

Report of San Diego Meetings, Consortium for the Barcode of Life

23-26 June 2008

The Consortium for the Barcode of Life (CBOL) held meetings of its Leading Labs Network (LLN), Data Analysis Working Group (DAWG), and Implementation Board (IB) at the San Diego Marriott Hotel and Marina on 23-26 June 2008. A total of 47 members from 15 countries participated in these meetings (see Agendas, Appendices 1-3, and participant lists, Appendices 4-6). All but a few of these participants took part in a half-day visit to the Community Cyberinfrastructure for Advanced Marine Microbial Ecology Research and Analysis at the University of California, San Diego. In addition, a representative of the Public Library of Science (PLOS), and two social scientists from Lancaster University, UK,¹ participated as observers in CBOL's meetings and the visit to CAMERA.

Leading Labs Network meeting.

LLN was created in 2007 as a network of barcoding labs around the world. The first full LLN meeting took place in Washington, DC in July 2007, at which participants agreed to create a website for sharing protocols for lab procedures and data management, other technical information on barcoding, and resources for training new barcoding researchers and labs. LLN also agreed to develop a system for offering several types of training opportunities, including individual training fellowships in advanced barcoding labs, and both *in situ* and *ex situ* short courses. LLN conducted a full-day short course on 18 September 2007 immediately before the Second International Barcode of Life Conference in Taipei.

Twenty-six representatives from 22 Leading Labs in 11 countries participated in the LLN meeting in San Diego on Monday and Tuesday, 23-24 June 2008 (see agenda, Appendix 1; participant list, Appendix 4). Representatives from three labs in the People's Republic of China were planning to attend but were unable to obtain visas.

LLN members discussed the most significant bottlenecks in their workflow and obstacles to the production and public release of BARCODE records on BOLD and GenBank:

- Lack of access to specimens with undegraded DNA;
- Delays in getting expert taxonomic identifications;
- Lack of clarity in the BARCODE data standards concerning minimum sequence length, coverage, and quality;
- Lack of funding for dedicated campaign managers and technicians;
- Lack of approved BARCODE regions for land plants and fungi (though efforts are underway to propose them to CBOL);
- Time- and labor-intensive steps needed to error-check data records;
- Lack of training opportunities & lab equipment; and
- Lack of IT support/advice on quality control steps.

¹ Anji Desai, a consultant to PLoS, participated in the visit to CAMERA and the IB meeting. Rebecca Ellis and Claire Waterton (Lancaster University, UK) are conducting a social science study of CBOL and they attended all three CBOL meetings and participated in the CAMERA visit.

The members of LLN agreed on the following immediate outcomes and developments:

1. A system for planning, prioritizing, and offering training opportunities; and
2. A timeline leading to training opportunities that includes:
 - a. Proposals for training fellowships for 1-2 individuals that can be submitted to CBOL at anytime by LLN members or major CBOL activities (FISH-BOL, ABBI, MBI and TBI);
 - b. Suggestions by 15 July to the LLN Steering Committee for short course plans, following the BOLI Case Study format (what, where, who, when, how much?) to be distributed to LLN by 31 July 2008;
 - c. Submission of plans for short courses to the LLN Steering Committee by 1 October 2008; and
 - d. Development of a priority list and timeline for short courses in Fall 2008.

Kristin Jett, a member of CBOL's Secretariat Office staff, presented a preliminary design for the LLN website through which protocols and training materials will be made available to the barcoding community. Based on the framework provided by this website design, LLN members developed:

3. A full outline for the content that will be needed to populate the website; and
4. An editorial board and workflow process for creating, vetting, and approving content contributed by the LLN members.

Data Analysis Working Group meeting.

DAWG was formed in February 2005 with a mission to develop new methods for analyzing, interpreting, and displaying DNA barcode data. DAWG has held four workshops (two at the DIMACS center at Rutgers University and two at the Museum National d'Histoire Naturelle in Paris), obtained grant support from the US National Science Foundation and the European Science Foundation, and organized a plenary session and a software demonstration area at the Second International Barcode of Life Conference in Taipei in September 2007. The presenters at the Taipei Conference agreed to publish their presentations in a proceedings volume in an appropriate journal and to move aggressively toward completion of their software packages and to make them available to the barcoding community as soon as possible.

Eleven DAWG members from six countries participated in the San Diego meeting (see agenda, Appendix 2; participant list, Appendix 5). They reported on approximately ten software analysis packages that they are developing and presented at the Taipei Conference. A proposal for a proceedings volume with their Taipei presentations is currently being considered by BMC Bioinformatics and a deadline for their manuscripts could be set for Fall 2008.

DAWG proposed a joint meeting with data managers from the Leading Labs Network, and 10 participants in the LLN meeting met with DAWG on Tuesday morning, 24 June 2008. This provided DAWG with an opportunity to gather user requirements for analytical software. It also provided LLN with current information on plans for the Barcode of Life Initiative (BOLI) Data Portal and the new analytical software that will be hosted there. The principal benefits and action items generated during this joint meeting were:

- Increased awareness by DAWG of normal barcode workflow, bottlenecks, sources of error, and currently used analytical tools;
- Advice on preferred formats for data input and output;
- User requirements for metadata to be included in analytical treatments;

- Linkages that would provide:
 - advice on analytical methods to the Plant Working Group as it selects the barcode regions for plants, and
 - Awareness among DAWG members concerning plans for barcoding using multiple gene regions; and
- The opportunity to establish working relationships with this important user group.

Michael Trizna, a member of CBOL's Secretariat Office, presented the design and first modules of the BOLI Data Portal, on which the analytical software developed by DAWG will be available. The Portal already includes an operational Aggregator which contains more than 80,000 public barcode records downloaded from BOLD and GenBank. Researchers will be able to add their private records in BOLD and from local computers to create customized datasets for submission to software developed by DAWG. DAWG also agreed to develop a set of model datasets with documented characteristics (varying levels of variability within and among species, varying sample sizes within and among local populations). These model datasets will enable developers to benchmark the performance of their software against standard datasets.

DAWG agreed on next steps that include:

1. Finalizing plans for a Conference Proceedings Volume from the DAWG presentations at the Taipei Conference;
2. Designing its timeline of activities and deliverables for the public launch of the BOLI Data Portal at the June 2009 e-Biosphere 09 Conference in London;
3. Submissions to an inventory of new analytical methods by DAWG members consisting of short descriptions of the software packages they intend to develop for the BOLI Data Portal by mid-July 2008;
4. Planning a 'Hack-a-Thon' for software programmers who will integrate their software into the BOLI Data Portal. The EOL BioSynthesis Center at the Field Museum of Natural History is a possible host for the meeting;
5. Planning a workshop for late 2008 that showcases the BOLI Data Portal and DAWG analytical packages for barcode researchers, population geneticists, evolutionary biologists, and biodiversity informaticians;
6. Holding a planning meeting in early 2009 to prepare DAWG deliverables for presentation at the e-Biosphere 09 Conference in London on 1-3 June 2009;

Visit to CAMERA.

The Community Cyberinfrastructure for Advanced Marine Microbial Ecology Research and Analysis (CAMERA; (see <http://camera.calit2.net/>) is a metagenomics project for marine microbial genomics located at the University of California, San Diego. CAMERA has many similarities to BOLI, both being distributed projects devoted to constructing a research infrastructure, including a public database and associated analytical software, for genetic data on biodiversity. The purpose of this visit was to exchange information and to begin exploring potential collaborations in the area of data display and visualization. Officials from CAMERA and UCSD gave presentations on the initiative, the metagenomic database, and the software capabilities it was developing. Staff members from CAMERA provided demonstrations of a tile-

wall graphic display system and a visualization cave. CBOL meeting participants gave brief presentations on barcoding, the BARCODE data standard, and DAWG.

CBOL invited two representatives from the Assembling the Tree of Life initiative (AToL) and the Encyclopedia of Life (EOL) to participate in the visit to CAMERA because of their shared interest in the visualization of biodiversity data (see Appendix 7 for a list of CAMERA, AToL and EOL participants). They gave presentations on the software that is being developed to manipulate and navigate through AToL and EOL data.

Implementation Board meeting.

The IB was created by CBOL's Executive Committee following the Taipei Conference in September 2007. IB's precursor, CBOL's Scientific Advisory Board (SAB), consisted of the chairs of four Working Groups and selected representatives from all geographic regions. The SAB recommended that it be replaced by a Board consisting of the chairs of all CBOL activities (Working Groups, campaigns, networks, demonstration projects) and important barcoding initiatives that are formally independent of CBOL. The San Diego meeting was the first IB meeting.

Sixteen IB members from nine countries participated in the San Diego meeting (see agenda, Appendix 3; participant list, Appendix 6). The first half-day of the meeting was devoted to bringing all IB members up to a common level of understanding of the status of CBOL and its component activities, and their goals for the 2008-2010 funding period under the current Sloan Foundation award. Several IB members reported on other activities and initiatives critical to CBOL, such as GenBank and iBOL, and major barcoding projects outside of CBOL.

The second day was devoted to operational and strategic issues of common concern to multiple activities, and to a discussion of how the IB preferred to operate internally and interact with CBOL's Executive Committee. The IB reached the following agreements:

1. IB members will provide the Secretariat Office with an annual program of work that outlines proposed activities and goals in no more than two pages. The Secretariat will develop and circulate a template for a quarterly report with which IB members will describe their progress and report unexpected obstacles. The Secretariat will use these annual and quarterly reports to develop summaries that will be provided to CBOL's Executive Committee, Member Organizations, and the other IB members.
2. The technical and training material being developed by LLN and the analytical software being developed by DAWG are of general interest to all IB members. They wish to be kept informed of these developments through a list-serve that the Secretariat Office will establish. The Secretariat Office will also circulate a list of activities and topics on which IB members can register their interest. The Secretariat will then try to design an alerting system that will keep IB members informed of relevant topics while avoiding excessive email traffic.
3. CBOL is integrally involved in the iBOL proposal to Genome Canada being prepared at the University of Guelph. Confirmation of matching funds from CBOL-related grants will be important documents in the proposal. CBOL will approach the Sloan Foundation for their agreement to submit a letter for the iBOL proposal confirming that the CBOL, MarBOL, TreeBOL, FunBOL, and DBIC grants can be counted as matching funds.
4. The Secretariat will establish a SharePoint capability through which IB members can share documents in a password-protected environment.

5. There would be significant mutual benefit from holding an advanced workshop on BOLD for the chairs of CBOL campaigns, demonstrator projects, and other barcoding initiatives. Project leaders would improve their ability to use BOLD and the BOLD IT team would get useful feedback from users. The Secretariat will work with BOLD staff to arrange this workshop as soon as feasible.
6. IB members would like to have access to PowerPoint slides for outreach. CBOL presentations are posted at http://www.barcoding.si.edu/barcoding_presentations.html and it would be useful to have more specific slides about LLN, DAWG, and training opportunities available only to IB members through the SharePoint.
7. CBOL has developed several brochures and posters that should be made available to IB members for their outreach activities. It would also be very useful to translate them into other languages. The IB suggested several strategies for obtaining translation services and urged the Secretariat to propose a process to the IB members.
8. IB members agreed that having brief outreach fliers on their respective activities would be useful for outreach and fund-raising. ABBI has already created one such document. The Secretariat will provide examples to the IB members.
9. The IB would like to operate as a Committee-of-the-whole and agreed not to select a Chair or other officers. The Executive Secretary will act as the IB's convenor and information channel to and from CBOL's Executive Committee.
10. The public release of BARCODE data on BOLD or GenBank is a principal BOLI goal, but there are traditions and concerns that promote the retention of potential BARCODE records as proprietary data. iBOL is convening a workshop on this topic and will need to include a policy statement on data release in its September 2008 proposal to Genome Canada. D. Schindel, R. Hanner, and S. Ratnasingham will participate in the workshop and will convey the IB's concern for graduate students and researchers into that discussion

Appendix 1: CBOL Leading Labs Network Agenda

23-24 June 2008: San Diego Marriott, Seaview Room

MONDAY, 23 JUNE 2008

8:00-8:30 am: Joint breakfast with DAWG in the *Seaview Room*

8:30-8:45 am: Review of goals and deliverables under 2008-2010 Sloan Foundation grant (*D. Schindel*)

8:45-8:50 am: DAWG members retire to Torrance Room, Leading Labs begins Session 1

Session 1: Goals for 2008 and this meeting

8:50-9:15 am: Goals for the Network and this meeting (*L. Weigt*)

What does it mean to be a Leading Lab? Definition, Responsibilities,

- Developing and sharing technical knowledge
- Developing and disseminating technical assistance
- Providing *in situ* and *ex situ* training opportunities

9:15-10:15 am: Barcoding developments and obstacles

Status report on top problems: Results of informal survey

- Where are the bottlenecks? Lab protocols? Data handling?
- Taxon-dependent issues
- Recovering degraded DNA
- Documenting Best Lab Practices

10:15-10:30 am: Coffee Break

Session 2: Increasing the flow of BARCODE records

10:30-11:00 am: Documenting protocols for high-throughput labs

11:00-11:30 am: Documenting protocols for smaller labs

11:30-12:00 noon: Data management processes

- Documentation for BOLD and BarSTool
- Potential use of Geneious or other systems

12:00-1:00 pm: Group lunch in Seaview Room

Session 3: In Situ and Ex Situ Training

1:00-1:30 pm: General issues to resolve:

- Relative need, benefits, costs, impact of *in situ* versus *ex situ* training
- Target audiences, background required, outcome goals for participants
- Instructors and support required
- General training or strictly linked to barcoding projects?
- Potential funding sources
- Optimal CBOL balance, criteria, for supporting *in situ* versus *ex situ* training

1:30-2:15 pm: Templates for *in situ* training:

- Agendas
- Class size
- Instrumentation, supplies, samples needed
- Curriculum
- Budgets

2:15-2:45 pm: *Ex situ* training:

- Which Leading Labs are hosting *ex situ* training? Lessons learned?
- Should there be a common curriculum?
- Sources of funding and criteria for selection?
- How to leverage existing fellowship/training opportunities?
- Should LL formalize and announce opportunities?

2:45-3:00 pm: Coffee Break

Session 4: Technical Resources for Training

3:00-3:15 pm: CBOL's goals and resources for 2008-2010 (*D. Schindel*)

3:15-3:35 pm: Presentation of proposed BOLI Data Portal (*M. Trizna*)

3:35-4:00 pm: Discussion of topics presented by BOLI Data Portal

4:00-4:20 pm: Presentation of proposed Leading Labs Network Website (*K. Jett*)

- Navigation and Content
- TaxTube
- WikiHow
- ToolBox

4:20-5:30 pm: Feedback and discussion

6:00 pm: Meet in hotel lobby for transportation to San Diego Zoo for Dinner Reception

TUESDAY, 24 JUNE 2008

8:00-8:30 am: Joint breakfast with DAWG in the *Seaview Room*

Session 5 and 6: Joint meeting with Data Analysis Working Group

8:30-9:00 am: Background for the joint DAWG-Leading Lab Meeting:

- DAWG discussion in Taipei
- Brief status report on DAWG software development activities

9:00-10:30 am: Issues on which DAWG wants input from Leading Labs:

- What's the current range of off-the-shelf analytical techniques available to barcoders?
- What's the 'ideal' interface for integrating information from heterogenous sources?
- What's the normal workflow for data analysis?
- How should current barcoding analysis tools be improved/extended to meet your needs?

- What technical limitations on data analysis have you encountered? (number of replicates, species, sequence length, number of regions)
- What are the data management challenges at the data analysis stage?
- What formats do you prefer for data input and output?
- What capabilities and approaches to data display and visualization would you like to see?

10:30-10:45 am: Mid-morning Break

10:45-12:00 noon: Desiderata from Leading Labs to DAWG

- For data management tools
- For data analysis tools
- For sample datasets
- For data display and visualization tools
- For software testing environment

12:00-1:00 pm: Group lunch in *Seaview Room*

Session 7: Roadmap for creating training resources 1: Website planning

- Missing topics
- Content providers
- Video opportunities
- Volunteer reviewers/ideas on how to handle vetting of information
- Contact/Volunteer representative from each participating lab
- Devo DreamTeam
- General comments

2:45-3:00 pm: Mid-afternoon Break

Session 8: Roadmap for creating training resources 2: *In Situ* Courses

- Calendar of potential training
- Training resources to assemble
- Fund raising

WEDNESDAY, 25 JUNE 2008

DAWG Participants will be meeting with staff of the Community Cyberinfrastructure for Advanced Marine Microbial Ecology Research and Analysis (CAMERA) at the University of California, San Diego. Leading Labs Network members may express their interest in attending this event.

Appendix 2: CBOL Data Analysis Working Group Agenda

23-24 June 2008: San Diego Marriott, Torrance Room

MONDAY, 23 JUNE 2008

8-8:30 am: Collective coffee and breakfast with Leading Labs in *Seaview Room*

8:30-8:45 am: Review of goals and deliverables under 2008-2010 Sloan Foundation grant in *Seaview Room (D. Schindel)*

8:45-9:00 am: Introductions (*N. Sarkar*)

9:00-9:15 am: Status of Proceedings Volume from Taipei Barcode Conference, and plan for publications of DAWG papers from Taipei (*N. Sarkar*)

9:15-9:45 am: Vladimir Pavlovic, "Alignment-free barcoding"

9:45-10:15 am: Saverio Vicario, "Phylogenetic Database Query System for Probabilistic Species Determination: Using a hierarchical approach to combine BLAST or HHM search with approximated Bayesian phylogenetic inference"

10:15-10:30 am: Coffee Break

10:30-11:00 am: Brian Golding, "Analyzing barcode sequences in a coalescent context"

11:00-11:30 am: Giovanni Felici, "Logic learning methods for barcode classification"

11:30-12:00 am: Donal Hickey, "Should DNA barcoding be both taxonomy-free and phylogeny-free?"

12:00-1:00 pm: Group lunch in Seaview Room

1:00-1:15 pm: Goals and plans for BOLI Data Portal (*D. Schindel*)

1:15-2:15 pm: Walk-through of BOLI Data Portal mock-up, followed by discussion (*M. Trizna*)

- Server location of Portal
- Major components of workflow
- Barcode data sources
- Data format standardization
- Identify software categories and analytical packages for Portal
- Identify model datasets and characteristics
- Multi-gene capabilities: What needs to be done for BOLI Data Portal to enable standard storage and analysis of multi-locus barcode data (e.g., pending 2 or 3 locus plant barcodes)?

2:15-2:45 pm: Henry Lu, "Information visualization and fusion for barcodes of life in environment and society"

2:45-3:30 pm: Robert Stones, "Bioinformatics Tools for Data Mining/Visualization" (Demo)

3:30-3:45 pm: Coffee Break

3:45-4:30 pm: Display and visualization of barcode data: discussion of needs and potential approaches

4:30-5:30 pm: Structured discussion of data portal development and implementation

6:00 pm: Meet in hotel lobby for transportation to San Diego Zoo for Dinner Reception

TUESDAY, 24 JUNE 2008

8:00-8:30 am: Collective coffee and breakfast with Leading Labs in *Seaview Room*

8:30-9:00 am: Background for the joint DAWG-Leading Lab Meeting:

- DAWG discussion in Taipei
- Brief status report on DAWG software development activities

9:00-10:30 am: Issues on which DAWG wants input from Leading Labs

- What's the current range of off-the-shelf analytical techniques available to barcoders?
- What's the 'ideal' interface for integrating information from heterogeneous sources?
- What's the normal workflow for data analysis?
- How should current barcoding analysis tools be improved/extended to meet your needs?
- What technical limitations on data analysis have you encountered? (number of replicates, species, sequence length, number of regions)
- What are the data management challenges at the data analysis stage?
- What formats do you prefer for data input and output?
- What capabilities and approaches to data display and visualization would you like to see?

10:30-10:45 am: Coffee Break

10:45-12:00 noon: Desiderata from Leading Labs to DAWG:

- For data management tools
- For data analysis tools
- For sample datasets
- For data display and visualization tools
- For software testing environment

12:00-1:00 pm: Group lunch in Seaview Room

1:00-1:30 pm: Recap of needs and priorities expressed by Leading Labs

1:30-3:00 pm: Discussion of:

- DAWG work-plan and timeline for construction of BOLI Data Portal
- Potential need for additional funding
- Design and leadership for possible grant proposals

3:00-3:30 pm: Coffee Break

3:30-5:00 pm: Final agreements on:

- Work-plan and division of labor
- Timeline and deliverables
- Project leadership
- Development of funding proposals
-

WEDNESDAY, 25 JUNE 2008

Optional on-site meeting with staff of the Community Cyberinfrastructure for Advanced Marine Microbial Ecology Research and Analysis (CAMERA) at the University of California, San Diego

Appendix 3: CBOL Implementation Board Agenda

25-26 June 2008: San Diego Marriott, Seaview Room

WEDNESDAY, 25 JUNE 2008:

7:30 AM-8:30 AM : Breakfast served in BAYSIDE SOUTH Room

Optional visit to CAMERA (Community Cyberinfrastructure for Advanced Marine Microbial Ecology Research and Analysis, UCSD)

12:00 – 12:45 pm: Lunch and preliminary discussion

1. Review of the role and responsibilities of Implementation Board (IB) in CBOL's Terms of Reference (David Schindel)
2. Self introduction by IB members
3. Report on CBOL's status, budget and goals (D. Schindel)
4. Goal of the first IB meeting

12:45 – 1:45 pm: Reports on barcode databases (15 minute presentations plus discussion)

5. GenBank/International Nucleotide Sequence Database Collaboration (Scott Federhen)
6. BOLD (Sujeevan Ratnasingham)

1:45 pm to 2:15 pm: Functioning of the Implementation Board

7. Discussion of mode of operation and leadership of Implementation Board
 - How will IB interact with CBOL's Executive Committee?
 - How IB will self-organize and interact (Listserve? Conference calls? Wiki?)

2:15– 3:15 pm: Status reports on activities/goals of Working Groups for 2008-2010

8. Plant WG (Peter Hollingsworth, 2:15-2:30; 5 min/5 slide presentation plus discussion)
9. Database WG (Robert Hanner, 2:30-2:45)
10. Data Analysis WG (Neil Sarkar, 2:45-3:00)
11. Leading Labs Network (Lee Weigt, 3:00-3:15)

3:15- 3:30 Coffee Break

3:30 – 4:30 pm: Status reports on activities/goals of Campaigns/Demonstration Projects

12. ABBI: All Birds Barcoding Initiative (Pablo Tubaro and Mark Stoeckle, 3:30- 3:45)
13. FISH-BOL (Robert Hanner, 3:45- 4:00)
14. MBI: Mosquito Barcoding Initiative (David Schindel, for Yvonne Linton, 4:00- 4:15)
15. TBI: Tephritid Barcoding Initiative (Marc de Meyer, 4:15-4:30)

4:30 – 5:45 pm: Plans and status of other initiatives

16. Moorea/BioCode (Chris Meyer, 4:30-4:45)
17. DBIC: DNA Barcoding Initiative for Conservation (George Amato, 4:45- 5:00)
18. MarBOL: Marine Barcoding (Dirk Steinke, 5:00-5:15)
19. TreeBOL (Damon Little, 5:15-5:30)
20. iBOL (Robert Hanner, for Paul Hebert, 5:30-5:45)

6:45 pm: Departure from hotel for group dinner

THURSDAY, 25 JUNE 2008:

8:00 am: Breakfast in conference room

8:30 – 10:30 am: Review and discussion of major CBOL goals and challenges in 2008-2010

21. BOLI Data Portal
22. e-Biosphere 09 Conference on Biodiversity Informatics, June 2009
23. Third International Barcode Conference, fall 2009
24. Participation and training
25. Public release of BARCODE data records
26. Funding for CBOL projects

10:30 – 10:45 am: Coffee break

11:00 – 12:00 noon: Informal reports of 23-24 June San Diego meetings

27. Leading Labs Network (Lee Weigt, 11:00-11:30)
28. Data Analysis WG (Neil Sarkarm 11:30-12:00)

12:00 – 12:30 pm: Discussion of IB leadership, possible selection of officers

12:30 – 2:00 pm: Working lunch: Informal introduction/discussion of potential initiatives

29. Fungi (Ursula Eberhart)
30. Algae (Line Le Gal)
31. Pollinators (David Schindel)

2:00 – 3:30 pm: Open discussion of potential integration/coordination across CBOL activities

3:30 – 4:00 pm: Coffee Break

4:00 – 5:30 pm: Final discussion to fix plans for:

32. Input to Leading Labs Network and BOLI Data Portal
33. Workshops and meetings of each CBOL activity
34. Cross-cutting activities to address shared challenges and opportunities
35. Proposal writing teams and targets for proposal submission
36. Interactions with CBOL Executive Committee
37. Next IB meeting

5:30 pm: Meeting adjourns

Appendix 4. Leading Labs Network

Name	Institution	Country	e-mail
Lee Weigt (Chair)	SI Laboratory of Analytical Biology	USA	weigt@si.edu
Rebecca Johnson	Australian Museum	Australia	Rebecca.Johnson@austrmus.gov.au
Janette Norman	Victoria Museum	Australia	jnorman@museum.vic.gov.au
Alex Smith	University of Guelph	Canada	salex@uoguelph.ca
Sujeevan Ratnasingham	University of Guelph	Canada	sratnasi@uoguelph.ca
Eric Pasquet	Museum National d'Histoire Naturelle	France	eric.pasquet@mnhn.fr
Sarah, Samadi	Museum National d'Histoire Naturelle	France	sarah@mnhn.fr
Daniel Masiga	African Insect Science for Food and Health	Kenya	dmasiga@icipe.org
Robert Skilton	Biosciences eastern and central Africa	Kenya	R.SKILTON@CGIAR.ORG
Kubata Bruno Kuilnga	International Livestock Research Institute	Kenya	b.kubata@cgiar.org
Johnson Ouma	Kenya Agricultural Research Institute	Kenya	joouma@gmail.com
Chang-Bae Kim	Korea College of Natural Sciences	Korea	cbk06511@hanmail.net
Ursula Eberhardt	Centraal Bureau voor Schimmelcultures (CBS)	Netherlands	u.eberhardt@cbs.knaw.nl
Biff Bermingham	Smithsonian Tropical Research Institute	Panama	Bermingham@si.edu
Oris Sanjur	Smithsonian Tropical Research Institute	Panama	SANJURO@si.edu
Ferozah Conrad	Kirstenbosch Research Centre	South Africa	conrad@sanbi.org
Tun-Yuan Cheng	Academia Sinica	Taiwan	guppy-fish@yahoo.com.tw
Wen-Bin Yeh	National Chung-Hsin University	Taiwan	wbyeh@nchu.edu.tw
Peter Hollingsworth	Royal Botanical Gardens, Edinburgh	UK	p.hollingsworth@rbge.ac.uk
Robyn Cowan	Royal Botanical Gardens, Kew	UK	r.cowan@rbgkew.org.uk
Matt Leslie	American Museum of Natural History	USA	mleslie@amnh.org
Sergios-Orestis Kolokotronis	American Museum of Natural History	USA	koloko@amnh.org
Joyce Gross	Berkeley Natural History Museums	USA	joyceg@berkeley.edu
Sarah Burgess-Herbert	CRES – San Diego Zoo	USA	SBurgess-Herbert@sandiegozoo.org
Damon Little	New York Botanical Garden	USA	Dlittle@nybg.org
Christopher Meyer	Smithsonian/ Moorea	USA	MeyerC@si.edu
Harim Cha	UCSD	USA	harim.cha@gmail.com
UNABLE TO ATTEND			
Yun Gao	Chinese Academy of Sciences	China	gaoy@mail.kiz.ac.cn
Min Feng	Institute of Botany, Chinese Academy of Sciences	China	fengmin@ibcas.ac.cn
Shi-liang Zhou	Institute of Botany, Chinese Academy of Sciences	China	slzhou@ibcas.ac.cn

Appendix 5. Data Analysis Working Group

Name	Institution	Country	e-mail
Neil Sarkar (Chair)	Marine Biological Lab	USA	sarkar@mbi.edu
Donal Hickey	Concordia College	Canada	dhickey@alcor.concordia.ca
Brian Golding	McMasters University	Canada	Golding@McMaster.CA
Chi Pang Li	The Chinese University of Hong Kong	Hong Kong	b492714@mailserv.cuhk.edu.hk ; albert2003@netvigator.com ;
Giovanni Felici	CNR Istituto di Analisi dei Sistemi ed Informatica	Italy	felici@IASI.CNR.IT ; giovanni.felici@gmail.com
Saverio Vicario	Istituto di Tecnologie Biomediche, sede di Bari, CNR	Italy	saverio.vicario@GMAIL.COM
Henry Horng-Shing Lu	National Chiao Tung University	Taiwan	hslu@STAT.NCTU.EDU.TW
Robert Stones	Central Science Laboratory	UK	r.stones@CSL.GOV.UK
Pavel Kuksa	DIMACS, Rutgers University	USA	pkuksa@CS.RUTGERS.EDU
Vladimir Pavlovic	DIMACS, Rutgers University	USA	vladimir@CS.RUTGERS.EDU
Kasper Munch	UC Berkeley, Department of Integrative Biology	USA	kaspermunch@berkeley.edu

Appendix 6. Implementation Board meeting participants

Name	CBOL Activity	Institution	Country	e-mail
Pablo Luis Tubaro	ABBI	Museo Argentino de Ciencias Naturales	Argentina	ptubaro@interlink.com.ar
Paul de Barro	Agriculture/trade	Long Pocket Laboratories	Australia	paul.debarro@csiro.au
Marc De Meyer	TBI	Royal Museum for Central Africa	Belgium	marc.de.meyer@africamuseum.be
Sujeevan Ratnasingham	BOLD	University of Guelph	Canada	sratnasi@uoguelph.ca
Dirk Steinke	MarBOL	University of Guelph	Canada	dsteinke@uoguelph.ca
Line Le Gall	Algae	Museum National d'Histoire Naturelle	France	legall@mnhn.fr
Daniel Masiga	Africa region	African Insect Science for Food and Health	Kenya	dmasiga@icip.e.org
Ursula Eberhardt	fungi	Centraal Bureau voor Schimmelcultures (CBS)	Netherlands	u.eberhardt@cbs.knaw.nl
Peter Hollingsworth	Plant WG	Royal Botanical Gardens, Edinburgh	Scotland	p.hollingsworth@rbge.ac.uk
Mark Stoeckle	ABBI	Rockefeller University	USA	MarkStoeckle@nyc.rr.com
George Amato	Conservation Initiative	American Museum of Natural History	USA	gamato@amnh.org
Neil Sarkar	Data Analysis WG	Marine Biological Lab	USA	sarkar@mbl.edu
Scott Federhen	GenBank/NCBI	NCBI	USA	federhen@ncbi.nlm.nih.gov
Lee Weigt	Leading Labs	SI Laboratory of Analytical Biology	USA	weigt@si.edu
Christopher Meyer	Moorea/BioCode	Smithsonian/ Moorea	USA	MeyerC@si.edu
Damon Little	TreeBOL	New York Botanical Garden	USA	Dlittle@nybg.org

Appendix 7. Participants in the CAMERA visit

Paul Gilna	CAMERA/UCSD	USA	E-mail: pgilna@ucsd.edu
John Wooley	UCSD	USA	jwooley@ucsd.edu
William H. Piel	Yale Peabody Museum/Assembling the Tree of Life	USA	william.piel@yale.edu
Richard Ree	Field Museum of Natural History/Encyclopedia of Life	USA	rree@fieldmuseum.org