeman arrives at iBOL with a record of successfully coordinating large multi-institutional, international research projects, networks and consortia in genomics, proteomics, stem cell research and population health.

He is currently based at the University of Calgary where he serves as Executive Director of the Calgary Institute for Population and Public Health. From 2004-2008, he was based at the University of Edinburgh as Project Manager of EuroStemCell, a European Commission Framework VI project integrating the efforts of 27 leading stem cell research groups in 13 academic centres and three biotech companies across eight countries. Prior to that, he served as Executive Director for the Alberta Network for Proteomics Innovation.

Dr. Freeman also has extensive private sector experience in R&D and operational executive roles in the international malting and brewing industry.

"I am delighted to be joining the International Barcode of Life Consortium," said Dr. Freeman. "This is a rare opportunity to contribute to a project with the vision, scope and resources to dramatically change the scientific landscape and inform diverse fields of human endeavor."

"Dr. Freeman's experience coordinating among multi-institutional and multinational research consortium partners will be particularly valuable to iBOL," said Dr. Burks (CEO and President, Ontario Genomics Institute). "His is an important and timely addition as iBOL moves towards its formal initiation in September 2010." ✤

New web site, new address

Change your bookmark. The International Barcode of Life web site has a new web address – www.iBOL.org, thanks to iBOL's outreach partners at the Consortium for the Barcode of Life (CBOL) who arranged the transfer of the url from its previous owner.

The change of address coincides with a radically changed look for iBOL's web presence. The web site redesign, undertaken in partnership with web consultants Threestone Studios, features dramatic pictures, up-to-date information about iBOL and its international network of collaborators, barcoding news and publications and links to iBOL content on social networking sites such as YouTube and Twitter.

Check it out and let us know what you think. Send your suggestions and comments to feedback@iBOLproject.org ✤

New building will be scientific hub for iBOL

Construction of the new C$18 million Centre for Biodiversity Genomics will start in August 2010 on the campus of the University of Guelph. Work is scheduled for completion in about 15 months.

This 3,500 square metre research facility will serve as the scientific and administrative hub of the International Barcode of Life project (iBOL). The building will house the iBOL Secretariat, server rooms and other bioinformatics infrastructure, plant and animal collections, offices for scientists, post-doc and grad students and support areas. ✤

$2m pilot project will promote barcoding in five developing countries

The International Development Research Centre (IDRC) has approved a C$2.16 million grant for a three-year pilot project to support the participation of researchers from five developing countries in the International Barcode of Life project.

The funds will also be used to promote the use of DNA barcoding to resolve serious problems impacting human health and wellbeing in the five countries.

The five countries involved are Argentina, Costa Rica, Kenya, Peru and South Africa.

IDRC Project Manager Greg Singer said that the grant will foster participation in iBOL by funding the training of postdoctoral fellows who will then move rapidly into the scientific workforce.

The funds will also be used to broaden exposure to barcode technology by delivering training sessions in Argentina and South Africa and by launching projects that validate the practical implications of barcode technology to:

- Identify agricultural and forestry pests, improving productivity.
- Identify saplings, helping to provide a better sense of forest dynamics.
- Identify arthropod species that are important vectors of human and livestock diseases.
- Aid marketplace surveillance of herbal medicines, bushmeat, seafood and other products.
- Identify species important in crop pollination.
- Identify endangered species intercepted at airports and other border points.

"Researchers from neighboring nations will be invited to training sessions, helping to energize understanding of DNA barcoding at a regional level," said Singer. "We also plan research activities that will build global understanding of the contributions that DNA barcoding can make to Access and Benefits Sharing under the Convention on Biological Diversity."

Singer said it is hoped that the collaborative research programs built with the IDRC grant will catalyze investments in support of DNA barcode programs in other developing nations by funders from other first world countries. ✤

What's the big idea?

According to one of the world's most famous magazines, the big idea is DNA barcoding. In the May edition of National Geographic, there is a three-page pictorial essay called "Scanning Life" in the regular feature section called The Big Idea. The story quotes iBOL Scientific Director Paul Hebert predicting that barcoding technology will follow the path of GPS and someone will invent a handheld barcoder. "I can imagine every kid getting on of these in his or her Christmas stocking," Hebert told the magazine. ✤


THE BIG IDEA | BIODIVERSITY

Life would be simpler if all species wore name tags. And in a way, they do. Turns out they've got a bar code in their DNA.

Scanning Life

IF YOU TURN ON A LIGHT AT NIGHT in the mountains of Papua New Guinea, says Paul Hebert, you will collect some 2,000 species of moth. Moving up the mountain a bit will net you a different but equally daunting crowd. As a young postdoc in the 1970s, Hebert, now an evolutionary biologist at the University of Guelph in Ontario, spent five years trying to make sense of that fluttering confusion, before finally deciding it was beyond his or any human's capacity. For two decades after that he retreated to water fleas, of which there are only 200 species. Then in 2003 he did something new. In a paper that year he began by describing the diversity of life as a "harsh burden" for biologists, and proceeded to suggest some relief: Every species on Earth could be assigned a simple DNA bar code, Hebert wrote, so it would be easy to tell them apart.

The bar code Hebert suggested is part of a gene called *CO1*, which helps produce the energy-carrying molecule ATP. *CO1* is so essential that every multicellular *(Continued on next page)*

30 NATIONAL GEOGRAPHIC • MAY 2010

Barcoding Polar Life

A scientist on the Arctic island of Spitsberg collects specimens for DNA barcoding.

The Polar Barcode of Life initiative (PolarBOL) was established at the first iBOL meeting in 2007 with the primary goal of providing the most efficient and accurate tool for mapping and monitoring polar biodiversity. PolarBOL collaborators coordinate barcoding efforts in ongoing bio-inventory projects and communicate through their website (http://www.ibolproject.org/polar/), newsletter and e-mail listserve.

DNA barcoding is an integral part of many research activities in both the Arctic and Antarctic and there are now 41,660 barcode records from specimens sampled in polar regions registered in Barcode of Life Data Systems (BOLD).

In the Arctic, marine, terrestrial and freshwater biodiversity is being investigated while the Antarctic focus has been in the terrestrial and marine realms. The largest single-project barcode contributors are the Polar Research Observatories for Biodiversity and the Environment (PROBE) documenting diversity in Churchill, Manitoba; and the Census of Antarctic Marine Life (CAML) focusing on the plethora of Antarctic marine animal taxa. ✤

PolarBOL contacts:

Antarctic -
Peter Smith (p.smith@niwa.co.nz),
Ian Hogg (hogg@waikato.ac.nz)

Arctic -
Sarah Adamowicz (sadamowi@uoguelph.ca),
Torbjørn Ekrem (torbjorn.ekrem@vm.ntnu.no)

Pakistan launches major insect project

Scientists in Pakistan are hoping to barcode more than 4,000 insect species over the next 16 months in a project that is expected to uncover a large number of previously unidentified and undescribed insects.

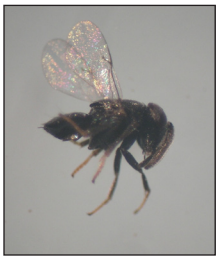
The PKR 1.74 million (US$207,000) project is being funded by the Higher Education Commission of Pakistan and will run for 18 months from April 2010.

Insect collection, taxonomy and preparation of samples will be undertaken by the National Institute for Biotechnology and Genetic Engineering (NIBGE) in Faisalbad under the direction of Dr. Muhammad Ashfaq. Universities and insect museums in Pakistan will provide taxonomic expertise while DNA sequencing, databasing and analysis will be performed at the Biodiversity Institute of Ontario in Guelph.

"A significant number of insect species of agricultural and economic importance will be barcoded," said Dr. Ashfaq. "Our major focus will be barcoding agricultural pests, insect pollinators and insect predators and parasitoids. If resources and time permit we will also include disease vectors."

In southern Pakistan, the project team will focus on the rich insect diversity found in the agricultural areas of Punjab and Sindh provinces. In the north, they will work in the foothills of the Himalayas in North Western Frontier Province, Kashmir and Punjab. With many trees and fruit orchards and lower exposure to pesticides than the plains provinces, these areas are expected to yield a rich haul of species, particularly pollinators and insect predators and parasitoids. ✤

Left: Dr. Muhammad Ashfaq will lead the project to barcode insects of agricultural and economic importance in Pakistan. Among them will be the treehopper and plant pest *Oxyrhachis taranda* (below), *Drosicha mangiferae* (bottom right), a problem species for mango growers, and (bottom left) *Aenasius bambawalei*, a parasitoid of the cotton mealybug.



Barcoding takes wing in Argentina

Argentina is now South America's first iBOL Regional Node. Pablo Tubaro says DNA barcoding is flying high in the land of silver.

Isabel Gómez Urquizo (left), the bird curator with the Colección Boliviana de Fauna (Animal Collection of Bolivia) at the Universidad Mayor de San Andrés in La Paz, and her team inspect specimens collected during an expedition to the Andean forest.

Argentina's involvement in DNA barcoding dates back to the creation of the Consortium for the Barcode of Life in 2004. But it wasn't until the following year, during the first International Barcode of Life Conference in London, that Argentina became an active participant in the first two global barcoding campaigns – the All Birds Barcoding Initiative (ABBI) and the Fish Barcode of Life Initiative (FishBOL). So far, the FishBOL campaign in Argentina has focused on marine species, mainly skates and species of commercial importance.

But Argentina's barcoding success story involves the project focused on birds, which was launched in 2006 as a collaboration between the National Science Museum of Argentina (MACN) and the Biodiversity Institute of Ontario/Canadian Centre for DNA Barcoding (BIO/CCDB).

The project received an early boost from a Richard Lounsbery Foundation grant that paid for the collection of specimens, student training at CCDB and assembling a DNA laboratory at MACN. A second Lounsbery grant is supporting the current phase of the project, which includes barcoding not only the birds of Argentina but also those of Bolivia and Peru.

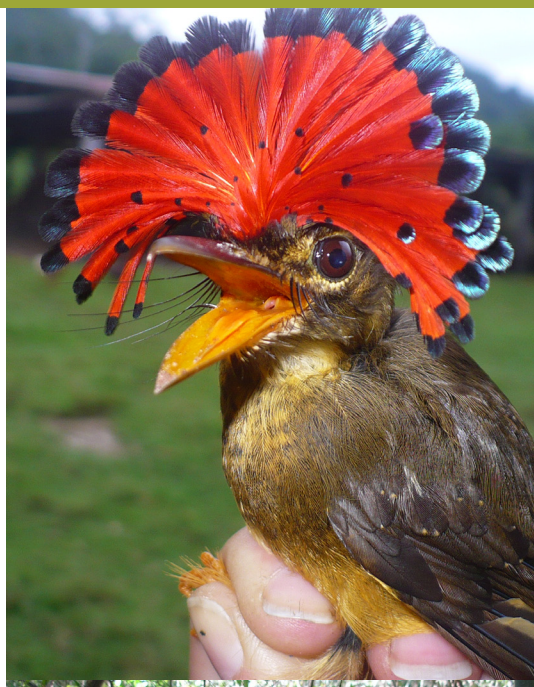
So far the project has produced almost 3,000 DNA barcodes from around 700 species of neotropical birds and has generated the second largest frozen tissue collection of birds in Latin America. It has also trained four PhD and post-doctoral students at BIO/CCDB. (Two other National Research Council [CONICET] researchers have also been trained at BIO/CCDB, one of them awarded with a grant from the Emerging Leaders of the Americas Program – ELAP.)

Remarkably, all this has been achieved in just five years – MACN had virtually no tissues before 2005 – and the collection is now growing at 1,000-1,500 new specimens a year.

The success of Argentina's involvement in the ABBI and FishBOL initiatives convinced CONICET that it should support the iBOL project. After a planning meeting at Guelph in 2007, CONICET representatives signed a Memorandum of Understanding that made Argentina a National Node of iBOL.

They also created a national committee and a special fund – the iBOL Argentina Fund, which also includes funds from private organizations such as Fundación Williams – to support the collection of barcode compliant materials. The fund currently pays for about 40 different collecting projects, including vertebrates, invertebrates, plants and fungi.

In 2009, CONICET's vice-president of



Cecilia Kopuchian, who will be one of the four post-docs working on the IDRC project in Argentina, takes a bird from a mist net during a field trip to Chaco National Park. She is framed by two of the beautiful specimens captured in Bolivia during the last collecting campaign for the bird barcoding project -- *Onychorhynchus coronatus* (above) and *Pipra chloromeros* (below).

technology, Dr. Faustino Siñeriz, joined the iBOL Board of Directors and CONICET approved Argentina's change of status to Regional Node of iBOL.

The upgrade will involve a series of substantial investments. These will include increasing the iBOL Argentina Fund, assembling several molecular labs dedicated to barcoding, employing technical staff and digitizing biological collections. Support will also come from Leading Labs Training Workshops and the oceanographic ship Puerto Deseado for sampling marine biodiversity.

Pictured during an expedition to Argentina's Iguazú National Park (from left): Pablo Tubaro, Argentina's representative on the iBOL Scientific Steering, Pilar Benites, Ana Barreira and Kevin Kerr.

Importantly, Argentina is one of five countries – the others are South Africa (Regional Node), Costa Rica, Peru and Kenya (National Nodes) – that will benefit from the IDRC grant entitled "Engaging Developing Nations in the International Barcode of Life Project". This collaboration with BIO/CCDB will support barcoding over the next three years by training young scientists, building the barcode reference libraries of species with the greatest socio-economic and environmental importance to each nation and developing and implementing an Access and Benefit Sharing policy for iBOL. ✤

Pablo Tubaro is head of ornithology and vice director of Argentina's National Science Museum. He is also Argentina's lead representative on the iBOL Scientific Steering Committee.

A catalyst for action in Kenya

iBOL has only two nodes in sub-Saharan Africa and both will benefit from the IDRC project. Helida Oyieke explains how that will play out in Kenya.

Kenya's scientific community really began to take notice of DNA barcoding and its utility as a taxonomic tool after I (as a member of the CBOL executive committee) convened the first East African Regional DNA Barcoding workshop in Nairobi in 2006.

As a result of that meeting, a steering committee was established to promote barcoding campaigns and develop a regional pilot project. With support from CBOL and BioNET, the committee came up with three concepts on priority projects for DNA barcoding. One of them, Barcoding Cyprinids, was developed into a full proposal by the National Museums of Kenya (NMK), the Kenya Marine and Fisheries Research Institute and the Nairobi-based International Centre for Insect Physiology and Ecology (ICIPE), in partnership with other regional institutions.

Dr. Dan Masiga, who is based at ICIPE and was a member of the CBOL Scientific Advisory Committee at that time, was instrumental in providing technical guidance on DNA barcoding discussions in the country.

The new funding from IDRC will help to jump-start long-term and large-scale DNA barcoding activities in Kenya. The Kenyan component of the Cyprinid project, which has received no funding to date, will be implemented along with other priority taxa currently being identified.

Laboratory facilities at the NMK, which will serve as the centre for the Kenya node, will be upgraded and training will be provided to a few key researchers who will lead implementation of the project.

Members of Kenya's DNA barcoding steering committee take a break during a recent meeting.

Given that programmes are currently under way to transform NMK into the Regional Centre of Excellence for Biodiversity Informatics, it is expected that the IDRC-funded iBOL activities in the Kenya Node will catalyze and promote DNA barcoding processes in the whole region. ✤

Helida Oyieke is Director of Research and Collection with the National Museums of Kenya. She is also Kenya's alternate representative on the iBOL Scientific Steering Commmittee.

No longer in the dark

By Elizabeth Clare

THIS PAGE
Above: A hearty meal of moth for this *Myotis septentrionalis*.
(photo by M. Brock Fenton)

Far right: Bat guano under the microscope.
(iBOL photo)

Right: A baby bat gets some nourishment.

FACING PAGE
Top: A red bat prepares to excrete for science.

Bottom: Young red bats in the wild.
(Photos by Erin Fraser except where noted)

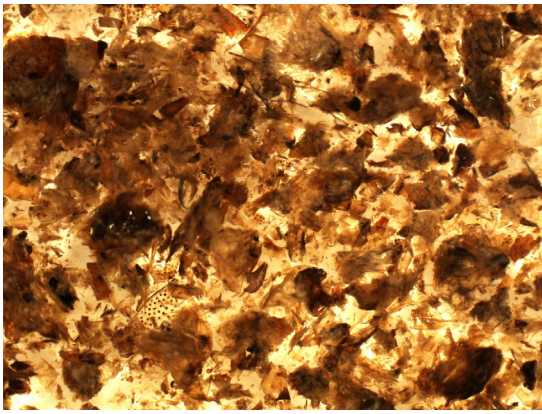
The rapid growth of the DNA barcode database continues to open up promising new areas for research, including the study of species' interactions in the wild. A recent article in the journal Molecular Ecology* describes how we used DNA barcoding to identify insect fragments found in the guano of the eastern red bat (*Lasiurus borealis*) linking this predator to its prey.

We know that many bats eat insects, but until now we have known little about the particular insect species that make up their diet. That's because bat feeding behaviour is almost impossible to observe directly. They hunt at night, fly fast and their insect prey is too small for a human observer to track.

For many years, researchers have collected and examined bat guano to reconstruct diet, carefully teasing apart the pellets of faeces under a microscope and examining the partially digested prey. But because prey items are often small and soft and bats tend to chew their food thoroughly, most of these identifications are not precise. We might know that the bats ate beetles or moths, but more detail is rarely possible.

However, insect fragments in the bat guano also contain tiny amounts of DNA. We were able to analyse those traces and identify exactly what was consumed.

Ph.D student Erin Fraser, of the University of Western Ontario (supervised by Dr. Brock Fenton), used mist nets to capture the bats and collect the guano they produced after a short time in captivity. Back in the lab we dissected the guano pellets under a microscope and extracted DNA from the tiny insect fragments (legs, wings, exoskeletal segments, etc.). Following PCR and sequencing of the barcode region, I used the Barcode of Life Data Systems (BOLD) database to match these unknown sequences to reference sequences and determine what the insect prey were.



Since the process of populating the BOLD database is ongoing, there weren't exact matches for every insect we encountered but overall, the survey was very successful. The guano of 56 bats yielded almost 800 separate barcode sequences, from which 127 distinct species of prey could be identified.

When we put names on them, we made a number of interesting discoveries. First, and least surprisingly, we showed that this bat species eats mainly moths, which is consistent with what was learned from morphological studies of guano. More intriguing were some of the specific kinds of moth which showed up on the list. These included a number of significant pest species, such as the gypsy moth (*Lymantria dispar*), tent caterpillars (genus *Malacosoma*) and coneworms (genus *Dioryctria*), as well as orchard and garden pests such as *Cydia*, *Acrobasis* and *Noctua pronuba*.

Also notable were the groups the bats do not eat. *Arctiids* are a family of moths

about what bats eat

which are abundant in the areas where these bats were hunting but were nearly absent from our DNA-based dietary survey. This is interesting because many moths in this family have multiple defensive adaptations which apparently help them avoid predation by bats, including the ability to produce high-frequency sounds in response to echolocation which may confuse the bats.

Besides moths, we found a few other groups not known as prey for eastern red bats, including Ephemeroptera (mayflies), Neuroptera (net-wings) and Trichoptera (caddisflies). We even found the occasional spider.

Now that we have shown the basic effectiveness of this technique for dietary analysis in bats, we are extending the survey to other insectivorous bats. DNA barcoding will continue to expand the list of identified species with recorded COI sequences, which will improve the accuracy and detail of matches we make from dietary samples. The same methods will also be applicable to other groups such as fruit eaters, pollinators, carnivores and sanguivores.

The ability to use DNA to identify organisms from fragmentary remains opens up many new avenues for research in ecology. It greatly increases the precision with which we are able to determine the structure of food webs – not simply saying that "bats eat insects" but that "these bats eat these particular insects". This type of information will be extremely valuable for unravelling species' interactions which are basic to our understanding of how ecosystems function. ✤

• • • • • • • • • • • • • • • • • • • •

Text adapted from an article by Robin Floyd and Elizabeth Clare published in the winter 2009 edition of BATS *magazine*.

* Clare EL, Fraser EE, Braid HE, Fenton MB, Hebert PDN. 2009. Species on the menu of a generalist predator, the eastern red bat (Lasiurus borealis): using a molecular approach to detect arthropod prey. Molecular Ecology. 18, 2532-2542.

Researchers Profiles.

- Elizbaeth Clare is a PhD candidate in the Department of Integrative Biology at the University of Guelph
- Erin Fraser is a PhD candidate in the Department of Biology at the University of Western Ontario
- Heather Braid has just completed her undergraduate degree at the University of Guelph.
- M. Brock Fenton is a professor of Biology at the University of Western Ontario.
- Paul Hebert is a Canada Research Chair in Molecular Biology at the University of Guelph and Scientific Director of iBOL.

Big funding boost for iBOL -

Continued from page 1

Liz Sandals, member of Provincial Parliament for the riding of Guelph, announces the Ontario Government's $8.1 million award to iBOL during an event at the Biodiversity Institute of Ontario.

Dr. Christian Burks, President and CEO of the Ontario Genomics Institute and interim chair of the iBOL Board of Directors, said: "The many international partnerships that comprise iBOL, and which are absolutely key to its success, will benefit tremendously from these funding commitments. We anticipate that the funded research and resources at the Canadian node for iBOL will be greatly leveraged by iBOL nodes in other countries that have already committed funds and research efforts to the project, and will help other potential international partners finalize their commitments to participate in iBOL."

Dr. Hebert said: "As the iBOL project approaches the end of its one-year preparatory phase, we look forward to official activation of this global undertaking with renewed optimism and determination. We are grateful for the vision shown by our federal and provincial governments and by their science funding agencies. Their leadership is enabling an initiative that will transform humanity's relationship with other living organisms."

Dr. Hebert announced that groundbreaking for the new Centre for Biodiversity Genomics will take place this summer at the University of Guelph. This $18 million facility, funded by the Canada Foundation for Innovation and the Ontario Ministry of Research and Innovation will house the iBOL Secretariat and key infrastructure needed to support iBOL research. The new centre, which will be the scientific hub for iBOL, is scheduled for completion in late 2011. (Story page 3)

Dr. Hebert also welcomed significant contributions from the Natural Sciences and Engineering Research Council of Canada, which awarded $1.2 million to support new DNA barcoding research programs, and from Canada's International Development Research Centre (IDRC), which has provided $2.2 million to enable researchers in five developing countries – Argentina, Costa Rica, Kenya, Peru and South Africa – to play key roles in iBOL. (Story page 4)

Dr. Faustino Siñeriz, Vice-President of Argentina's National Council of Scientific and Technical Research, welcomed the additional support provided by IDRC, noting that it reinforced his organization's recent decision to upgrade Argentina's participation in iBOL to a Regional Node.

Paul Skelton, Director of the South African Institute for Aquatic Biodiversity, said that the IDRC funding would make African biodiversity a much more significant part of the iBOL research program. "It will provide a major incentive for South African institutions to meet the requirements of Regional Node participation in iBOL while giving organizations and researchers from a wide range of African nations the opportunity to contribute to and benefit from IBOL," Dr. Skelton said. ✤

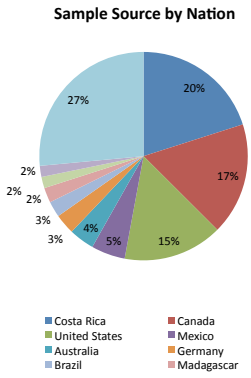
iBOL Releases 150K barcode records -

Continued from page 1

The data were released in two stages:

- January 28, 2010 – 102,772 sequences for 101,124 specimens analyzed at the Canadian Centre for DNA Barcoding during the first two quarters of the iBOL pre-activation period (July1- December 31, 2009).
- May 28, 2010 – 54,415 sequences from 53,889 specimens, for the third quarter of the iBOL pre-activation period (January 1- March 31).

The specimens derived from more than 350 institutions in 192 countries. Most samples came from Costa Rica (30,876) followed by Canada (26,785) and the United States (23,739). Filling out the top 10 were Mexico, Australia, Germany, Brazil, Madagascar, Antarctica and Indonesia.



Barcode data generated by iBOL will be released quarterly with each barcode record seeing release in two stages. The initial stage will include barcode sequence(s), trace files, high-level (ordinal) taxonomic assignment, GPS co-ordinates and the country of collection. The second stage will include a more precise taxonomic assignment and ancillary data, such as images, for each specimen. ✤

Ebbe Nielsen Prize -

Continued from page 1

Ratnasingham is Informatics Director for the Biodiversity Institute of Ontario at the University of Guelph and informatics lead for the International Barcode of Life project. He is the first Canadian winner of the €30,000 Ebbe Nielsen Prize, named for the late Danish entomologist who helped to create the Global Biodiversity Information Facility.

Ratnasingham's development of the BOLD system "is a major and innovative landmark in bringing genomic data on biodiversity to research and research applications for science and society," said Leonard Krishtalka, chair of the GBIF science committee.

Established in 2001, the prize is given to a promising, early-career researcher using biosystematics and biodiversity informatics in novel ways.

"The prize acknowledges the value of genetic data to biodiversity science and recognizes the important work that Sujeevan and his colleagues have been doing," said iBOL Scientific Director Paul Hebert.

Ratnasingham has overseen the growth of BOLD into a system that combines barcode data with images and other information about species' genetic and morphological traits, geographical data and taxonomy.

"The focus of my work is developing systems that allow many researchers to work together and share information on a global scale," says Ratnasingham.

Scientists use the system to enter, share and analyze information about hundreds of different barcoding projects around the world and to contribute to a barcode library that now contains hundreds of thousands of specimen barcodes from 70,000 species. "It's impossible to work with this volume and diversity of data without novel computing platforms," says Ratnasingham. ✤

A full finish with just a hint of worminess

Researchers at the Biodiversity Institute of Ontario have discovered that even if you don't swallow the worm at the bottom of a bottle of the Mexican liquor mescal, you haven't avoided the worm's DNA. They've determined that the distilled alcoholic beverage itself contains the DNA of the agave moth caterpillar – the famous "worm" that most people avoid consuming.

The mescal worm test was part of a study to test the hypothesis that DNA from a preserved specimen can leak into its preservative liquid. When they tested a sample of liquid from a bottle of mescal, the researchers found that the liquor contained DNA, which they amplified and sequenced to obtain a DNA barcode.

Comparing the sample to thousands of records of Lepidoptera DNA barcodes stored in the Barcode of Life Data Systems (BOLD) confirmed that the mescal liquid contained DNA related to the family Cossidae, of which the agave moth is a member.

"This is a surprising result," said research team member Mehrdad Hajibabaei, an assistant professor at the University of Guelph and co-lead of iBOL Working Group 4.1 (Environmental Barcoding).

"Mescal contains only 40 percent ethanol and potentially many impurities that can degrade DNA," he said. "Showing that the DNA of a preserved specimen can be obtained from the preservative liquid introduces a range of important possibilities.

"We can develop inexpensive, high-throughput, field-friendly and non-invasive genetic analysis protocols for situations where the original tissue cannot be touched or when there is simply no sample left for analysis."

The scientists also successfully identified other preserved specimens by analyzing the preservative ethanol, including whole insects and plant leaves. Their findings were published in the March issue of the journal BioTechniques. ✤

Solution to Crossword in Issue 1

ACROSS: 3 GUELPH, 6 FIVE, 9 CANADA, 11 AUSTRALIA, 15 MITOCHONDRION, 17 PANAMA, 21 HEBERT, 23 FISHBOL, 24 GGTCAACAAA, 25 MATK.

DOWN: 1 CHINA, 2 INDIA, 4 PROTISTA, 5 WOLBACHIA, 7 POLLINATORS, 8 FUNGI, 10 ARGENTINA, 12 LEPIDOPTERA, 13 SOUTH AFRICA, 14 BIODIVERSITY, 16 TAIPEI, 18 COSTA RICA, 19 LEGS, 20 RUSSIA, 22 BOLD

In the last issue, we profiled the all taxa biotic inventory of Churchill in the Canadian Arctic. This time, we head to the island of Moorea to examine the tropical half of iBOL Working Group 2.1 (Barcoding Biotas).

Barcoding Paradise:

By Matt Hawes

Moorea is just your everyday South Pacific island paradise – jungle-cloaked mountains, impossibly blue lagoons ringed by coral reefs, palm-fringed beaches kissed by tropical breezes. To the travel writer Arthur Frommer, this heart-shaped gem 17km west of Tahiti was simply "the most beautiful island in the world." It even inspired Charles Darwin to come up with a theory about how coral atolls are formed.

Moorea's warm climate supports a diverse tropical ecosystem…but just how diverse is it? Researchers are shedding new light on this biota as the Moorea Biocode Project undertakes an ambitious inventory of all species and complex relationships that exist within the island's ecosystem.

Heading the project is Dr. Christopher Meyer, research zoologist and curator in the Department of Invertebrate Zoology at the Smithsonian National Museum of Natural History. He and his team of scientists are barcoding all of the island's non-microbial species with the goal of compiling an inventory for the entire macrobiota found on the small island.

"It's one thing to go around and collect specimens for all the well known groups, the 'low hanging fruit' of diversity," says Meyer, "but it's quite another to get all the players in an ecological system or community."

Moorea was chosen for a complete biotic inventory for a variety of reasons. The island itself is in the middle of the world's largest ocean, relatively isolated from biotas of the rest of the world. It contains most tropical groups and there's a manageable number of plant, animal and fungal species.

Additionally, two long-standing research stations there ensure continued research activity that will take advantage of the barcode library generated by the project. Researchers are already using the data to analyze the role of invasive species and to create biodiversity monitoring methods to track effects of climate change, such as sea level rise or ocean acidification.

The total number of animal, plant and algae species on Moorea is estimated to range from 8,000 to 16,000 species, but the final total is still uncertain. In contrast, preliminary estimates for fungi were about 5,000 species -- but in the first year researchers have already raised that estimate to more than 50,000 species. Clearly, a significant amount of research is still to be conducted and more surprises may be in store.

The project is focused on three areas: inventory, informatics and sustainability, all playing an integral role in determining the project's success.

For example, the informatics team is creating tools to record and track all research data generated by the collecting teams. Over 1,000 specimens can be examined in a day. Global unique identifiers (GUIDs) help them effectively locate and track the specimens and their related objects (photos, tissues, DNA extractions) as they're created and dispersed among partner institutions.

The barcoding process starts by extracting the specimen's DNA at its highest possible quality – in the field – and then preserving voucher specimens. One quarter of the extracted DNA is shipped to the Smithsonian Institution for sequencing and barcoding.

During the pilot phase, approximately one-third of the estimated macrobiota of the island have been inventoried, focusing on the "easier" groups but deliberately cutting across a broad taxonomic gamut (e.g. plants to animals) and

the Moorea Biocode project

across multiple habitats (marine, freshwater and terrestrial). This process was designed to identify limitations and bottlenecks in all field aspects including collection, tissue preparation, photo-vouchering, curation, and identifying appropriate repositories.

Moreover, as an All Taxa Biotic Inventory (ATBI), the Moorea scientists also faced the analytical challenges (DNA extraction, primer development and marker choice) of producing sequence data that cut across phyla, including little-studied but perhaps ecologically important "orphan phyla". They have successfully met all these challenges and, in the process, provided template material for CBOL, BOLD, LABS, and others to test and establish protocols that go beyond standard taxoncentric barcoding campaigns.

If this project is successful, Meyer believes other tropical locations, such as the neighboring Marquesas Islands or American Samoa, will be next in line as potential Biocode sites. "If we document all the pieces and show why they're important, and we're successful in accomplishing those goals, then doors will open for future projects," says Meyer. "We still have a long way to go to truly capture and circumscribe the diversity of these regions. I think the exciting stuff is still to come." ✤

Matt Hawes is a writer for the Students Promoting Awareness of Research Knowledge (SPARK) program at the University of Guelph

FACING PAGE
Another day at the office for Moorea Biocode project leader Chris Meyer.

THIS PAGE:
Above: Lab work on Moorea looks more like a seafood buffet. Top and middle right: The waters surrounding Moorea support life of astonishing diversity and beauty, like this reef fish and the colourful *Lophozozymus incisus* crab.

Bottom right: Collecting bugs on one of Moorea's craggy slopes. (All photos courtesy of Moorea Biocode)

Chris Meyer's Moorea Biocode collaborators are Neil Davies, Matteo Garbelotto, Rosemary Gillespie, Jean-Yves Meyer, Brent Mishler, Craig Moritz, Gustav Paulay, Claude Payri, Serge Planes and George Roderick. Partners include University of California-Berkeley, Centre National de la Recherche Scientifique (France), Smithsonian Institution, Florida Museum of Natural History, French Polynesia Research Department, Institute of Research for Development in New Caledonia and Association for Marine Exploration in Hawaii. The informatics team includes researchers from UC Berkeley and Biomatters (New Zealand). The research is funded by a grant from the Gordon and Betty Moore Foundation.

Barcode Bulletin Q&A with Michelle van der Bank -

Continued from page 16

BB: What will be the next African plant biodiversity hotspots to be barcoded?

MvdB: We've recently started a project on the woody vegetation of the Cheringoma district in central Mozambique. It is a botanically exciting area with about 300 tree species and many narrow endemics. During the civil war in Mozambique (1975-1994), botanical collecting came to a halt. If any plants were collected, it was mostly from roadsides because of the danger of landmines. There has been some collection since then but mostly in the south. For this project, we are working closely with Meg Coates Palgrave, who is an enthusiastic dendrologist and author of the book *Trees of Southern Africa*. Meg knows the Cheringoma district like the palm of her hand – it's a fantastic experience and treat to be in the field with her. Her enthusiasm, dedication and passion for trees are infectious. In fact it was Meg that introduced us to Cheringoma. We'll go to Cheringoma for two weeks in June on our second collecting trip.

BB: Southern Africa is an area of significant activity for plant barcoding generally but particularly for the TreeBOL project, where it's reported you have barcoded about 80 percent of the region's tree species. How did you achieve so much so quickly.

MvdB: My project manager Olivier Maurin and I spend huge amounts of time in the field collecting. We also receive specimens from other people collecting such as Tony Abbott, an indigenous plant specialist who is collecting rare plants for us in KwaZulu Natal. We have a wish list on the web and we encourage taxonomists to collect samples for us when they go out in the field. We supply the silica, all they have to do is bring back the plants. It is working well but we need more taxonomists who are interested in barcoding and aware of the benefits it can bring to them. The media have also been very useful in advertising the TreeBOL project and getting the public interested. We get amateur botanists calling us offering their help, which is fantastic.

BB: Tell us about the current University of Johannesburg projects to barcode protected timber and traded trees in Africa to assist custom officials at ports/borders.

MvdB: We've started to build a DNA barcode database for about 50 CITES [Convention on International Trade in Endangered Species of Wild Fauna and Flora] tree species that the Department of Water Affairs and Forestry (DWAF) expect to encounter at South African ports and borders. Customs officials are trained in basic macroscopic identification of timber samples but when there is doubt, the samples have to be sent to a specialist for microscopic identification. A consignment suspected of being illegally harvested from CITES or protected tree products can only be impounded for 72 hours so they need to get samples to the specialist very rapidly. Currently, there are only two of these in South Africa, in Stellenbosch and Pietermaritzburg. Also bark harvested from trees cannot be identified using conventional macro or microscopic processes which use core samples. Thus, there is an urgent need for fast and accurate identification that can be used as evidence in court cases involving protected trees and we hope that we can provide this service to DWAF.

BB: Are you working on any other plant barcoding applications, such as invasive species etc.?

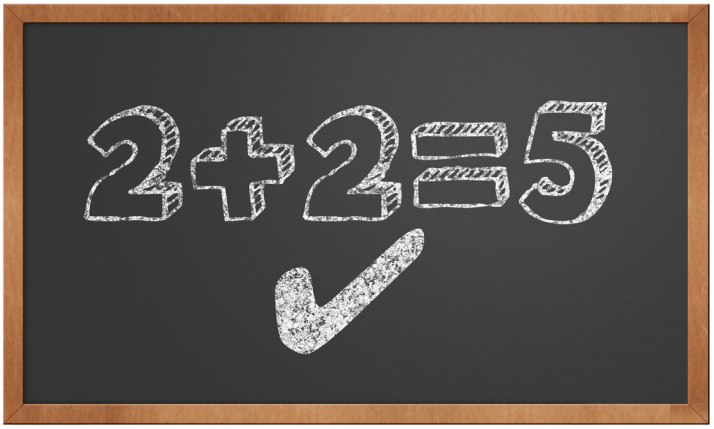
MvdB: Yes, we have a project on DNA barcoding of African cycads as well as medicinal plants of South Africa.

BB: At the moment, most DNA barcoding seems to happening in southern and central Africa. What is being done to get the west and north involved?

MvdB: Every year, CBOL funds short courses on DNA barcoding for 20 African researchers/students. So far, we've attracted people from Ethiopia, Cameroon, Ghana, Nigeria, Zimbabwe and Kenya. Our hope is that by transferring these skills and technology to other African countries, we'll get more institutions involved in the TreeBOL project. Currently we have formal collaborations with nine countries in Africa – Egypt, Nigeria, Ghana, Ethiopia, Kenya, Mauritius, Mozambique, Zimbabwe and Namibia. We are immensely excited to contribute in this way to the global database of biodiversity and, ultimately, to conservation. ✤

The figures don't lie... or do they?

iBOL-Canada Project Manager Greg Singer discovers a mathematical oddity in DNA barcoding data.



In 1973, a study was undertaken to find out whether there was gender bias in admissions to graduate programs at the University of California at Berkeley. The initial results certainly seemed to suggest that women weren't getting a fair shake:

	Applicants	% Admitted
Male	8442	44%
Female	4321	35%

But remarkably, when the statistics were broken down by department, it was found that the rate of admission for women was actually higher than that of men in four of the six departments studied:

Department	Men		Women	
	Applicants	% Admitted	Applicants	% Admitted
A	825	62%	108	82%
B	560	63%	25	68%
C	325	37%	593	34%
D	417	33%	375	35%
E	191	28%	393	24%
F	272	6%	341	7%

This counterintuitive outcome, known as Simpson's Paradox, was first described by E. H. Simpson in a 1951 paper and has since been observed in a variety of data, but particularly in studies from the biological or social sciences. For U.C. Berkeley, the solution to this conundrum came with the realization that more women were applying to departments that were much harder to get into. So admission to Department B was relatively easy but carried little weight because only 25 women applied. Deparment E, on the other hand, was very difficult to get into and carried a lot of weight because 393 women applied.

What does all this have to do with DNA barcoding? The next table shows real data from the Canadian Centre for DNA Barcoding in the past year, compiled with the help of Justin Schonfeld and members of the BOLD team at the Biodiversity Institute of Ontario. It illustrates the PCR success rate for two different primer combinations for two different insect orders:

	LepF1 – MLepR1	LCO1490_t1 – MLepR1
Coleoptera	145/169 (85.8%)	626/822 (76.2%)
Trichoptera	284/578 (49.1%)	88/185 (47.6%)

It is clear from the data that the LepF1-MLepR1 primer pair is better than the LCO1490_t1-MLEPR1 pair: for the Coleoptera there is 86% PCR success when using the former versus 76.2% when using the latter. Similarly, the first primer pair performs marginally better in Trichoptera than the second (49.1% versus 47.6%).

But look what happens when we aggregate the data:

	LepF1 – MLepR1	LCO1490_t1 – MLepR1
Coleoptera + Trichoptera	429/747 (57.4%)	714/1007 (70.9%)

Not only is the LCO1490_t1–MLepR1 primer combination superior overall, but by a wide margin!

This finding is not unique. I identified some 35 order and primer combinations that produce similar counterintuitive results. Like the U.C. Berkeley example, the solution to this puzzle lies in understanding that the LepF1-MLepR1 primer combination is used more frequently on samples that are unlikely to be successful (the Trichoptera), whereas the LOC1490_t1-MlepR1 pair is used more frequently on samples that are likely to be successful (the Coleoptera). These underlying biases cause the counterintuitive reversal of conclusions to occur when the data are combined.

Statistical analysis suggests that there's a 1-in-60 chance of encountering Simpson's Paradox in the wild, or about 1.7%. In this study I found the pattern in 1.4% of the primer-order combinations that I studied.

So it's a rare problem, but it does illustrate the potential pitfalls of using aggregate data to draw conclusions without understanding the underlying details. ✤

Michelle van der Bank

is a Lecturer in the Department of Botany and Plant Biotechnology at the University of Johannesburg. She is a leading authority and practitioner in plant barcoding and was a member of the CBOL Plant Working Group that recommended the standard plant barcode.

Barcode Bulletin: How did you first become interested in plant diversity and can you describe the path that led you to focus on plant barcoding?

Michelle van der Bank: I grew up with a love for biodiversity. My father was a naturalist and I spent a lot of my childhood close to nature in a small village in Namibia. I used to collect everything (except spiders because I have a fear of them) until I was told we had no more space to store things. I started out as a zoologist working on *Trichodina*, tiny parasites in the bladders of frogs. Then I took a position in the botany department at the University of the Witwatersrand and this is where my real career began. I decided to continue my PhD in botany, not zoology. I also had very strong influences in my life, like Prof. Mark Chase from the Royal Botanical Gardens, Kew. As for plant barcoding, I was on a walking trail with my colleague Vincent Savolainen in the Kruger Park early in 2005 just after the first barcoding meeting in London. We spent a lot of time talking about barcoding during this trip. I know it sounds crazy, being in Kruger and talking about work but this is always how it is with Vincent and me. Most of our projects are planned this way. We then went back and wrote a proposal to the Park to barcode its flora.

BB: You were a member of the study group that developed the recommendation leading to last year's adoption of rbcL and matK as the required barcode regions for land plants. What impact has adoption of the standard plant barcode had on your work in southern Africa and on plant barcoding generally?

MvdB: Being part of this group of scientists recommending the plant barcodes was an exciting experience and credit has to go to Pete Hollingsworth, who did a incredible job to get us all together. The adoption of the barcode has enabled us to accelerate rapidly. We had collected hundreds of plant samples over the past few years and were just waiting for the acceptance of the barcode to proceed with sequencing.

BB: You mentioned the project to barcode the flora of Kruger National Park. How much progress has been made and what has your work revealed so far? Have there been any big surprises or major discoveries?

MvdB: Since the start of the project in 2006 there has been considerable progress. Our target of producing a phylogenetic tree for all genera of flora within the KNP has progressed well, with most taxa of trees and shrubs (93 percent) completed for the rbcL and matK regions. The remaining seven percent are extremely rare in the park and a dedicated trip to collect these taxa is planned for 2010. We've collected and barcoded 91 of the 250 grass species occurring in the park. All data are available on BOLD Systems (http://www.boldsystems.org/). We've collected about 1,200 herbaceous species and the aim for this year is to barcode all of them. The park's trees and shrubs have been well studied so I think that it's the herbaceous plants that will produce the big surprises. Having said that, we are currently busy describing a new tree species, *Combretum nwambiyana* that O. Maurin, M. Jordaan and A.E. van Wyk collected in the Nwambiyana sandveld in KNP bordering Mozambique. The Nwambiyana sandveld is a unique area with high plant diversity but mostly inaccessible to the public. At first, we identified the tree as *Combretum mkuzense*, but barcoding data plus the two additional regions recommended by CBOL (ITS and trnH-psbA) confirm this to be a new species.

BB: What will be main impact of your work in KNP?

MvdB: We have completed a phylogenetic diversity analysis and will submit a conservation proposal based on the findings to the park authorities towards the end of this year.

Continued on page 14

The Barcode Bulletin is published quarterly by the International Barcode of Life project (iBOL).

Send your contributions, comments, suggestions, corrections to newsletter@iBOLproject.org or to:

International Barcode of Life Secretariat,
Biodiversity Institute of Ontario, University of Guelph,
50 Stone Road West, Guelph, ON, Canada N1G 2W1
Tel: 1-519-780-5483 Fax: 1-519-824-5703
www.iBOL.org