Consortium for the Barcode of Life: Support for Consortium Activities and Secretariat Office

Draft for Review

A request for continuation of support by the Alfred P. Sloan Foundation
submitted by the Smithsonian Institution on behalf of

Dr. Scott E. Miller
Principal Investigator
Consortium for the Barcode of Life

Office address: National Museum of Natural History, Room CE-119
10th & Constitution Avenue, NW, Washington, DC 20560

Mailing address: National Museum of Natural History, Smithsonian Institution
P.O. Box 37012, MRC-105, Washington, DC 20013-7012

Phone: 202/633-0808; Fax 202/633-2938
CBOL website: http://www.barcoding.si.edu

22 January 2006

Amount Requested: $1,554,732
Year One (1 July 2006 – 30 June 2007): $844,208
Year Two(1 July 2007 – 30 June 2008): $710,524

Calendar Years: 2006: $372,104; 2007: $827,366; 2008: $355,262

_____ _____
Scott E. Miller Date
Principal Investigator (202/633-1036)

_____ _____
Ardelle G. Foss, CRA Date
Director, Smithsonian Office of Sponsored Projects (202/275-0840)

EXECUTIVE SUMMARY

Molecular biomarkers have become invaluable tools in taxonomic research and in systematic and evolutionary biology. A variety of gene regions have been used in different plant and animal groups, and when used together, different gene regions provide detailed information on the evolutionary relationships of taxonomic groups. In general, more sequence data means better resolution. However, many circumstances require fast and cost-effective ways of assigning a specimen to a known species. In these cases, a short standardized diagnostic gene region may be sufficient, making a diverse array of regions unnecessary. DNA barcoding uses this standardized minimalist approach to specimen identification, and the Consortium for the Barcode of Life (CBOL), supported by the Alfred P. Sloan Foundation, is promoting DNA barcoding as the global standard for species-level identification.

CBOL was launched in May 2004 and has undergone rapid growth since that time. Its organizational structure has been established. More than 100 museums, herbaria, zoos, research organizations, government agencies and companies have joined CBOL. Four active Working Groups have been established, two global barcoding campaigns have been launched, and a public repository for standardized barcode data has been established.

CBOL requests renewal of its support from the Sloan Foundation for the 24 month period starting on 1 July 2006. CBOL plans to achieve four goals by the end of this period: (1) Provide the research community with the tools needed to launch successful barcoding projects, including software for data management and analysis, lab protocols, and global standards for data quality; (2) Create a fully operational demonstrator that can assign any specimen from any known species in a significant taxonomic group to its correct species; (3) Promote global participation in the Barcode Initiative through four regional meetings in the developing world and by doubling the

number of CBOL Member Organizations in developing countries, and (4) Increase the number of barcode records in public databases from the current level of 50,000 specimens in 10,000 species to 500,000 records in 50,000 species, primarily, but not exclusively, through CBOL's taxon-specific global barcoding "campaigns".

BACKGROUND: DNA BARCODING AND CBOL

Modern biological taxonomy began in the 18th century with the Linnean system of binomial nomenclature and higher classification. Since that time, taxonomists have been describing species on the basis of their morphology, ecology, behavior, and most recently on their biochemistry (in the form of electrophoretic data) and genetic make-up (e.g., restriction fragment length polymorphisms, protein and nucleotide sequences). Out of more than two centuries of research, knowledge of approximately 1.7 million valid species has emerged. This knowledge resides in a vast, distributed and unwieldy set of information resources: natural history collections in museums, herbaria, zoos and botanical gardens; frozen tissue collections; cell cultures; databases of raw observational information; taxonomic keys; and the printed taxonomic literature. As our understanding of biodiversity grows, we seem less and less able to answer the most common question that is likely to arise: ***To what species does this specimen belong?***

In late 2002, presentations and publications began to propose that a short gene sequence from a standardized position in the genome could be used as a diagnostic tool for assigning specimens to known species. Hebert et al. (2003a and 2003b) proposed that the mitochondrial gene COI (cytochrome c oxidase 1) could serve, in essence, as a "barcode" analogous to the Universal Product Codes (UPCs) that differentiate commercial products and enable their rapid, universal and unambiguous identification. The analogy between the "DNA barcode" and the UPC is meant to be illustrative rather than literal; all products of the same type have identical UPCs, but

the members of a species have some limited variability in their DNA barcodes. Nevertheless, COI is proving effective as a diagnostic tool within many taxonomic groups.

The Alfred P. Sloan Foundation took an early interest in the notion of a DNA barcode and provided support for two exploratory workshops, held in March and September 2003 at the Banbury Conference Center, Cold Spring Harbor Laboratory, New York. These workshops led to a proposal to the Sloan Foundation for creation of a Consortium for the Barcode of Life (CBOL) and support for its Secretariat Office. The proposal was submitted by the Smithsonian Institution on behalf of Dr. Scott Miller, Principal Investigator. A launch grant for $669,000 was awarded, effective 1 April 2004. An inaugural meeting was held at the Smithsonian in May 2004 and the Secretariat Office was opened in September 2004 under the direction of Dr. David Schindel, CBOL's Executive Secretary.

CBOL is devoted to exploring and promoting the potential of DNA barcoding as the global standard for species identification. It pursues specific goals such as:

- Creation of a public database of DNA barcode reference records;
- Determination of the DNA barcode sequences for important taxonomic groups, and eventually for all species;
- Promoting the development of faster, cheaper and more portable instruments and methods for obtaining DNA barcode data from biological samples; and
- Encouraging global participation in the "Barcode Initiative."

THE SCIENTIFIC AND SOCIETAL VALUE OF DNA BARCODING

Taxonomy is the domain of a few tens of thousands of professionals who possess the skills and experience necessary to distinguish closely related species, and to recognize, describe and document new species. In contrast with this limited supply of taxonomic expertise, the need

to assign specimens to known species arises every day and everywhere, for diverse practical reasons. Farmers need to know which pests are attacking their crops. Border inspectors need to know whether the life forms about to enter a country are dangerous or benign. Food inspectors need to verify that products are accurately labeled and do not carry parasites or pathogens. Public health officials need to identify organisms that might be carrying diseases. Taxonomists create the knowledge needed to identify specimens, but there are too few taxonomists to meet the demand for this service. Using DNA barcoding, non-specialists can assign specimens to known species – even specimens that can confound specialists (e.g., eggs, larvae, incomplete adults). Barcoding can therefore address "the taxonomic impediment" and can free taxonomists from routine identifications for the companion task of documenting new species.

CBOL'S LONG-TERM STRATEGY

CBOL's Executive Committee has agreed upon its goals and strategy for the proposed funding period (1 July 2006 to 30 June 2008) and for the years beyond this period. CBOL plans to make DNA barcoding a ready-to-use tool by the end of this grant period, and to create a demonstrator system that will provide a model for species identification using DNA barcodes. This demonstrator will integrate the software, data standards, protocols, and database capacity developed by CBOL's Working Groups, thereby providing researchers anywhere in the world with a practical lesson in how barcoding projects can be implemented.

A second goal is to expand greatly the opportunities for all taxonomists and consumers of taxonomic data to use DNA barcoding. With CBOL's encouragement, biotech companies will develop new processes and instruments that will increase the number of labs that can produce barcode data and reduce the time and cost needed to do so. CBOL plans to develop new protocols that will also expand the numbers and types of museum specimens that can contribute

barcode data, such as formalin-fixed specimens and dried specimens that are decades old.

The coming two years may see the peak of activity for CBOL. CBOL's Executive Committee anticipates a smaller role and budget for the Consortium beyond 2008, as barcoding becomes normal scientific practice. CBOL's principal goal after the coming two years might be disseminating the tools of barcoding, promoting their use in the everyday work of taxonomy, and institutionalizing support for barcoding. CBOL's Working Groups will be disestablished as their goals are achieved, and the long-term responsibilities for data curation will be distributed to appropriate permanent groups. CBOL plans to make DNA barcoding a standard practice in taxonomy by 2010, and a practical, reliable approach to specimen identification for non-taxonomists.

RESULTS UNDER SLOAN FOUNDATION SUPPORT

The original proposal to the Sloan Foundation specified a number of outputs and outcomes that were expected during the first 2.5 years of CBOL's existence.

OUTPUTS TO DATE: All of the outputs promised in the 2004 proposal to the Sloan Foundation have been produced in less than two years:

Inaugurate CBOL formally by 30 June 2004. CBOL's inaugural meeting was held at the Smithsonian in May 2004;

Expand membership in CBOL, with members to include at least 6 of the top 10 repositories of both animals and plants. CBOL now has more than 100 Member Organizations from more than 40 countries (see Appendix 1, Table A). These organizations include 7 of the top 10 botanical and 7 of the top 10 zoological repositories (see Appendix 1, Table B). It is important to recognize that although much of CBOL's goal to build a comprehensive barcode library could be accomplished by a small number of large repositories, the larger goal of promoting barcoding as

a global standard requires broad participation.

Create a CBOL Secretariat Office and hire an outstanding person to serve as Executive Secretary by 30 September 2004. Dr. David Schindel, an experienced executive from the National Science Foundation, joined CBOL and opened the Secretariat Office at the Smithsonian in September 2004. His CV is presented at the end of this proposal; and

Conduct the first International Barcode Conference by the end of 2004. The Natural History Museum, London, hosted a conference of 220 participants from 44 countries on 7-9 February 2005. The meeting was heavily oversubscribed and therefore streamed directly onto the internet. Bursaries for 30 developing country participants were given. This goal was achieved slightly later than planned but was highly successful and visible in the international popular and professional press. The Royal Society published a conference proceedings volume in its Philosophical Transactions: Biological Science series. CBOL has made the contents of the volume available without charge on its website through a non-exclusive non-competitive agreement with the Royal Society.

OUTCOMES TO DATE. Excellent progress has also been made toward the outcomes specified in the 2004 proposal to the Sloan Foundation:

Begin development of a long-term funding stream for CBOL from sources other than the Sloan Foundation. Engage government agencies and other potential funders. CBOL has conducted an aggressive program of outreach to government departments (including agencies for international development) and private companies (see Appendix 2, history of outreach efforts). These efforts are already producing tangible results in the form of co-funding for CBOL activities and funding for barcoding projects (see Appendix 3, leveraging of support from the Sloan Foundation). Nearly all of the CBOL activities proposed for 2006-2007 will be co-funded

from non-Sloan sources, and CBOL is confident in its ability to attract significant new contributions.

Promote public awareness of DNA barcoding. The popular press has given significant coverage to DNA barcoding. Coverage was particularly active following the publication of two barcode papers that uncovered cryptic species of birds (Hebert et al. 2004a) and butterflies (Hebert et al. 2004b) and after the February 2005 London conference. Articles on barcoding have appeared in the *NY Times, The Economist, National Geographic News, Science News,* and dozens of regional newspapers. In addition, the science news sections of *Science* and *Nature* magazines have each carried multiple stories on barcoding. (see References Cited, and news clippings posted on the CBOL and Rockefeller University websites[1]). CBOL has created and maintains a website[2] at the Smithsonian Institution which will begin compiling statistics on visitor traffic in January 2006.

Other accomplishments during 2004-2005 award period. Significant progress has been made during the 20 months following the May 2004 launch of CBOL:

Growth in participation. CBOL now has more than 100 Member Organizations from 40 countries on six continents (see Appendix 1), including the world's largest museums and herbaria (see Appendix 2). About thirty Member Organizations are from 20 developing countries.

Governance established. The original CBOL Steering Committee created a simple and straightforward organizational structure for the Consortium, consisting of an Executive Committee, Scientific Advisory Board, Secretariat Office and Working Groups (see Appendix 4). The structure was approved by the Member Organizations following the London conference

[1] http://www.barcoding.si.edu/DNAMediaCoverage.htm, http://phe.rockefeller.edu/BarcodeConference/press.php
[2] www.barcoding.si.edu

in February 2005. The Steering Committee has been replaced by a seven-member Executive Committee (EC) that includes the Principal Investigator of the Sloan Foundation grant (and Chair of the EC), CBOL's Executive Secretary, and representatives from Member Organizations in Europe, Africa and South America (see Appendix 4, Table A, for the Executive Committee's member list). The EC has created a 15-member Scientific Advisory Board (SAB; see Appendix 4, Table B, for member list) that includes the Chairs of CBOL's four Working Groups and 11 representatives from nominees from Member Organizations in different geographic regions.

Research results. A series of publications have demonstrated the effectiveness of COI as a diagnostic molecular marker for species of many animal groups. In some cases, COI has revealed heretofore hidden variations that are likely to be new species. Publications using COI as a species-level diagnostic appear at an increasing rate, adding continually to the global database of barcode records and increasing confidence in DNA barcoding. Appendix 5 includes the abstracts of four significant papers on barcoding.

Tens of thousands of barcode data records. Approximately 50,000 specimens representing 10,000 species have been barcoded to date. These data records reside in the Barcode of Life Database (BoLD) at the Biodiversity Institute of Ontario at the University of Guelph.

Creation of the barcode data standards. The Database WG has established data standards for DNA barcode records submitted to the International Nucleotide Sequence Database Collaboration (INSDC, which includes GenBank, the European Molecular Biology Lab, and the DNA Data Bank of Japan). These data standards (see Appendix 6) were developed in response to taxonomists involved in the Barcode Initiative and in collaboration with the leading taxonomic and biodiversity initiatives such as the Global Biodiversity Information Facility (GBIF), the Taxonomic Database Working Group (TDWG), Species2000, and the International Plant Names

Index. These data standards have broken new ground by connecting databases of gene sequences with voucher specimens and species names found in authoritative sources.

As part of a pilot project, BoLD data records for approximately 1500 specimens representing 300 species have been deposited as GenBank records that adhere to the BARCODE data standards. Many thousands of additional records in BoLD are being prepared for submission to GenBank.

New barcoding activities. CBOL has supported the design and launch of several large-scale barcoding initiatives. The Fish Barcode of Life (FishBOL) and All Birds Barcoding Initiative (ABBI) have set the goal of obtaining barcodes for all 35,000 species of fishes and 10,000 species of birds by 2010. Inaugural workshops for each initiative were held with support from the Sloan Foundation. CBOL supported these workshops during their planning phase and by providing travel stipends for participants from developing countries.

New areas of application. DNA barcoding is becoming a useful tool in research areas other than taxonomy, and CBOL has reached out actively to potential users (see Appendix 2). Ecologists, environmental scientists, agricultural inspectors, public health officials and others with a need to identify specimens are exploring barcoding as a new approach to applied problems.

An active Program of Work. CBOL launched four Working Groups (WGs) at or soon after the May 2004 inaugural meeting at the Smithsonian. These WGs are devoted to: (1) creating the database of DNA barcode records; (2) recommending lab protocols for obtaining barcode sequence data; (3) establishing protocols for the analysis, interpretation and visualization of barcode data; and (4) determining the optimal barcode region for plants, among which COI varies too little to act as a barcode region. CBOL has also established an International Network for the Barcoding of Invasive and Pest Species with more than 50 members.

Attacking "the formalin problem". Millions of well-curated and identified museum specimens

(primarily fish and marine invertebrates) were fixed in formalin, which up until now has made their DNA inaccessible. CBOL has asked the National Academies of Science to organize and hold a workshop, to be held in early 2006, at which chemists, molecular biologists, biomedical researchers and others will review past approaches to DNA extraction from formalin-fixed specimens. They will then suggest new research directions that CBOL will promote.

Improving the research environment for taxonomy. CBOL's Database Working Group is promoting linkages among barcode sequences, voucher specimens in museums and other repositories, valid species names and taxonomic publications. This tighter integration of sequences, specimens and species names will improve the research environment for all taxonomists, not just barcoders. With this goal in mind, CBOL organized and held a meeting of 80 participants on electronic access to the taxonomic literature in conjunction with the London barcode conference. This meeting was co-sponsored by the National Biological Information Infrastructure of the US Geological Survey and led directly to the establishment of the Biodiversity Heritage Library Network (BHL). BHL is part of the Open Content Alliance and includes major libraries of taxonomic literature that are exploring a major digitization initiative.

An active debate in taxonomy. The introduction and growth of DNA barcoding have not been without apprehension and controversy. Some taxonomists are rightly concerned that gene sequence data will displace morphological, ecological and other classes of data as the bases for taxonomy (proposed as "DNA taxonomy" in Tautz, et al., 2003, for example). Over the past two years, barcoding has become the most widely and intensively debated topic in taxonomy. *Science*, *Nature*, *Systematic Biology*, *Bioscience*, *Trends in Ecology and Evolution* and other leading journals have carried a steady stream of articles, commentaries and letters either for or against barcoding (see examples in References Cited). Not since the cladistics debate during the

1960s and 70s has taxonomy had a higher profile in the scientific community. CBOL has worked to clarify misconceptions (the principal one being that DNA barcoding is the same as DNA taxonomy) and to promote the use of barcoding as a global standard for species identification and a tool for integrative taxonomy (Schindel and Miller, 2005).

Engaging biotech companies in developing barcode technologies. DNA barcoding is a multi-step process that currently takes several hours and costs $1-2 per specimen in consumables. One of CBOL's goals is to expand the use of barcoding by reducing the time and cost of the process, and enabling field use by developing more portable instruments. CBOL has approached several instrument and reagent manufacturing companies to make them aware of the potential market for barcoding technology. Two of these companies are co-funding CBOL activities (see Appendix 3). Appendix 7 outlines unit costs of barcoding and where savings might be achieved.

PROPOSED PROGRAM OF WORK AND GOALS FOR 2006-2008

During its first two years, the Barcode Initiative has reached an important milestone. Research projects based on the analysis of well-determined voucher species in a variety of taxonomic groups have shown that limited within-species COI variation and significant between-species differences. These are necessary, but not sufficient, conditions for the long-term success of DNA barcoding as a tool for species identification. During 2006-2008, CBOL will pursue tangible goals in four areas (see Appendix 8, CBOL Status in 2006, Ambitions for 2008, and Appendix 9, CBOL Milestones for 2008). First, CBOL will continue to pursue its core mission of ***enabling the universal application of DNA barcoding*** through its Working Groups and Committees.

Second, CBOL proposes a new activity: ***developing a complete and operational demonstrator system for species identification in a group of interest***. The demonstrator will

operate through a new Barcode of Life Initiative (BoLI) Data Portal. CBOL's Working Groups will remove the major technical obstacles involving laboratory protocols, data analysis procedures, and data management. Their products will be integrated into the BoLI Data Portal.

Third, CBOL will continue to pursue its goal of ***populating a global database of reference DNA barcode records***, primarily but not exclusively through two global barcoding campaigns. Regional working groups have been organized to locate and barcode tens of thousands of bird and fish specimens, and CBOL will support their efforts during the proposed renewal period. Several new barcoding efforts will also be launched, greatly increasing the number of species for which DNA barcodes have been determined.

Fourth, CBOL will expand its mission of ***promoting global participation in the Barcode Initiative*** by engaging researchers and users in developing countries. Their participation will grow through CBOL's major barcoding campaigns (FishBOL and ABBI, described below), networks, and through initiatives launched through CBOL's regional outreach efforts.

ENABLING UNIVERSAL APPLICATION OF BARCODING. One of CBOL's core missions is to enable barcoding projects by addressing shared technical obstacles. The following proposed activities will provide the research community with the tools needed to implement barcoding projects.

Database Working Group. This WG proposes to create an interoperable data environment in which barcode data are integrated with the most relevant other biodiversity data resources. This environment will be based on protocols that the WG will develop with its collaborators for the management and curation of barcode data records deposited in the International Nucleotide Sequence Database Collaborative (INSDC), comprising GenBank, the European Molecular Biology Laboratory (EMBL) and the DNA Databank of Japan (DDBJ). These protocols would establish and implement linkages among (1) barcode records in INSDC, (2) relevant data records

in related biodiversity data resources such as the Global Biodiversity Information Facility (GBIF), Species2000, ITIS, BoLD and the Zoological Information Management System (ZIMS), and (3) taxonomic journals and other publications. These protocols will provide the foundation for the proposed BoLI Data Portal (see below).

Deliverable for the award period. The Database WG will disseminate information on the new Barcode data standards and will produce a user's guide to the submission, curation and management of barcode records in INSDC, and their linkage to other data resources.

Background. The Database WG has made extraordinary progress in the 18 months following CBOL's inaugural meeting in May 2004. Four meetings have been held at which a productive collaboration has been formed with NIH's National Center for Biotechnology Information (NCBI, host to GenBank) and the leading biodiversity informatics initiatives (GBIF, Species2000, ITIS, among many others). The Working Group developed standards for DNA barcode data submitted to GenBank and its two collaborating databases, EMBL and DDBJ (see Appendix 6). The proposed data standards have been approved by the three international databases, and by CBOL's Scientific Advisory Board and Executive Committee.

DNA Working Group. The overarching goal for 2006-2007 will be to develop and disseminate the information necessary to expand the source of barcode data beyond freshly collected specimens, and increasing the number of laboratories that are producing barcode data.

Deliverable for the award period. The DNA WG will develop, compile and disseminate information that will enable labs to initiate barcoding projects, and will promote the creation of one leading barcode lab in each region of the world.

Background. There have been four constituent types of participants in the DNA Working Group to date: the two leading barcoding labs (University of Guelph and the Smithsonian); an

additional 10-20 labs in museums, herbaria, and zoos that have started to collect barcode data; private-sector biotechnology companies that develop and sell enzymes, primers, and other reagents necessary for barcoding; and individual researchers with an interest in developing capabilities to collect barcode data.

The DNA Working Group proposes to hold one full meeting of the Working Group per year, at which the participating labs and companies would:

- Agree on and implement a distributed program of work to develop and test new barcoding protocols, especially for older museum specimens and formalin-fixed material;
- Discuss techniques for reducing the cost and increasing the speed of barcoding;
- Exchange information on new equipment and technical approaches to barcoding;
- Explore possible discounts for CBOL Member Organizations for high-volume purchases;
- Develop models and guidance for developing new barcoding labs; and

The WG will produce reports that disseminate their findings to the barcoding community.

Data Analysis Working Group. This WG proposes to develop protocols and software for the sampling, analysis, interpretation and visualization/display of barcode data. Protocols will include guidance to barcoding researchers on:

- necessary sample sizes;
- choice of appropriate analytical routines;
- issues of effective population size, gene flow, lineage sorting, and other population genetics factors that bear on the interpretation of barcode data for species-level taxonomy; and
- ways to display barcode data most effectively for the purpose at hand.

The WG will reach this goal by engaging statisticians, population geneticists, computer scientists, machine learning specialists and applied mathematicians in open competitions.

Deliverable for the award period. The Data Analysis WG will produce protocols and software that will be presented at the Second International Barcode Conference (planned for February 2007), submitted as manuscripts for the conference proceedings volume, and will be incorporated into the proposed BoLI Data Portal.

Background. The Data Analysis WG has held two planning meetings that produced a list of technical/analytical challenges and a set of research questions associated with barcode data. The WG has reached out to the population genetics, statistics and computer science communities to promote their participation in these challenges and questions.

The WG proposes to hold one workshop in May 2006. Participants will be selected on a competitive basis and will present preliminary results on the new analytical protocols and software they propose to develop. Participants will receive feedback and will then go on to prepare final deliverables. Ten dissertation improvement awards of $5,000 each will be given by the WG Steering Committee to the most promising graduate student projects presented at the workshop. Funds for these awards will be raised from non-Sloan sources (discussions are underway with the National Science Foundation and the European Science Foundation). The award process will be managed by the Rutgers University Center for Discrete Mathematics and Theoretical Computer Science (DIMACS), which has managed similar competitions over the past 15 years. The WG will organize and propose a session on data analysis for the Second International Barcode Conference planned for February 2007. The data analysis protocols and software presented at the conference will be incorporated into the BoLI Data Portal.

Plant Working Group. Following the London Barcode Conference in February 2005, Dr. Robyn Cowan (Royal Botanical Gardens, Kew) submitted proposals to the Alfred P. Sloan and Gordon and Betty Moore Foundations on behalf of the Plant Working Group. These proposals

($400,000 each) were successful and they are now supporting a two-year effort to identify the optimal barcode regions for plants. The first phase, in which effort was concentrated at Kew and the University of Reading, explored 100 candidate regions and identified five likely barcode regions in December 2005. The second phase, involving approximately ten herbaria, museums, universities and biodiversity research organizations, will identify the best one or few barcode regions. Final results will be presented at a workshop at the NY Botanical Gardens in December 2006 and at the Second International Barcode Conference in February 2007.

BoLI Data Portal. CBOL proposes to construct a data portal for the Barcode of Life Initiative that is distinct from but linked to the CBOL website. This Portal would complement the BoLD system at the Biodiversity Institute of Ontario, University of Guelph, by providing the broader research community with one-stop access to barcode data. CBOL is promoting BoLD as the workbench on which barcode records are assembled prior to submission to GenBank. The Canadian Barcoding Network will also be relying on BoLD as its workbench and repository, along with many other barcoding projects. BoLD is therefore designed for participants in barcoding projects. The BoLI Data Portal would need strong interoperability with BoLD as well as GenBank, but it will be designed for users who are not participants in barcoding projects.

Deliverables for the award period. CBOL will create a Portal that hosts:

- The data standards and user's guide to BARCODE records;
- Websites for the FishBOL and ABBI campaigns, INBIPS, and the demonstrator development project (see below);
- Protocols and software developed by the Data Analysis Working Group; and
- Linkages to barcode data in GenBank, BoLD and other relevant repositories.

DEMONSTRATOR DEVELOPMENT PROJECT. CBOL's Executive Committee proposes to develop

a complete and operational system for species identification of a single taxon of interest during 2006-2008. This system will demonstrate that DNA barcoding can be effective and reliable as an identification system for unidentified specimens and for specimens that cannot be identified readily by other methods (e.g., juvenile stages, partial or damaged specimens, gut contents). The Executive Committee is in the process of deciding between mosquitoes (3200 known species) and tephritid fruit flies (4400 known species) as the subject taxon of the demonstrator. This project would be managed by CBOL in a more "top down" centralized manner than longer-term distributed "bottom-up" projects such as CBOL barcoding "campaigns" (see below).

Deliverables for the award period. CBOL will create a global identification system for the selected group by (1) obtaining barcode records for all species in the selected group, (2) submitting them to INSDC, and (3) providing user-friendly access to these records and identification tools through the BoLI Data Portal.

Committees. The CBOL Executive Committee has agreed to form two committees that will oversee small and focused programs of work. A ***Technology Development Committee*** will advise CBOL on potential new directions for barcoding technologies, especially those leading to a portable barcoding device. This committee will also identify private companies that are exploring these directions, with which CBOL should make contact. A ***Legal/IPR Committee*** will advise CBOL on intellectual property rights issues related to barcoding and on the legal services that CBOL will need. This committee will also consider the ethical implications of barcoding, the relationship between barcoding and unethical practices such as biopiracy, and the potential misuse of DNA barcoding and barcode data.

POPULATING A GLOBAL BARCODE DATA BASE. FishBOL and ABBI are CBOL's two largest barcoding "campaigns". They have held inaugural workshops at which distributed management

plans were accepted, involving Steering Committees and Regional Working Groups (RWGs).

Following the London Barcode Conference in February 2005, representatives of agencies responsible for the control of invasive and pest species self-organized the International Network for the Barcoding of Invasive and Pest Species (INBIPS). They have exchanged information informally and have planned a more structured and ambitious program of work for 2006-2007.

Deliverables for the award period. Each of the campaign RWGs will finalize its strategic plan for obtaining specimens, arranging for the necessary lab capacity (either within region or through inter-regional collaboration), identifying the collecting and export permits that will be needed, and developing strategies and proposals for raising the necessary funds. The Steering Committee of each campaign will develop a plan for internal and external communications and will launch a public website. ABBI has set the goal of barcoding 10,000 specimens representing 2,000 bird species during the award period. FishBOL has set the goal of 30,000 specimens representing 6,000 fish species.

INBIPS will create and maintain a website that will act as a clearinghouse for information on barcoding activities that focus on invasive and pest species, and will conduct an active program of outreach to national and international agencies involved with the control of invasive and pest species. The INBIPS Steering Committee will develop plans for two barcoding projects and will hold one workshop each year to launch these initiatives.

PROMOTING GLOBAL PARTICIPATION. Biodiversity research, by its very nature, is a global enterprise that affects all geographic regions and sectors of society. CBOL is devoted to global participation in the Barcode Initiative and exploring the use of DNA barcoding for sustainable development, in accordance with the Convention on Biological Diversity. CBOL's strategy for promoting participation in developing countries is presented below. The strategies for engaging

partners in non-academic sectors of society are presented in Appendix 10.

Deliverables for the award period. Approximately one-third of CBOL's Member Organization is from developing countries. CBOL's Executive Committee has set a goal of doubling the number of developing country Member Organizations by the beginning of 2007. Toward that end, CBOL is working with four local organizing committees to plan the following regional workshops during the next two years:

- A southern African regional meeting hosted by the South African Biodiversity Institute (SANBI) in April 2006 (using funds in the present Sloan Foundation grant to CBOL);
- A South American regional meeting organized by the Brazilian National Institute of Amazonian Research (INPA) in the second quarter of 2006;
- An east Africa regional meeting organized by the National Museums of Kenya and ICIPE, in conjunction with BOZONET, a nascent capacity building network, in October 2006; and
- A south Asian regional meeting (tentative host and date: University of Peradeniya, Sri Lanka in August 2006).

Participants in the regional meetings will include taxonomists with and without experience in molecular biomarkers, representatives of biodiversity research and conservation organizations, government officials in agriculture, environment, public health and research, and private sector companies in health, food/agriculture, and biotechnology. The regional meetings will have five goals: (1) raise awareness of barcoding among researchers, research organizations, and potential users; (2) explore the potential applications of DNA barcoding to the highest priority problems facing the countries represented at the meetings; (3) assess the greatest needs and opportunities for DNA barcoding in the region, (4) identify the highest priorities in the form of national and regional action plans; and (5) initiate the formation of intra-regional networks and

intercontinental partnerships that will implement the action plans.

CBOL will not provide funding for the action plans, but CBOL plans to cooperate with regional partners in their efforts to obtain funding and implement barcoding through:

- In-country training activities such as short courses for researchers and advanced students on technical aspects of DNA barcoding and associated specimen curation;
- Research training fellowships that will allow researchers and technicians to spend longer periods of time (weeks to months) in partner labs for advanced training and research projects;
- Infrastructure improvement such as equipment acquisition; and
- Other forms of high-priority capacity-building identified during the regional meeting.

Second International Barcode Conference. CBOL's Executive Committee plans to convene a second international barcode conference in February 2007, two years after the first conference. The conference will be held in southeast Asia to promote participation of that region in barcoding activities. The 2005 conference had 220 participants and was heavily oversubscribed. The Organizing Committee for the 2007 conference will plan for 400-500 participants.

As of January 2006, there were only 10 CBOL Member Organizations from six Asian countries. CBOL considers Asia as an important participant in the Barcode Initiative and recognizes the need to expand greatly the participation in Asia. In preparation for the 2007 international conference, the CBOL Secretariat will identify and contact important potential Member Organizations and will contact them with information on barcoding. The proposed budget includes support for outreach visits by members of the CBOL Executive Committee and Scientific Advisory Board, and Asia will be a high priority for the use of these funds.

REFERENCES CITED

Hebert, P.D.N., Cywinska, A., Ball, S.L., and deWaard, J.R. (2003a). "Biological identifications through DNA barcodes." *Proceedings of the Royal Society London B* 270:313-321.

Hebert, P.D.N., Ratnasingham, S., and deWaard, J.R. (2003b). "Barcoding animal life: cytochrome c oxidase 1 divergences among closely related species." *Proceedings of the Royal Society London B* 270:S596-S599.

Hebert, P.D.N., Stoeckle, M.Y., Zemlak, T.S., and Francis, C.M. (2004a). "Identification of birds through DNA barcodes." *Public Library of Science, Biology* 2: e312.

Hebert, P.D.N., Penton, E.H., Burns, J.M., Janzen, D.H., and Hallwachs, W. (2004b). "Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*." *Proceedings of the National Academy of Science* 101:14812-14817.

Schindel, D.E. and S.E. Miller. (2005). Barcoding a useful tool for taxonomists, letter to Nature letter, 435:17

Tautz, D., Arctander, P., Minelli, A., Thomas, R. H. & Vogler, A. P. 2003 A plea for DNA taxonomy. *Trends Ecol. Evol.* 18, 70–74.

DNA barcoding has generated an active exchange of views in the professional literature, including the following:

Heber, P.D.N, and Gregory, T.R., 2005, The Promise of DNA Barcoding for Taxonomy, Systematic Biology 54 (5):852-859.

Lipscomb, D., Platnick, N. & Wheeler, Q. D. 2003 The intellectual content of taxonomy: a comment on DNA taxonomy. *Trends Ecol. Evol.* **18**, 65–66.

Malte C. Ebach, M.C. and Holdrege, C. (2005) DNA barcoding is no substitute for taxonomy. Nature 434:697

Moritz C. and Cicero C. (2004) DNA barcoding: Promise and pitfalls. PLoS Biol 2(10): e354.

Stoeckle M (2003) Taxonomy, DNA and the barcode of life. BioScience 53: 2–3.

Vincent S. Smith, V.S., 2005, DNA Barcoding: Perspectives from a "Partnerships for Enhancing Expertise in Taxonomy" (PEET) Debate, Systematic Biology 54 (5):841-844

Will, K.W. and Rubinoff, D. (2004) Myth of the molecule: DNA barcodes for species cannot replace morphology for identification and classification. Cladistics 20, 47–55

Will, K.W., Mishler, B.D., and Wheeler, Q.D., 2005, The Perils of DNA Barcoding and the Need for Integrative Taxonomy, Systematic Biology 54 (5):844-851

Some of the reports of DNA barcoding in newspapers and magazines are:

Brownlee, C. DNA Barcodes: Life under the scanner. Science News. December 4, 2004. 166:360-361

Marshall, E. Will DNA Bar Codes Breathe Life Into Classification? SCIENCE February 18, 2005. 307:1037

Wade, N. A Species in a second: Promise of DNA "Bar Codes". NY Times. December 14, 2004

Barcode Me. NewScientist. 26 June 2004. 132-135

Pearson, H. DNA barcodes tag species. Published online 27 September 2004

There may be more species on Earth than previously imagined. *The Economist.* 30 September *2004*

Appendix 1. CBOL Member Organizations

Table A. CBOL Member Organizations (as of 20 January 2006)

Museo Argentino de Ciencias Naturales	Buenos Aires	Argentina
Australian Museum	Sydney	Australia
Botanic Gardens Trust, National Herbarium of NSW	Sydney	Australia
CSIRO Marine Sciences and Entomology	Hobart	Australia
MacQuarie University, Department of Chemistry and Biomolecular Sciences & Biotechnology Research Institute	Sydney	Australia
National Herbarium of Victoria at Royal Botanic Gardens Melbourne	Melbourne	Australia
University of Rajshahi, Department of Botany		Bangladesh
Royal Belgian Institute for Natural Sciences	Brussels	Belgium
Royal Museum for Central Africa	Tervuren	Belgium
Instituto Nacional de Pesquisas da Amazonia	Manaus	Brazil
Universidad Federal de Minas Gerais		Brazil
Acadia University	Nova Scotia	Canada
Canadian Museum of Nature	Ottawa	Canada
Connell Memorial Herbarium, University of New Brunswick	New Brunswick	Canada
University of Guelph		Canada
Beijing Genomics Institute, Chinese Academy of Sciences	Beijing	China
Department of Biology & Simon F.S. Li Marine Science Lab, The Chinese University of Hong Kong		China
Guizhou University, Department of Entomology	Guizhou	China
Institute of Zoology, Chinese Academy of Sciences	Beijing	China
Plant Protection College, Shandong Agricultural University		China
Research Center for Biodiversity, Academia Sinica	Taiwan	Chinese Taipei
Universidad de los Andes	Bogota	Colombia
Area de Conservacion Guanacaste	Santa Rosa	Costa Rica
National Biodiversity Institute (INBio)	San Jose	Costa Rica
Laboratoire d'Ecologie et Gestion des Ressources Animales (LEGERA), University of Kisangani	Kisangani	Democratic Republic of Congo

Global Biodiversity Information Facility	Copenhagen	Denmark
Natural History Museum of Denmark	Copenhagen	Denmark
Institut National de la Recherche Agronomique (INRA)	Paris	France
Laboratoire d'Ecologie et gestion des Ressources Animales (LEGERA)		France
Museum National d'Histoire Naturelle	Paris	France
Botanic Garden Berlin (Botanischer Garten und Botanisches Museum)	Berlin	Germany
Center for Experimental and Evolutionary Biodiversity	Hannover	Germany
Knebelsberger and Miller Bioservices	Munich	Germany
Staatliche Naturwissenschaftliche Sammlungen Bayerns	Munich	Germany
Zoological Institute and Museum	Bonn	Germany
University of Ghana	Legon	Ghana
Natural History Museum of Crete	Crete	Greece
St Peter's College & MEWS	Mumbai	India
Vector Control Research Centre (Indian Council of Medical Research), Min. of Health & Family Welfare, Govt. of India	Pondicherry	India
Plant Pest and Disease Research Institute		Iran
Institute of Biomedical Technologies	Bari	Italy
International Centre for Genetic Engineering and Biotechnology	Trieste	Italy
Istituto Zooprofilattico Sperimentale del Piemonte	Genova	Italy
National Institute of Genetics, ROIS	Shizuoka	Japan
National Museums of Kenya	Nairobi	Kenya
Instituto de Biología, Universidad Nacional Autónomo de México	Mexico City	México
UCB Gump South Pacific Research Station	Moorea	French Polynesia
Scientific Institute (NMNH)		Morocco
CBS Fungal Biodiversity Centre	Utrecht	Netherlands
Consortium of Netherlands Taxonomic Facilities		Netherlands
National Museum of Natural History Naturalis	Leiden	Netherlands
Zoological Museum of the University of Amsterdam	Amsterdam	Netherlands
Allan Wilson Centre for Molecular Ecology and Evolution, Massey University		New Zealand

Landcare Research	Lincoln	New Zealand
National Centre for Advanced Bio-Protection Technologies		New Zealand
Forestry Research Institute of Nigeria		Nigeria
Natural History Museum, Obafemi Awolowo University	Ile-Ife	Nigeria
University of Bergen, the Natural History Collections, Bergen Museum	Bergen	Norway
New Guinea Binatang Research Center (PNG)	Madang	Papua New Guinea
Philippine Rice Research Institute (PhilRice)	Maligaya	Philippines
Museum and Institute of Zoology- Polish Academy of Sciences		Poland
Institute of Marine Research (IMAR)		Portugal
Zoological Museum of Moscow University	Moscow	Russia
Raffles Museum of Biodiversity Research, National University of Singapore	Singapore	Singapore
South African National Biodiversity Institute (SANBI)	Pretoria	South Africa
University of Fort Hare	Fort Hare	South Africa
University of Johannesburg	Johannesburg	South Africa
Jardin Botanico Canario Viera Y Clavijo		Spain
University of Peradeniya	Peradeniya	Sri Lanka
Goteborg University, Department of Zoology	Goteborg	Sweden
Swedish Museum of Natural History	Stockholm	Sweden
Agroscope FAW Waedenswil	Waedenswil	Switzerland
BioNET-International		UK
Fisheries Research Service	Aberdeen	UK
Royal Botanic Garden Edinburgh	Edinburgh	UK
Royal Botanic Garden Kew	London	UK
School of Biological Sciences, Univ. of Wales	Swansea	UK
The Natural History Museum	London	UK
The University of Reading	Reading	UK
University of Hull	Hull	UK
University of Wales	Bangor	UK

Academy of Natural Sciences, Philadelphia	Philadelphia, PA	USA
American Museum of Natural History	New York, NY	USA
Bishop Museum	Honolulu, Hawaii	USA
California Academy of Sciences	San Francisco, CA	USA
Coriell Institute for Medical Research	New Jersey	USA
Florida Fish and Wildlife Conservation Commission	St. Petersburg	USA
Florida Museum of Natural History	Gainesville, FL	USA
Integrated Taxonomic Information System (ITIS)	Washington, DC	USA
Missouri Botanical Garden	St. Louis, MO	USA
Museum of Comparative Zoology, Harvard University	Cambridge, MA	USA
National Biological Information Infrastructure (NBII)	Reston, VA	USA
New England Biolabs, Inc.	Ipswich, MA	USA
NSF International (formerly National Sanitation Foundation)	Ann Arbor, MI	USA
Ocean Genome Legacy Foundation	Ipswich, MA	USA
Olive Natural Heritage Society, Inc	West Shokan, New York	USA
San Diego Supercomputer Center at UCSD	La Jolla, CA	USA
Scripps Institution of Oceanography	La Jolla, CA	USA
Smithsonian Institution	Washington, DC	USA
The New York Botanical Garden	New York, NY	USA
University of Connecticut Marine and Technology Center	Groton, Connecticut	USA
University of Kansas Biodiversity Institute	Lawrence, Kansas	USA
University of New Hampshire	Durham, NH	USA

Table B. CBOL Member Organizations among top museums/herbaria.

The following is a compilation of the leading repositories of plants and animals. CBOL Member Organizations are shown in *red italics*.

Institutional collections, ranked by # specimens

	Institution	Specimens
Plants		
1	*Musee National d'Histoire Naturelle, Paris*	8,877,300
2	*Royal Botanical Garden, Kew*	7,000,000
3	*New York Botanical Garden*	7,000,000
4	Komarov Botanical Institute	5,770,000
5	*Swedish Museum of Natural History, Stockholm*	5,600,000
6	Conservatoire et Jardin Botaniques, Geneva	5,500,000
7	*Missouri Botanical Garden*	5,219,216
8	*Natural History Museum, London*	5,200,000
9	Harvard University Herbarium	5,000,000
10	*National Museum of Natural History, Washington, DC*	4,500,000
Animals		
1	*National Museum of Natural History, Washington, DC*	78,500,000
2	*Natural History Museum, London*	55,000,000
3	*Musee National d'Histoire Naturelle, Paris*	50,000,000
4	*American Museum of Natural History, New York*	21,000,000
5	Humboldt Museum, Berlin	21,000,000
6	*Museum of Comparative Zoology, Harvard*	21,000,000
7	*Bishop Museum, Honolulu*	20,000,000
8	*Academy of Natural Sciences, Philadelphia*	17,000,000
9	Field Museum, Chicago	17,000,000
10	Zoological Institute, St. Petersberg	15,000,000

Appendix 2. CBOL outreach activities during current grant period

The CBOL Executive Secretary gave the following outreach presentations:

1. IUBS US National Committee
2. Department of Homeland Security, Washington, 3 November 2004
3. European Commission, Directorate General for Research, Brussels, 8 November 2004
4. White House Office of Science and Technology Policy, Washington, 3 December 2004
5. US Department of Agriculture, Beltsville Agricultural Research Center, 10 December 2004
6. National Agricultural Library and USDA Agricultural Research Service, Beltsville, MD, 28 February 2005
7. Duke University, Durham, NC, 28 March 2005
8. New England Biolabs, Inc., Beverley, MA, 4 April 2005
9. Belgian Platform for Biodiversity Research Strategy, Brussels, 18 April 2005
10. University of California, Berkeley, 9 May 2005
11. Affymetrix, Inc., Santa Clara, CA, 10 May 2005
12. California Department of Food and Agriculture, Sacramento, CA, 11 May 2005
13. Natural Sciences Collections Alliance annual meeting, Santa Barbara, CA, 13 May 2005
14. American Type Culture Collection, Manassas, VA, 26 May 2005
15. Smithsonian Ornithology Group, Suitland, MD, 3 June 2005
16. Interagency BioEco group, Washington, 13 December 2004; 1 June 2005; 21 July 2005
17. NOAA National Marine Fisheries, Silver Spring, MD, 7 January 2005
18. European Commission-US Biotechnology Task Force Workshop on Food Safety, Brussels, 24 January 2005
19. US Environmental Protection Agency, Washington, 1 February 2005
20. National Biological Information Infrastructure, US Geological Survey, Reston, VA, 1 February 2005
21. Ocean Biogeographic Information International Committee, San Francisco, 23 June 2005
22. Fogarty International Center, NIH, Bethesda, MD, 28 June 2005
23. Canon Life Sciences, Canon Corp., Arlington, VA, 21 July 2005
24. Animal-Plant Health Inspection Service (APHIS), USDA, Washington, 21 July 2005
25. Division of Oceans, Environment and Science, US Department of State, Washington, 26 July 2005
26. International Foundation of Science, Stockholm, 10 October 2005
27. Swiss Agency for Development Cooperation, Bern, 17 October 2005
28. Biodiversity Forum of the Swiss Academy of Sciences, Bern, 17 October 2005
29. Directorate General for Research, European Commission, Brussels, 18 October 2005
30. Netherlands Research Organization (NWO), The Hague, 19 October 2005
31. Netherlands Agency for Development (RAWO), The Hague, 19 October 2005
32. Norwegian Agency for Development (NORAD), Oslo, 20 October 2005
33. Museum of Zoology, University of Copenhagen, 21 October 2005
34. Subsidiary Body for Scientific, Technical and Technological Affairs (SBSTTA) of the Convention on Biological Diversity, Montreal, 28 November 2005

The following outreach presentations were made by members of CBOL's Executive Committee:

35. James Edwards, 77th Annual Meeting of the Genetics Society of Japan, Shinjuku, 29 September 2005

36. Scott Miller, Smithsonian Botany Symposium, Washington, DC, September 2005
37. Scott Miller, Kenya
38. Scott Miller, Papua New Guinea

The following outreach presentations were made by members of CBOL's Scientific Advisory Board:

39. Robyn Cowan, GfBS/Society for Biological Systematics annual meeting, Basel, Switzerland, 13-16 September 2005
40. Karen Armstrong, APEC Workshop on Invasive and Alien Species, Beijing, 19 September 2005
41. Pablo Luis Tubaro, CYTED (Ciencia y Técnica para el Desarrollo) and National Net of Biological Collections, Buenos Aires, 22 November 2005
42. Freek Bakker, Society of French Systematics, Paris, 9 November 2005
43. Pablo Luis Tubaro, National Academy of Sciences of Argentina, Buenos Aires, 14 December 2005

By others associated with CBOL:

44. Lee Weigt, Manager of Smithsonian Institution's Laboratory for Analytical Biology, "DNA at the SI: facilities, science, and DNA Barcoding", presented at the Chespeake Biological Laboratory, Solomons MD on 6 October:
45. Daniel Janzen, University of Pennsylvania, presentation at the National Botanical Garden, Washington, DC, 7 April 2005
46. Daniel Janzen, Keynote at the annual meeting of Society for Conservation Biology, Brasilia, Brazil, 17 July 2005; to be published as a policy paper in the April 2006 issue of "Conservation in Practice"
47. Daniel Janzen, Keynote, McGuire Center for Lepidoptera Biodiversity and Conservation, and University of Florida at Gainesville, Annual meeting of Society for Tropical Lepidoptera, 2 October 2005
48. Daniel Janzen, Princeton University Department of Biology colloquium, 30 November 2005
49. Daniel Janzen, Keynote to the annual meeting of Discover Life in America, the NGO supporting the All Taxa Biodiversity Inventory (ATBI) of Great Smokey Mountains National Park, Gatlinburg, NC, 7 December 2005.

Appendix 3. Leveraging of Sloan Foundation Support

Cash Contributions. Despite its early stage of development, CBOL has already been successful in attracting $155,000 in co-funding of its activities from sources other than the Sloan Foundation. CBOL has received or has been promised the following contributions toward its sponsored activities:

First International Barcode Conference:

- Cash and in-kind support from The Natural History Museum, London representing a total investment of $50,000
- In-kind contribution of staff time, Royal Botanic Gardens, Kew; estimated value $10,000

Library-Laboratory Meeting in London on electronic access to taxonomic literature:

- $10,000 from the Integrated Taxonomic Information System (ITIS), National Biological Information Infrastructure (NBII), US Geological Survey; and
- $20,000 from the Smithsonian Institution's Under Secretary for Science for a follow-on meeting at which the Biodiversity Heritage Library initiative was organized

Database Working Group:

- Travel and in-kind support from the National Center for Biotechnology Information (NCBI), US National Institutes of Health; estimated value $3,000
- Commitment of $10,000 from the Global Biodiversity Informatics Facility (GBIF) for a workshop on data standards

DNA Working Group Formalin Workshop:

- $12,000 from the US Environmental Protection Agency;
- $10,000 from the US Department of Agriculture;
- $10,000 from New England Biolabs, Inc.;
- $10,000 from Sigma-Aldrich;
- $10,000 from the National Evolutionary Synthesis Center, Duke University; and
- $10,000 from the Museum of Comparative Zoology, Harvard University

In-Kind and Cash Contributions. CBOL has catalyzed several important Consortium activities by providing partial funding for planning meetings. The meeting hosts provided 25-50% of the meeting costs (either as cash or in-kind support) in each of the following cases:

- ABBI Steering Committee meeting, Harvard University (5 April 2005)
- Data Analysis Working Group planning meetings at the Rutgers University Center for Discrete Mathematics and Advanced Computer Science (DIMACS, 26 September 2005);
- Data Analysis Working Group planning meetings at the National Museum of Natural History, Paris (15 November 2005);
- All Mosquitoes Barcoding Initiative Steering Committee meeting, Natural History Museum, London (21-22 November 2005); and
- Plant Working Group Steering Committee meeting, Royal Botanic Gardens, Kew, (12 December 2005)

CBOL has added value to Consortium activities by ensuring travel support for participants from developing countries. Travel stipends have been provided by CBOL for:

- 30 developing country participants at the First International Barcode Conference, London; February 2005

- 4 developing country participants at the inaugural FishBOL workshop, University of Guelph, June 2005
- 4 developing country participants at the inaugural ABBI workshop, Harvard University, September 2005

Other Investments in DNA Barcoding. CBOL's activities and support have also contributed, directly or indirectly, to several successful requests for funding. The following successful proposals were submitted with letters of support from CBOL, and they indicate the generally increasing level of support for barcoding activities:

- John Kress, Chairman of Botany in the National Museum of Natural History, Smithsonian Institution has received:
  - $30K from the museum for support of research leading to the identification of a barcode region for flowering plants, and
  - $50K from the museum and the United States Botanic Garden to "Barcode the Medicinal Plants of the World"
- The UK's Economic and Social Research Council (ESRC) has awarded a grant of £250,000 for a three-year project entitled 'Taxonomy at crossroads: Science, publics and policy in biodiversity,' which will study, in real time, the impact of barcoding (and innovations in general) on Taxonomy and Systematics.
- A proposal to the Netherlands Research Organization (NWO) for the formation of a national barcoding network, support for barcoding infrastructure, and for barcoding initiatives on fungi, insects, birds and other taxonomic groups;
- A proposal to the USDA for support of a barcoding project on food-borne nematodes and other parasites;
- Proposal to Genome Canada for support of barcoding initiatives on several taxonomic groups by the Canadian Barcode Network; and
- Proposals to the Sloan Foundation for support of planning workshops for barcoding initiative on fish (FishBOL), birds (ABBI), and marine life (Census of Marine Life).

The Barcode Initiative has attracted significant funding to taxonomy from traditional and new sources of support. The following grants are now supporting major barcoding activities:

- The Biodiversity Institute of Ontario (BIO) at the University of Guelph has received the following grants:
  - CAN$10M from the Canada Foundation for Innovation/Ontario Innovation Trust for construction of a new 20,000 sq.ft. BIO building, in which all new barcoding facilities will be located;
  - $2.4M from the Gordon and Betty Moore Foundation for three years of staff, equipment and consumables associated with the production of 225,000 barcode records from various taxonomic groups;
  - CAN$4.96M from the Natural Sciences and Engineering Research Council of Canada to establish the Canadian Barcode of Life Network and to support its activities to barcode 10,000 Canadian species. This five year grant will provide consumables and salary support to 40 researchers in several Canadian universities and government labs; and
  - CAN$5M from Genome Canada for support of barcoding initiatives by the Canadian Barcode of Life Network on fungi, plants and protists.

- The Laboratory for Analytical Biology (LAB) of the National Museum of Natural History, Smithsonian Institution, has received a $420K increase in its annual Congressional appropriations. These new funds will support two new FTEs (a project manager and technician for barcoding) and three annual acquisitions of robotic systems for high-throughput barcoding (ca. $200K each). Starting with FY 2008, the $200K previously used for instrument acquisition will be devoted to barcoding projects.
- The bird identification lab in the Smithsonian's National Museum of Natural History received:
  - $125K from the US Air Force in 2003 for acquisition of a capillary sequencing system ABI 3100 which is being used in the LAB for barcoding birds and other taxonomic groups. In addition, the bird identification lab receives
  - $100K per year for staff and consumables from the Federal Aviation Administration (FAA), part of which is devoted to barcoding of bird remains following airplane strikes. The Air National Guard provided an additional $10K in 2004 to increase the number of bird barcodes that could be obtained
  - $85K from the FAA to resolve the taxonomy of approximately 50 gull species using COI and additional mitochondrial gene regions
- The Directorate General for Research of the European Commission has awarded €11.8M for a five year effort to create the European Distributed Institute of Taxonomy, a European Network of Excellence that includes 16 major museums and herbaria in the EU, two in Russia and two in the US. One of the eight EDIT Work Packages has a large barcoding component.
- The Royal Belgian Institute of Natural Sciences has received €136K from the Belgian Lottery for a 4-capillary sequencer and €12K per year from the Belgian Development Agency for barcoding projects in Africa.
- Dr. Daniel Janzen of the University of Pennsylvania received $250K from the Wege Foundation for support of barcoding initiatives on the insect fauna in Costa Rica.

Based on requests for supporting letters from CBOL and other evidence, we have reason to believe that many other proposals involving barcoding have been submitted to research funding agencies during 2005.

Appendix 4. CBOL Governance

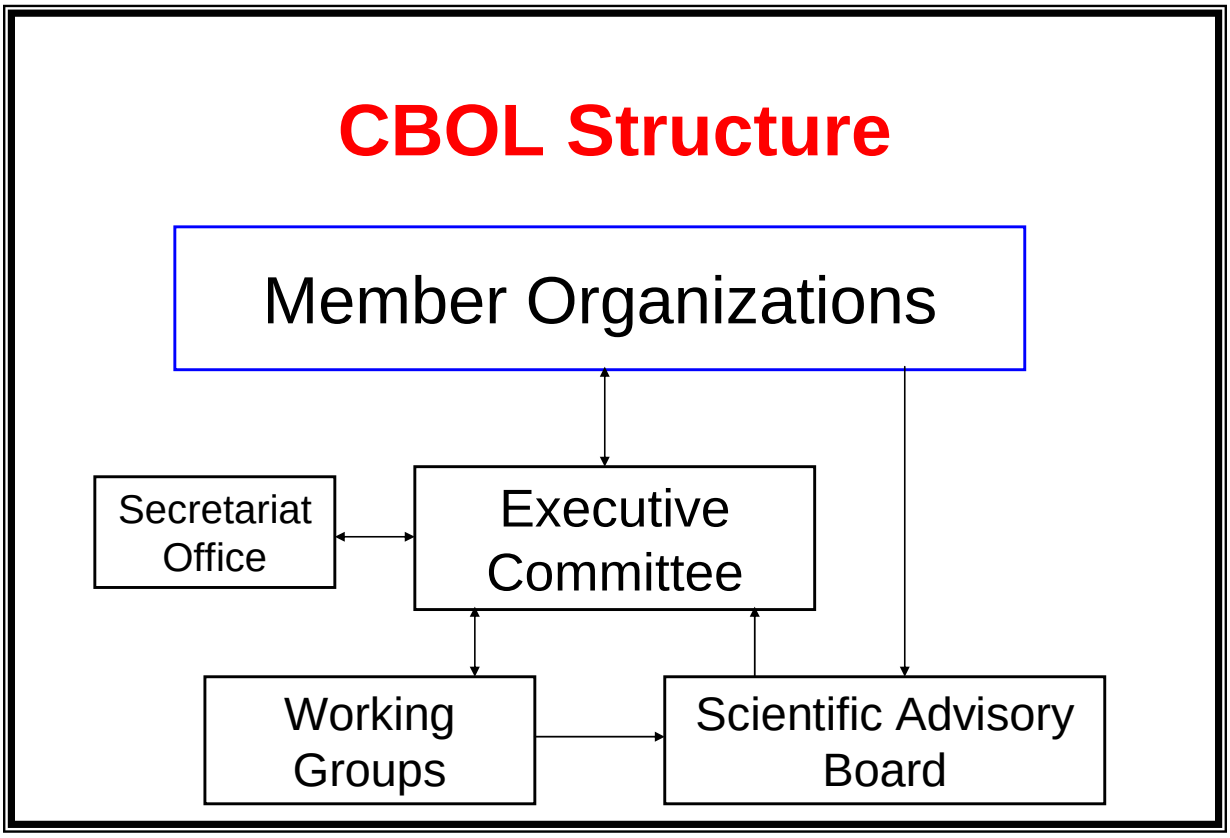


Table A. Members of the CBOL Executive Committee

Scott Miller, Chair	Associate Director for Science, National Zoological Park	Smithsonian Institution, Washington, DC	USA	
José Alves-Gomes	Director	National Institute of Amazonian Research, Manaus	Brazil	Appointed 7/1/05 Term ends 6/19/07
James Edwards	Executive Director	Global Biodiversity Informatics Facility, Copenhagen	Denmark	Appointed 7/1/05 Term ends 6/19/06
Richard Lane	Director of Science	The Natural History Museum, London	United Kingdom	Appointed 7/1/05 Term ends 6/19/06
Helida Oyieke	Director of Research	National Museums of Kenya, Nairobi	Kenya	Appointed 7/1/05 Term ends 6/19/07
Simon Tillier	Coordinator, European Distributed Institute of Taxonomy	National Museum of Natural History, Paris	France	Appointed 7/1/05 Term ends 6/19/07
David Schindel	Executive Secretary	CBOL, Washington, DC	USA	

Table B. Members of the CBOL Scientific Advisory Board

James Hanken, Chair	Harvard University	USA	amphibians	Appointed 8/1/05 Term ends 7/31/06
Freek Bakker, Vice Chair	National Herbarium	Netherlands	plants, molecular techniques	Appointed 8/1/05 Term ends 7/31/07
Karen Armstrong	Lincoln University	New Zealand	invasive species, molecular markers	Appointed 8/1/05 Term ends 7/31/07
Sandro Luis Bonatto	Pontifícia Universidade Católica do Rio	Brazil	Molecular evolution	Appointed 8/1/05 Term ends 7/31/06
Robyn Cowan	Royal Botanic Gardens Kew	UK	plants	Chair, Plant Working Group
Robert Hanner	University of Guelph	Canada	bioinformatics, insects	Chair, Database Working Group
Paul Hebert	University of Guelph	Canada	molecular techniques, insects	Chair, DNA Working Group
Athula L.T. Perera	University of Peradeniya	Sri Lanka	agricultural biology, plant breeding	Appointed 8/1/05 Term ends 7/31/07
Cecilia Saccone	CNR Institute of Biomedical Technologies	Italy	biochemistry, bioinformatics	Appointed 8/1/05 Term ends 7/31/07
Ole Seberg	University of Copenhagen	Denmark	plants	Appointed 8/1/05 Term ends 7/31/06
Pablo Luis Tubaro	Museum of Natural Sciences	Argentina	birds	Appointed 8/1/05 Term ends 7/31/06
Michel Veiulle	National Museum of Natural History, Paris	France	population genetics	Chair, Data Analysis Working Group

The Executive Committee is in the process of selecting the final three members of the Scientific Advisory Board to represent Africa, southeast Asia and Australia.

Appendix 5. Abstracts of Four Significant Barcode Publications

Hebert, P.D.N., Cywinska, A., Ball, S.L., and deWaard, J.R. (2003a). "Biological identifications through DNA barcodes." *Proceedings of the Royal Society London B* 270:313-321.

Although much biological research depends upon species diagnoses, taxonomic expertise is collapsing. We are convinced that the sole prospect for a sustainable identification capability lies in the construction of systems that employ DNA sequences as taxon 'barcodes'. We establish that the mitochondrial gene cytochrome *c* oxidase I (COI) can serve as the core of a global bioidentification system for animals. First, we demonstrate that COI profiles, derived from the low-density sampling of higher taxonomic categories, ordinarily assign newly analysed taxa to the appropriate phylum or order. Second, we demonstrate that species-level assignments can be obtained by creating comprehensive COI profiles. A model COI profile, based upon the analysis of a single individual from each of 200 closely allied species of lepidopterans, was 100% successful in correctly identifying subsequent specimens.When fully developed, a COI identification system will provide a reliable, cost-effective and accessible solution to the current problem of species identification. Its assembly will also generate important new insights into the diversification of life and the rules of molecular evolution.

Hebert, P.D.N., Stoeckle, M.Y., Zemlak, T.S., and Francis, C.M. (2004a). "Identification of birds through DNA barcodes." *Public Library of Science, Biology* 2: e312.

Short DNA sequences from a standardized region of the genome provide a DNA barcode for identifying species. Compiling a public library of DNA barcodes linked to named specimens could provide a new master key for identifying species, one whose power will rise with increased taxon coverage and with faster, cheaper sequencing. Recent work suggests that sequence diversity in a 648-bp region of the mitochondrial gene, cytochrome c oxidase I (COI), might serve as a DNA barcode for the identification of animal species. This study tested the effectiveness of a COI barcode in discriminating bird species, one of the largest and best-studied vertebrate groups. We determined COI barcodes for 260 species of North American birds and found that distinguishing species was generally straightforward. All species had a different COI barcode(s), and the differences between closely related species were, on average, 18 times higher than the differences within species. Our results identified four probable new species of North American birds, suggesting that a global survey will lead to the recognition of many additional bird species. The finding of large COI sequence differences between, as compared to small differences within, species confirms the effectiveness of COI barcodes for the identification of bird species. This result plus those from other groups of animals imply that a standard screening threshold of sequence difference (103 average intraspecific difference) could speed the discovery of new animal species. The growing evidence for the effectiveness of DNA barcodes as a basis for species identification supports an international exercise that has recently begun to assemble a comprehensive library of COI sequences linked to named specimens.

Hebert, P.D.N., Penton, E.H., Burns, J.M., Janzen, D.H., and Hallwachs, W. (2004b). "Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*." *Proceedings of the National Academy of Science* 101:14812-14817.

Astraptes fulgerator, first described in 1775, is a common and widely distributed neotropical skipper butterfly (Lepidoptera: Hesperiidae). We combine 25 years of natural history observations in northwestern Costa Rica with morphological study and DNA barcoding of museum specimens to show that *A. fulgerator* is a complex of at least 10 species in this region. Largely sympatric, these taxa have mostly different caterpillar food plants, mostly distinctive caterpillars, and somewhat different ecosystem preferences but only subtly differing adults with no genitalic divergence. Our results add to the evidence that cryptic species are prevalent in tropical regions, a critical issue in efforts to document global species richness. They also illustrate the value of DNA barcoding, especially when coupled with traditional taxonomic tools, in disclosing hidden diversity.

Gary W. Saunders, G.W. 2005. Applying DNA barcoding to red macroalgae: a preliminary appraisal holds promise for future applications. Phil. Trans. R. Soc. B (2005) 360, 1879–1888.

Marine macroalgae, especially the Rhodophyta, can be notoriously difficult to identify owing to their relatively simple morphology and anatomy, convergence, rampant phenotypic plasticity, and alternation of heteromorphic generations. It is thus not surprising that algal systematists have come to rely heavily on genetic tools for molecular assisted alpha taxonomy. Unfortunately the number of suitable marker systems in the three available genomes is enormous and, although most workers have settled on one of three or four models, the lack of an accepted standard hinders the comparison of results between laboratories. The advantages of a standard system are obvious for practical purposes of species discovery and identification; as well, compliance with a universal marker, such as cox1 being developed under the label 'DNA barcode', would allow algal systematists to benefit from the rapidly emerging technologies. Novel primers were developed for red algae to PCR amplify and sequence the 50 cox1 'barcode' region and were used to assess three known species-complex questions: (i) Mazzaella species in the Northeast Pacific; (ii) species of the genera Dilsea and Neodilsea in the Northeast Pacific; and (iii) Asteromenia peltata from three oceans. These models were selected because they have all caused confusion with regards to species number, distribution, and identification in the field, and because they have all been studied with molecular tools. In all cases the DNA barcode resolved accurately and unequivocally species identities and, with the enhanced sampling here, turned up a variety of novel observations in need of further taxonomic investigation.

Appendix 6. Data Standards for Barcode Records

Proposed Standards for BARCODE Records in INSDC (BRIs)

By Robert Hanner, Chair
Database Working Group, Consortium for the Barcode of Life

November 6, 2005

Background. The Consortium for the Barcode of Life (CBOL) formed a Database Working Group (DWG) at its inaugural meeting, held at the Smithsonian Institution in May 2004. The DWG was created to pursue one of CBOL's principal goals: a global reference library of DNA barcode sequences that is integrated with other systems of biodiversity information (e.g., databases of specimens, species, biogeographic information). At this inaugural meeting, the DWG participants and Chair agreed that DNA Barcode data should be archived in the public domain, preferably by the International Nucleotide Sequence Database Collaboration (INSDC)[3]. At this initial meeting, the DWG also endorsed the need to link barcode records to voucher specimens and valid species names. In September of 2004, the DWG convened a meeting on the campus of the US National Institutes of Health, hosted by GenBank at the National Center for Biotechnology Information (NCBI). This meeting outlined a proposal for new data standards that would apply to DNA barcode records submitted to INSDC members in the future. In April 2005, the DWG consulted with representatives of leading taxonomic initiatives[4] and refined its data standards proposal based on their input. In May 2005, GenBank presented the proposal at the INSDC annual meeting where it was greeted with strong support and swift approval. The DWG subsequently met with representatives of major museum database initiatives to discuss implementation of the proposed data standards[5]. Participants at this meeting endorsed the proposed standards without reservation.

If the following proposal is approved, the DWG would work with NCBI to develop a more detailed set of user guidelines, to be posted on the CBOL and INSDC websites.

Proposal. The proposed standards include three major components:

1) **Creation of a reserved keyword ("BARCODE")**. NCBI and its collaborators will add the BARCODE 'flag' to new submissions that meet the standards established in consultation

[3] GenBank, the European Molecular Biology Laboratory (EMBL), and the DNA Data Bank of Japan (DDBJ)

[4] DWG meeting at the Smithsonian Institution's Center for Research and Conservation, Front Royal, Virginia, 27-29 April 2005. Participants represented: The University of Guelph Barcode of Life Database (BoLD); Species2000; Integrated Taxonomic Information System (ITIS); the Global Biodiversity Information Facility (GBIF); the Duke University National Evolutionary Synthesis Center (NESCENT); NCBI; the Ocean Biogeographic Information System, the Census of Marine Life; ZooRecord of Thomson Publishing; International Plant Names Index (IPNI); iPlants; the International Commission on Zoological Nomenclature (ICZN); uBio of the Marine Biological Lab, Woods Hole; the National Biological Information System of the US Geological Survey; the US Department of Agriculture's GRIN database; the Natural History Museum, London; the Royal Botanic Gardens, Kew; the Smithsonian Institution; and CBOL.

[5] DWG meeting at NCBI on 3 October 2005. Participants represented: the Global Biodiversity Information Facility (GBIF); the Zoological Information Management System (ZIMS); BoLD; the Taxonomic Database Working Group (TDWG); NESCENT; CBOL; and database initiatives at the University of California, Berkeley Museum of Vertebrate Zoology; the University of Kansas Biodiversity Research Center; University of Alaska Museum.

with CBOL. Data records that meet these criteria will be known as BARCODE records in INSDC (BRIs);

2) **Required data elements.** DWG proposes that the following data elements be required of all BRIs. In requiring these data, DWG seeks to provide the user community with reliable, retrievable and verifiable information concerning the barcode sequence itself, the specimen from which it was obtained, and the species name that was applied by the submitter.

DWG proposes that each BRI must:

a) Include a link to a voucher specimen using a structured field specified by CBOL and NCBI[6], and to the metadata associated with that specimen and contained in the public database of the voucher specimen's repository.
b) Include a link to a documented species name found in one of the sources specified by CBOL and NCBI[7];
c) Include Country-Code, using the controlled vocabulary used by GenBank;
d) Come from a gene region accepted by CBOL as an effective barcode (see process for approving candidate barcode regions, 3b, below). Initially, only cytochrome c oxidase 1 is approved as a barcode region, defined relative to the mouse mitochondrial genome as the 648 bp region that starts at position 58 and stops at position 705.
e) Include at least 500 contiguous unambiguous base-pairs from bidirectional sequencing within the approved barcode region. However, if requested, GenBank could assign the BARCODE flag to records with shorter sequences following guidelines defined by CBOL (see 3a, below);
f) Include no more than 1% ambiguous sites for the entire submitted sequence;
g) Include the name of the gene region used;
h) Be associated with trace file submitted to the NCBI Trace Archive or the Ensembl Trace Server; and
i) Include the sequences of all forward and reverse primers used. For records in which the contiguous sequence was assembled from more than one amplicon or when a cocktail of multiple primers was used for amplification, multiple sets of primer pairs must be provided. In addition, submission of the names of the forward and reverse primers with the primer sequences is strongly recommended.

Strongly recommended data elements. The following data elements have been added to the INSDC at CBOL's request for validation of the voucher specimen, and will be strongly recommended but not required:

j) Latitude and longitude;
k) Name of the identifier;
l) Name of the collector; and
m) Date of collection.

3) **Governance rules**. The INSDC provides an archive of records that can only be changed by the submitter. In the case of BRIs, the following modifications to the rules governing changes to data records are proposed to assure and maintain data quality and consistency:

[6] The voucher specimen identifier uses a triplet structure (institution|collection|item) as used in the DarwinCore, advocated by GBIF. This triplet field is parallel to the Life Science Identifier (LSID) that is an Object Management Group (OMG) standard.

[7] DWG proposes a hierarchy of sources of species names, including vetted checklists such at Catalog of Life, nomenclators such as IPNI and the Zoological Record; lists of all published names such as uBio and the proposed NameBank; recent publications that have not yet been incorporated into compilations; and pre-publication data resources.

a) CBOL will define the circumstances under which records shorter than the 500 base-pair minimum could be BRIs. These might include sequences from type specimens or specimens of extinct or extremely rare species;
b) CBOL will be responsible for establishing, implementing, and offering a process whereby research groups could propose and justify a non-COI gene region to which the BARCODE flag can be given;
c) BRIs that are assembled on and submitted to GenBank from the University of Guelph's Barcode of Life Database (BoLD) will be considered by GenBank to be submitted jointly by the individual researcher and BoLD. These records can be modified by either party;
d) BRIs that are submitted to GenBank directly by individual researchers may only be modified by the submitter. However, GenBank will remove the BARCODE flag from these records at CBOL's recommendation. These records would remain in GenBank as non-BARCODE records; and
e) DWG and NCBI will develop a proposal to CBOL for attaching third-party comments, criticisms, and suggested corrections to BRIs, thereby providing the research community with additional quality indicators. These third-party comments would also support CBOL's review of BRIs from which the BARCODE flag might be removed.

Appendix 7. The DNA Barcoding Process

DNA barcoding involves the following steps. The time and cost associated with each step is shown below, along with possible improvements that CBOL is promoting through partnerships with private biotech companies. The time and cost per step shown below apply to fresh specimens. The process is slower and more expensive for specimens with degraded DNA, which is the case for older museum specimens and specimens that were fixed in formalin.

Step in Barcoding Process	Time required (and cost)	Potential Improvement
1. DNA extraction: dissolving or grinding biological sample and isolating DNA from residue	24-48 hours for fresh specimens by traditional methods, 12-24 with robotic extractors; faster reagents reduce processing time to a few hours; ($.34 per sample for quick extractions, $.75-1.00 for extractions that produce DNA samples that can be banked and used for other analyses)	Principal challenge is extraction of DNA from formalin-fixed specimens and museum specimens older than 5-10 years
2. PCR amplification: thermal cycling produces millions of copies of the target region flanked by "primers"[8]	Normally 2-3 hours ($.24); 30 minutes on robotic systems ($.10 - .20/specimen)	Experimental microfluidic systems are reducing the PCR process to minutes. Highly portable PCR systems have been developed and could become hand-held in the next few years
3. Verification of PCR product: this step is needed to check the success of the PCR process during primer design and for non-fresh specimens. In other cases, this step is not necessary	Several hours using gel electrophoresis (pennies to $.34/specimen); robotic, miniaturized, and DNA-chip systems are becoming available that work in less than one hour ($.10/specimen)	Further miniaturization and robotization could streamline this step significantly
4. Clean-up of amplification products: removal of unwanted primer nucleotides and dyes that will interfere with sequencing	Several commercial kits are available that perform this step in 30 minutes ($.32-64/specimen with traditional equipment, $.05-.15 with robotic systems)	High volume purchases of kits through CBOL and use of robotics could reduce the cost of this step significantly for barcoding lab

[8] Primers are short sequences that flank the barcode gene region. "Universal" primers work in many cases but customized primer design is often needed when starting barcoding in new taxonomic groups. This process can cost $100 in materials per group.

5. Sequencing reaction: inserting labeled nucleotide bases that can be read by the sequencer	2-4 hours, $.50/specimen)	Commercial kits are available but expensive, so reducing the volume of reagent used is critical to containing costs. High volume purchases and development of smaller, faster robotic systems
6. Sequencing: wide variety of techniques for determining nucleotide sequence, ranging from single gels to high-throughput capillary sequencers	Oldest technologies produce a few sequences per day; best available capillary sequences can process 384 specimens in 1-3 hours ($.25-.80/specimen)	Table-top and hand-held sequencers are under development and prototype systems will appear in next 3-5 years
Mimimum time (and cost)	32 hours ($1.35)	

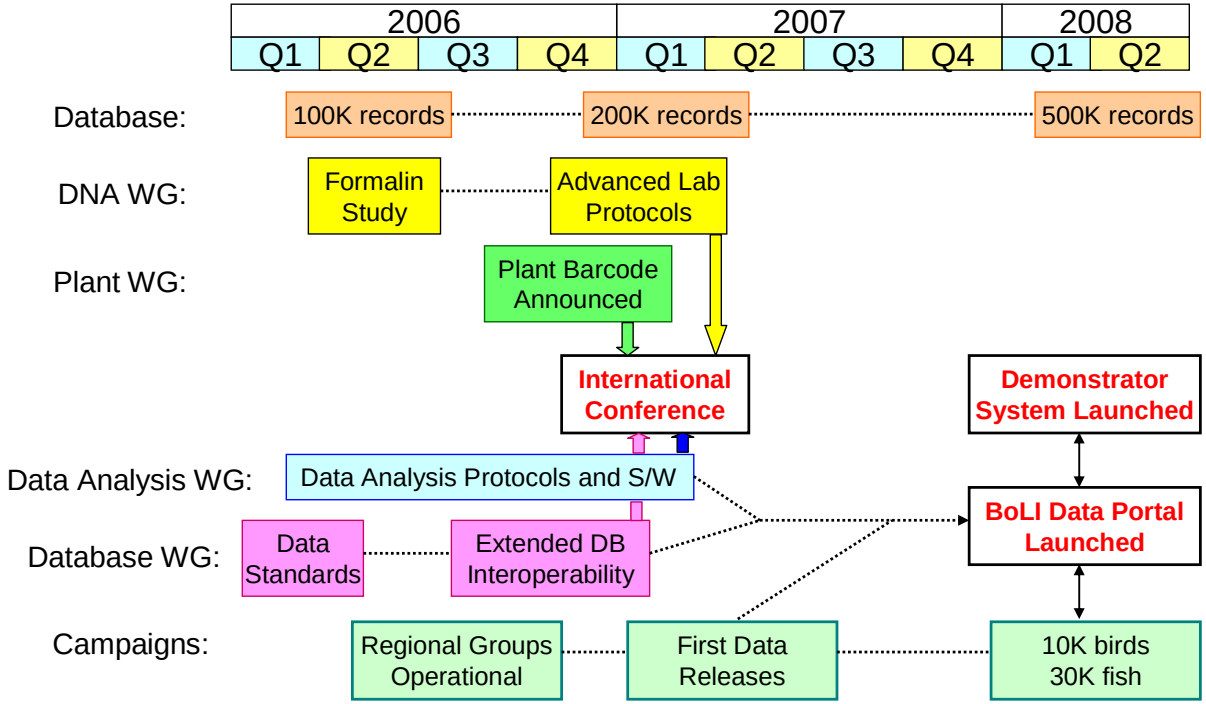
Appendix 8. CBOL Status in 2006, Ambitions for 2008

The following table summarizes the status of CBOL and its activities in 2006, and the new activities it seeks to implement by the end of the proposed support period on 30 June 2008.

	2006	2008
Major Participants	High-throughput barcoding labs at Univ. Guelph, Smithsonian; national barcoding networks in Canada, Netherlands; 4 CBOL Working Groups (ca. 100 participants); 220 attendees at 2005 London conference	Barcoding labs, regional and national barcoding networks on every continent; 500 attendees at 2007 international conference
Major projects	FishBOL and ABBI campaigns (20 regional Working Groups); International Network for Barcoding Invasive and Pest Species; project to identify optimal plant barcode region	Major projects involving public health (e.g., mosquitoes), agriculture (fruitflies), conservation (primates and bushmeat)
CBOL Membership	100+ Member Organizations from 40 countries; 30 MOs from 20 developing countries	200+ Member Organizations from 60+ countries; 60+ MOs from 40+developing countries
Production of barcode data	High-throughput labs produce 50K records per year, $2-4/specimen for materials	Several labs produce 75-100K records/year at under $1 per sample; first table-top barcoding systems developed
Sources of barcode data	Newly collected specimens; museum specimens less than 5 years old; frozen tissue/cells	Older museum specimens; formalin-fixed specimens
Analysis of barcode data	Primarily by K2 sequence similarity and Neighbor Joining trees	New protocols for sampling, analysis, interpretation, and display/visualization, including both similarity and character-based barcode data
Barcode Database	Data standards developed for barcode records in INSDC	Data standards lead to greater interoperability of databases, especially GenBank, GBIF, and Catalog of Life
Access to barcode data	BoLD is primary workbench for barcode projects; barcode records in BoLD are submitted to GenBank and are accessible on both databases	BoLD and its mirror sites on different continents exchange data with GenBank, EMBL and DDBJ; BoLI Data Portal provides user access to analytical tools and barcode data in all four databases

Appendix 9.

Milestones for 2008



Reaching milestones such as 500K total barcode records, and 10K bird and 30K fish barcodes is contingent on the availability of research funding from non-Sloan sources. Numbers of database records refer to public barcode records in GenBank, EMBL and DDBJ, plus public and private barcode records in BoLD. BoLD records are made public (either on BoLD or by submission to GenBank) when the investigator who generated these data approves their release following publication of related papers. Participants in CBOL-initiated projects are urged to make their barcode records public as quickly as possible, and many have agreed to release their data immediately.

Appendix 10: Outreach Strategies

CBOL has initiated partnerships with organizations in several different sectors of society on several continents, and these efforts will be accelerated greatly during the two years of activity proposed here. The following three strategies have proven effective and will be used in the future:

- Presentations by members of CBOL's Executive Committee and Scientific Advisory Board will be an important outreach mechanism. A small fund is requested for continued support of travel by CBOL's representatives to give presentations to a variety of audiences.
- CBOL's Executive Secretary gave 34 presentations to academic, government and industrial audiences during his 16 months of service. Travel funds are requested to continue these outreach efforts.
- CBOL will continue to seek opportunities to organize barcoding sessions in larger meetings and to schedule meetings back-to-back with related meetings. These strategies have proven effective in disseminating and gathering information to/from related communities, and engaging them in the Barcode Initiative.

The following specific strategies will be used to strengthen links with and promote activities in specific sectors.

Academic Taxonomic Community. CBOL's Scientific Advisory Board will take the lead in identifying international organizations with which CBOL should establish closer links. SAB members will identify international meetings at which CBOL should propose and organize sessions on DNA Barcoding.

Private Sector Biotech Companies. Barcoding will accelerate through the development of more advanced instruments, reagents and processes. CBOL intends to continue its efforts to promote technology development by contacting private biotech companies and informing them of the potential market for their products for species identification. A list of government and academic barcode researchers with special interests in technology development has also been compiled. CBOL will convene these researchers as an electronic discussion community with the goal of gathering information on companies that should be contacted, new technologies that should be explored, and new opportunities for outreach that should be pursued.

Agriculture and Environment. CBOL's International Network for Barcoding Invasive and Pest Species (INBIPS) has done an electronic survey of research institutes and government agencies, and has compiled a database of more than 50 relevant contacts. This contact list will be used to disseminate information on barcoding, gather information on how barcoding is being used or could be used, and to organize international activities. CBOL is also convening a Steering Committee that will provide advice on a possible fruit fly barcode project. The US Department of Agriculture's Animal and Plant Health Inspection Service (APHIS) and several counterpart agencies in other countries have expressed keen interest in the project.

Public Health. CBOL supported a meeting of a Steering Committee for a possible mosquito barcoding project. This committee included leading experts from several continents, and they are formulating a plan for a two-year initiative. This plan will include the major partners that will be needed to provide specimens, lab capacity, and funding. The Wellcome Trust and the Bill and Melinda Gates Foundation are two possible sources of funding for the initiative.

Conservation. CBOL has convened a Steering Committee on Barcoding for Endangered Species Conservation that will provide recommendations on potential future directions. These will include possible barcoding projects (such as African bushmeat or Asian turtles) and the agencies and organizations with which CBOL should build partnerships.

SCOTT EVERETT MILLER

Associate Director for Science, National Zoological Park, and Curator of Entomology, National Museum of Natural History, Washington, DC 20560-0105, USA. Phone: 202-633-1036. Fax: 202-786-3141, millers@si.edu

Professional Preparation:

University of California, Santa Barbara, B.A. Biology, 1981
Harvard University, Cambridge, Massachusetts, Ph.D. Biology, 1986

Professional Appointments:

Smithsonian Institution, Washington, DC
Associate Director, NZP, 2004-
Supervisory Research Entomologist, NMNH, 1998-2004 (on leave 1998-1999)
Collaborator, NMNH, 1983-1987; Research Associate, NMNH, 1987-1997
International Centre of Insect Physiology and Ecology (ICIPE), Nairobi
Principal Scientist and Programme Leader, 1998-1999
Bernice Pauahi Bishop Museum, Honolulu, Hawaii
Research Associate, 1998-
Chair, Department of Natural Science, 1993-1997; Chair, Department of Entomology, 1986-1993
University of Hawaii, Honolulu; Affiliate Graduate Faculty, Dept. of Entomology, 1987-
Santa Barbara Museum of Natural History, Santa Barbara, CA; Research Associate, 1975-
Natural History Museum of Los Angeles County, Los Angeles, CA; Research Associate, 1977-

Grants, Fellowships and Contracts (over 6 million dollars total):

National Science Foundation, Research Grants
DEB 0515678 (coPI 2005-2008) [insect herbivore ecology in Papua New Guinea]
DEB 0211591 (coPI 2002-2005) [insect herbivore ecology in Papua New Guinea]
DEB 9707928 (PI 1997-2002) [insect herbivore ecology on Rubiaceae in Papua New Guinea]
DEB 9629751 (PI 1996-1997) [insect herbivore ecology on euphorbs in Papua New Guinea]
DEB 9528025 (PI 1996-1999) [systematics and biogeography of aquatic insects in New Guinea]
DEB 9407297 (PI 1994-1996) [insect herbivore ecology on *Ficus* in Papua New Guinea]
BSR 8413586 (Co-PI 1985-1987) [systematics of Dalceridae (Dissertation)]
National Science Foundation, Research Collections in Systematics and Ecology Grants,
BSR 8511669, 8706426, 8913871, DEB 9423814 (PI 1986-1997) [Bishop Museum entomology]
BSR 8912364 and supplement (PI 1993-1994) [Bishop Musuem herbarium]
BSR 8313189 (PI 1988-1989) [Catalog of Australasian and Oceanic Flies]
National Science Foundation, Database Activities Grants,
DBA 9631091 (co-PI with John Helly *et al.* 1996-1999) [WWW based data management]
National Institutes of Health, ICBG, Co-PI with Louis Barrows et al., 2003-2008
Conservation and sustainable use of biodiversity in Papua New Guinea
Sloan Foundation, Consortium for the Barcode of Life Secretariat [2004-2006]
Other academic grants: Hawaii Bishop Research Institute, New England Biolabs Foundation, Santa Cruz Island Research Fund, Smithsonian Institution, Harvard University, Mellon Foundation, Nando Perreti Foundation, National Academy of Sciences, American Philosophical Society, Global Biodiversity Information Facility
Applied conservation grants: MacArthur Foundation, Mellon Foundation, UNEP, GEF, IDRC
Environmental survey and research contracts: US Fish & Wildlife Service, Space Biospheres Ventures, PT Freeport Indonesia, US Department of Agriculture

Representation on significant committees and review panels:

International (representing US Government and/or Smithsonian): Convention on Biological Diversity (especially Global Taxonomy Initiative), Global Biodiversity Information Facility, Global Invasive Species Program, BioNet International

Federal government: Committee on Environment and Natural Resources, National Invasive Species Council working groups, other interagency groups related to invasive species and emerging diseases

Private: National Geographic Society, Travelers Conservation Foundation

Board of Directors: American Entomological Institute, Consortium for the Barcode of Life, Indo-Pacific Conservation Alliance, Xerces Society

Honorary members in professional organizations:

American Association for the Advancement of Science (Fellow); Royal Entomological Society of London (Fellow)

Related publications:

Novotny, V., S. E. Miller, Y. Basset, L. Cizek, K. Darrow, B. Kaupa, J. Kua, and G. D. Weiblen. 2005. An altitudinal comparison of caterpillar (Lepidoptera) assemblages on Ficus trees in Papua New Guinea. *Journal of Biogeography* 32: 1303-1314.

Basset, Y., V. Novotny, S. E. Miller, G. D. Weiblen, O. Missa, and A. J. A. Stewart. 2004. Conservation and biological monitoring of tropical forests: The role of parataxonomists. *Journal of Applied Ecology* 41: 163-174.

Miller, S. E., and L. M. Rogo. 2002. Challenges and opportunities in understanding and utilisation of African insect diversity. *Cimbebasia* 17: 197-218 ("2001").

Miller, S. E., W. J. Kress, and C. Samper K. 2004. Crisis for biodiversity collections. *Science* 303: 310.

Novotny, V., Y. Basset, S. E. Miller, R. L. Kitching, M. J. Laidlaw, P. Drozd, and L. Cizek. 2004. Local species richness of leaf-chewing insects feeding on woody plants from one hectare of a lowland rainforest. *Conservation Biology* 18: 227-237.

Other Significant Publications (143 in total):

Sekhran, N. & Miller, S.E. (eds) 1995. *Papua New Guinea Country Study on Biological Diversity*. Papua New Guinea Department of Environment and Conservation, Waigani. xl + 438 pp.

Keast, A. & Miller, S.E. (eds) 1997. *The origin and evolution of Pacific island biotas, New Guinea to Eastern Polynesia: Patterns and processes*. SPB Publications, Amsterdam. viii + 531 pp.

Holloway, J.D., Kibby, G. Peggie, D., Carter, D. & Miller, S.E. 2001. *Families of Malesian moths and butterflies*. Fauna Malesia Handbook Series. Brill, Leiden. xii + 456 pp.

Basset, Y., V. Novotny, S. E. Miller, and R. L. Kitching, eds. 2003. Arthropods of tropical forests: Spatio-temporal dynamics and resource use in the canopy. Cambridge University Press, Cambridge. xvi + 474 pp.

Novotny, V., Basset, Y., Miller, S.E., Weiblen, G., Bremer, B., Cizek, L. & Drozd, P. 2002. Low host specificity of herbivorous insects in a tropical forest. *Nature* 416: 841-848.

David E. Schindel, Ph.D.

Education

University of Michigan, B.S. in Geological Sciences	1973
Harvard University, Ph.D. in Geological Sciences	1979
Federal Executive Institute: Leadership for a Democratic Society	1995

Experience

National Science Foundation, July 1986 to present:

Executive Secretary, Consortium for the Barcode of Life (on detail from NSF to the Smithsonian Institution's National Museum of Natural History)	September 2004-present
Head, National Science Foundation's Europe Office	July 1998-August 2004):
Executive Associate, Office of Integrative Activities	January-June 1998
Legislative Fellow for Senator Jeff Bingaman (D-NM) on detail from NSF to the United States Senate during a one-year developmental assignment	January-December 1997:
Senior Science Advisor, Office of Science & Technology Infrastructure	1993-1996
Program Director, Elementary, Secondary, and Informal Education	1991-1993
Program Director, Biological Research Resources	1989-1991
Associate Program Director, Systematic Biology	1986-1989

Yale University, 1978-1986

Associate Professor, Department of Geology & Geophysics	1984-1986
Assistant Professor Department of Geology & Geophysics	1978-1984
Curator of Invertebrate Fossils, Yale Peabody Museum	1978-1986

Smithsonian Institution, 1977-1978

Predoctoral Fellow in Paleobiology, US National Museum of Natural History

Other Relevant Experience

Advisor on elementary school science to Montgomery County Public Schools (Maryland)	1990
Member of a Maryland State Department of Education task force on graduation requirements	1992

Publications

Most important articles:

Schindel, D.E. and S.J. Gould. 1977. Biological interaction between fossil species: character displacement in Bermudian land snails. Paleobiology 3:259-269.

Schindel, D.E. 1980. Microstratigraphic sampling and the limits of paleontologic resolution. Paleobiology 6:408-426.

Vermeij, G.J., D.E. Schindel and E. Zipser. 1981. Predation through geologic time: evidence from gastropod shell repair. Science 214:1025-1026.

Schindel, D.E. 1982. Resolution analysis: A new approach to the gaps in the fossil record. Paleobiology. 8:340-353.

Schindel, D.E. 1982. Punctuations in the Pennsylvanian evolutionary history of Glabrocingulum (Mollusca: Archaeogastropoda). Bull. Geol. Soc. Amer. 39:400-408.

Schindel, D.E., G.J. Vermeij and E. Zipser. 1982. Frequencies of repaired shell fractures among the Pennsylvanian gastropods of north-central Texas. Jour. Paleontol., 56:729-740.

Schindel, D.E. 1982. The Gaps in the Fossil Record. Nature., 297:282-284.

Schindel, D.E., 1991, Unoccupied Morphospace and the Coiled Geometry of Gastropods: Architectural Constraint or Geometric Covariation?, in Allmon, W.D. and Ross, R., eds., Causes of Evolution, Univ. Chicago Press, 270-304.

Most recent articles:

Schindel, D.E., 2001, The American Academic Marketplace, in "German and American Higher Education: Educational philosophies and Political Systems", conference convened by the Bavarian-American Academy, Transaction Press, Rutgers University, Piscataway, NJ, pp.163-172.

Schindel, D.E. and S.E. Miller, 2005, Barcoding a useful tool for taxonomists, letter to Nature letter, 435:17

Schindel, D.E., in press, Consortium for the Barcode of Life: A rapid, cost-effective system for species identification, Proceedings of the International Conference on Biodiversity, Paris, 24-28 January 2005 under the patronage of the Ministry delegated to Research and of UNESCO.

Budget Justification

The Sloan Foundation's initial grant to CBOL was for $669,000 over the period of 1 April 2004 to 30 June 2006. The present proposal requests two years of support and a significant increase in the annual budget during that interval. The overall justification for this increase is straightforward. CBOL has been very successful during the initial award period in organizing and launching a variety of activities. The next phase of the Barcode Initiative will be devoted to implementing, promoting participation in, managing and ensuring the success of these activities. Achieving these goals will require an increase in the staff of CBOL's Secretariat and increased support for many activities.

Cost-Sharing

With the exception of staff resources in the Secretariat Office, nearly all of the activities supported under the proposed budget will rely on partial support from sources other than the Sloan Foundation. In many cases, CBOL's ability to conduct proposed activities will be contingent on the success of fund-raising efforts. In deciding which costs to include in this proposal, and which funds should be sought from non-Sloan sources, CBOL has been guided by the following principles:

- CBOL's core mission is to promote and facilitate barcoding activities by others, not to conduct barcoding projects.
- The Sloan Foundation's support for the CBOL Secretariat and its core functions (meetings, conferences, working groups) is a critical complement to funding from agencies that support basic and applied research, and international development. CBOL recognizes and respects the distinction among these different sources of funding and seeks funds from each source, according to their respective missions. For many of the items described in the following sections, CBOL is requesting only that part of the funds for an activity that will be needed for administration and as a catalyst. The remaining portions will be sought from research agencies, international development funds, private companies, or other sources that are appropriate to that activity.
- The long-term success of barcoding will rely on support from the users of barcode data (especially government agencies). CBOL's past activities have set the stage for their involvement and investments during the coming two years.
- Barcoding labs represent a significant potential market for instruments, reagents, and other lab consumables. Private sector companies will want to know about barcoding, and they have already shown their willingness to invest in the initiative.
- Barcoding projects can be highly distributed, involving researchers, collections and laboratories in different geographic regions. CBOL tries to impress on participants the need for each of them to seek funding for their component of the project, rather than relying on someone else to raise funds for all components.

Staff Resources

CBOL Executive Secretary. The proposed responsibilities of the Executive Secretary will change during the second grant period. CBOL now has more than 100 Member Organizations, an Executive Committee, a Scientific Advisory Board, four Working Groups, a variety of barcoding campaigns and projects, and a plan to produce a working demonstrator system for species identification in two years. The Executive Secretary will need to:

- oversee and coordinate all these activities,
- promote international participation in each,
- play the leading role in organizing and implementing CBOL's ambitious program of outreach to developing countries,
- continue as a member of CBOL's Executive Committee,
- act as the primary point of contact for the Executive Committee and Scientific Advisory Board, and
- help raise the funds necessary to reach the goals of each component of CBOL's Program of Work.

During the initial grant period, the Executive Secretary was also involved in the operations of most component activities (e.g., organizing and attending Working Group meetings and workshops, assisting in the development of new activities). Now that these activities are established and underway, CBOL proposes to pass these responsibilities to a Project Manager, as described below.

The incumbent, Dr. David Schindel, is currently serving in this capacity on a two-year detail from the National Science Foundation. The initial grant from the Sloan Foundation included salary for this position at a GS-14 level. NSF agreed to pay the difference between this amount and Dr. Schindel's actual salary and benefits (approximately $50,000 annually). CBOL is exploring an extension of Dr. Schindel's detail and NSF's willingness to continue its co-funding of his salary and benefits. Anticipating that NSF will agree to the detail but not necessarily to the co-funding, versions both with and without the full salary and benefits are included in this request.

Administrator. CBOL's Administrator plays the lead role in managing the finances of the Sloan Foundation grant and of other funds received by CBOL, and in overseeing the daily operations of the Secretariat Office. This involves:

- helping to organize CBOL-sponsored meetings and other CBOL events, including the four proposed regional outreach meetings in developing countries;
- working closely with the organizers to plan and conduct the Second International Barcode Conference;
- paying for meetings and travel of meeting participants,
- maintaining the CBOL website,
- managing all interactions with the administrative systems of the Smithsonian Institution and the US National Museum of Natural History (CBOL's host organizations),
- supporting the Executive Secretary's fund-raising efforts through prospect research and acting as liaison with the Smithsonian's Office of Sponsored Projects and the Museum's development officer;
- overseeing the work of CBOL's interns, and
- maintaining all financial and other records (e.g., mailing lists, agreements with Member Organizations).

CBOL is requesting continued support for the CBOL Administrator at the GS-9 level.

Project Manager. As CBOL's Program of Work has developed, the need for hands-on involvement by the Secretariat Office has grown beyond the resources provided by the first Sloan Foundation grant. For this reason, CBOL is proposing to create the new position of

Project Manager at the GS-9 level. The Project Manager's efforts would be divided roughly among the following responsibilities:

- Overseeing and managing the Demonstrator Development Project, working closely with its Steering Committee. CBOL proposes to use a more centralized management approach to this effort, in contrast with the distributed management of barcoding campaigns (FishBOL and ABBI) and CBOL's Working Groups;
- Facilitating the activities of CBOL's Working Groups, networks and other activities. This will involve working directly with the WG chairs and members, managing their communications through the NBII Portal and other means, helping to arrange and conduct their meetings and activities, and coordinating fund-raising needs with the Executive Secretary;
- Facilitating barcoding campaigns and other barcoding projects. This will involve maintaining and/or coordinating websites for each, helping to plan meetings, providing support for data management and monitoring of progress, and other responsibilities that arise as the campaigns and projects develop.

Interns. CBOL's Secretariat Office employed two summer interns in 2005 and they proved extremely valuable. CBOL now requests support for four intern positions, each of which would be supported for 40 person-weeks per year. These intern positions would be shared among the CBOL Secretariat Office, the coordinator of the Demonstrator Development Project, FishBOL and ABBI. The duration and assignment of intern appointments to the Secretariat Office and distributed activities will depend on needs and the availability of suitable candidates.

Meetings

CBOL proposes to continue its previous "baseline" schedule of meetings for the Executive Committee (two per year) and one international conference during the two-year award period. New support is being requested for one annual meeting of CBOL's Scientific Advisory Board. CBOL and DNA barcoding are graduating from a period of introduction to one of rapid growth and implementation. Accordingly, CBOL is proposing an ambitious schedule of new meetings and activities that are responsible for a significant increase in the overall budget. An annual budget of $50,000 each is requested for meetings of the Steering Committees and Regional Working Groups of the major barcoding campaigns (FishBOL and ABBI). These meetings will all be co-funded from non-Sloan sources by campaign participants. CBOL also requests $50,000 per year for meetings related to the demonstrator development project. Support is also requested for partial funding of three regional meetings in developing regions. International development agencies will be asked to provide approximately two-thirds of the funds needed to hold these outreach meetings. CBOL is requesting $20,000 per year for smaller ***planning meetings*** on emerging topics that may lead to the launch of new barcoding activities.

Working Groups and Other Activities

The ***Database Working Group*** proposes to hold three meetings per year, at which different combinations of partners will be assembled for the purpose of developing shared protocols. Each meeting will last for two days and will include approximately 30 participants. This proposal includes a request for $30,000 per year for these meetings, which represents approximately 50% of their anticipated costs. The balance will be sought from partner

organizations such as GBIF and the National Biological Information Infrastructure of the US Geological Survey.

The ***DNA Working Group*** proposes to hold an annual workshop with representatives from approximately 10 leading barcoding labs around the world and several biotech companies. This proposal includes a request for $10,000 in co-funding for these meetings, with the balance to be provided by the participants.

The ***Data Analysis Working Group*** proposes to hold one workshop in May 2006 and to offer then dissertation improvement awards ($5,000 each) to graduate students participating in a competition to develop new software tools. The Center for Discrete Mathematics and Theoretical Computer Science (DIMACS) at Rutgers University will manage the competition and award process for these awards. DIMACS has more than ten years of experience in managing this kind of activity, having been an NSF-funded Science and Technology Center from 1989 to 2000. This proposal includes a request for $20,000 for partial support of these activities. The WG is preparing proposals to the National Science Foundation and the European Science Foundation for the balance of funds needed for these activities.

No funds are being requested in this proposal for the ***Plant Working Group***. Their work is being supported by other grants from the Sloan and Moore Foundations.

The ***International Network for Barcoding Invasive and Pest Species*** (INBIPS) plans to hold one Steering Committee meeting per year and one workshop per year. CBOL would contribute $15,000 towards these meetings per year, with a far greater sum being provided by the participating agencies involved.

CBOL requests $3000 per year to convene two annual meetings of the ***Technology Development Committee*** at the Smithsonian Institution. Private companies will be invited to attend part of these meetings and will be asked to contribute toward meeting costs.

The sum of $5000 is requested for legal services to be obtained following the advice of the ***Legal/IPR Committee*** during the first year. These services will include a limited patent search to ensure that the DNA barcoding technique is in the public domain, and to advise CBOL on issues of liability.

Secretariat Operations

The budget request includes support for normal office operations based on the expenses incurred during the first 14 months of operations of the Secretariat Office (September 2004 – November 2005. Staff travel includes travel expenses for the Executive Secretary and Project Manager, and for travel of the Principal Investigator and CBOL Administrator to CBOL meetings. Voice communications includes conference calls and cell phone.

The Smithsonian has provided the Secretariat Office with four desktop computers for the Executive Secretary, Administrator, and two interns from its IT inventory. For the proposed award period, CBOL requests support for the purchase of one replacement desktop computer per year ($1,000 each), and for a color laser printer that will allow the Secretariat to produce brochures in-house. Higher than normal support for printing and publications is requested for the second year of the grant in anticipation of the Second International Barcode Conference that is planned for February of that year. This will cover the printing cost of a new CBOL poster and for costs associated with producing a conference proceedings volume.

CBOL will hire an independent contractor for the ***Data Portal Development***. One possible candidate contractor is the firm used by BoLD, which would provide strong interoperability between the two systems. The database that underlies the Portal would also provide the Secretariat Office with much-needed knowledge management capabilities for participants, organizations, documents, meetings, etc. Separate task orders for the ABBI and FishBOL websites will be created under one umbrella contract, and these task orders will be co-managed with the FishBOL and ABBI campaign coordinators.

Review #1

Consortium for the Barcode of Life
Proposal for continuation of support by the
Alfred P. Sloan foundation

Summary

CBOL has made a tangible, positive impression in the arena of animal identification. Taxonomy is currently experiencing a renaissance, and the contribution of DNA/RNA sequencing technology to that rebirth cannot be overestimated. At the same time, imaging technology, large-scale relational databasing capacity and the internet are poised to create a global forum for a revolution in organismal taxonomy and systematics.
This reviewer thoroughly endorses the continuation of the Alfred P. Sloan Foundation's support of the CBOL initiative, but urges increased collaboration, cooperation and complimentarity with existing initiatives that address broader aspects of organismal characterisation and databasing. In particular the Global Biodiversity Information Facility (GBIF) and the International Commission on Zoological Nomenclature (ICZN) need to work closely with the CBOL initiative. Such collaboration is essential to maintain the credibility of CBOL among the broader community of systematists, professional and amateur.

1. Harmonizing taxonomic initiatives

The taxonomic renaissance is almost entirely due to the IT revolution. Taxonomy is a discipline that cries out for web technology to express itself at its best. Pictures paint thousands of words, but are expensive to publish in hard copy – the web completely overcomes this constraint. Graphics software allows for the production and dissemination of images of a quality hitherto unimaginable. Databasing of taxonomic information in a variety of inter-accessible formats is now possible on an unprecedented scale. IT technology has encouraged several new initiatives in taxonomy, especially animal taxonomy, and for CBOL to succeed in the long-term, the project needs to harmonise with the most important of these initiatives.

GBIF, the Global Biodiversity Information Facility, has set out to produce ECAT, the Electronic Catalogue of organism names. The uBIo initiative is setting out to database every organism name – including all vernacular and mis-spelled names; the Catalogue of Life project represents the combined effort of Species 2000 and ITIS – major efforts to catalogue recently revised taxa. Within the commercial sector we have *Zoological Record*'s Index of Organism Names (ION), possibly the most complete record of names available.

Most recently is a new initiative, ZooBank, a register of all the scientific names of animals, to be made retrospective and thus eventually pick up every scientific name of every animal ever coined.

ZooBank is a joint initiative between ICZN, GBIF and *Zoological Record* (Thomson Zoological Ltd). In order for CBOL to maintain any credibility regarding the scientific animal names associated with its gene sequences, integration with ZooBank is essential.
This idea was first proposed at the Front Royal meeting on data standards for CBOL data in GenBank in April 2005, and needs to be pursued in consultation with the above partners. It would be very helpful if ICZN and the ZooBank initiative could be added to the list of partners involved in developing data standards (bottom of p. 8 of the proposal).

2. Allocation of effort / Targets and timetables

The CBOL consortium has already a proven track record of being able to realise its aims within a proposed time-frame. With a broader approach and wider consultation I have no doubt that the aims of the current proposal are realistic in terms of allocation of effort and resources.
Appendix 9 "milestones for 2008" is in this respect an extremely useful overview of the next three years.

3. Management / individuals – skills and capacities

I have known Scott Miller for many years as a fellow entomologist and taxonomist. I have no doubt in his ability to manage a large-scale project such as the present CBOL initiative, and consider him ideal to lead such an important venture. David Schindel has proved himself to be skilled at organising and presenting the results of the CBOL project at several meetings.

4. Other aspects

Apart from the need for increased collaboration with ZooBank and ICZN, it is not entirely clearly how CBOL is going to forge links with Developing country institutions. Although numbers of institutions based in under-resourced countries are listed (University of Ghana, Obafemi Awolowo University, Nigeria), it is not clear how these institutes and their staff will benefit from the new proposal. I have direct, personal experience of working in both of these institutes, as well as many of the other developing country organisations listed on pp. 22-25 of the proposal, and can testify that it is extremely difficult to produce any productive collaboration without allocating resources to those institutes. It would be very helpful to see more tangible strategies for addressing such developing country needs for training in barcoding and data management.

I therefore fully endorse the continuation of the Alfred P. Sloan's invaluable support of the CBOL initiative, and sincerely hope that we can work more closely together in the near future.

Review #2

Consortium for the Barcode of Life (CBOL) Grant Proposal

Stated goals are exploring and promoting the potential of DNA barcoding as the global standard for species identification, with specific goals such as:

- Creation of a public database of DNA barcode reference records
- Determination of the DNA barcode sequences for important taxonomic groups, and eventually for all species
- Promoting the development of faster, cheaper and more portable instruments and methods for obtaining DNA barcode data from biological samples
- Encouraging global participation in the "Barcode Initiative"

For the period for which funds are requested (from the summary):

1) Provide the research community with the tools needed to launch successful barcoding projects, including software for data management and analysis, lab protocols, and global standards for data quality
2) Create a fully operational demonstrator that can assign any specimen from any known species in a significant taxonomic group to its correct species
3) Promote global participation in the Barcode initiative through 4 regional meetings in the developing world and by doubling the number of CBOL Member Organizations in developing countries
4) Increase the number of barcode records in public databases from the current level of 50,000 specimens in 10,000 species to 500,000 records in 50,000 species, primarily, but not exclusively, through CBOL's taxonomic global barcoding "campaigns".

Proposed program of work and goals for 2006-2008 (pp. 11-12):

1) Continuing to pursue core mission of enabling the universal application of DNA barcoding
2) Developing a complete and operational demonstrator system for species identification in a group of interest (p.16)
3) Continuing to populate a global database of reference DNA barcode records (primarily bird and fish, but new barcoding efforts will be launched)
4) Expand its mission of promoting global participation by engaging researchers and users in developing nations

Based on the information provided, CBOL achieved the goals set forth in the initial proposal (2004). This new grant proposal is essentially a continuation of those efforts. Much of this grant is aimed at funding the structure and organization of CBOL, meetings, and the organization of records into a cohesive, accessible database, with overall rules on how to input new data. The substance of this new proposal is the development of a complete and operational demonstration system for an insect group or groups by the end of the grant period (June, 2008). Either Culicidae or Tephritidae will be the target. Although obvious to this reviewer why those two groups were selected, it might help others for CBOL to provide additional information justifying their choice, and to provide details on how a complete COI barcode library for either group will be obtained in 2 years. It would also help if data were provided from previous work to further

support its usefulness and effectiveness. It was not clear exactly how the funds will be utilized. Are they to go towards increasing infrastructure and organization? How much will go towards lab tests, etc.

In general, the CBOL concept is appears to be sound and should prove to be very useful for diagnostics for a number of purposes as well as a powerful resource for the research community. The deliverables stated appear plausible and attainable. There has been a good level of productivity from the last round of funding from the Sloan Foundation, and every reason to expect that this level of productivity will continue. Thus, overall, this is a strong proposal and merits funding.

Notes:

- p.4 under CBOL's long-term strategies, the first paragraph is about the two-year funding period (short term)
- p. 8 the research results section seems pretty weak. Just one paragraph, and four significant papers?
- Goals as outlined on pp 1-2 and pp 11-12 (noted above): goals 3 and 4 are juxtaposed and goal 1 is worded differently. (I don't know how important this is….)
- The barcoding process is well described in the table beginning on pg. 39. What is important about this table is that the CBOL group is recognizing the limitations imposed by current technology or methods. One of the most powerful and practical applications of this technology could be to meet the safeguarding mission of national plant protection organizations. For this technology to be truly useful from this regulatory standpoint – e.g., identification of materials entering a country at a port of entry – it will be essential that most of these limitations be addressed. This would mean streamlined sample preparation and results in probably less 6 hours and almost certainly in less than 24 hours. This said, it does appear that the CBOL group is well aware of these needs.

Review # 3

Review of proposal to A. P. Sloan Foundation for continued support to operate the Consortium for the Barcode of Life (CBOL) Secretariat Office and activities.

DNA barcoding is the one topic that recently has most profoundly affected biosystematics and taxonomic research worldwide, both in terms of scholarly discussions as well as in changing methodological approaches and daily work practises in this scientific discipline. A large part if not most of this extraordinary, sudden development with the clear potential to further revolutionize large areas of taxonomy can be credited to the efforts of CBOL and its Secretariat over the past two years. By facilitating such a successful scientific initiative leading to global impact, the CBOL Secretariat can strongly be recommended for continuous support purely on the basis of its past, recent accomplishments.

The present strong proposal, however, provides in itself sufficient justification for sustained funding, and I would have no hesitation to grant the CBOL Secretariat a solid financial basis for continued successful, expanded operations over the next two years. Based on a sound strategy and clear vision, the proposal presents realistic goals, which the current CBOL Secretariat staff with demonstrably outstanding leadership qualities will surely be able to reach. The proposers have an overall impressive track record of previous achievements, particularly during the start up period of CBOL operations, which leave little doubt as to their ability to successfully implement the currently proposed work plan for CBOL phase II. Considering the momentum already gained and the need for further expansion of the Secretariat's activities during the operational phase of CBOL, the increased budget appears fully justified, including the request for additional staff.

As for the detailed content of the proposal, all four main goals are highly significant in terms of the overall CBOL strategy, and the deliverables indicated are tangible and appear well within reach. The first two goals (1: tools, protocols and standards; 2: operational demonstrator project) would appear to deserve even higher priorities, as they could conceivably accelerate significantly the overall acceptance and implementation of the barcoding approach by the broader scientific and user communities. For the further promotion of global participation and outreach (goal 3) in DNA barcoding, the four proposed workshops will surely represent important opportunities, especially towards the developing world. Considering the general state of need and the usual hesitant approach taken by many developing countries towards new trends and developments, any promotion of barcoding would find more immediate acceptance if it could be directly linked to new or additional opportunities for funding of such activities. Such an approach could perhaps be explored jointly with potential donors specialised in bi- or multilateral developmental and/or scientific cooperation with developing countries, which then could join up directly with the organisation of the proposed workshops.

As an additional strength of the proposal, CBOL is already well linked to other relevant international projects and initiatives (e.g., GBIF, TDWG, GenBank, etc), which present additional opportunities for synergies and should be further explored. Similarly, opportunities for co-operation with industry and the private sector are clearly identified, but could be expanded beyond the possible support for development of barcoding hardware and refined molecular techniques. By focussing on taxa with known or suspected vector or pest properties for specific

demonstrator projects, it should be possible to solicit interest from the health or agriculture sector for the establishment of specific barcode libraries for such organisms, especially if those are otherwise difficult to identify.

In order to also come up with some critical remarks on the proposal, the need to operate and support a number of related, but distinct CBOL committees and working groups could perhaps be questioned. Without neglecting the effectiveness of committees and expert panels for ensuring an active involvement of larger parts of the scientific community, the resources and administrative costs required from the Secretariat for the successful operation of such bodies will increase with their number. Furthermore, assuming a growing number of different bodies and panels, it becomes more difficult to explain why these groups are needed and what their objectives are towards third parties. In particular, the need for a separate "data analysis working group" in addition to a "database working group" as well as support for a "technology development" working group might not be easily explained to an outside audience. Although the work and expected outputs defined for these different bodies appear well founded, their exact mandate, inter-relationship, and mode(s) of operation is not fully evident from the CBOL governance structure as depicted in appendix 4.

Finally, some questions and suggestions for CBOL perhaps to consider in the longer term, without questioning thereby this overall excellent proposal:

- how can CBOL be still co-ordinated at the international level, when increasing barcoding efforts come underway in many parts of the World and needs for international promotion decreases?
- will CBOL and the relevant institutions involved be able to develop a long term strategy to secure and sustain the operation of the barcode database(s)?
- how best to ensure the availability of sufficient taxonomic experts and resources for providing reliably determined/verified specimens for barcoding, especially if the need and numbers of requests will increase considerably?
- how best to entertain the challenge to keep barcode data freely and openly available in the light of growing potential commercial interests and applications?

With the expectation that these questions will provide additional arguments to grant support for this important activity, I wish the proposal best of luck.

Review # 4

Review of Proposal Entitled "Consortium for the Barcode of Life: Support for Consortium Activities and Secretariat Office".

The Consortium for the Barcode of Life (CBOL) was begun in 2004, with funding from the Alfred P. Sloan Foundation, with the mission to promote DNA barcoding as the global standard for species-level identification. The current proposal requests funding for an additional two years to achieve four goals, including providing the research community with analytical tools, demonstrating the utility of the method, doubling the number of CBOL member organizations and rapidly increasing the number of specimens and species represented by barcode records in public databases.

The idea of a single genetic sequence that can be used as a barcode to identify specimens to the species level is compelling and the authors of this proposal have done a great job of making the case for the utility of such a method. The authors have also shown proof-of-concept evidence that the CO1 gene could serve as a locus for such a barcoding effort, although the use of a single gene as a species-level barcode has yet to be widely adopted within the scientific community. The comments below are focused less on the question of whether the underlying science is likely to progress and more on the question of whether the strategy proposed by CBOL to promote DNA barcoding as a global standard is effective and efficient.

During the first two years of CBOL, the bulk of effort was devoted to outreach efforts and establishing the organizational structure of the consortium. Expected outcomes that were attained included expanding membership in CBOL, promoting public awareness of barcoding, and gaining financial support for the activities of CBOL. Through various publications and presentations, CBOL has been successful in promoting awareness and expanding the consortium membership, especially among public sector organizations. Private sector outreach efforts have been less successful and fundraising efforts have resulted in modest commitments of funding. While the expansion in membership indicates that there is an increasing interest in the consortium, the lack of enthusiasm by the private sector and funding agencies may indicate that CBOL must make a more compelling case for investment to these audiences.

The current proposal represents a shift in strategy from building awareness to developing a product and demonstrating its value. Such a change is the right thing to do. Once the utility of the product is demonstrated and made obvious, then investment and adoption by others should follow.

The proposed goals for the next two years are consistent with the new strategy and should be able to be accomplished within the time frame proposed. Enabling the application of barcoding by developing protocols and standards is an appropriate activity of CBOL. Developing a demonstration project will focus CBOL efforts and produce an outcome of practical utility. Populating a database and making it readily available for use on the internet will provide the tool required for users to realize the potential use of the barcoding method. Successful completion of each of these 3 goals should enable the consortium to "sell" the idea to others and increase participation in the consortium. Lastly, the limited program for international outreach is

appropriate to the scope CBOL and may bring fresh ideas for implementing the standard on a global scale.

Although the strategy seems to be the right one and the targets/deliverables are consistent with the strategy, the plan for implementing the strategy and meeting the targets should be improved in two important aspects.

1. The management plan, including the resource allocation plan should be tied more closely to accomplishing the goals. In the current proposal there appears to be a disconnect between some of the key goals and resources dedicated to accomplishing the goals. For example, the plan for enabling universal application of barcoding (major goal #1) is to utilize three working groups (Database, DNA, and Data Analysis) to develop and disseminate standard protocols. However, only $100 K, 6.4% of the funds requested, is dedicated to these working groups. Of course some of the funds allocated to other areas are likely to contribute to major goal #1, but making this explicit helps to ensure accountability for the resources.
2. The plan for outreach to industry should be more clearly thought through. The best approach is to have people with industry experience involved in such planning. Outreach to industry could be more effective if someone with experience in industry and/or experience working closely with industry is hired as a staff member or a long term consultant. Additionally, it might be useful to appoint someone from industry to the executive committee. These people should be empowered to directly affect CBOL's strategy with regard to industry outreach.

If the CBOL can structure itself more closely around the key goals that it wishes to accomplish, then it is reasonable that the goals can be accomplished in the time frame proposed. The proposed strategy, with goals and deliverables outlined, should contribute directly to the mission of CBOL, to promote DNA barcoding as the global standard for species-level identification. Global standards can only be established if there is widespread adoption and CBOL can best facilitate widespread adoption if it develops the tools to demonstrate the value of barcoding.

Review # 5

RE: REVIEW OF PROPOSAL ON "CONSORTIUM FOR THE BARCODE OF LIFE (CBOL)"

The present proposal for an extension of the funding for CBOL makes very interesting reading. I will try and give my frank comments as best as I can.

When the initiative was first annouced a few years ago and the "media circus" took hold, I was very sceptical to say the least. Being more of a traditional morphology-based taxonomist/systematist of fish and crustaceans, I was seriously concerned that the misconception that the COI gene alone will solve all the problems of taxonomy and the dearth of taxonomists simultaneous, would not be good for the science. Having used various genes to help my work nevertheless, I appreciated the power and new directions the molecular tools permited, especially in discerning patterns and helping fingerprint common taxa. Reading this proposal, I am glad the CBOL initiative has adapted, considered the criticism, and "evolved".

To get to the main proposal, my specific comments are:

1. I concur that it is important to continue supporting the exercise. Especially in building up a database of the more common, more economically important taxa, where fingerprinting using the COI gene offers many fruitful possibilities. It is possible to build a strong database of such species and help managers and researchers quickly identify samples, even when the whole plant/animal is no longer intact. I believe it is possible to have such a system operational with a high degree of accuracy. To this effect, a demonstrator model is much desired. The target of using mosquitoes or food fish for example, is eminently achievable. That the COI tool can link morphologically very different life-stages is to be encouraged. I would also urge development of microarrays and DNA chips to facilitate this process and move towards even greater automation.

2. The current CBOL proposal for the future is in line with the ZOOBANK proposal recently launched by several senior biologists like Ed Wilson and many members of the International Commission of Zoological Nomenclature (including myself) – I foresee that the ZOOBANK, together with GENEBANK, will add a whole new dimension to classifying the planet's diversity. Personally – I am not too happy with the way GENEBANK works – the lack of proper control, lack of matching vouchers with DNA samples etc., has created various problems. There are too many inaccuracies and incorrect species assignments to make it that valuable. A better managed system under CBOL is to be encouraged, closely linked to ZOOBANK perhaps – has great potential. I encourage this to be further developed in CBOL.

3. As discussed in point 2, the current codes in BoLD should eventually be in the public domain like GENEBANK. I am encouraged by the controls implemented to minimise problems.

4. The proposal to try to find a way to use formalin material MUST BE encouraged. It is very important. This to me is a major impediment – especially with fish, herptiles etc. which have been traditionally preserved in formalin. The problem with CBOL – the challenge rather – is that it may not be possible to collect a good number of species these days. They may no longer be as common a century ago, be very rare now, or simply uneconomical to try and get; but formalin material exists. To

rely just on fresh alcohol material is not the best solution. So if the tools can be developed to somehow reconstitute and/or reconstruct the damaged DNA, from both the biochemical and bioiformatics realms – then a major hurdle in molecular phylogenies and barcoding will be crossed.

5. I note with great interest that a major fish exercise will be implemented. I believe the CBOL can help the taxonomic challenges now faced by ichthyologists especially with regards to wild-caught fisheries, legal protection of endangered species, the aquarium trade and aquaculture. The taxonomic impediment is real here. This of course is a personal bias, but being the most speciose group of vertebrates on the planet, they are too important to sideline.

6. I am heartened to see that more developing countries will be engaged. I encourage the transfer of more technology and training and expertise to such countries, notably to Africa and Asia, where a great deal of the biodiversity resides. I know of many keen biologists in such countries where the sad state of their economies make any molecular exercise financially prohibitive. That a major meeting will be planned for SE Asia is great!

Overall, my prognosis is that the current proposal deserves support and funding. To be blunt – will CBOL be the "magic bullet" that will solve all the nightmares and woes a traditional taxonomist like me have? No – but it will be a very helpful tool if it is applied carefully and judiciously. Will it make solve the challenge of recording Earth's high biodiversity? No. But it will help in the documentation, make it go faster and bring us new ideas, ask new questions and reach new vistas. To this effect, the proposals here are good, well thought out and need support. The funding requested is not as high I would have expected – so it would seem to be a good "bargain".

Please feel free to email me if there are any more queries.

Consortium for the Barcode of Life (CBOL)
Responses to Reviews

We wish to thank the reviewers for the effort and thought they invested in their evaluations. It is clear to us that they read and analyzed our proposal very thoroughly and have understood our intent completely. We do not dispute any of their findings and we will forward their suggestions to CBOL's Executive Committee for consideration. All five reviewers have identified key challenges that CBOL will face during the coming two years. In each of the following four categories, we will agree with the concerns raised by the reviewers, admit that these are the most challenging issues that we will face, and describe the strategies we plan to use and the steps we plan to take.

1. <u>Barcoding Projects</u>. Reviewer 2 noted the ***lack of extensive research results*** in the proposal, pointing to a single paragraph and four 4 significant papers. The Sloan Foundation's support to CBOL is for administrative and catalytic activities, not for research projects or other direct support for barcoding. For that reason the proposal concentrates on CBOL and its catalytic role, rather than barcoding research itself. We could have described dozens of other studies in which the effectiveness of COI as a species-level diagnostic marker has been demonstrated, were it not for the length limits set by the proposal guidelines.

Reviewer 3 noted the challenge that barcoding will face in obtaining the **taxonomic expertise to needed to identify voucher specimens.** The "taxonomic impediment" applies to barcoding as well as other activities that require taxonomic identification. As one reviewer noted, barcoding will not solve the lack of adequate numbers of experts, but it will make the work they do more efficient and effective. CBOL hopes to minimize the need to identify voucher specimens in the next two years by concentrating on taxonomic groups that have: (1) well-established taxonomies, (2) communities of taxonomic experts that are well-organized and

willing to participate, and (3) collections from which DNA can be obtained (frozen tissue, cell cultures, zoo specimens, collections in alcohol that were not formalin-fixed, and recently collected dried museum specimens) for a significant portion of species.

Reviewer 2 raised several questions (though no real concerns or objections) concerning the ***demonstrator development project*:**

- How were candidate groups chosen? CBOL's Executive Committee plans to make a top-down decision between tephritids and culicids in July 2006, based on proposals that are being developed by two Steering Committees of taxonomic experts. CBOL's Scientific Advisory Board identified these two candidate groups from among the "bottom-up case studies" of potential barcoding projects submitted by the research community.
- How will a complete reference library be assembled in two years? CBOL's Executive Committee has put this question to the two Steering Committees. The taxonomists and the barcoding experts who are serving on these Steering Committees have been charged with developing proposals to obtain the necessary voucher specimens and determine their barcodes within two years. Their proposals will be submitted to CBOL's Executive Committee in June 2006.

2. <u>Database and Data Standards</u>. Reviewers 1 and 5 stressed the need for ***collaboration with GBIF/ECAT, ICZN/ZooBank, uBio, and other.*** We agree completely. Jim Edwards, GBIF's Executive Director, is a member of CBOL's Executive Committee. Andrew Polaszec, ICZN's Executive Secretary, is a member of the Database Working Group and participated in CBOL's Front Royal meeting at which data standards for barcode records were finalized. (Reviewer 1 noted correctly that neither ICZN nor ZooBank were mentioned in the body of the proposal, but they were noted as partners in Appendix 6, which described the Barcode Data Standards.) There

is no technical plan for ZooBank currently, so it is too early to know how barcode data would fit into a registration system of taxonomic names, but CBOL intends to work closely with ICZN on this issue.

Reviewers 3 and 5 asked about CBOL's plans to ensure the ***long-term sustainability of the barcode database as a public resource,*** and how CBOL would ***resist the trend toward commercialization of research data.*** CBOL does not see itself as a permanent organization, and we are therefore trying to promote ownership of barcode data by permanent organizations and the research community in general. CBOL has no plans to construct and operate an independent data repository for barcode records. The data standards for barcode data have been adopted by INSDC (GenBank, EMBO and DDBJ), and barcode records are beginning to flow into them. The CBOL-INSDC agreement sets the stage for INSDC to become the permanent repository of barcode data. Over the next two years, CBOL will be working with INSDC and other partners (e.g., GBIF, Catalog of Life), to establish a system for data quality that will check for errors as barcode records are being assembled in BoLD, and at the point they are submitted to INSDC. During this time, CBOL's Database WG plans to establish a community-based system of data curation and third-party commentary. Once established, this system will allow users of barcode data to attach comments to BARCODE records in INSDC whose reliability is questionable. Over time, this community-based comment system will replace CBOL's role as the steward of quality assurance.

One reviewer suggested that "BoLD's code" should be made public, which we take to mean their data management source code. The University of Guelph and the Canadian government have made major investments in BoLD, and CBOL must respect their regulations. CBOL can only encourage BoLD to consider an open architecture.

3. <u>Outreach Activities</u>. This section will respond to reviewer comments concerning ***outreach to developing countries*** (reviewers 1 and 3)***, private industry, and potential users/supporters of DNA barcoding*** (reviewers 4 and 5). CBOL feels that before funds can be obtained for training and capacity building in developing countries, the countries must first set their own priorities. CBOL is catalyzing these planning efforts. BioNET INTERNATIONAL, a CBOL Member Organization and the co-organizer of the regional meetings described in the proposal, has already obtained US$50K toward the first two meetings from the Swiss Development Agency (SDC). US$10K will be used to support development of follow-on proposals. SDC plans to evaluate the impact of the first two meetings and will consider awarding an additional US$50K for the next two meetings.

CBOL has already started to approach donors with small requests for planning activities, and we have had an extremely high success rate thus far. Within the past six months, we have obtained small contributions (US$10-20K) each from three US government agencies, two universities, and two private sector companies. CBOL has concentrated on several governmental and intergovernmental agencies (USDA, USEPA, NSF, European Commission, European Science Foundation) and we believe that as large research projects are planned, they will be receptive to funding proposals.

We agree completely with the reviewers who stressed the need to engage directly with industrial partners. As noted in Appendix 2, CBOL has made presentations to a number of important companies (e.g., Affymetrix, New England Biolabs). Establishing close relations with private sector companies is a real challenge. Instrument manufacturers, in particular, are not willing to work in a group setting because of trade secrets. We take on board the excellent suggestion from one reviewer to add an industry representative to CBOL, and the Scientific

Advisory Board would probably be the most appropriate. CBOL has compiled a list of company contacts and researchers with interest in technology development, and we will form an online discussion forum to gather advice and promote activity with industrial partners.

4. <u>Management Issues</u>. Reviewer 3 expressed concern about the potential for the ***proliferation of Working Groups and Committees.*** CBOL agrees and has already declined to add activities that have been proposed by Member Organizations and will not start others. We hope to ramp down Working Groups and Committees as they complete their missions. CBOL is developing Terms of Reference for all CBOL groups, which addresses reviewer 3's suggestion to clarify mandates and modes of operation.

Reviewer 4 asked if there is a ***disconnect between priority of goals and level of funding***. Specifically, this reviewer noted that three Working Groups will receive only $100K, or 6% of requested funds, even though their work is central to CBOL's goals. These WGs have been in operation for almost two years and they are already attracting non-Sloan funding for of their activities (e.g., non-Sloan funding for 80% of the formalin workshop). CBOL is therefore seeking to catalyze new activities such as the demonstrator development project.