

International Journal of the Commons

Vol. 13, no 1 2019, pp. 205–243

Publisher: Uopen Journals

URL: <http://www.thecommonsjournal.org>

DOI: 10.18352/ijc.859

Copyright: content is licensed under a Creative Commons Attribution 3.0 License

ISSN: 1875-0281

## Governance of a global genetic resource commons for non-commercial research: a case-study of the DNA barcode commons

Janis Geary

Department of Medicine, University of Alberta, Canada

[janisgeary@gmail.com](mailto:janisgeary@gmail.com)

Tania Bubela

Faculty of Health Sciences, Simon Fraser University, Canada

[tbubela@sfu.ca](mailto:tbubela@sfu.ca)

**Abstract:** Life sciences research that uses genetic resources is increasingly collaborative and global, yet collective action remains a significant barrier to the creation and management of shared research resources. These resources include sequence data and associated metadata, and biological samples, and can be understood as a type of knowledge commons. Collective action by stakeholders to create and use knowledge commons for research has potential benefits for all involved, including minimizing costs and sharing risks, but there are gaps in our understanding of how institutional arrangements may promote such collective action in the context of global genetic resources. We address this research gap by examining the attributes of an exemplar global knowledge commons: The DNA barcode commons. DNA barcodes are short, standardized gene regions that can be used to inexpensively identify unknown specimens, and proponents have led international efforts to make DNA barcodes a standard species identification tool. Our research examined if and how attributes of the DNA barcode commons, including governance of DNA barcode resources and management of infrastructure, facilitate global participation in DNA barcoding efforts. Our data sources included key informant interviews, organizational documents, scientific outputs of the DNA barcoding community, and DNA barcode record submissions. Our research suggested that the goal of creating a globally inclusive DNA barcode commons is partially impeded by the assumption that scientific norms and expectations held by researchers in high income countries are universal. We found scientific norms are informed by a complex history of resource misappropriation

and mistrust between stakeholders. DNA barcode organizations can mitigate the challenges caused by its global membership through creating more inclusive governance structures, developing norms for the community that are specific to the context of DNA barcoding, and through increasing awareness and knowledge of pertinent legal frameworks.

**Keywords:** DNA barcoding, genetic resources, global, heterogeneity, knowledge commons, non-commercial research

**Acknowledgement:** The authors would like to express their sincere gratitude to interviewees who graciously volunteered their time, candor, and expertise to this work. We would also like to thank Drs. Paul Hebert and David Castle for supporting our work within the DNA barcoding community, and numerous other individuals who provided additional support. We thank Drs. Ashok Kumbamu and Harley Ferreira de Cerqueira in helping conduct interviews, Mark Bieber and Westerly Luth for their work on the bibliometric analysis, and Mark Bieber and Jennifer-Ann McGetrick for their work on the medicinal plant record analysis. We would also like to thank individuals who reviewed drafts of this manuscript and provided expert feedback including: Drs. Cindy Jardine, Trish Reay, Karen Goodman, and Felix Sperling. The author JG's work was supported by the following: Canadian Institutes of Health Research Doctoral Award (201010); Alberta Innovates Health Solutions Studentship; and Izaak Walton Killam Memorial Scholarship.

## 1. Introduction

Life sciences research can be enhanced through collective action to create and manage genetic resources and their supporting infrastructure. Infrastructure includes databases for sequence and associated metadata and bio-repositories for biological samples. Collectively, data, metadata and biological samples comprise a knowledge commons – resources managed according to terms that encourage efficiency, equitable use, and sustainability. Collective action by stakeholders to create and use knowledge commons has potential benefits for all involved, including minimizing costs and sharing risks. However, gaps remain in understanding how institutional arrangements may promote collective action in a global context.

The natural resources-based Institutional Analysis and Development (IAD) framework to analyse commons governance has been modified to account for differences with knowledge commons (KC-IAD) (Berge and Laerhoven 2011; Hess and Ostrom 2011; Frischmann et al. 2014). Knowledge commons are generally non-rivalrous and boundless, which makes the exclusion of some users difficult. Many knowledge commons are built to solve a particular problem; their governance must encourage both creation and use of the resource. This burden of creating knowledge commons and distributing benefits derived from their utilization is not necessarily equitably shared among participants in the commons (Bubela et al. 2012; Strandburg et al. 2014).

To date, analyses have focused on high income country (HIC) knowledge commons (Dedeurwaerdere 2010b; Bubela et al. 2012; Contreras 2014). Little research has focused on knowledge commons that engage a globally heterogeneous community, with institutional and individual participants from regions with historical power and economic imbalances. In addition to economic, language, and cultural differences, global participation makes communication to facilitate collective action difficult.

We address this research gap with a case study of a global knowledge commons: The DNA barcode commons. The DNA barcode commons facilitates large-scale documentation of life on earth and identification of unknown specimens. Identification compares the barcode sequence of unknown specimens against a comprehensive barcode reference database. DNA barcoding proponents have led international efforts to make DNA barcodes a standard species identification tool for taxonomic and biodiversity research and to incorporate their use into regulatory practices that require species identification.

With coordination efforts, the DNA barcoding community has built or adapted infrastructure (databases and biorepositories), produced millions of barcode records (Ratnasingham 2015), and published thousands of scientific papers (Bubela et al. 2015a,b). Despite these key indicators of a successful commons, the global DNA barcoding environment is characterized by an inequitable distribution of risks and benefits from the use of the genetic resources that comprise the commons.

Our research examined if and how DNA barcode commons attributes, including resource governance and infrastructure management, facilitate global participation in DNA barcoding efforts. To answer these questions, we employed a case-study approach guided by the KC-IAD framework, which enabled the identification of how the attributes of the DNA barcode commons led to governance challenges. We conclude with recommendations that promote collective action and further the goals of this global research commons.

### **1.1. DNA barcoding**

In January 2003, Paul Hebert and colleagues proposed DNA barcodes as a standardized species identification system (Hebert et al. 2003). By December, proponents had developed DNA barcode standards, worked to overcome opposition to using DNA barcodes for species identification, and begun building a global DNA barcode network (Stoeckle 2003). The first formal organization created in 2004 for DNA barcoding was the Consortium for the Barcode of Life (CBOL), Smithsonian, Washington, DC.

DNA barcoding gained global momentum because it enabled a range of practical applications (Hebert et al. 2003). An open access, comprehensive database of DNA barcode records facilitated rapid identification of unknown specimens in situations where morphological identification was impossible, for example, where traditional taxonomic expertise is unavailable or the specimen lacks distinguishing features, such as butchered meat or insect larvae. Proponents envisaged

shipping unknown specimens to a laboratory equipped to produce low-cost DNA barcodes, which could then be matched against known barcodes (Pennisi 2003).

Paul Hebert led an international initiative to build a comprehensive barcode reference database. Canadian funders supported infrastructure development, including the Canadian Centre for DNA Barcoding within the Centre for Biodiversity Genomics, and the Barcode of Life Data System (BOLD) in 2007 (Ratnasingham and Hebert 2007) at the University of Guelph. The iBOL Project launched in 2010, funded through Genome Canada's International Consortium Initiative. iBOL included 28 nations as 'nodes' partnered through formal agreements (iBOL 2015b).

Barcoders at the 2015 6th International Barcode of Life Conference participated in a workshop to establish the International Society for the Barcode of Life (ISBOL). ISBOL will "coordinate completion of the [barcode] registry, to facilitate the development of barcode applications and to communicate with stakeholders" (Castle et al. 2015). An interim governance council to initiate ISBOL was created, comprising the authors of the Kunming Declaration on the Promotion of DNA Barcoding and Biodiversity Science (Li et al. 2013) and representatives from key regions and organizations. The council is seeking feedback on proposed structure and governance from the broader DNA barcoding community.

## 2. Methods

### 2.1. Case study approach

We conducted a mixed-method case study to analyse how the factors outlined in the KC-IAD framework influence DNA barcode commons governance (Figure 1). The use of the KC-IAD framework is one of the main sources of rigor in our research (Mayan 2009), because our analysis and interpretations of data were drawn from previous knowledge of how factors and variables in a knowledge commons relate to each other. Data derived from a document and literature search, key informant interviews, bibliometric analysis of barcoding publications, and an analysis of barcode record submissions to BOLD.

Our use of quotes in reporting provides confirmability and demonstrates to the reader that the results are grounded in data (Morse et al. 2002). Additionally, we actively sought input and feedback from the DNA barcoding community beyond their participation in formal interviews to add to the credibility of our research (Given and Saumure 2013). The lead author, JG, visited the Biodiversity Institute of Ontario (BIO), which leads barcoding efforts, in May 2012 to learn about the facility and its workflow for producing barcode records. We shared interview guides with barcode leaders and organizational administrators and invited feedback to ensure that questions were relevant to the barcoding community. We presented preliminary findings at three international DNA barcoding conferences (Bubela 2013; Bubela et al. 2015a; Geary and Bubela 2015; Geary et al. 2016) and invited feedback from conference attendees and interviewees. We co-organized a workshop in February 2013 that discussed medicinal plant barcoding and

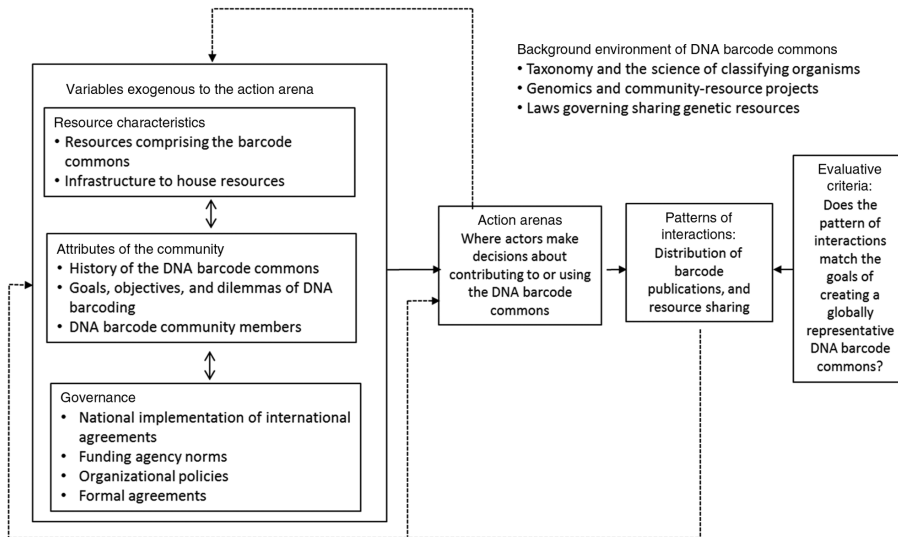


Figure 1: The DNA barcode commons described within the KC-IAD Framework (Frischmann et al. 2014).

issues related to sharing genetic resources. The workshop resulted in a publication with leaders in the barcoding community (Schindel et al. 2015).

### 2.1.1. Document search and analysis

We collected publicly available documents about DNA barcoding procedures, protocols, and history from the iBOL and CBOL websites in 2012, with a repeat search in 2015 (iBOL 2012; Consortium for the Barcode of Life 2015). We obtained additional document from iBOL staff during a visit to the Canadian Centre for DNA Barcoding in 2010. We reviewed key publications detailing: the science of DNA barcoding (Hebert et al. 2003); controversies about the science (Moritz and Cicero 2004; Gregory 2005; Dupuis et al. 2012; Collins and Cruickshank 2013); the international efforts (Schindel et al. 2008; Schindel 2010; Vernooy et al. 2010; Schindel et al. 2015); potential applications (Wong and Hanner 2008; Yancy et al. 2008; Gross 2012); organizational efforts of DNA barcoding proponents (Adamowicz 2015; Castle et al. 2015); and database-building efforts (Ratnasingham and Hebert 2007; Sonet et al. 2013).

### 2.1.2. Key informant interviews and analysis

The authors and two research assistants interviewed expert key informants from 14 countries, including 35 individuals who participated in DNA barcoding projects, three policy makers involved in funding and DNA barcoding project oversight, and 12 individuals involved in genetic resource governance. This research received ethical approval from the University of Alberta Research Ethics

Board – Health Panel. We conducted half the interviews at iBOL conferences in Adelaide, Australia (2011) and Kunming, China (2013). Other interviews were by phone (n=8) or in-person (n=17). We used a semi-structured interview guide developed based on subject matter knowledge.

We analysed interview transcripts and documents using the KC-IAD framework as an *a priori* frame to guide our content analysis in NVivo qualitative analysis Software (QSR International Pty Ltd. Version 10, 2012). Prior, we listened to each interview to verify the transcription, and make notes about central concepts to inform subsequent data collection and analysis. Based on KC-IAD categories, we assigned descriptive codes to each statement, grouped codes to form themes, and examined the themes within and between each interviewee group. When reporting direct quotes, we edited quotes for grammar and clarity.

We grouped interviewees based on whether their main work affiliation was in a Like-Minded Mega Diverse Country (LMMC) or a non-LMMC. LMMC are a group of countries established in 2002 to promote their similar interests in protecting biodiversity (LMMC 2002). LMMCs included: Brazil, China, Colombia, Ghana, India, Indonesia, Kenya, Mexico, and South Africa, and non-LMMCs included: Australia, Canada, New Zealand, United Kingdom (UK), and the United States (US). Despite not being a member of the LMMC, we included Ghana because its interests aligned with other African countries. In cases where identifying the country would risk identifying an individual participant, we have referenced the individual's region.

### **2.1.3. Bibliometric analysis of DNA barcode publications**

We searched the Scopus database for peer-reviewed literature that referenced any of four seminal barcode papers (Hebert et al. 2003; Stoeckle and Hebert 2008; Hollingsworth et al. 2009; Schoch et al. 2012). We compiled a database of information about each article including: publication source, publication year, number of citations, author names, and institutional affiliations of authors. The research team's data specialist and programmer, Mark Bieber, developed a customized author-name disambiguation program that combined synonymous names of single individuals and separated identical names of different individuals.

From the resultant author-publication database, we identified authors with institutional affiliations in high, middle, or low income countries (World Bank 2016); and whether or not the paper was published in a highly ranked journal. The large number of authors in our publication dataset allowed us to use the four-category Gross National Income (GNI) per capita levels from the World Bank (upper income,  $\geq$ \$12,476; upper middle income, \$4036 to \$12,475; lower middle income \$1026 to \$4035; and low income,  $\leq$ \$1025) rather than the dichotomous categorization used for our qualitative analysis of interviews (World Bank 2016). We used InCites Journal Citation Reports (Thomson Reuters 2016) to identify the top 10 ranking journals in each field category relevant to DNA barcoding. We used Stata v. 11 to calculate odds ratios (ORs) and 95% confidence intervals (CI) as measures of association between authors'

country income levels and biodiversity status and the outcome of publication in highly ranked journals.

We used Gephi 0.8.2 Beta (Bastian et al. 2009) to geographically display the location of the primary affiliation of each author in the database and to visualize links between authors based on co-authorship on a single paper. We described the author sets (the set of authors on a single paper) based on the proportions of papers that span across different geographic regions.

#### **2.1.4. Analysis of barcode record submissions to BOLD**

We examined patterns of sharing barcode records and biological specimens across different country income levels (World Bank 2016), using two exemplars: barcodes of medicinal plants and barcodes of mosquitoes. Sharing with respect to medicinal plants raises heightened concerns among barcoding participants because of the potential for misappropriation of benefits from commercially valuable medical applications. On the other hand, the potential to use barcodes to rapidly identify mosquitoes has public health implications. Each of the mosquito genera, *Anopheles*, *Aedes*, and *Culex* include species that are distributed worldwide and transmit diseases (including malaria, yellow fever, and West Nile fever, respectively) (WHO 2016).

We accessed barcode record information from two user interfaces within BOLD: The taxonomy browser and the public data portal. The taxonomy browser allows users to search the database for information about a specific taxonomic category (genus to phylum), and it includes summary information for published (i.e. the record producer has made it available to view or download) and unpublished records (i.e. the record producer has not made it available to view or download). Users can view basic information about the taxonomic group, how many specimen records are in the database, how many of those records have been published, and the country-of-origin of specimens. The public data portal allows users to search based on a variety of factors (e.g. geographical identifiers, name of specimen collector, taxonomic groups), and download published barcode records individually or in batches. Users can download custom datasets, including sequence trace files, all available taxonomic information, where each specimen was collected (global positioning system (GPS) coordinates and/or country) and stored (institution name), and other metadata, such as time of collection. The public data portal also mines barcode gene region sequences from GenBank.

We downloaded plant records from the BOLD public data portal in January 2013 (150,220 records). We created a list of 17,895 medicinal plant records on BOLD by using a table look up function in our database to cross-reference the BOLD plant records with a list of 1300 known medicinal plant species names (obtained from <http://www.ars-grin.gov/duke/ethnobot.html>). We were unable to search each of the 1300 plants in the taxonomy browser, so we did not estimate unpublished medicinal plant records. Of the identified public medicinal plant records, 5788 included the latitude and longitude where the specimen was collected, and an additional 8151 included specimen's country of origin.

Of 3956 records without any specimen collection information, 2036 were mined from GenBank. We created a variable that indicate if the specimen was stored in the country where it was collected, and we used SPSS v.19 to tabulate the published medical plant records separately by the country income level, which allowed us to determine the proportion of materials that are stored outside of the country of origin for different country income levels.

We downloaded barcode records for each mosquito genus from the public data portal. Because we were only interested in three mosquito genera, we were able to search each one using the taxonomy browser. We could therefore count the number of unpublished records for each mosquito genera. By tabulating published and unpublished records separately by country income level, we were able to approximate the number of barcode records that were produced in different countries, as well as the proportion of barcode records not shared via publication.

## 2.2. Developing recommendations for the DNA barcode community

After we completed the above analyses, we situated the results within the KC-IAD framework (detailed in Figure 2). We used existing knowledge of how factors within the framework impact each other and influence action arenas to infer how our observations contribute to challenges in collective action and overall governance of the DNA barcode commons.

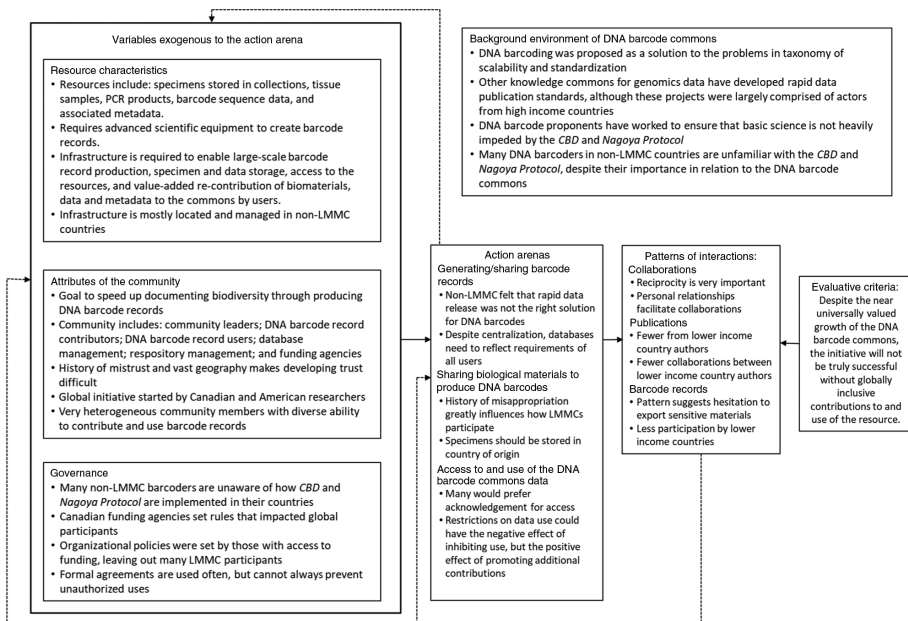


Figure 2: Summary of findings situated in the KC-IAD.



### 3. Results

The results of our case-study are summarized in Figure 2. In the following description of results, we first describe the background environment of the DNA barcode commons, followed by a description of its resources and infrastructure, attributes of its community, governance, action arenas, patterns of interactions, and evaluative criteria.

#### 3.1. Background environment

##### 3.1.1. Taxonomy and the science of biological identifications through DNA barcodes

For most of taxonomy's history, taxonomists differentiated species based on morphological distinctions, which is slow and requires expertise. Only 10% of an estimated 10–20 million species have been described over the last 250 years (Wilson 2003). Nevertheless, documenting global biodiversity is critical for mitigating anthropogenic and other threats, including climate change and habitat destruction (Hebert et al. 2003). Taxonomy experienced a resurgence in the early 2000s with the advance of DNA sequencing technology and bioinformatics infrastructure (Waterton et al. 2013); these methods are faster, cheaper, and less dependent on human resources than traditional taxonomy (Hebert et al. 2003). However, proposals to expand DNA-methods for taxonomy were met with resistance (Tautz et al. 2003), because the use of non-standardized gene regions to differentiate species prevented automated analyses at the scale needed to document biodiversity (Moritz and Cicero 2004).

In 2003, Paul Hebert proposed DNA barcodes, which are short and ubiquitous gene sequences, as a solution to the scalability and standardization issues. While DNA barcoding was not accepted by taxonomists without controversy and debate (Will and Rubinoff 2004; Ebach and Holdrege 2005; Dupuis et al. 2012), it has nevertheless gained prominence as a taxonomic tool. As of August 2017, the BOLD barcode database included 5625K barcode records (published and unpublished) (BOLD Systems 2015), and GenBank, the globally recognized open access repository for genetic sequences, contained 1,846,059 sequences labeled as “barcodes”.

##### 3.1.2. Knowledge commons for genomics data

The DNA barcoding initiative was influenced by another coordinated effort to generate sequence databases: the Human Genome Project (HGP) (Collins et al. 2003). Starting in 1991, the International Human Genome Sequencing Consortium at the National Human Genome Research Institute (NHGRI) published rapid data release standards for the HGP. The rapid data release principles were reconfirmed in the 1996 Bermuda Accord (Bermuda Sequence Policies Archive 2016). In 2003, genomics leaders convened in Fort Lauderdale to discuss updating standards (NHGRI 2003) based on the assumption that rapid and free sequence data release would promote scientific and public interests. The principles were updated and expanded in 2009 (Toronto International Data Release Workshop Authors 2009).

Our case study points to a notable omission from current genomics data release principles, namely, consideration of whether the benefits of creating open access databases could accrue equitably in both high and lower income countries. While there has been no explicit exclusion of developing country stakeholders in large-scale genomics projects, few such stakeholders have been involved in policy setting (Helmy et al. 2016). For example, 96% of the 71 authors of the Toronto principles were from the US, Canada, and European countries (Toronto International Data Release Workshop Authors 2009).

Furthermore, the genomics databases are located in HICs. NCBI, which manages GenBank in the US, is part of the International Nucleotide Sequence Database Collaboration (INSDC) with the European Bioinformatics Institute and the DNA Data Bank of Japan. The three databases mirror each other and exchange data daily (Benson et al. 2013).

### 3.1.3. Laws governing genetic resource sharing and utilization

The DNA barcode commons comprises genetic resources defined as “genetic material of actual or potential value” (United Nations 1992). As such, it is subject to international legal instruments that govern genetic resources, their derivatives, and associated traditional knowledge. However, several interviewees involved in various BOL activities were unaware of the legal instruments or did not believe in their applicability. At the international level, genetic resource sharing is addressed by the *Convention on Biological Diversity (CBD)* and the related *Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization (Nagoya Protocol)*. The *CBD* sets out three objectives, including the fair and equitable sharing of benefits arising from the utilization of genetic resources. The *Nagoya Protocol* provides a legal framework to implement the access and benefits sharing (ABS) objectives of the *CBD*. Its development was driven by biodiversity-rich countries to combat misappropriation of genetic resources.

While the *CBD* and the *Nagoya Protocol* grant national sovereignty over genetic resources and mechanisms to protect such resources, they also encourage countries to provide access. LMMCs came together in 2002 to protect legitimate interests in how they govern access to their biodiversity (LMMC 2002). Researchers in HICs were concerned the *Nagoya Protocol* would negatively impact non-commercial biodiversity research that supported the objectives of the *CBD* (Schindel 2010). As stated by one interviewee:

*To be honest, I haven't been updated on the [CBD]. But from what information I have, I do have some serious concerns about the way biological resources are being treated because I have no commercial interest in using biodiversity to apply for a patent and stuff like that – Researcher, Canada*

In 2008, CBOL co-hosted a workshop to address the challenges that an overly-restrictive ABS agreement could create for non-commercial research. The group put forward a statement to the *CBD Conference of the Parties 2010 (COP10)*

to suggest provisions for simplified measures to access genetic resources for non-commercial research (Schindel et al. 2008), and such provisions were included when the *Nagoya Protocol* was adopted at COP10 (United Nations 2010).

Patterns of national implementation of the *CBD* and *Nagoya Protocol* pose additional challenges for research. While the *CBD* has 196 parties since entering into force in 1993, the US is notably not a party to either the *CBD* or the *Nagoya Protocol* (United Nations 1992; United Nations 2010). As a major participant in global biodiversity research and development, the non-participation of the US results in the perception that the *CBD* has less impact than it should. Since entering into force in 2014, the *Nagoya Protocol* has 109 parties (in effect as of October 2018), although only 30 of the 109 are high-income (World Bank 2016). Because few high income countries are parties to this agreement, it is difficult to enforce unauthorized genetic resource use. Eleven of the 28 countries that participate in iBOL are not party to the *Nagoya Protocol* (United Nations 2010).

### 3.2. Resources and infrastructure

#### 3.2.1. The DNA barcode production pipeline

Barcode records are a unique type of archived genetic information in that they comprise short, standardised DNA sequences that are linked to a stored specimen and other metadata, which provides the necessary information to use the record for taxonomic identification. The process for creating a barcode record (Figure 3) begins with a taxonomically-identified specimen. Specimens can be derived from collections, or collected from the field and subsequently stored as a reference. Only a small sample is needed to extract DNA, and amplify the barcode gene region(s) using polymerase chain reaction (PCR). Unlike whole specimens or extracted DNA, PCR products only contain the small barcode region, not the entire genome. Individuals who wish to create barcode records for specimens, but do not have access to PCR equipment, can ship whole specimens or extracted DNA to sequencing facilities. A DNA barcode sequence becomes a barcode record once it is produced, quality controlled and linked to its metadata (information that describes the data). Metadata include dates on which specimens were collected and by whom, where the reference specimens are stored, and primer sequences used to generate the barcode sequences. The barcode records may include photographs of the specimens. Barcode records enable scientific, curiosity-based, and regulatory uses by others.

In sum, the resources that comprise the DNA barcode commons include: specimens stored in collections, tissue samples, PCR products, barcode sequence data, and associated metadata.

#### 3.2.2. Infrastructure to house DNA barcode resources

The DNA barcode commons requires infrastructure to enable large-scale barcode record production (A-C of Figure 3), specimen and data storage (D-F of Figure 3), access to the resources (H of Figure 3), and value-added re-contribution of biomaterials, data and metadata to the commons by users (I of Figure 3).

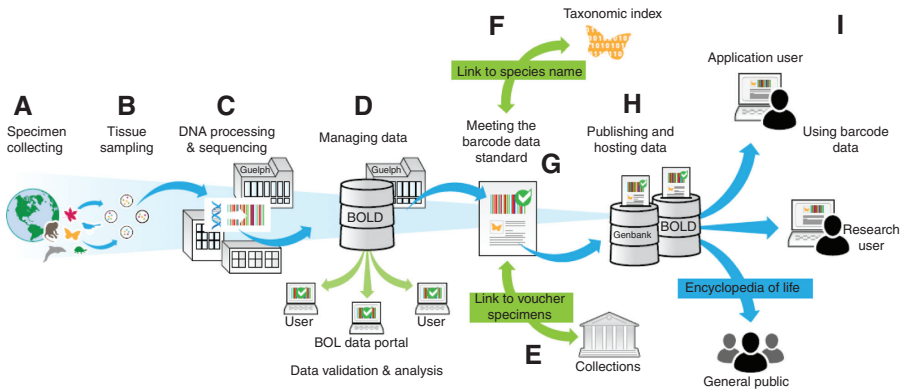


Figure 3: The DNA Barcode Pipeline (pipeline image reproduced with permission from CBOL (Consortium for the Barcode of Life 2015)).

Since 2003, barcoders have used existing data infrastructure (e.g. GenBank) to store barcode sequences (Hanner 2009). In 2005, CBOL formed a working group to develop data standards for barcode records stored in international nucleotide databases. Researchers predicted the barcode database size and specialized informatics would necessitate independent data infrastructure (Ratnasingham and Hebert 2007). In 2007, the Canadian Centre for DNA Barcoding launched BOLD, which included 14,000 users from 94 countries in 2015 (Ratnasingham 2015).

BOLD is now established as the main barcode record curator (5624K as of Aug 2017, (BOLD Systems 2015)), and it includes open access and privately-held data. The online system allows individuals to work with their barcode sequences on a private “workbench”; they can later publish the sequence to the open access database (accessible through BOLD’s public data portal). Interviewees expressed preference for BOLD over other databases like GenBank. As one interviewee explained,

*One of the nice things about the BOLD database is that it allows you to include a bunch of other data, than just the genetic data like in GenBank. That’s especially important for doing biodiversity studies. It adds capacity to what one might want to do with the data after it has been collected and is made available – Researcher, US*

The BOLD platform allows researchers to curate and analyse their barcode data before the records are published. Researchers can view raw sequence outputs and metadata, and download barcode record compilations from the open access database in several formats, enabling statistical comparisons (.xml, .tsv) and phylogenetic analyses (FASTA, TRACE). BOLD enables anyone to search and view data; it provides a taxonomy browser that allows users interested in specific taxonomic

groups to view the progress of DNA barcoding efforts and read descriptions. BOLD communicates with other platforms, and sequences published within BOLD are copied to GenBank. Barcode records, once published, are not subject to any restrictions on their use.

In addition to barcode sequencing facilities, barcoders need infrastructure to house at least one reference specimen for each unique barcode record. Specimens must be stored in a repository where they can be re-examined, if necessary, to verify the taxonomic identification (Moritz and Cicero 2004; DeSalle et al. 2005). Specimens may be housed as part of museum collections (e.g. the Natural History Museum of the Smithsonian in Washington, DC), in botanical gardens (e.g. Kew Royal Botanic Gardens in the UK), in bio-repositories, such as seed repositories (e.g. Svalbard Global Seed Vault in Norway), and as part of private collections held at research institutions. BIO includes units to manage the reference specimens it receives.

### 3.3. Attributes of the DNA barcoding community

#### 3.3.1. Goals and dilemmas

The DNA barcode commons goals are to: speed up the documentation of global biodiversity, facilitate monitoring, and enable a broad array of applications based on an open access, globally-representative DNA barcode record database (iBOL 2015e). Similar to other knowledge commons, the value of the DNA barcode commons increases as more people contribute to the resource, use it for intended and novel uses, and re-contribute value-added data (network effect) (Schofield et al. 2009; Dedeurwaerdere 2010a; Bubela et al. 2012). Individuals might stop contributing to the commons if they feel others are utilizing the resource without contributing to it (“free-riding”) (Dedeurwaerdere 2010b). In the research context, this translates to a fear of “being scooped” in publication priority; this fear is common to many scientific disciplines (Contreras 2010; Joly et al. 2012) and was frequently cited by interviewees in our study. LMMC interviewees preferred data release be delayed until after publication. One interviewee described the extent of a colleague’s concerns about data release prior to publication:

*She had the [publication] proofs and some email came telling her to release the data. And she didn’t want to. I had to speak with her and I had to tell her, “[It’s] no problem if you release the data” and then “No, no, but I don’t want to” although the paper is accepted, I had to tell her “nobody is going to steal your data” – Researcher, Mexico*

In addition, the success of the barcode commons relies on *global* participation, representative of *global* biodiversity. Other well-characterized research commons comprise resources that can feasibly be obtained and managed by less diverse research communities, such as the biomedical research commons of mouse-related research models and reagents (Bubela et al. 2012; Mishra and Bubela 2014; Bubela et al. 2017). A researcher from South Africa expressed this sentiment:

*To me [having formal participation by African countries] is hugely important, it's actually central. If the goal of iBOL is a global database of biodiversity, you can't speak of a global database if you've left out Africa because Africa is a major continental mass with a major coast line.*

Global participation presents additional collective action challenges because of the concerns held by LMMCs about genetic resource misappropriation. Although legal frameworks protect against barcode record misappropriation for commercial research, some LMMC interviewees explained that lack of trust remained a significant barrier to shipping specimens to out-of-country sequencing centres or storing barcode records on foreign servers. Non-LMMC interviewees were less cognizant of challenges related to trust and, in some cases, brushed off the issue:

*I think the international community is way past the sort of mid-20th century colonial style attitude where samples were harvested from biodiversity and permanently relocated into technology rich countries. I think the mentality of the global research community has gotten over it. – Researcher, Canada*

Sharing genetic resources presents a challenge for non-commercial research, because the resources that comprise the commons are not evenly distributed among global actors; LMMC participants in the DNA barcode commons are more likely to have access to biodiversity to build the commons, and non-LMMC participants are more likely to have access to research funding and infrastructure to use the commons. These inequities present a variation of the free-rider dilemma: the barcoding effort may be perceived to free-ride on LMMC biodiversity, because it inadequately provides benefits associated with the use of genetic resources. Thus governance must ensure that differential participation and resource commitments (e.g., research infrastructure vs biological specimens) merit an equitable distribution of benefits and burdens.

Scholars have demonstrated commons participants often develop the necessary trust to overcome these challenges through face-to-face communication (Ostrom 2003). This communication is hampered by the distance between global actors, resulting in a significant dilemma as to how to effectively govern this global knowledge commons.

### **3.3.2. Community members**

Our analysis identified six categories of actors in the DNA barcode community: community leaders; contributors of DNA barcode records (contributors); DNA barcode record users; databases; repositories; and funding agencies.

#### **3.3.2.1. Community leaders**

Many interviewees spoke about individuals who influenced DNA barcoding and developing this research commons. Paul Hebert led barcode infrastructure funding initiatives at the University of Guelph (iBOL 2015c). Scott Miller and David Schindel led the Consortium for the Barcode of Life, which shaped DNA

barcoding policies to create standards for barcode records (Consortium for the Barcode of Life 2015).

Community leaders had influence beyond the DNA barcoding commons. David Schindel advocated to the *CBD* during negotiations for the *Nagoya Protocol*, arguing for simplified measures for non-commercial research (Schindel 2010). He also promoted standard ABS agreements for non-commercial research to engage LMMC countries in barcoding efforts (Vernooy et al. 2010; Schindel et al. 2015). Further, prominent biologists, such as Dan Janzen and Winnie Hallwachs, adopted DNA barcoding starting in 2003, thereby accelerating its acceptance within the scientific community (Janzen 2004; Burns et al. 2008).

### **3.3.2.2. Contributors of DNA barcode records**

Most contributors are researchers (including taxonomists, ecologists, evolutionary biologists, systematists, and bio-informaticians), working at universities or other research-intensive institutions, including museums and herbariums. Their contributions include: adding data or specimens to the commons; developing quality-control measures; refining methods for producing or utilizing barcodes (Meusnier et al. 2008); and/or studying barcode utility. In addition to researchers, lay contributors may suggest changes to taxonomic identifiers and highlight errors in the dataset. Further, the LifeScanner program allows individuals to collect specimens for DNA barcoding (including whole specimens or tissue samples) and receive information about the specimen (Biodiversity Institute of Ontario 2015). The resulting barcode records are then deposited into an open access database.

### **3.3.2.3. DNA barcode record users**

In addition to researchers, other users include individuals who work for agencies reliant on specimen identification, such as food and drug regulatory agencies (Yancy et al. 2008) or border control agencies (Johnson et al. 2014). High school students have used the barcode database for science experiments (Wong and Hanner 2008), and LifeScanner enables non-experts with no access to specialized equipment to use the barcode database to identify unknown animal specimens (Biodiversity Institute of Ontario 2015). Interviewees emphasized that DNA barcode records should be openly available online to enhance public biodiversity knowledge.

### **3.3.2.4. Databases**

As the requirement to publish sequence data with scientific articles predates DNA barcoding efforts, DNA barcoders initially deposited their sequence data into genomics repositories like GenBank. As part of INSDC, Genbank's policy is to provide open access to all records (Nakamura et al. 2013).

The minimum standards for submitting barcode records to BOLD and receiving a barcode identifier on GenBank are: reference specimen information (including unique identifiers and the institution storing the specimen); the taxonomic phylum; and the country in which the specimen was collected. However, barcode sequences may be submitted to GenBank without the required metadata, and BOLD mines

GenBank for barcode sequences to broaden its database of sequences for phylogenetic analyses. The BOLD data policies initially stipulated that a complete barcode record should include GPS coordinates (Ratnasingham and Hebert 2007). However, interviewees felt sharing specific GPS coordinates enabled unauthorized specimen collection, especially for endangered species. The data standards for barcode records suggest sharing GPS coordinates, but do not require it (Hanner 2009).

### **3.3.2.5. Repositories**

Specimen collections are stored in a range of facilities, including collections in individual research laboratories, research institution repositories, and national or regional collections housed in herbariums and museums. Each repository sets policies for specimen access, which may be modified to meet the specific requirements of depositors or to conform with national laws. Deposit and use are mediated by material transfer agreements (MTA), for example, repository staff and users may not use the specimens for unauthorized work and may not share the specimens with third parties without permission of the depositor (Bubela et al. 2015b).

### **3.3.2.6. Funding agencies**

Funding agencies distribute the financial resources needed to develop, maintain, and enable use of the DNA barcode commons. They are influential in promulgating rules for commons governance, such as data and materials sharing policies. In general, barcoding funds are distributed to two project types: large-scale resource-building initiatives (national or international) and smaller country-level projects that generate barcode data based on institutional or individual research grants.

Many agencies, internationally, have funded large-scale initiatives, beginning with the Alfred P. Sloan Foundation that funded CBOL for over \$6 million between 2003 and 2010 (Consortium for the Barcode of Life 2015). In addition, Canadian funding agencies have provided substantial funding for barcoding initiatives. The Canada Foundation for Innovation, the Ontario Research Foundation, and Genome Canada provided almost \$30M to develop infrastructure at BIO, including the Centre for Biodiversity Genomics, the Canadian Centre for DNA Barcoding, and BOLD (iBOL 2015d), and initiated the iBOL Project.

Many other funders support barcoding projects that generate barcode records, which expand the taxonomic coverage of the reference database. In 2015, iBOL listed 35 funders (from 15 countries) that each provided more than \$100,000 to support iBOL research (iBOL 2015d). The Canadian International Research and Development Centre (IDRC) provided \$2.2 million to support the barcoding efforts of developing countries.

## **3.4. Governance**

### **3.4.1. National implementation of *CBD* and *Nagoya Protocol***

National laws and regulations that implement the *CBD* and *Nagoya Protocol*, if they exist, govern the access and utilization of genetic resources. They impose



bureaucratic requirements for export permitting, place limits on utilization, and generally impose a system of ABS. Researchers who import genetic resources from countries with national ABS laws should conform with their substantive and procedural requirements.

National implementation of the *CBD* and *Nagoya Protocol* includes the designation of a competent national authority to provide access to genetic resources and administer policies to govern their use. Countries may also implement policies to encourage research that contributes to bioconservation, including simplified measures for accessing genetic resources. For example, Australia has implemented a process to allow for simplified measures, and other countries (Mexico, Indonesia, and Brazil) distinguish between commercial and non-commercial research (UNEP/CBD/SBSTTA/16/INF/37 2012).

Most of the iBOL partner nations are party to the *CBD* (with the only exception of the US), and over half (17/28) of partner nations are party to the *Nagoya Protocol*. Despite the relevance of these legal instruments to our study participants, non-LMMC interviewees could not describe the implementation of the *CBD* or *Nagoya Protocol* in their own countries. In contrast, LMMC interviewees were more aware of these legal instruments and how national implementation impacted their own work. Many LMMC and non-LMMC interviewees spoke with frustration about government policies that restrict access to genetic resources without a realistic understanding of their utilization and value. Several interviewees mentioned that their government viewed genetic resources as analogous to mineral resources that could be mined:

*They seem to believe that the genetic resource is like gold, and that you will sell [it], and that everyone everywhere is going to exploit our biodiversity. It's really so hard to have a clear dialogue with them because I have the feeling that they don't really understand what genetic resources are – Researcher, South America*

#### **3.4.2. Indirect governance by funding agencies**

Funding agencies promulgate policies on data and materials sharing as conditions of award, with varied capacity to enforce these policies (Mishra et al. 2016). For example, Genome Canada promulgates rules about data release to which all its funded projects must adhere, including iBOL (Genome Canada 2008). Genome Canada's policy is based on the principle of rapid data release with the intention of accelerating translational research benefits. Despite the wide range of funders of iBOL, individuals who participated in iBOL through the Canadian Centre for DNA Barcoding services were bound by Genome Canada policies, and iBOL administrators reported progress via a corporate board of three senior Genome Canada staff (iBOL 2015c).

In the early phases of iBOL, Genome Canada provided the majority of funding for DNA barcode sequencing of specimens sent to BIO with the condition that the data generated be openly released. Interviewees held a wide range of opinions

on the appropriate delay prior to data release, although most supported a limited delay to respect publication priority. Interviewees did not, however, suggest mechanisms for enforcement, although one policy maker from Canada emphasized the importance of rules to govern behaviour within large-scale projects:

*And if a scientist doesn't like the rules he can go play in his own pen, right? I mean we have to grow up a little bit. We're not working in that solitary confinement that we used to work and it didn't matter. We're dealing with large collaborative cooperative projects that you have to play by the rules. And the whole thing won't work if you don't have rules.* – Policy Maker, Canada

A Research Oversight Committee appointed by the iBOL board also provided guidance (iBOL 2015c). Perspectives from outside of this structure were represented by the International Scientific Steering Committee (ISSC), which advised the Scientific Director (Paul Hebert) on research plans and deliverables. Genome Canada set the rules for membership on the ISSC, which included active barcoding projects, a commitment to the iBOL data release policy, and barcode research funding over \$250,000. However, there was no structure in place for the funding agencies themselves to coordinate policies. One policy-maker interviewee cited this lack of coordination as a significant challenge in crafting effective policies. Overall, the policy-making structure contributed to decision-making inequities and a lack of representation from lower income countries. One researcher explained the impact of the centralized global organizations:

*That's why some people think that there should be another organization. Because you see, [iBOL and CBOL] are national organizations. And therefore probably we need a neutral one, which would then listen to other countries. But [Canada and the US] are now more or less being selfish, "Well, this is what we are doing as individual countries". If we have a neutral body, then probably they will listen more to others. I think that they should listen more to voices from Africa, in particular. Because you see, there are no funding agencies [in my country]* – Researcher, Africa

### **3.4.3. Formal agreements to govern actions**

iBOL and CBOL both influenced the legal instruments that govern genetic resource exchange within the barcode commons. iBOL developed a standard MTA for materials (specimens, tissue samples, PCR products) sent to the Canadian Centre for DNA Barcoding. The MTA was between the Canadian Centre and the institution of the individual providing the specimen. It included terms that the material was on permanent loan and that the provider deposit the data into open access databases.

At the international level, CBOL members were involved in developing ABS agreements for non-commercial research. Such agreements establish how benefits and risks are shared between partners, and provide reassurance to provider countries that there will be no un-approved commercial use of their genetic resources.

Benefits-sharing may include requirements for collaborations and access to training and new technologies (Schindel et al. 2015).

While one non-LMMC interviewee stated a preference to “not worry about the legal things because as soon as you get the lawyers involved then there are all kinds of issues that they want to deal with” (Researcher, US), most interviewees from LMMC and non-LMMC used MTAs to set the terms of access to and utilization of genetic resources. LMMC interviewees emphasized that MTAs were essential for ensuring specimens were not used for commercial research or research beyond the original MTA scope without re-negotiation. Several interviewees favoured standard MTAs and ABS Agreements for convenience and to minimize “paperwork”. An LMMC policy maker confirmed standard agreements for ABS reduced the burden on under-resourced countries:

*One of the more difficult things you can do as a regulator, if you're an under-resourced country, is having to negotiate case by case ABS agreements again and again and again. Because the people you negotiate with have got the money and the ability to draw in good lawyers, whereas here we don't have the budget. So, it would really suit us to have a sort of standardized benefit sharing arrangement that wasn't to be left negotiated every time – Policy Maker, Africa*

Interviewees did, however, point out that once genetic resources had been shipped, there was no guarantee for how they would be used, even with an executed MTA, due to lack of monitoring of the terms of the MTA and enforcement.

### 3.5. Action arenas

#### 3.5.1. Generating and sharing DNA barcode records

Individuals and institutions in countries with advanced scientific infrastructure and access to funding sources often favour rapid and open sharing of genetic resources (Field et al. 2009). iBOL policies reflected this preference, supported by many interviewees. Interviewees explained that the benefits to science from sharing outweighed the risks to individual researchers; data release requirements were the best way to increase the coverage of the DNA barcode record database; and data release requirements were justified when CCBD provide free sequencing.

While LMMC interviewees appreciated the history of genomics data release policies, they felt the unique circumstances of biodiversity research warranted a different approach. Therefore, most did not approve of rapid, pre-publication data release. Researcher interviewees indicated that generating barcodes is labour intensive, and too much value was placed on where the DNA was sequenced. One researcher from South America said “the real hard work nowadays is not sequencing; it's going to the field, collecting samples, preserving, shipping. All of that should not be underestimated”. LMMC researchers also felt they were disadvantaged by requirements to release data before publication.

Interviewees who produced barcode records preferred the enhanced, barcode-specific capabilities of BOLD over GenBank. They highlighted ease of use and

the ability to view metadata and raw sequence files. Some interviewees, however, were not able to share all the metadata required by BOLD for a DNA barcode record, and so appreciated the option to submit sequence data to GenBank,

*In some cases you have to [submit to GenBank] because sometimes you get material, you are working on a phylogenetic group, you have systematic research but you didn't get reference specimen or pictures so you can't really submit it to BOLD, so then you have to go through the GenBank, which is painful to submit, where BOLD is a delight – Researcher, South Africa*

While researchers acknowledged the value of the central databases such as BOLD and GenBank over local databases, LMMC researchers felt the central databases should have more involvement in policy-setting and management from international stakeholders so that the interests of contributors and users from lower-resourced settings could be appropriately considered. Some researcher interviewees cited lack of trust in North American and European research institutions as a reason to duplicate national-level data from BOLD on local servers.

### **3.5.2. Sharing biological materials to produce DNA barcodes**

DNA barcoding proponents were concerned about the ramifications of the *Nagoya Protocol* on biodiversity research, fearing that restrictive agreements for accessing genetic resources would have the unintended consequence of slowing biodiversity science. Many interviewees spoke of the need for researchers to access genetic resources. They argued that the misappropriation threat was overstated, because only the ubiquitous barcode region with no commercial value would be sequenced:

*I think some of the representatives of developing countries don't understand [that] the barcoding gene that we use is not really of any commercial value, because it goes everywhere and it doesn't actually code for any particular product that you might want to develop commercially. (UK iBOL Project Participant)*

This common perspective, however, failed to acknowledge mistrust stemming from a long history of misappropriation of genetic resources, and many LMMC interviewees described nuanced challenges for governing how genetic resources for barcoding are accessed and shared. The LMMC researchers and policy makers understood that genetic resources shared for DNA barcoding projects were intended for biodiversity science. However, when a specimen or tissue sample is shipped internationally, it includes the whole genome. LMMC researchers and policy makers often do not trust recipients to use the materials for DNA barcoding only, as one African researcher explained: “The thing is that we don't trust them. I mean, three years from now [BOL project leaders] will say, ‘Oh, now this is what we want to do.’ Meanwhile, you have given them the specimen already and you can't prevent them from using [all of its genetic information]”.

One interviewee mentioned there was more protection for genetic resources when there was a potential for commercialization, because laws in some countries are clear on proving the source of the materials, whereas most scientific publications do not have the same requirement.

Some interviewees preferred to share genetic resources for barcoding only on the condition the specimens and extracted DNA would be destroyed after generating the barcode sequence. Other participants stated that the storage of specimens and DNA extracts was necessary to allow for quality control and future research. This view gave greater consideration to the value of the resource for research than to the potential for misappropriation:

*Creating the repository is a huge resource to the future. There may be potential research avenues we haven't even thought of yet. So I think [storing genetic resources] is a really good idea. I would be very sad if, for example, due to concerns over property or potential commercialization problems that we were required to destroy the genetic [resources] – Researcher, Canada*

Despite willingness to export genetic resources under the right conditions, many interviewees felt specimens should be stored in the country of origin. For many interviewees from LMMCs, this would require developing expensive storage infrastructure. To mitigate associated costs, the iBOL model enabled countries without the necessary infrastructure to export genetic resources for barcoding to countries with existing infrastructure. However, LMMC interviewees were frustrated when collections from their countries were housed in foreign repositories, as explained by one African researcher:

*[My country] was a colony of Great Britain for some time. As a result of that, most of our systematic work being done on collections made from our country was then taken overseas, and that's where the typed specimens are. As a result of that, I've got to now spend a lot of my time and money extracting from those institutions scattered around the world at enormous difficulty – Researcher, Africa*

A few LMMC researchers expressed the opinion that the only way to develop equitable partnerships is to build infrastructure to conduct research and store genetic resources locally. In addition to enabling access to and control over specimens in LMMCs, interviewees pointed out that local infrastructure would help build research capacity in their countries. Capacity building is one form of benefit that may be returned to countries of origin in exchange for access to and utilization of genetic resources.

### **3.5.3. Access to and use of the DNA barcode data**

Databases used to store DNA barcode records were designed to allow open access to data and unrestricted use, under the assumption that this benefits the greatest number of potential users (Ratnasingham and Hebert 2007; Nakamura

et al. 2013). The open access requirement was largely informed by the standards created after the HGP and enforced by Genome Canada through oversight of the iBOL project. Many interviewees from both LMMCs and non-LMMCs expressed their support of open access principles for genomics research. One Mexican researcher stated, “[The barcode record database] should stay open access. Because barcodes cannot be used to do any harm, I think. It’s just too little DNA”. Another Canadian explained: “I like the idea that somebody in India ... has access to my data, and they can do things that I would never have imagined doing with it”.

While BOLD and Genbank are designed to encourage access and place no restrictions on data use or distribution (Ratnasingham and Hebert 2007; NCBI 2016), many interviewees supported controlling access to sensitive data, such as geographic coordinates of protected species or information about newly invasive species. An Australian researcher explained: “There are data sensitivity issues. We have rare and endangered species; you wouldn’t want to tell people where their precise location is [to avoid disturbance or illegal/unethical collection]”. LMMC interviewees emphasized different levels of access could be granted to certain types of users.

Interviewees were divided on potential data use restrictions, particularly whether data users should be required to acknowledge or cite data contributors. Some felt that collecting specimens and uploading data were not activities that warranted acknowledgement or benefits sharing. However, other interviewees felt data users should, at a minimum, acknowledge data contributors:

*Bioinformaticians, maybe they don’t understand the value of the fieldwork and making the data available. If they just instantly get the data and they got a publication, it’s good but they should also respect those who contributed to the data – Researcher, China*

Other interviewees stated their belief that there should be no restrictions on data use; one interviewee explained commercial applications were the main benefit of the open database:

*Once you get that [barcode record] database then there are commercial applications that will be developed and there are academic applications that will be developed. Three-fourths of the motivation of doing a barcode database are commercial application so if you somehow think that that’s a bad thing then you ought not to participate – Researcher, US*

### **3.6. Patterns of interactions and outcomes**

#### **3.6.1. Collaborations**

Collaborations within the DNA barcode commons define who participates, and how the commons is built, maintained, and used. Researcher interviewees identified reciprocity as a key factor when entering a collaboration, but the definition

of reciprocity varied. While many interviewees mentioned mutual scientific goals and complementary research programs in response to questions about how they form collaborations, LMMC interviewees placed value on relationships in which partners had equal opportunities to make meaningful contributions beyond specimen collection, “You have to treat each other as equals. You don’t want to be seen in the bottom of the list in small-print acknowledgement.” (Researcher, Africa).

Both non-LMMC and LMMC researchers emphasized the importance of professional reputation in selecting collaborators and in deciding on the nature of collaborative activities. One South American researcher succinctly stated his “no assholes” rule. However, the reliance on personal relationships can exacerbate inequities, because personal connections are often developed at scientific meetings unaffordable by researchers from lower income countries.

LMMC researchers expressed apprehension about sharing genetic resources with international collaborators based on the risk of misappropriation of genetic resources. Researcher interviewees explained, however, that personal relationships mitigated this fear:

*The people don’t want [genetic resources] to be stolen by [the US and Canada] again. But every history is different [for] each person. In my case, I have no problem because I know [non-LMMC Researcher], so I can work with him and no problem. But most of the people that are working with us [in our institution] – they don’t want to [share genetic resources with researchers from other countries] – Researcher, Mexico*

### 3.6.2. Publications

Peer-reviewed publications are the primary outcome of basic research, are used as a metric to evaluate researchers (Nelkin 1998), and are a key benefit for academic users of the DNA barcode commons. Many arguments for open access database management structures include the claim that researchers in lower income countries would benefit from access to data, which supports their own research publications and enhances their professional profile:

*If you’re in a poor developing country, [if] a lot of the sequences of organisms in [your] area have all been put into the common database, you can get all that stuff for nothing, because someone else has paid for it – iBOL project participant, United Kingdom*

Our search for publications that referenced seminal DNA barcode papers yielded 3557 scientific journal publications from 2003 to 2014. This large sample size enabled us to delineate more finely the characteristics of authors. While we categorised our interviewees as working in LMMCs or non-LMMCs, for our quantitative bibliometric analyses, we used four World Bank categories for national income (see methods section). In general, the first three income categories coincides with our LMMC category, and the category of “High Income Countries (HIC)” coincides with non-LMMC.

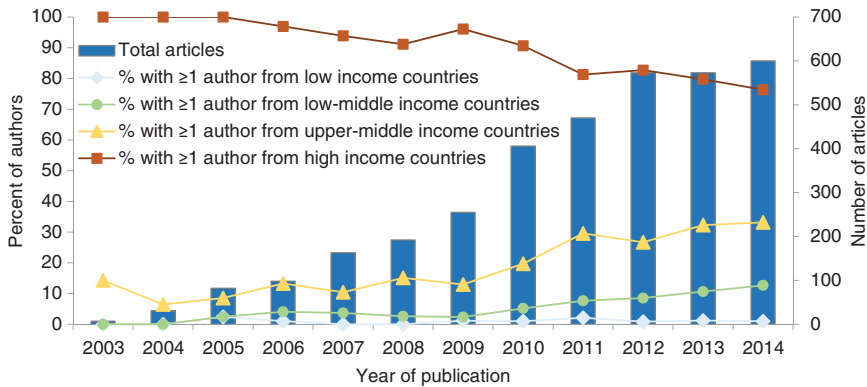


Figure 4: The number of articles citing four seminal barcode papers, published each year during 2003–2014, and percent of articles with at least one author from the specified income group of countries. Income levels are as defined by the World Bank Country and Lending Groups (World Bank 2016).

The number of publications in our dataset increased in each year from 2003 to 2011, and plateaued to approximately 600 publications per year from 2012 to 2014 (Figure 4). This asymptote is expected as a field matures and citation to original publications declines (Barnett 1992; Bouabid 2011). From 2003 to 2005, every article in our dataset had at least one author from a HIC (Figure 4), which suggests early barcoding activity was driven by HIC researchers. While the proportion of articles with authors from low-middle and upper-middle income countries has risen, the majority of DNA barcode publications have been produced solely by authors in HICs. Only 1% of publications included authors from low income countries. These data suggest the growing DNA barcoding community has not expanded to include low income country researchers at the same pace as low-middle and upper-middle income country researchers.

In addition, co-authorship was most frequent between authors from Western HICs (defined by the United Nations as Canada, US, Western Europe, Australia and New Zealand (United Nations DGACM 2016)). We identified comparatively few co-authorship links between authors from Western and non-Western countries (Figure 5: grey versus coloured lines).

We counted the number of publications with authors from the following regions: Western HICs, Eastern Europe, Latin America and the Caribbean, Africa, and Asia/Pacific, as well as the number with author sets that spanned more than one region. Over half (54%) of the 3557 articles in our dataset had authors only from Western HICs (Table 1). Because only 2% (80/3557) of the publications had author sets that spanned more than two regions, we excluded these from Table 1. Regions rich in biodiversity, such as Africa and Latin America, had few author sets within or across these regions. For example, compared to the 54% of articles with author sets restricted to the Western HICs, only 2.5% had author sets



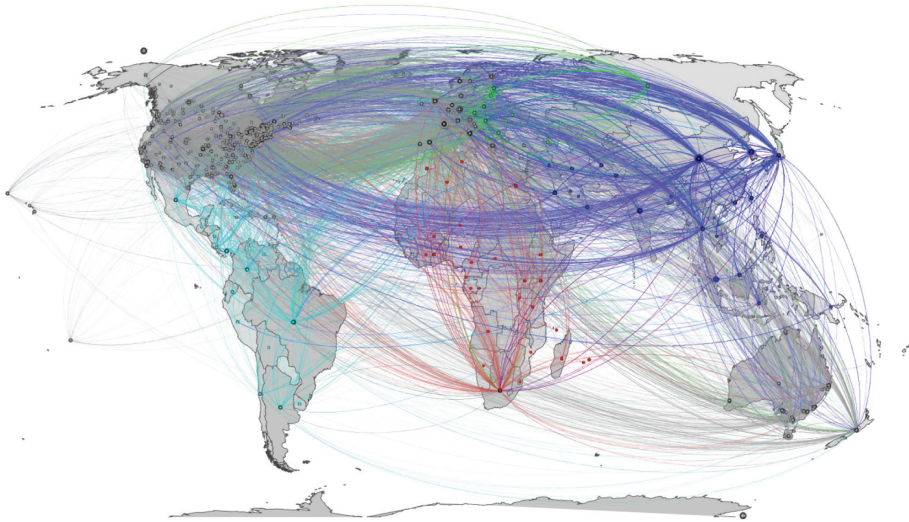


Figure 5: Co-authorship in the DNA barcoding publication database. Each node represents an author, and size of node indicates relative number of times the author has been mentioned in the database. Each line between nodes indicates that the authors co-authored a publication. Lines in grey indicate collaborations restricted to Western HICs. The coloured lines represent collaborations with other regions (United Nations DGACM 2016).

Table 1: Percent of 3557 publications by geographic regions of residence of author sets (excluding 80 publications with author sets spanning more than 2 regions).

	West <sup>^</sup>	East Europe	Latin <sup>*</sup>	Africa	Asia <sup>+</sup>
West <sup>^</sup>	<b>53.9</b>	3.5	5.3	2.6	7.5
East Europe	3.5	<b>2.3</b>	0.4	0.0	0.2
Latin <sup>*</sup>	5.3	0.4	<b>4.0</b>	0.0	0.1
Africa	2.6	0.0	0.0	<b>0.9</b>	0.1
Asia <sup>+</sup>	7.5	0.2	0.1	0.1	<b>17.3</b>

Darker shading indicates a higher proportion of publications with some or all authors from the region.

Bolded numbers indicate author sets restricted to a particular region.

<sup>^</sup>Canada, US, Western Europe, Australia, New Zealand.

<sup>\*</sup>Latin American and the Caribbean.

<sup>+</sup>Asia and the Pacific.

confined to Eastern Europe, 3.9% to Latin America and the Caribbean, 0.8% to Africa, and 17% to Asia and the Pacific (Table 1). Articles with authors from more than one region that did not include western countries only made up 3.2% of the articles in our database.

Researchers are evaluated by their institutions by both the quantity and the quality of their publications. One measure of the latter is the impact factor of

*Table 2: Odds Ratios (ORs) for the association between income level of an authors' country of residence and publication in a high impact journal for 3557 publications that cited four seminal DNA barcoding papers.*

	N	OR	95% CI
Income level			
Only high income country authors	2386	1.0	
Mix of high and middle or low income country authors	615	0.76	0.54–1.1
Only middle or low	556	0.09	0.04–0.23

OR, odds ratio.

journals in which they publish (Callaham et al. 2002). Authors from low, low-middle and upper-middle income countries had lower odds of publication in high impact journals compared to articles with only HIC authors (Table 2). Publications with author sets from a mix of country income levels had 76% of the odds of publication in highly ranked journals compared to articles with author sets restricted to HICs (Table 2). Publications with author sets restricted to low, low-middle and upper-middle income countries had 9% of the odds of publication in highly ranked journals compared to articles with author sets restricted to HICs (Table 2).

In summary, the pattern of interactions evidenced in publications suggests that the DNA barcode community is dominated by actors whose institutional affiliations are located in high-income, Western countries. Publications are a major outcome of research commons. The KC-IAD framework suggests that the outcomes of interactions inform how actors will behave in action arenas within the commons. Next, we further analyse outcomes of specimen collection, barcode record generation and data sharing, using two exemplar species groups.

### **3.6.3. BOLD records for exemplar species groups: medicinal plants and mosquito disease vectors**

We chose two exemplars to examine the outcomes of specimen collection, barcode record generation and data sharing in BOLD: medicinal plants and mosquito disease vectors. Medicinal plants and their derivative natural health products exemplify a potentially commercializable genetic resource. Mosquitos represent globally distributed genetic resources that are relevant to public health, although the disease burdens of mosquito borne diseases, such as malaria, are greatest in lower income countries.

We identified 17,895 published medicinal plant records in BOLD as of February 2013, of which 11,685 specified specimen origin (Table 3). Fifty-four percent (6297/11,685) of published medicinal plant records with origin data on BOLD were collected in HICs, while only 0.4% (50/11,685) were collected in low income countries. For the 9477 records with metadata on where the reference specimen was stored, only 3% (280/9477) were stored outside of the country of origin. This pattern may indicate legal constraints and/or the unwillingness of some researchers to share genetic resources with foreign collaborators, as described by LMMC interviewees.

Table 3: Number of published medicinal plant records in BOLD by income level of country where the specimen was collected.

Income level of country where specimen was collected (World Bank 2016)	Total published records on BOLD	Total records indicating voucher storage site	Data mined from GenBank (no voucher storage information)	Voucher is stored outside of origin country
		n (% of total published records on BOLD)	n (% of total published records on BOLD)	n (% of records with storage site)
Low income	50	3 (6%)	47 (94%)	3 (100%)
Low-middle income	640	395 (62%)	245 (38%)	33 (8%)
Upper-middle income	4698	4226 (90%)	472 (10%)	142 (3%)
High income	6297	4853 (77%)	1444 (23%)	102 (2%)
Total	11,685	9477 (81%)	2208 (19%)	280 (3%)

Table 4: Number of published *Aedes* sp., *Anopheles* sp. or *Culex* sp. records in BOLD by income level of country where the specimen was collected.

Income level of country where specimen was collected (World Bank 2016)	Total published records on BOLD	Total records indicating voucher storage site	Data mined from GenBank (no voucher storage information)	Voucher is stored outside of origin country
		n (% of total published records on BOLD)	n (% of total published records on BOLD)	n (% of records with storage site)
Low income	313	1 (0%)	312 (99%)	1 (100%)
Low-middle income	2817	1577 (56%)	1240 (44%)	88 (6%)
Upper-middle income	3312	1251 (38%)	2061 (62%)	944 (76%)
High income	10,855	9414 (87%)	1441 (13%)	1488 (16%)
Total	17,297	12,243 (71%)	5054 (29%)	2521 (21%)

We identified 17,297 published barcode records for mosquito disease vectors from genera *Anopheles*, *Aedes*, and *Culex* as of May 2016 (Table 4). Even fewer records were published in BOLD for the three mosquito species compared to medicinal plants, and only three were linked to specimens originating from low income countries. Twenty-one percent (2521/12,243) of mosquito reference specimens were stored in collections outside of the origin country, which is higher than the 3% of medicinal plant specimens stored outside the country of origin. This suggests fewer constraints and more willingness to share mosquito records and specimens, which have little commercial value, compared to medicinal plants.

In addition, we compared the number of published and unpublished mosquito records on BOLD (this comparison was not possible for medicinal plants – see methods). We identified 47,355 total records for the three genera. Of these, only 35% of *Anopheles* sp., 25% of *Culex* sp., and 62% of *Aedes* sp. records have been made available for anyone to view or download. This implies that many more individuals participate in DNA barcoding efforts and use DNA barcoding data infrastructure to manage their barcode data than contribute to the commons.

These two exemplars suggest individuals from HICs contribute more data and specimens to the DNA barcode commons, contrary to the goals of having a globally representative database. While many interviewees expressed that open access databases would provide the most benefits for potential users, and some emphasized the benefits for researchers in LMMCs, our analysis of the patterns of interactions and outcomes demonstrates global participation should not be assumed.

### 3.7. Evaluative criteria

Evaluative criteria are used by participants and observers of action arenas to assess the success of processes and their outcomes. Different stakeholders within the DNA barcode commons will have different criteria for evaluating whether or not the effort is successful. Interviewees, from both LMMC and non-LMMC, stated that the continual growth of the barcode record database was a marker of success. However, LMMC interviewees argued that the initiative could not achieve its global goals without increased participation of actors from LMMC. The global commons criterion is consistent with the explicit goals of the community as expressed in iBOL and CBOL documents, for example through the call to “make every species count”.

## 4. Discussion

The DNA barcode commons has achieved success in some areas, as evidenced by its large and growing number of publicly-accessible barcode records and scientific publications. However, our analysis demonstrates that the attributes of the DNA barcode commons have created challenges for global participation. The goal of achieving a *globally representative* barcode commons is hindered by its inequitable governance structures and inequitably distributed resources, infrastructure and rewards for participation, such as publications in high impact journals. With respect to governance, many DNA barcode commons actors from LMMC were concerned about the lack of knowledge and implementation of ABS laws in the countries that were leading barcoding efforts. Further, the infrastructure to produce and store barcode records is concentrated a few HIC, which discouraged LMMC participation due to the need to transfer genetic resources outside their country of origin. Our exemplars on medicinal plants and mosquito genera highlight that researchers rarely export voucher specimens. In the following section, we present recommendations to improve equity in governance of the DNA barcoding commons if the goal is to make it a global initiative for the benefit of the world’s biodiversity and science.

### 4.1. Recommendations for the establishment of a global DNA barcode commons

Our recommendations fall into three categories: inclusive governance structures, development of new norms, and greater emphasis on relevant legal instruments.

#### 4.1.1. Create inclusive and equitable governance structures

The difficulties that heterogeneous communities face in establishing trust (Ruttan 2006), can be overcome through representative governance structures (Poteete and Ostrom 2004, Ostrom 2003). Indeed it is well-established that those who are affected by rules and norms in a commons should have a role in developing the rules of the commons (Frischmann et al. 2014; Ostrom 2005). However, we found the governance in DNA barcoding has been dominated by the norms and standards of HIC, whose stakeholders provide financial resources and infrastructure. For example, the central barcoding project, iBOL, implemented policies on participation in governance as well as data and materials sharing established by Canadian funding agencies. As described above, participation in governance required substantial funding commitments, which by default excluded the perspectives of LMMC researchers and stakeholders. As a result, many concerns about participating in the DNA barcoding commons shared by interviewees from LMMCs in our study were not reflected in iBOL policies or governance structure.

The DNA barcode community has already begun to develop a new governance body: The International Society for the Barcode of Life (ISBOL) (Castle et al. 2015). We encourage ISBOL to be representative of the diverse and global barcoding community, particularly in its leadership.

#### 4.1.2. Develop new norms for genetic resource sharing specific to the DNA barcode commons

Governing bodies of the DNA barcode commons have promulgated community data-sharing norms to promote wide-spread use and re-contribution of value-added data (Schofield et al. 2009; Dedeurwaerdere 2010a; Bubela et al. 2012). These norms have a historical precedent in large-scale genomics projects (Field et al. 2009; Toronto International Data Release Workshop Authors 2009) and require rapid data sharing (Genome Canada 2008; iBOL 2015a). They are based on the assumption that norms of open access and unrestricted use are universally held across the globe and will best facilitate a network effect.

Our study suggests, however, that the application of these norms to a *global* knowledge commons may inhibit global participation, thereby limiting both participation and the network effect of use and re-contribution of value-added data. Previous research on knowledge commons supports this, and has emphasized the importance of fit between formal institutional arrangements and the norms of the specific community to which they apply (Dedeurwaerdere 2010b). Interviewees from LMMC were hesitant to share data and materials when they received limited benefits from participation (e.g. scientific credit, increased capacity). Setting restrictions on use, such as requiring citation, attribution, or an embargo period for first use of the data by the contributor, may enhance LMMC research participation, as will resources directed to building scientific capacity and barcoding efforts.

Restrictions on use of the DNA barcode commons is also necessary to comply with the *CBD* and the *Nagoya Protocol*. Barcoding proponents have argued for access to genetic resources under the “simplified measures” for non-commercial use set out in the *Nagoya Protocol* (Schindel et al. 2008). Yet the barcode database does not include restrictions on commercial use of the records, for example in developing species identification tools for food and drug regulators and for law enforcement agencies (Wong and Hanner 2008; Rehman et al. 2015). Barcode records stored in BOLD should be accompanied by terms that outline restrictions on use of the data for commercial purposes, or at a minimum specify the benefits that should accrue to contributors or their countries. Further, the protection of sensitive information, such as geo-location data for endangered species, is in keeping the goals of these international instruments to protect global biodiversity.

Finally, strategies to manage specimens for barcoding should be modified to encourage participation in the barcode commons. As detailed in our section on enclosure arenas related to sharing biological materials, the expectation that specimens can be internationally distributed and freely accessed discourages many LMMC participants. Institutions that house barcoding infrastructure and contribute to its pipeline should enable the destruction of specimens and DNA extracts following the generation of the DNA barcode if requested by participants. This step will provide some confidence that LMMC genetic resources will not be misappropriated and build trust. However, destruction needs to be balanced against the need for data linked to reference specimens. A solution is to support the development of LMMC infrastructure to store and manage reference specimens. This would reduce the need to store the specimens in HICs and would provide a benefit in exchange for access to genetic resources.

#### **4.1.3. Emphasize the importance of existing legal frameworks**

Current barcoding governance documents have not adequately referenced international legal instruments or national laws that govern genetic resources, nor do these legal instruments adequately inform the actions of many DNA barcode community members. This diversity in the understanding and application of laws may lead to conflict between participants, due to a lack of shared expectations about access, utilization and equitable distribution of benefits. Improved compliance with the legal framework for genetic resources would allay many concerns of LMMC participants and facilitate their participation in barcoding efforts. Enhanced education about the legal framework would help the global community of barcoders understand LMMC concerns.

BOL organizations should a) develop governance documents that explicitly consider and comply with the legal and policy frameworks of global sharing and utilization of genetic resources, including the *CBD* and the *Nagoya Protocol*; and b) co-develop with LMMC partners educational materials for the community on the legal and policy context of DNA barcoding activities to enhance understanding and compliance. Completion of a short online training module could be a prerequisite to use of databases, such as BOLD.

## 4.2. Implications for global research commons

Our study used the KC-IAD framework to develop a comprehensive description of the DNA barcode commons and opportunities to strengthen its collective action with the goal of *global* participation in the development of DNA barcode resources. Our research confirms the utility of the KC-IAD framework for understanding knowledge commons and making recommendations for their governance (Cole 2014; Frischmann et al. 2014).

Our study demonstrates the importance of defining the attributes of the community by distinguishing between international knowledge commons, where like-minded participants span international borders, and global knowledge commons, where outcomes depend on global participation. Our observation was that the governance of the DNA barcoding commons was based on institutional arrangements developed for an international knowledge commons – the HGP. Previous work on knowledge commons has also not explicitly accounted for this distinction between international and global (Dedeurwaerdere 2010b; Contreras 2014; Bubela et al. 2017). However, others have emphasized the importance of being attentive to the social organization of participants (Madison 2014).

Nevertheless, similarities exist between international and global knowledge commons. For example, researcher motivations to participate in commons are dominated by reputation and social identity influences in the scientific community (Dedeurwaerdere et al. 2016). While we identified similar motivations, these were tempered by the negative history of the commons, grounded in colonialism and resource misappropriation, which result in wealth and power inequities. These factors are important in the study of all knowledge commons (Frischmann et al. 2014).

Researchers and other stakeholders interested in building knowledge commons can draw three broad lessons from our case-study. First, an open-access/unrestricted use model may not enhance the goals of all knowledge commons. Other institutional arrangements that involve some restrictions on access and use could enhance trust, collection action and participation in the commons. Second, contributions to the building of the commons should be equally valued, regardless of their technical nature. In our case, the collection of specimens should be as valued as the sequencing of their DNA and the bioinformatics involved in their transformation to a DNA barcode record. Third, in-kind contributions, such as sample collection and biodiversity, should be as valued as funding and infrastructure contributions; the lack of the latter should not serve as a barrier to participation in governance of the knowledge commons.

Finally, BOL organisations, with equitable and representative participation in governance, should develop and implement evaluative criteria that are reflective of the collective goals of the actors. Evaluative criteria are an under-developed component of the IAD framework (Cole 2014) and could be the focus of future studies.

### 4.3. Limitations

Our study has several limitations. We mainly interviewed individuals with direct involvement in BOL organizations and efforts, meaning our analysis did not represent perspectives of those who independently participate in DNA barcoding activities. Similarly, we did not analyse all barcoding scientific publications, only those that referenced seminal papers. We only examined a small subset of BOLD records relevant to our two exemplars; other exemplars may have revealed different patterns of use. Our data interpretation was limited to our own perspectives, and it is possible another individual might have different views. However, our use of an established theoretical framework reduced the reliance on our individual interpretation of data. Finally, while our findings are not necessarily generalizable to other global knowledge commons, the rich description of our case study provides the necessary contextual details to enable transferability to future studies of *global* knowledge commons governance.

## 5. Conclusions

We have demonstrated that the goal of creating a globally inclusive DNA barcode commons has not yet been fully realised, using the KC-IAD framework. Our research provides evidence of the risks and benefits of commons participation that are not equitably shared across a set of heterogeneous global participants. It offers suggestions of how to improve equity and increase collective action. The newly created ISBOL could mitigate some of the challenges of global participation through representative governance and consideration of access and benefit sharing and legal instruments that may enhance participation in the DNA barcoding commons.

### Literature cited

- Adamowicz, S. J. 2015. International Barcode of Life: Evolution of a Global Research Community. *Genome* 58(5):151–162. <http://dx.doi.org/10.1139/gen-2015-0094>.
- Barnett, R. E. 1992. Conflicting Visions: A Critique of Ian Macneil's Relational Theory of Contract. *Virginia Law Review* 78(5):1175–1206. <http://dx.doi.org/10.2307/1073395>.
- Bastian, M., S. Heymann, and M. Jacomy. 2009. Gephi: An Open Source Software for Exploring and Manipulating Networks. <https://gephi.org/users/publications/>.
- Benson, D. A., M. Cavanaugh, K. Clark, I. Karsch-Mizrachi, D. J. Lipman, J. Ostell, and E. W. Sayers. 2013. GenBank. *Nucleic Acids Research* 41(D1):D36–D42. <http://dx.doi.org/10.1093/nar/gks1195>.
- Berge, E. and F. van Laerhoven. 2011. Governing the Commons for Two Decades: A Complex Story. *International Journal of the Commons* 5(2):160–187. <http://dx.doi.org/10.18352/ijc.325>.



- Bermuda Sequence Policies Archive. 2016. Accessed September 8. [http://web.ornl.gov/sci/techresources/Human\\_Genome/research/bermuda.shtml#2](http://web.ornl.gov/sci/techresources/Human_Genome/research/bermuda.shtml#2).
- Biodiversity Institute of Ontario. 2015. Lifescanner. *Lifescanner*. <http://lifescanner.net/>.
- BOLD Systems. 2015. Public Data Portal – Record List. *BOLD Systems*. [http://www.boldsystems.org/index.php/Public\\_BINSearch?searchtype=records](http://www.boldsystems.org/index.php/Public_BINSearch?searchtype=records).
- Bouabid, H. 2011. Revisiting Citation Aging: A Model for Citation Distribution and Life-Cycle Prediction. *Scientometrics* 88(1):199. <http://dx.doi.org/10.1007/s11192-011-0370-5>.
- Bubela, T. 2013. International Collaborative Barcoding Research: Moving the Debate Forward on Access and Benefit Sharing of Genetic Resources to Promote Biodiversity. Plenary presented at the 5th International Barcode of Life Conference, Kunming, China, October 27.
- Bubela, T., R. Adams, S. Chandrasekharan, A. Mishra, and S. Liu. 2017. Governance of Biomedical Research Commons to Advance Clinical Translation: Lessons from the Mouse Model Community. In: *Governing Medical Knowledge Commons*, eds. K. J. Strandburg, B. M. Frischmann, and M. J. Madison, 222–258. Cambridge, UK: Cambridge University Press. <https://doi.org/10.1017/9781316544587.011>.
- Bubela, T., J. Geary, M. Bieber, and D. S. Hik. (2015a). Diffusing Barcoding: The Global Spread of a Good Idea. *Scientific Abstracts from the 6th International Barcode of Life Conference*. 58:201–202. Genome. <http://www.nrcresearch-press.com/doi/10.1139/gen-2015-0087>.
- Bubela, T., J. Guebert, and A. Mishra. 2015b. Use and Misuse of Material Transfer Agreements: Lessons in Proportionality from Research, Repositories, and Litigation. *PLOS Biology* 13(2):e1002060. <http://dx.doi.org/10.1371/journal.pbio.1002060>.
- Bubela, T., P. N. Schofield, C. D. Ryan, R. Adams, and D. Einhorn. 2012. Managing Intellectual Property to Promote Pre-Competitive Research: The Mouse as a Model for Constructing a Robust Research Commons. *Journal of Law, Information and Science* 22(1):98. <http://search.informit.com.au/documentSummary;dn=050814668635987;res=IELHSS>.
- Burns, J. M., D. H. Janzen, M. Hajibabaei, W. Hallwachs, and P. D. N. Hebert. 2008. DNA Barcodes and Cryptic Species of Skipper Butterflies in the Genus *Perichares* in Area de Conservación Guanacaste, Costa Rica. *Proceedings of the National Academy of Sciences*. 105(17):6350–6355. <http://dx.doi.org/10.1073/pnas.0712181105>.
- Callaham, M., R. L. Wears, and E. Weber. 2002. Journal Prestige, Publication Bias, and Other Characteristics Associated With Citation of Published Studies in Peer-Reviewed Journals. *The Journal of the American Medical Association* 287(21):2847–2850. <http://dx.doi.org/10.1001/jama.287.21.2847>.
- Castle, D., R. Hanner, P. M. Hollingsworth, and D. LI. 2015. Discussion Document: The International Society for the Barcode of Life (ISBOL). <http://ibol.org/wp-content/uploads/2015/08/ISBOL150819-Final.pdf>.

- Cole, D. H. 2014. Learning from Lin: Lessons from the Natural Commons for the Knowledge Commons. In: *Governing Knowledge Commons*, eds. B. M. Frischmann, M. J. Madison, and K. J. Strandburg, 45–68. New York: Oxford University Press.
- Collins, R. A. and R. H. Cruickshank. 2013. The Seven Deadly Sins of DNA Barcoding. *Molecular Ecology Resources* 13(6):969–975. <http://dx.doi.org/10.1111/1755-0998.12046>.
- Collins, F. S., M. Morgan, and A. Patrinos. 2003. The Human Genome Project: Lessons from Large-Scale Biology. *Science*. 300(5617):286–290. <http://dx.doi.org/10.1126/science.1084564>.
- Consortium for the Barcode of Life. 2015. What Is CBOL? *Barcode of Life*. <http://www.barcodeoflife.org/content/about/what-cbol>.
- Contreras, J. 2010. Data Sharing, Latency Variables, and Science Commons. *Berkeley Technology Law Journal* 25(4):1601–172. <http://dx.doi.org/10.15779/Z38GW9F>.
- Contreras, T. L. 2014. Constructing the Genome Commons. In: *Governing Knowledge Commons*, eds. B. M. Frischmann, M. J. Madison, and K. J. Strandburg, 99–136. New York: Oxford University Press.
- Dedeurwaerdere, T. (2010a). Self-Governance and International Regulation of the Global Microbial Commons: Introduction to the Special Issue on the Microbial Commons. *International Journal of the Commons* 4(1):390–403. <http://dx.doi.org/10.18352/ijc.215>.
- Dedeurwaerdere, T. (2010b). Global Microbial Commons: Institutional Challenges for the Global Exchange and Distribution of Microorganisms in the Life Sciences. *Research in Microbiology*. 161(6):414–421. <http://dx.doi.org/10.1016/j.resmic.2010.04.012>.
- Dedeurwaerdere, T., P. Melindi-Ghidi, and A. Broggiato. 2016. Global Scientific Research Commons under the Nagoya Protocol: Towards a Collaborative Economy Model for the Sharing of Basic Research Assets. *Environmental Science & Policy* 55(1):1–10. <http://dx.doi.org/10.1016/j.envsci.2015.08.006>.
- DeSalle, R., M. G. Egan, and M. Siddall. 2005. The Unholy Trinity: Taxonomy, Species Delimitation and DNA Barcoding. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 360(1462):1905–1916. <http://dx.doi.org/10.1098/rstb.2005.1722>.
- Dupuis, J. R., A. D. Roe, and F. A. H. Sperling. (2012). Multi-Locus Species Delimitation in Closely Related Animals and Fungi: One Marker Is Not Enough. *Molecular Ecology* 21(18):4422–4436. <http://dx.doi.org/10.1111/j.1365-294X.2012.05642.x>.
- Ebach, M. C. and C. Holdrege. 2005. DNA Barcoding Is No Substitute for Taxonomy. *Nature*. 434(7034):697–697. <http://dx.doi.org/10.1038/434697b>.
- Field, D., S.-A. Sansone, A. Collis, T. Booth, P. Dukes, S. K. Gregurick, K. Kennedy, P. Kolar, E. Kolker, M. Maxon, S. Millard, A. M. Mugabushaka, N. Perrin, J. E. Remacle, K. Remington, P. Rocca-Serra, C. F. Taylor, M. Thorley, B. Tiwari, and J. Wilbanks. 2009. Omics Data Sharing. *Science (New York, N.Y.)* 326(5950):234–236. <http://dx.doi.org/10.1126/science.1180598>.

- Frischmann, B. M., M. J. Madison, and K. J. Strandburg. 2014. Governing Knowledge Commons. In: *Governing Knowledge Commons*, eds. B. M. Frischmann, M. J. Madison, and K. J. Strandburg, 1–43. New York: Oxford University Press.
- Geary, J. and T. Bubela. 2015. Global Perspectives on Participating in the International Barcode of Life Project. *Genome Scientific abstracts from the 6th International Barcode of Life Conference* 58(May):219–220. <http://dx.doi.org/10.1139/gen-2015-0087>.
- Geary, J., E. Camicioli, and T. Bubela. 2016. DNA Barcoding in the Media: Does Coverage of Cool Science Reflect Its Social Context. *Genome* 59: 738–750. <http://dx.doi.org/10.1139/gen-2015-0210>.
- Genome Canada. 2008. Genome Canada Data Release and Resource Sharing. <https://www.genomecanada.ca/sites/default/files/pdf/en/DataReleaseandResourceSharingPolicy.pdf>.
- Given, L. M. and K. Saumure. 2013. Trustworthiness. In: *The SAGE Encyclopedia of Qualitative Research Methods The SAGE Encyclopedia of Qualitative Research Methods*. 2455 Teller Road, Thousand Oaks California 91320 United States: SAGE Publications, Inc. Accessed June 2. <http://knowledge.sagepub.com/view/research/n470.xml>.
- Gregory, T. R. 2005. DNA Barcoding Does Not Compete with Taxonomy. *Nature* 434(7037):1067–1067. <http://dx.doi.org/10.1038/4341067b>.
- Gross, M. 2012. Barcoding Biodiversity. *Current Biology* 22(3):R73–R76. <http://dx.doi.org/10.1016/j.cub.2012.01.036>.
- Hanner, R. 2009. Data Standards for BARCODE Records in INSDC (BRIs). Consortium for the Barcode of Life.
- Hebert, P. D. N., A. Cywinska, S. L. Ball, and J. R. deWaard. 2003. Biological Identifications through DNA Barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*. 270(1512):313–321. <http://dx.doi.org/10.1098/rspb.2002.2218>.
- Helmy, M., M. Awad, and K. A. Mosa. 2016. Limited Resources of Genome Sequencing in Developing Countries: Challenges and Solutions. *Applied & Translational Genomics* 9:15–19. Accessed June 2. <http://dx.doi.org/10.1016/j.atg.2016.03.003>.
- Hess, C. and E. Ostrom. 2011. Introduction: An Overview of the Knowledge Commons. In: *Understanding Knowledge as a Commons: From Theory to Practice*, eds. C. Hess and E. Ostrom, 3–26. Cambridge: MIT Press.
- Hollingsworth, P. M., L. L. Forrest, J. L. Spouge, M. Hajibabaei, S. Ratnasingham, M. van der Bank, M. W. Chase, R. S. Cowan, D. L. Erickson, A. J. Fazekas, S. W. Graham, K. E. James, K. J. Kim, W. J. Kress, H. Schneider, J. van AlphenStahl, S. C. Barrett, C. van den Berg, D. Bogarin, K. S. Burgess, K. M. Cameron, M. Carine, J. Chacón, A. Clark, J. J. Clarkson, F. Conrad, D. S. Devey, C. S. Ford, T. A. Hedderson, M. L. Hollingsworth, B. C. Husband, L. J. Kelly, P. R. Kesanakurti, J. S. Kim, Y. D. Kim, R. Lahaye, H. L. Lee, D. G. Long, S. Madriñán, O. Maurin, I. Meusnier, S. G. Newmaster, C. W. Park, D. M. Percy, G. Petersen, J. E. Richardson, G. A. Salazar, V. Savolainen, O. Seberg, M. J.

- Wilkinson, D. K. Yi, and D. P. Little. 2009. A DNA Barcode for Land Plants. *Proceedings of the National Academy of Sciences* 106(31):12794–12797. <http://dx.doi.org/10.1073/pnas.0905845106>.
- iBOL. 2012. iBOL. *iBOL*. <http://ibol.org/>.
- iBOL. 2015a. Data & Resource Sharing Policies. *International Barcode of Life*. <http://ibol.org/resources/data-release-policy/>.
- iBOL. 2015b. Partner Nations. *International Barcode of Life*. <http://ibol.org/about-us/partner-nations/>.
- iBOL. 2015c. Project Governance. The Organization of the iBOL Team. *International Barcode of Life*. <http://ibol.org/about-us/project-governance/#>.
- iBOL. 2015d. Sponsors | iBOL. *International Barcode of Life*. <http://ibol.org/about-us/sponsors/>.
- iBOL. 2015e. What Is iBOL | iBOL. *International Barcode of Life*. <http://ibol.org/about-us/what-is-ibol/>.
- Janzen, D. H. 2004. Now Is the Time. *Philosophical Transactions of the Royal Society B: Biological Sciences* 359(1444):731–732. <http://dx.doi.org/10.1098/rstb.2003.1444>.
- Johnson, R. N., L. Wilson-Wilde, and A. Linacre. 2014. Current and Future Directions of DNA in Wildlife Forensic Science. *Forensic Science International: Genetics* 10(May):1–11. <http://dx.doi.org/10.1016/j.fsigen.2013.12.007>.
- Joly, Y., E. S. Dove, K. L. Kennedy, M. Bobrow, B. F. F. Ouellette, S. O. M. Dyke, K. Kato, and B. M. Knoppers. 2012. Open Science and Community Norms Data Retention and Publication Moratoria Policies in Genomics Projects. *Medical Law International* 12(2):92–120. <http://dx.doi.org/10.1177/0968533212458431>.
- Li, D., D. Castle, R. Hanner, and P. Hollingsworth. 2013. Kunming Declaration on the Promotion of DNA Barcoding and Biodiversity Science. Kunming, China. [http://www.innogen.ac.uk/downloads/Kunming-Declaration\[2\].pdf](http://www.innogen.ac.uk/downloads/Kunming-Declaration[2].pdf).
- LMMC. 2002. Cancun Declaration of Like-Minded Megadiversity Countries. Cancun, Mexico. [https://www.environment.gov.za/sites/default/files/docs/cancun\\_lmhc\\_declaration1.pdf](https://www.environment.gov.za/sites/default/files/docs/cancun_lmhc_declaration1.pdf).
- Madison, M. 2014. Commons at the Intersection of Peer Production, Citizen Science, and Big Data: Galaxy Zoo. In: *Governing Knowledge Commons*, eds. Frischmann, Brett M., M. J. Madison, and K. J. Strandburg, 209–254. New York: Oxford University Press.
- Mayan, M. J. 2009. *Essentials of Qualitative Inquiry*. New York: Routledge Taylor & Francis Group.
- Meusnier, I., G. A. Singer, J.-F. Landry, D. A. Hickey, P. D. N. Hebert, and M. Hajibabaei. 2008. A Universal DNA Mini-Barcode for Biodiversity Analysis. *BMC Genomics* 9:214. <http://dx.doi.org/10.1186/1471-2164-9-214>.
- Mishra, A. and T. Bubela. 2014. Legal Agreements and the Governance of Research Commons: Lessons from Materials Sharing in Mouse Genomics. *OMICS: A Journal of Integrative Biology* 18(4):254–273. <http://dx.doi.org/10.1089/omi.2013.0158>.

- Mishra, A., P. Schofield, and T. Bubela. 2016. Sustaining Large-Scale Infrastructure to Promote Pre-Competitive Biomedical Research: Lessons from Mouse Genomics. *New Biotechnology* 33(2):280–294. <http://dx.doi.org/10.1016/j.nbt.2015.10.002>.
- Moritz, C. and C. Cicero. 2004. DNA Barcoding: Promise and Pitfalls. *PLoS Biol* 2(10):e354. <http://dx.doi.org/10.1371/journal.pbio.0020354>.
- Morse, J. M., M. Barrett, M. Mayan, K. Olson, and J. Spiers. 2002. Verification Strategies for Establishing Reliability and Validity in Qualitative Research. *International Journal of Qualitative Methods* 1(2):13–22. <http://dx.doi.org/10.1177/160940690200100202>.
- Nakamura, Y., G. Cochrane, and I. Karsch-Mizrachi. 2013. The International Nucleotide Sequence Database Collaboration. *Nucleic Acids Research* 41(D1):D21–D24. <http://dx.doi.org/10.1093/nar/gks1084>.
- NCBI. 2016. NCBI Website and Data Usage Policies and Disclaimers. <https://www.ncbi.nlm.nih.gov/home/about/policies.shtml>.
- Nelkin, D. 1998. The Performance of Science. *The Lancet* 352(9131):893–894. [http://dx.doi.org/10.1016/S0140-6736\(05\)60031-1](http://dx.doi.org/10.1016/S0140-6736(05)60031-1).
- NHGRI. 2003. Reaffirmation and Extension of NHGRI Rapid Data Release Policies. <https://www.genome.gov/10506537/>.
- Ostrom, E. 2003. Toward a Behavioral Theory Linking Trust, Reciprocity, and Reputation. In: *Trust and Reciprocity: Interdisciplinary Lessons from Experimental Research*, ed. E. Ostrom, 19–79. New York: Russell Sage Foundation Publications.
- Ostrom, E. 2005. *Understanding Institutional Diversity*. Princeton, New Jersey: Princeton University Press.
- Pennisi, E. 2003. Modernizing the Tree of Life. *Science* 300(5626):1692–1697. <http://dx.doi.org/10.1126/science.300.5626.1692>.
- Poteete, A. R. and E. Ostrom. 2004. Heterogeneity, Group Size and Collective Action: The Role of Institutions in Forest Management. *Development and Change* 35(3):435–461. <http://dx.doi.org/10.1111/j.1467-7660.2004.00360.x>.
- Ratnasingham, S. 2015. Informatics Tools for the Second Decade of DNA Barcoding. Presented at the 6th International Barcode of Life Conference, Guelph, August 20. <http://dnabarcodes2015.org/wp-content/uploads/2015/10/SujeevanRatnasingham.pdf>.
- Ratnasingham, S. and P. D. N. Hebert. 2007. Bold: The Barcode of Life Data System (<http://www.barcodinglife.org>). *Molecular Ecology Notes* 7(3):355–364. <http://dx.doi.org/10.1111/j.1471-8286.2007.01678.x>.
- Rehman, A., S. Jafar, N. Raja, and J. Mahar. 2015. Use of DNA Barcoding to Control the Illegal Wildlife Trade: A CITES Case Report from Pakistan. *Journal of Bioresource Management* 2(2):19–22. <http://corescholar.libraries.wright.edu/jbm/vol2/iss2/3>.
- Ruttan, L. M. 2006. Sociocultural Heterogeneity and the Commons. *Current Anthropology* 47(5):843–853. <http://dx.doi.org/10.1086/507185>.

- Schindel, D. 2010. Biology without Borders. *Nature*. 467(7317):779–781. <http://dx.doi.org/10.1038/467779a>.
- Schindel, D., T. Bubela, J. Rosenthal, D. Castle, P. du Plessis, R. Bye, and Pmcw. 2015. The New Age of the Nagoya Protocol. *Nature Conservation* 12(August):43–56. <http://dx.doi.org/10.3897/natureconservation.12.5412>.
- Schindel, D., C. Hauser, S. E. Miller, and Participants in an international workshop. 2008. Preserving International Access to Genetic Resources for Non-Commercial Biodiversity Research. In: Bonn, Germany. <https://69.90.183.227/abs/doc/cbol-dfg-bonn-abs-article-en.pdf>.
- Schoch, C. L., K. A. Seifert, S. Huhndorf, V. Robert, J. L. Spouge, C. A. Levesque, W. Chen, and Fungal Barcoding Consortium. 2012. Nuclear Ribosomal Internal Transcribed Spacer (ITS) Region as a Universal DNA Barcode Marker for Fungi. *Proceedings of the National Academy of Sciences* 109(16):6241–6246. <http://dx.doi.org/10.1073/pnas.1117018109>.
- Schofield, P. N., T. Bubela, T. Weaver, L. Portilla, S. D. Brown, J. M. Hancock, D. Einhorn, G. Tocchini-Valentini, M. H. de Angelis, and N. Rosenthal. 2009. Post-Publication Sharing of Data and Tools. *Nature* 461(7261):171–173. <http://dx.doi.org/10.1038/461171a>.
- Sonet, G., K. Jordaens, Y. Braet, L. Bourguignon, E. Dupont, T. Backeljau, M. De Meyer, and S. Desmyter. 2013. Utility of GenBank and the Barcode of Life Data Systems (BOLD) for the Identification of Forensically Important Diptera from Belgium and France. *ZooKeys* 365(December):307–328. <http://dx.doi.org/10.3897/zookeys.365.6027>.
- Stoeckle, M. 2003. Taxonomy, DNA, and the Bar Code of Life. *BioScience* 53(9):796–797. [http://dx.doi.org/10.1641/0006-3568\(2003\)053\[0796:TDATB C\]2.0.CO;2](http://dx.doi.org/10.1641/0006-3568(2003)053[0796:TDATB C]2.0.CO;2).
- Stoeckle, M. and P. D. N. Hebert. 2008. Barcode of Life. *Scientific American* 299(4):82–88. <http://dx.doi.org/10.1038/scientificamerican1008-82>.
- Strandburg, K. J., B. M. Frischmann, and C. Cui. 2014. The Rare Diseases Clinical Research Network. In: *Governing Knowledge Commons*, eds. Frischmann, Brett M., M. J. Madison, and K. J. Strandburg, 155–207. New York: Oxford University Press.
- Tautz, D., P. Arctander, A. Minelli, R. H. Thomas, and A. P. Vogler. 2003. A Plea for DNA Taxonomy. *Trends in Ecology & Evolution* 18(2):70–74. [http://dx.doi.org/10.1016/S0169-5347\(02\)00041-1](http://dx.doi.org/10.1016/S0169-5347(02)00041-1).
- Thomson Reuters. 2016. 2016 Journal Citation Reports Science Edition. *InCites Journal Citation Reports*. <https://error.incites.thomsonreuters.com/error/Error?DestApp=IC2JCR&Alias=IC2&Domain=.thomsonreuters.com&Src=IP&Params=DestApp%3DIC2JCR&RouterURL=https%3A%2F%2Flogin.incites.thomsonreuters.com%2F&Error=IPError>.
- Toronto International Data Release Workshop Authors. 2009. Prepublication Data Sharing. *Nature* 461(7261):168–170. <http://dx.doi.org/10.1038/461168a>.

- UNEP/CBD/SBSTTA/16/INF/37. 2012. Implementation of the Nagoya Protocol on Access and Benefits Sharing and the Global Taxonomy Initiative. <https://www.cbd.int/kb/record/meetingDocument/87319?Event=SBSTTA-16>.
- United Nations. 1992. UN Convention on Biological Diversity [Online]. Available at <https://www.cbd.int/Web/Default.shtml> [Accessed Nov 16, 2015].
- United Nations. 2010. Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from Their Utilization (ABS) to the Convention on Biological Diversity. 20 Oct 2010. C.N.782.2010.
- United Nations DGACM. 2016. United Nations Regional Groups of Member States. <http://www.un.org/depts/DGACM/RegionalGroups.shtml>.
- Vernooy, R., E. Haribabu, M. R. Muller, J. H. Vogel, P. D. N. Hebert, D. E. Schindel, J. Shimura, and G. A. C. Singer. 2010. Barcoding Life to Conserve Biological Diversity: Beyond the Taxonomic Imperative. *PLoS Biology* 8(7):e1000417. <http://dx.doi.org/10.1371/journal.pbio.1000417>.
- Waterton, C., R. Ellis, and B. Wynne. 2013. *Barcoding Nature: Shifting Cultures of Taxonomy in an Age of Biodiversity Loss*. Milton Park, Abingdon, Oxon: Routledge.
- WHO. 2016. Vector-Borne Diseases. WHO. <http://www.who.int/mediacentre/factsheets/fs387/en/>.
- Will, K. W. and D. Rubinoff. 2004. Myth of the Molecule: DNA Barcodes for Species Cannot Replace Morphology for Identification and Classification. *Cladistics* 20(1):47–55. <http://dx.doi.org/10.1111/j.1096-0031.2003.00008.x>.
- Wilson, E. O. 2003. The Encyclopedia of Life. *Trends in Ecology & Evolution* 18(2):77–80. [http://dx.doi.org/10.1016/S0169-5347\(02\)00040-X](http://dx.doi.org/10.1016/S0169-5347(02)00040-X).
- Wong, E. H.-K. and R. H. Hanner. 2008. DNA Barcoding Detects Market Substitution in North American Seafood. *Food Research International* 41(8):828–837. <http://dx.doi.org/10.1016/j.foodres.2008.07.005>.
- World Bank. 2016. Country and Lending Groups | Data. <http://data.worldbank.org/about/country-and-lending-groups>.
- Yancy, H. F., T. S. Zemlak, J. A. Mason, J. D. Washington, B. J. Tenge, N.-L. T. Nguyen, J. D. Barnett, W. E. Savary, W. E. Hill, M. M. Moore, F. S. Fry, S. C. Randolph, P. L. Rogers, and P. D. Hebert. 2008. Potential Use of DNA Barcodes in Regulatory Science: Applications of the Regulatory Fish Encyclopedia. *Journal of Food Protection* 71(1):210–217. <https://doi.org/10.4315/0362-028X-71.1.210>.