How the DNA Barcode Commons are Governed: Understanding how a Heterogeneous Global Community Shares Genetic Resources for Non-Commercial Use

by

Janis Dawn Geary

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Public Health

School of Public Health
University of Alberta

Abstract

Life sciences research that uses genetic materials is increasingly collaborative and global. Research partnerships between researchers in high income countries and low and middle income countries have potential benefits for all involved, including minimizing costs and sharing risks. Such partnerships are supported through the development of research infrastructure for sharing genetic resources, including databases and bio-repositories. These shared resources are a type of "knowledge commons" - sets of information-based resources available on terms that encourage efficiency, equitable use, and sustainability and that are managed by groups of various sizes. My empirical research aims to guide development of evidence-based best practices for the rules and institutional structures to govern global knowledge commons. It is grounded in a case-study of the DNA barcoding commons - an international effort to facilitate biodiversity monitoring through standardization of DNA-based species identification. This global resource comprises genetic resources including biological specimens, DNA barcode sequence data and associated metadata.

Genetic resource commons share governance challenges with other types of knowledge commons, notably the need to encourage the re-contribution of value-added data arising from use of the resource, and discourage "free-riders", or individuals who use the resource without ever contributing to it. These challenges are exacerbated in genetic resource commons because of the differential capacity to use and contribute to the commons. Genetic resources are disproportionately located in the mega-diverse lower income countries, while the infrastructure necessary to generate, store, manage and utilize those genetic resources is disproportionately located in high income countries. This unequal distribution gives rise to complex and intertwined social and scientific concerns relating to the equitable distribution of benefits arising from the

commons, the protection of genetic resources from unsanctioned commercial development, and the progress of research. These concerns are significant stumbling blocks towards building shared resources, and create the need to understand which rules, structures, and incentives promote fair rules to govern contribution to and use of the commons.

My research advanced theory on the development of knowledge commons and provides recommendations for best practices. Using a case-study approach, I examined how research stakeholders in global partnerships established governance structures for the global DNA barcoding commons, and identified ways to improve them. My empirical research answers three broad questions:

1) How do factors identified by the Institutional Analysis and Development (IAD) framework influence effective governance of a global knowledge commons?

2) How does heterogeneity inform the rules used by actors to govern their behaviours?

3) How are topics and issues relevant to governing the global DNA barcode commons presented in newspaper coverage?

I found that the IAD factors relevant to the DNA barcoding commons created challenges for the collective action required for effective governance. While the DNA barcode commons is functioning to generate and share barcode records, the strategy to openly share data and materials has not yet produced a globally representative barcode resource. To achieve this goal, governance structures should promote an equitable distribution of burdens and benefits for contribution, access, and use. The challenges in governance are related to other findings from my research; the bases of the rules used by DNA barcoding participants suggested a lack of shared understanding for crafting rules between heterogeneous stakeholders. Nevertheless, my research

pointed to mechanisms to develop suitable rules for participation in a global knowledge commons based on shared expectations in contexts where the heterogeneous participants might otherwise choose rules that drive conflicting behaviours. Lastly, my research demonstrated that critical issues, such as fair and equitable access and benefits sharing, were omitted from public barcoding discourse in countries where influential policies and guidelines are being developed. The considerable media coverage focused on positive aspects of barcoding science. There exists an opportunity, therefore, for leaders of the barcoding community to generate more awareness of the social and policy context of DNA barcoding activities and their conservation/regulatory goals.

Overall, my empirical research aimed to inform best practices for the governance of global knowledge commons, and I have outlined several strategies based on the exemplar DNA barcode commons. Through facilitating representative governance, collective rules-making, and consideration of the complex social and policy context, global knowledge commons can enhance partnerships between researchers in high and lower income countries with benefits for all involved.

Preface (Mandatory due to research ethics approval and collaborative work)

This thesis is an original work by Janis Geary (JG). The research that informs this work received ethics approval from the University of Alberta Health Research Ethics Board, Project Name "Building Robust Research Commons to Support Large-Scale International Research Projects in Genomics and Life Sciences", No. Pro00016720, October 6, 2010. Some of the research in this thesis is the result of collaborations. Collaborators are: Dr Tania Bubela (TB), Dr Cindy Jardine (CJ), Dr. Trish Reay (TR), Dr Karen Goodman (KG), Dr Ashok Kumbamu (AK), Dr Harley De Cerqueira (HC), Mark Bieber (MB), Jennifer Ann McGetrick (JM), Westerly Luth (WL), and Emma Camicioli (EC).

Chapters 1 and 5 are the original, unpublished work of the author. JG conceptualized and wrote each chapter with guidance from TB, and input from CJ and KG.

Chapter 2 is original research by JG in preparation for submission for publication in a peer-reviewed journal. TB provided oversight to JG for developing and conducting the study. JG adapted an interview guide previously developed by TB and AK. AK and HC conducted interviews in India and Brazil, respectively. JG verified all transcripts, other than the Portuguese transcripts that were translated and verified by HC. JG analyzed the interview transcripts. JG and TB developed the iBOL participant survey questions in partnership with barcoding stakeholders, and JG analyzed the data collected from the survey. JG developed the strategy to identify relevant publications for the bibliometric analysis, and MB retrieved the articles. MB developed the computer program to disambiguate author names in the database of DNA barcoding publications. WL and MB helped draft the description of the computer program in the Appendix. MB produced Figure 4: Co-authorship in the DNA barcoding publication database. JG developed the strategy to identify medicinal plant and mosquito records from the Barcode of Life Data System (BOLD). JM produced the list of medicinal plant names, and MB retrieved medicinal plant records from BOLD using an automated script. JG and JM analyzed the retrieved data. JG retrieved the mosquito records from BOLD and conducted the analysis. JG wrote the manuscript with input from TB, KG, and CJ.

Chapter 3 is original research by JG in preparation for submission for publication in a peer-reviewed journal. JG used a subset of the interviews collected and verified as described above. JG developed the analytical framework under the supervision of TB and TR, and

conducted the data analysis. WL provided secondary coding to verify JG's interpretation of the data. JG wrote the manuscript with input from TB, TR, KG, and CJ.

A version of chapter 4 has been published as J Geary, E Camicioli, and T Bubela, "DNA Barcoding in the Media: Does Coverage of Cool Science Reflect Its Social Context", 2016, *Genome* 59(9): 738-750, doi:10.1139/gen-2015-0210. JG designed the study with input from EC, under the supervision of TB. JG and EC collected the data, and JG cleaned the data set, conducted analysis, and took a primary role in preparing the first draft of the manuscript. EC provided feedback on describing data collection procedures, and TB edited the text and provided expert knowledge of literature on scientific communication through the media.

Acknowledgements

I am very grateful for all of the things that had to come together for my work to be possible: Funding that provided me the means to pursue my own interests; mentors that allowed me the freedom to be an independent scholar, and the support to do so successfully; and a research community that was as willing to talk to me about their challenges as they were their accomplishments.

I received doctoral support from: the Canadian Institutes of Health Research Frederick Banting and Charles Best Canada Graduate Scholarship; the Alberta Innovates: Health Solutions Health Research Studentship; the Killam Trusts' Izaak Walton Killam Memorial Scholarship; and the University of Alberta President's Doctoral Prize of Distinction. My research costs were also supported by the International Barcode of Life Project, funded through Genome Canada.

I have had immense support from my supervisory committee. I would like to thank my primary supervisor, Dr. Tania Bubela, for her overall help and support in shaping my research, and especially her work to help improve my writing. I also acknowledge my co-supervisor, Dr. Cindy Jardine, for providing key feedback at crucial junctions during my PhD. Dr. Trish Reay was an immense support in helping me work through how to develop my theoretically-grounded analytical plans, and how to interpret my data. I am grateful to Dr. Reay for her encouragement of my academic pursuits, and support of my belief that organizational theory can have practical applications. I would like to specifically acknowledge Dr. Karen Goodman for her mentorship over the last decade. I am especially grateful for her ability to see the value in different approaches to science, and the skill sets needed to accomplish them (and her patient teaching of the approaches to research that do not come naturally to me). Dr. Goodman's influence on my academic career started well before my PhD, and it will surely continue afterwards.

I would like to thank Dr. Paul Hebert, Dr. David Schindel, and Dr. David Castle for their support in researching DNA barcoding as a case-study, as well as the staff at the Biodiversity Institute of Ontario. I am very grateful to each individual who not only agreed to be interviewed, but enthusiastically shared with me their perspectives.

I have had the great benefit of being involved in not one, but two large research groups throughout my PhD. I am very lucky to have such a fantastic group of colleagues. I want to specifically acknowledge Mark Bieber, Emma Camicioli, Westerly Luth and Freja McGetrick

(from Law and Risk Communication in Health – LaRCH) for their contributions to this work. I would like to acknowledge their overall moral support of our research team, along with Steph Kowal, Shelly Benjaminy and the rest of the LaRCH team. I would like to acknowledge the support of Emily Walker for her help with statistics, and the support from her and the rest of my colleagues from the Canadian North *Helicobacter pylori* Working Group.

Throughout my PhD, I have been very lucky to have the support of a fantastic group of scholars and friends in my PhD program. I would like to acknowledge Fatima Al Sayah, Anushka Ataullahjan, Amy Colquhoun, Elaine Hyshka, Nikki Kain, and Jacqueline Torti for their thoughtful and ongoing support of me and other students. I also have the wonderful support of academic friends not in my PhD program who constantly inspire me to think about science and health in different ways, including April Gislason, Amanda Jones, and Anna Seifried.

I would like to thank my exceptional family. I am grateful to my parents, Hugh and Judy Geary, for always encouraging me and my sister to question the world around us. For those who maybe get annoyed at my diligent scrutiny, you can blame them for instilling in me a staunch belief that I am not required to think things simply because other people have already thought them. I am grateful to my sister Dr. Laina Geary who, in addition to being an inspiring and creative organic chemist, got a PhD before me and therefore made it impossible for me to not get one too. For those who also have four-legged moral support in their lives, I want you to know my 6 years of PhD scholarship were greatly enhanced by my disgruntled toy poodle Zeus, and my happy-go-lucky shiba Oki. And finally, I am especially grateful to my patient and wonderful husband, Kam Shergill, who has somehow graciously put up with me being a student for the entire time he has known me. Kam is one of the smartest people I know, and I am so lucky to have him as my partner.

Table of Contents

List of Tables

List of Figures

Glossary of Terms

Term	Definition
BOLD	Barcode of Life Data System; "The Barcode of Life Data Systems (BOLD) is an informatics workbench aiding the acquisition, storage, analysis, and publication of DNA barcode records. By assembling molecular, morphological, and distributional data, it bridges a traditional bioinformatics chasm. BOLD is freely available to any researcher with interests in DNA barcoding." (Ratnasingham and Hebert 2007, 355)
CBD	*Convention on Biological Diversity* is an international agreement with the three main objectives of "1) The conservation of biological diversity, 2) The sustainable use of the components of biological diversity, and 3) The fair and equitable sharing of the benefits arising out of the utilization of genetic resources (United Nations 1992)
CBOL	Consortium for the Barcode of Life. An organization established in 2004 at the Smithsonian Institution that is "an international initiative devoted to developing DNA barcoding as a global standard for the identification of biological species." (Consortium for the Barcode of Life 2015b)
Commons	Resources shared by a group of people (Hess and Ostrom 2011)
DNA barcode commons	The set of resources (biological materials, genetic material, genetic sequences, and metadata) and infrastructure (databases and biorepositories) need to support DNA barcoding activities
DNA barcoding	The process of using short, standardized gene regions to differentiate and identify species (Hebert et al. 2003)
Exclusion	A characteristic of a resources that refers to the costliness/difficulty of preventing use of a resource (McGinnis 2011)
Genetic resource	"Genetic resources (GRs) refer to genetic material of actual or potential value. Genetic material is any material of plant, animal, microbial or other origin containing functional units of heredity. Examples include material of plant, animal, or microbial origin, such as medicinal plants, agricultural crops and animal breeds." (United Nations 1992)
Grammar of institutions	A five-component "syntax for analyzing and expressing institutional statements that can be used to distinguish systematically among rules, norms, and shared strategies". The components (ADICO) are: Attribute (the participant), Deontic (may, must, or must not), aIm (outcome or action), Condition (when/where the aim is permitted or forbidden), Or else (consequence for not following) (Crawford and Ostrom 2005, 139)
Heterogeneity	Refers to the cultural and financial differences between actors in the commons. (Vedeld 2000)
IAD Framework	An analytical tool for studying institutions that "assigns all relevant explanatory factors and variables to categories and locates these categories within a foundational structure of logical relationships" (McGinnis 2011, 169)
iBOL	International Barcode of Life. An organization established in 2010 at the University of Guelph that brings together researchers with a main goal of

	"extending the geographic and taxonomic coverage of the barcode reference library" (iBOL 2015i)
Institution	"Human-constructed constraints or opportunities within which individual choices take place and which shape the consequences of their choices" (McGinnis 2011, 170)
Institutional logics	"Institutional logics are both material and symbolic--they provide the formal and informal rules of action, interaction, and interpretation that guide and constrain decision makers in accomplishing the organization's tasks and in obtaining social status, credits, penalties, and rewards in the process. These rules constitute a set of assumptions and values, usually implicit, about how to interpret organizational reality, what constitutes appropriate behaviour, and how to succeed" (Thornton 2004, 69)
Institutional statement	A *Rule*, *Norm*, or *Strategy* that sets out "constraints or opportunities to prescribe, permit, or advise actions or outcomes for participants in action situations" (Crawford and Ostrom 2005, 137)
ISBOL	International Society for the Barcode of Life; An international coordinating organization tasked with implementing the eight objectives of the Kunming Declaration (Castle et al. 2015)
ISSC	International Scientific Steering Committee of iBOL; advises the Scientific Director on research plans and deliverables (iBOL 2015e)
KC-IAD Framework	The IAD framework as adapted for the study of knowledge commons (Frischmann et al. 2014)
Knowledge commons	"Commons arrangements for overcoming various social dilemmas associated with sharing and producing information, innovation and creative works" (Frischmann et al. 2014, 1)
Like-minded mega diverse country (LMMC)	"The Like-Minded Megadiverse Countries (LMMC) is a group of countries that harbour the majority of the Earth's species and are therefore considered extremely biodiverse. They are rich in biological diversity (60-70% of the world's biodiversity) and associated traditional knowledge. These countries have effectively joined efforts in negotiating the development of the Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits arising from their utilisation to the Convention on Biological Diversity (CBD), which was adopted in Japan in 2010. 18 in total, these countries are located in, or partially in, tropical or subtropical regions. " (LMMC 2002)
Nagoya Protocol	*Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization "*is a supplementary agreement to the Convention on Biological Diversity. It provides a transparent legal framework for the effective implementation of one of the three objectives of the CBD: the fair and equitable sharing of benefits arising out of the utilization of genetic resources" (United Nations 2010)
Norm	A type of institutional statement that includes four components of the grammar, ADIC. As such, not following *Norms* has no defined consequence (Crawford and Ostrom 2005)

Open access	A property rights system where anyone may access resources, although there may be restrictions on use of the resources (Suber 2011)
Polymerase chain reaction	"Polymerase chain reaction, or PCR, is a laboratory technique used to make multiple copies of a segment of DNA. PCR is very precise and can be used to amplify, or copy, a specific DNA target from a mixture of DNA molecules." (Nature Education 2014)
Primer	A short strand of DNA used to initiate a chain reaction for amplification or sequencing reactions.
Public Good	Good that are non-subtractable and have high-costs associated with excluding users (McGinnis 2011)
Rivalrous	A characteristic of a resource that refers to the results when one individual's use lowers another's potential use (McGinnis 2011)
Rule	A type of institutional statement that includes all five components of the grammar, ADICO (Crawford and Ostrom 2005)
Strategy	A type of institutional statement that includes three components of the grammar, AIC. As such, strategies are the least prescriptive of institutional statements (Crawford and Ostrom 2005)
Voucher specimen	A specimen linked to a DNA barcode record and held in a repository to enable verification of species identification

References

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.

Consortium for the Barcode of Life. (2015). What Is CBOL? *Barcode of Life*. http://www.barcodeoflife.org/content/about/what-cbol.

Crawford, S., and E. Ostrom. (2005). A Grammar of Institutions. In: *Understanding Institutional Diversity*. ed. E. Ostrom, 137–74. Princeton University Press.

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.

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): 313–21. http://dx.doi.org/10.1098/rspb.2002.2218.

iBOL. (2015a). Project Governance. The Organization of the iBOL Team. *International Barcode of Life*. http://ibol.org/about-us/project-governance/#.

iBOL. (2015b). What Is iBOL | iBOL. *International Barcode of Life*. http://ibol.org/about-us/what-is-ibol/.

LMMC. (2002). Cancun Declaration of Like-Minded Megadiversity Countries. Cancun, Mexico. https://www.environment.gov.za/sites/default/files/docs/cancun_lmmc_declaration1.pdf.

McGinnis, M. D. (2011). An Introduction to IAD and the Language of the Ostrom Workshop: A Simple Guide to a Complex Framework. *Policy Studies Journal*. 39(1): 169–83. http://dx.doi.org/10.1111/j.1541-0072.2010.00401.x.

Nature Education. (2014). Polymerase Chain Reaction / PCR | Learn Science at Scitable. http://www.nature.com/scitable/definition/polymerase-chain-reaction-pcr-110.

Ratnasingham, S., and P. D. N. Hebert. (2007). Bold: The Barcode of Life Data System (Http://www.barcodinglife.org). *Molecular Ecology Notes*. 7(3): 355–64. http://dx.doi.org/10.1111/j.1471-8286.2007.01678.x.

Suber, P. (2011). Introduction: An Overview of the Knowlege Commons. In: *Understanding Knowledge as a Commons: From Theory to Practice*. eds. C. Hess and E. Ostrom, 171–208. MIT Press.

Thornton, P. H. (2004). *Markets from Culture: Institutional Logics and Organizational Decisions in Higher Education Publishing*. Stanford University Press.

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.

Vedeld, T. (2000). Village Politics: Heterogeneity, Leadership and Collective Action. *Journal of Development Studies*. 36(5): 105–34. http://dx.doi.org/10.1080/00220380008422648.

Chapter 1: Introduction

Overview

Life sciences research that uses genetic resources is increasingly collaborative and global and can be enhanced through partnerships between researchers in high and lower income countries (Ray et al. 2009). Global research partnerships have benefits for all involved, including minimizing costs and sharing risks (Padma 2005), enabling access to important resources or markets (Melon et al. 2009; Thorsteinsdottir et al. 2004), and bringing diverse perspectives to research questions of global importance (Freeman and Huang 2014). Such partnerships are supported through the development of research infrastructure for genetic resources, including databases and bio-repositories. If managed appropriately, infrastructure for genetic resources can add value to research globally by enhancing the quality of the resources and enabling their distribution to the research community. These shared resources are a type of "knowledge commons" - sets of information-based resources available on terms that encourage efficiency and equitable use and that are managed by groups of varying sizes and interests (Ostrom et al. 1999; Frischmann et al. 2014). My empirical research aims to inform best practices for the governance of global knowledge commons.

Empirical research to study knowledge commons has built on the body of research on institutional arrangements used to govern natural resource commons by Nobel Laureate and political economist, Elinor Ostrom. Ostrom (1990), proposed the Institutional Analysis and Development (IAD) framework for understanding the governance of natural resource commons which has been tested in settings as diverse as community policing (Kahan 2002), fisheries (Imperial and Yandle 2005), forestry (Andersson 2006) and pastoral land (Behnke 1994). In 2006, Ostrom and colleagues began to apply the lessons from natural resource governance to study knowledge commons (Hess and Ostrom 2006; Hess and Ostrom 2011b; Hess 2012). Frischmann and colleagues adapted the IAD framework to account for the differences between natural resource and knowledge commons (Frischmann et al. 2014). I refer to their adaptation as the Knowledge Commons IAD (KC-IAD) framework. My research expands our understanding of institutional arrangements for governing knowledge commons by applying the KC-IAD framework to a *global* genetic resource knowledge commons: The DNA barcode commons.

Prior analyses of knowledge commons have focused on those situated in high income countries, whose activities were based on shared infrastructure (Bubela et al. 2012; Contreras 2014; Dedeurwaerdere 2010c). Little research has focused on the establishment and ongoing governance of knowledge commons that engage institutional and individual participants from disparate regions with historical power and economic imbalances. In addition to economic, language, and cultural differences, wide geographic distribution of participants makes the communication to facilitate collective action difficult. These factors add to the heterogeneity of the participants, which itself has been extensively studied as a factor that increases the challenges in establishing collective action in commons. Communities may vary based on any number of factors, making heterogeneity a difficult factor to study empirically.

Introduction to the DNA barcode commons

In 2003, a team led by Paul Hebert at the University of Guelph in Canada proposed DNA barcoding as a tool to accelerate documenting life on earth – a pre-requisite for the study of anthropogenic and other impacts on biodiversity (Hebert et al. 2003). The group demonstrated the potential to differentiate species by way of sequencing a small region of an organism's DNA. Previous efforts at classifying organisms were inefficient (requiring highly specialized taxonomists), expensive (requiring extensive DNA sequencing), and not scalable due to lack of standardization (Tautz et al. 2003). Hebert et al suggested using short, highly conserved, and relatively ubiquitous DNA sequences. These sequence characteristics make high-throughput analyses possible, enabling a barcoding pipeline with attendant economies of scale.

While DNA barcoding is promoted as a tool to help identify life on earth, the process to create a DNA barcode begins with taxonomically classified specimens. Therefore, biological materials are an integral component of the DNA barcode commons. Specimens must be shipped to locations capable of performing other tasks associated with creating barcode records, including isolating DNA, sequencing the relevant barcode, and storing the voucher specimen. The result of the barcoding process is an open access, comprehensive database of DNA barcodes linked to metadata and a reference specimen.

The effort to build the DNA barcoding database quickly gained momentum not only because of its potential contributions to the field of taxonomy but also its enablement of a range

of practical applications (Hebert et al. 2003). Access to such a database facilitates rapid identifications of unknown specimens in situations where morphological identification is impossible. Identifications are made by matching the DNA barcode from an unknown specimen to the known barcode record linked to the voucher specimen.

Creating a comprehensive database of DNA barcode records was necessarily a global endeavor, and scientists quickly came together to generate international participation through two main organizations: The Consortium for the Barcode of Life (CBOL) and the International Barcode of Life Project (IBOL). CBOL was founded in 2004 at the Smithsonian Institution, and focused on promoting the DNA barcode system and developing global standards. In 2010, Paul Hebert launched an international initiative to build a barcode reference library with the International Barcode of Life (iBOL). Leaders within the DNA barcode community began the process to establish a new organization in 2015, the International Society for the Barcode of Life (ISBOL), to continue governance of the endeavor at the international level. The development of ISBOL is ongoing as of this writing; therefore, recommendations for governance structures for DNA barcoding are timely.

The goals of the DNA barcoding community require sharing genetic resources across borders, which creates additional challenges to the collective action required to develop the commons. Genetic resources are disproportionately located in lower income countries (Myers et al. 2000), and the research infrastructure necessary to utilize those genetic resources is disproportionately located in high income countries. The problems created because of differential capacity to contribute to and use the research commons are exacerbated by the history of colonialism and its exploitation of lower income countries to the disproportionate benefit of high income countries. Given the history of exploitation of lower income countries by high income countries, the differential distribution of genetic resources and the capacity to make use of them gives rise to complex and intertwined social and scientific challenges relating to the equitable distribution of benefits from genomics research. These challenges are a formidable barrier to building global genetic resource commons that effectively promote the progress of life sciences research (Bubela and Gold 2012).

Objectives

My interdisciplinary research aims to advance theory and practice of governing knowledge commons through a case-study of the DNA barcode commons that addresses three research questions: 1) How do factors included in the IAD framework influence effective governance of a global knowledge commons; 2) How does heterogeneity inform the rules used by actors to govern their behaviours; and 3) How are topics and issues relevant to governing the global DNA barcode commons presented in newspaper coverage? Each question is addressed in a scholarly paper intended to be readable as a stand-alone journal article. The three papers present the content of my thesis. Before describing each paper, I provide a brief overview of the background of empirical work on the commons and on genetic resources. As this work crosses several academic disciplines, I have included a glossary of key terms in the preceding sections.

Background

History of empirically studying commons

Early empirical work on commons focused on natural resources such as fisheries, water, and land. One of the most widely cited theories on governing natural resource commons was written in 1968 by Hardin, who speculated that if shared resources were left unmanaged, they would be quickly depleted by individuals seeking to maximize their own benefit, and would no longer be available to any member of the community. He termed this phenomenon "the tragedy of the commons" (Hardin 1968), and suggested private property as a solution. Despite the popularity of the concept, Hardin's ideas were refuted by many scholars, notably because he did not provide any evidence or data to support his claims (Feeny et al. 1990). Scholars pointed out that a more accurate term for the phenomenon Hardin described would be "the tragedy of the *unmanaged* commons". Research efforts have provided examples of informal community institutions, which successfully manage shared resources without formal government or private property schemes (Bromley 1992; Ostrom and Gardner 1993; Ostrom 1992).

Ostrom made a significant contribution to refuting Hardin's commons theory with her 1990 work *Governing the Commons* (Ostrom 1990), which provided in-depth analyses of successful commons across the globe, including Swiss grazing pastures, Japanese forests, and

Philippine irrigation systems. Although she stressed the importance of understanding the specific context of each individual type of commons, she outlined design principles that were common to each of the regimes she studied. She refined these principles over time through subsequent studies. As outlined in her 2005 book, *Understanding Institutional Diversity* (Ostrom 2005a), the principles are: clearly defined physical and social boundaries; benefits that are proportional to contributions; active participation of users in rule making; a system for self-monitoring behaviour; a graduated system of sanctions; conflict resolution mechanisms; and recognition of rights to organize. In the same work, Ostrom provided the fullest description of her IAD Framework (Ostrom 2005a), which provides key variables for the systematic study of how institutions shape social interactions and decision-making for managing common resources. Applying this systematic framework to various types of commons has greatly facilitated comparisons across studies, and has led to well-established principles and concepts related to managing natural resource commons (Poteete et al. 2010c). However, the framework has not yet been applied extensively to other types of commons, such as knowledge commons used for research (Dedeurwaerdere 2010c; Hess and Ostrom 2011b; Hess and Ostrom 2006; Bubela et al. 2012). Knowledge commons are similar to natural resource commons in that they are jointly used and managed by groups of varying sizes and interests. These similarities make it useful to apply the same framework to their study. However, there are several key differences (Bubela et al. 2012) that make it necessary to theoretically expand and enrich the IAD framework if it is to be applied to knowledge commons effectively.

The two characteristics that are frequently used to define the nature of the resources in a commons are exclusion and subtractability of use (Ostrom 2005b). Exclusion refers to how difficult it is to restrict use, and subtractability refers to how much one person's use prevents another from also using the resource. Knowledge commons are not subtractable, and it can often be difficult to develop exclusions to prevent "freeloading", which occurs when users benefit from the commons without contributing to it (Ostrom 2005b). Indeed, a fundamental difference between natural resource commons and knowledge commons is how increased use by individuals impacts the commons. For natural resource commons, unsuccessful management can allow overuse by individuals and depletion of the resource. In contrast, for knowledge commons, extensive use by individuals does not result in depletion. Instead, the value of knowledge

commons is enhanced with use of the resources (data or biological specimens), and unsuccessful management may result in under-use of the resources (Schofield et al. 2009). Indeed, the value of a knowledge commons is enhanced with use and re-contribution of new knowledge and materials (Schofield et al. 2009; Bubela et al. 2012). Further, knowledge commons are not as limited in geographic scope as natural resource commons, and have the potential to be global.

These differences between types of commons led scholars to develop an adapted framework for studying knowledge commons. (Frischmann et al. 2014). We refer to the adapted framework as the KC-IAD, and detail its components in Chapter 2.

Heterogeneity and the Commons

Heterogeneity is one aspect of the attributes of a commons community and broadly refers to characteristics that are different among actors. Variables that represent socio-cultural and economic characteristics are two categories that are frequently studied to understand heterogeneity (Ruttan 2008; Ruttan 2006). However, communities may vary based on any number of factors, making heterogeneity a difficult factor to study empirically. Nonetheless, substantial research has focused on understanding the impact of heterogeneity on commons governance. The emerging theoretical consensus is that increased heterogeneity is a challenge for communities to overcome, but does not inevitably lead to failure (Ostrom 2005a).

Looking beyond the question of whether heterogeneity generally influences outcomes of a commons, scholars have increasingly focused on the different impacts of specific types of heterogeneity. Most of such studies are quantitative and include heterogeneity as a narrowly defined variable. A review of these studies suggested that how heterogeneity impacts a commons depends on the type of heterogeneity that is examined as the independent variable, and whether the study measured collective action or the provision of a collective good as the dependent variable (Ruttan 2008).

Most studies of heterogeneity have examined rivalrous resources in natural resource commons, and few studies have explored the relationship between heterogeneity and knowledge commons governance. Knowledge commons require separate study and theory development because the goals of successful knowledge commons governance are different from natural resource commons: Rather than preventing overuse and ensuring sustainability, goals are to

increase use of the resource, promote active re-contribution of value-added information, and achieve a fair distribution of the costs and benefits of the resource among users spread across a potentially unlimited geography. Many institutional arrangements for these non-subtractable resources provide free and open access. In theory, open access makes use more efficient, but in practice, it creates additional challenges. Open access leaves the resource vulnerable to free riders - those who take from the commons without contributing to it – and free riding reduces trust and therefore the ability to establish shared norms within the community. Thus, heterogeneity in a knowledge commons may include many dimensions that differ from the previously studied natural resource commons communities.

Overview of genetic resources

Most of the world's genetic resources are located in lower income countries in tropical regions, with some estimates of the proportion located in these regions as high as 95% (McDougall 1995). Because many countries wish to protect their sovereignty over genetic resources in their borders, extensive work has been done at the international level to promote fair and equitable access to genetic resources and the sharing of benefits that arise from research and development, collectively termed access and benefit sharing (ABS). A group of highly-biodiverse countries joined to create the Like-Minded Mega Diverse Country (LMMC) group, which promotes the interests of its member countries in international agreement negotiations related to genetic resources.

Genetic resources in countries that are signatory to the *Convention on Biological Diversity* (*CBD*), such as Canada, are now subject to its definitions and terms pertaining to genetic resources and ABS. Article 2 of the *CBD* defines genetic resources as "genetic material of actual or potential value" and genetic material is defined as "any material of plant, animal, microbial or other origin containing functional units of heredity" (United Nations 1992). The objectives of the *CBD* are "the conservation of biological diversity, the sustainable use of its components and the fair and equitable sharing of the benefits arising out of the utilization of genetic resources" (United Nations 1992). The *CBD* presumes countries will make genetic resources accessible for research (United Nations 1992). Article 15(7) of the *CBD* calls for countries to take measures aimed at sharing with indigenous communities "in a fair and equitable

way the results of research and development and the benefits arising from the commercial and other utilization of genetic resources" (United Nations 1992).

As a supplement to the ABS provisions in the *CBD*, the *Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization to the Convention on Biological Diversity* (hereinafter referred to as *The Protocol)* was adopted on October 29, 2010 (United Nations 2010). *The Protocol* clarifies and emphasizes the importance of ensuring that genetic resources and associated traditional knowledge are accessed and utilized in a fair and equitable way. The basic principles for gaining access require obtaining prior informed consent of traditional knowledge holders as well as negotiating agreements with mutually agreed terms (Kamau et al. 2010). Article 8(1) of the protocol suggests simplified procedures for those who wish to access genetic resources for non-commercial use (United Nations 2010). *The Protocol* entered into force in October 2014 after the required 50 ratifications, and as of December 2016, 93 countries have ratified it (Convention on Biological Diversity 2012). Therefore, researchers and other individuals wishing to access and utilize genetic resources are increasingly required to be aware of *The Protocol* and how it has been implemented in countries they work in.

Overview of thesis chapters

Chapter 2: Global inequities and sharing genetic resources for non-commercial research: A case-study of the DNA barcode commons

Chapter 2 provides a comprehensive description of DNA barcoding science and the DNA barcode commons. In this chapter, I used multiple methods to empirically examine how factors identified in the KC-IAD framework influence governance of DNA barcode resources and management of infrastructure, and facilitate global participation in DNA barcoding efforts. I used qualitative content analysis to analyze semi-structured interviews with DNA barcoding stakeholders, and DNA barcoding organizational and procedural documents. I used descriptive statistics to analyze a database of DNA barcoding publications, and two types of barcode submission records. I used a logistic regression to measure the association between barcoding

author characteristics and article publication, and used a network analysis to display co-authorship patterns of DNA barcoding publications.

This paper also includes a detailed description of the KC-IAD framework. My analysis contributes to my overall goal of producing best practices for governing global knowledge commons by describing how factors identified in the KC-IAD framework may encourage or inhibit successful governance in the unique context of a global research environment.

I have written this paper for an audience of knowledge commons scholars and intend to submit it for publication in the International Journal of the Commons.

Chapter 3: The impact of heterogeneity in a global knowledge commons: Implications for governance of the DNA Barcode Commons

Chapter 3 is an in-depth examination of how one factor (heterogeneity) relates to one component of governance (rules-in-use). My aims of this paper are twofold: 1) to empirically examine how heterogeneity relates to rules-in-use in a global knowledge commons, and 2) to contribute an expanded, theoretically-grounded method for the study of heterogeneity and knowledge commons, using the concept of institutional logics. I detail the relevant theories and the conceptual framework used to accomplish the aims. I outline how rules are studied within a commons using the grammar of institutions developed by Elinor Ostrom and colleagues (Crawford and Ostrom 2005), and introduces institutional logics theory to expand the understanding of heterogeneity in relation to rules. Using a mixed methods approach, I developed an analytical framework that combines the grammar of institutions with the concept of institutional logics, and I used that framework to analyse interviews with DNA barcode researchers and barcode of life organizational documents.

One of the benefits of focusing on rules in an institutional analysis is that unlike many factors of a knowledge commons, rules are modifiable. This analysis, therefore, contributes to my overall research goals through suggesting rules that could be modified or new rules that could be adopted by the DNA barcode community to improve governance and increase global participation.

I have written this paper for an audience of knowledge commons scholars and intend to submit it for publication in the International Journal of the Commons.

Chapter 4: DNA barcoding in the media: Does coverage of cool science reflect its social context

Chapter 4 describes an empirical examination of how efforts to establish the DNA barcode commons have been represented in public discourse. I present my analysis of English-language newspaper coverage of DNA barcode projects and, specifically, whether newspaper coverage reflected the breadth of the scientific and social mandates of BOL organizations. Given that mass media coverage is a proxy for public discourse and sheds light on government policy directions (McCombs 2004), this paper contributes to my overall goal to identify which DNA barcode topics and issues relevant to governance of the commons are underrepresented and may benefit from targeted communication efforts. In this paper, I use the first-person plural "we" as it was published with multiple authors, although I have outlined my specific contributions in my thesis preface.

This paper was written for an audience of DNA barcode researchers and stakeholders, and was published in the journal *Genome* (Geary et al. 2016).

Rigor

One of the main sources of rigor in my research is the credibility gained from using well-established theoretical frameworks to guide my analysis (Mayan 2009). While my data analysis was largely done independently, my conclusions were drawn from the previous knowledge of how factors and variables in a knowledge commons related to each other. The rich descriptions provided in a case-study such as this also provide transferability, as others who may study global knowledge commons are aware of the important contextual factors that have impacted my conclusions. The use of quotes in my reporting provides confirmability, demonstrating to the reader that my results are grounded in my data (Morse et al. 2002). My use of qualitative analysis software provides confirmability as well, as it provides ongoing opportunities to examine data, and review how I made analytical decisions.

Conclusion

My research enhances existing scholarship on knowledge commons through examining a global and heterogeneous knowledge commons, the DNA barcode commons. Specifically, my research contributes an understanding of how factors in a global research environment impact knowledge commons governance, how heterogeneity relates to rules used by commons participants, and how topics related to governance are reflected in public discourse. Together, my three empirical papers contribute to the overall theory on how global knowledge commons can be governed successfully, and suggest tangible governance strategies for the DNA barcoding community to facilitate achieving their goal of a globally representative DNA barcode commons.

Personal Perspectives

As much of my data analysis was conducted by me alone, it is necessary for me to outline my own personal training and perspectives that may have influenced the lens through which I viewed my data. My disciplinary training, content areas, and personal experiences have all shaped how I approach answering research questions, and producing results from data.

My undergraduate training was in microbiology, and the work experience I gained during this period and shortly after was in molecular genetics. I subsequently received graduate level training in public health, including global health and epidemiology. I have worked as part of a large interdisciplinary team of researchers conducting community-driven public health research with indigenous communities since 2007. My PhD training has been based in public health, although my coursework has included organizational theory, risk communication, and public policy. As my training has been based on different research paradigms, I do not identify with a single one, although my awareness that there *are* paradigms is typical of grounding in non-positivist research approaches. I believe that a person's epistemology need not be fixed, and the nature of what can be known is a function of what you are interested in learning about. However, I most closely align with a constructivist paradigm.

The content areas I have worked in have also informed my understanding of the value and purpose of scientific inquiry, as I have mainly worked in areas burdened by significant power imbalances. My undergraduate experience was in a research group that studied natural

immunity to HIV infection in a cohort of commercial sex workers in a slum in Nairobi, Kenya. My Master's research examined the potential to involved traditional medicine practitioners in expanding coverage of AIDS care programs in rural Uganda. Most noteworthy in my training is my experience in community-driven research in partnership with Northern Canadian First Nations and Inuit communities. These experiences provided the foundation for my understanding of how science can and should benefit non-academic partners. In all of these experiences I developed an appreciation for how the process of science and engaging with partners can shape its outcomes, and this piqued my interest in studying science policy to pursue not only equity in partnerships, but mechanisms for the best possible science in the context of power imbalances.

My personal experiences within academia have also informed my views on power and equity. I grew up in rural Manitoba, in a non-academic family. Neither of my parents had University degrees (my mother never completed high school), and I have observed a number of disadvantages for young researchers from similar situations. Since my undergraduate work experience, I have been increasingly aware of power and privilege within academia, and what I perceive to be a concentration of opportunities for those who have gained entry into the system, rather than for those who have demonstrated creativity or excellence. I am wary of the "status quo" in academia, and am interested in systems that reward excellence rather than privilege.

Finally, some may argue that my analysis could be biased from my friendly relationship with members of the DNA barcoding community. However, prolonged engagement with a community you are studying is a mechanism to enhance rigor in qualitative research, and there is increasing emphasis on engagement with knowledge users to improve the quality and applicability of research. Additionally, I have felt no pressure from the community to represent their efforts in any particular manner. Therefore, I do not believe my engagement has led to any bias in my analysis, but rather has enhanced it.

References

Andersson, K. (2006). Understanding Decentralized Forest Governance: An Application of the Institutional Analysis and Development Framework. *Sustainability: Science, Practice, & Policy*. 2(1): 25–35. http://search.proquest.com/openview/50b2eef538348d33635cad951e68368c/1?pq-origsite=gscholar&cbl=136137.

Behnke, R. (1994). Natural Resource Management in Pastoral Africa. *Development Policy Review*. 12(1): 5–28. http://dx.doi.org/10.1111/j.1467-7679.1994.tb00053.x.

Bromley, D. W. (1992). The Commons, Common Property, and Environmental Policy. *Environmental and Resource Economics*. 2(1): 1–17. http://dx.doi.org/10.1007/BF00324686.

Bubela, T., and R. E. Gold, eds. (2012). *Genetic Resources and Traditional Knowledge: Case Studies and Conflicting Interests*. Cheltenham, UK and Northampton, MA: Edward Elgar.

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.

Contreras, J. (2014). Constructing the Genome Commons. In: *Governing Knowledge Commons*. eds. B. M. Frischmann, M. J. Madison, and K. J. Strandburg, 99–136. Oxford University Press.

Convention on Biological Diversity. 2012. List of Parties. *Convention on Biological Diversity*. Accessed November 25. http://www.cbd.int/convention/parties/list/.

Crawford, S., and E. Ostrom. (2005). A Grammar of Institutions. In: *Understanding Institutional Diversity*. ed. E. Ostrom, 137–74. Princeton University Press.

Dedeurwaerdere, T. (2010). Global Microbial Commons: Institutional Challenges for the Global Exchange and Distribution of Microorganisms in the Life Sciences. *Research in Microbiology*. 161(6): 414–21. http://dx.doi.org/10.1016/j.resmic.2010.04.012.

Feeny, D., F. Berkes, B. J. McCay, and J. M. Acheson. (1990). The Tragedy of the Commons: Twenty-Two Years Later. *Human Ecology*. 18(1): 1–19. http://dx.doi.org/10.1007/BF00889070.

Freeman, R. B., and W. Huang. (2014). Collaboration: Strength in Diversity. *Nature*. 513(7518): 305. http://dx.doi.org/10.1038/513305a.

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., E. Camicioli, and T. Bubela. (2016). DNA Barcoding in the Media: Does Coverage of Cool Science Reflect Its Social Context. *Genome*. http://dx.doi.org/10.1139/gen-2015-0210.

Hardin, G. (1968). The Tragedy of the Commons. *Science*. 162(3859): 1243–48. http://dx.doi.org/10.1126/science.162.3859.1243.

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): 313–21. http://dx.doi.org/10.1098/rspb.2002.2218.

Hess, C. (2012). The Unfolding of the Knowledge Commons. *St Antony's International Review*. 8(1): 13–24. http://www.ingentaconnect.com/content/stair/stair/2012/00000008/00000001/art00003.

Hess, C., and E. Ostrom. (2006). A Framework for Analysing the Microbiological Commons. *International Social Science Journal*. 58(188): 335–349. http://dx.doi.org/10.1111/j.1468-2451.2006.00622.x.

Hess, C., and E. Ostrom. (2011). *Understanding Knowledge as a Commons: From Theory to Practice*. MIT Press.

Imperial, M. T., and T. Yandle. (2005). Taking Institutions Seriously: Using the IAD Framework to Analyze Fisheries Policy. *Society & Natural Resources*. 18(6): 493–509. http://dx.doi.org/10.1080/08941920590947922.

Kahan, D. M. (2002). Reciprocity, Collective Action, and Community Policing. *California Law Review*. 90(5): 1513–39. http://dx.doi.org/10.2307/3481363.

Kamau, E. C., B. Fedder, and G. Winter. (2010). The Nagoya Protocol on Access to Genetic Resources and Benefit Sharing: What Is New and What Are the Implications for Provider and User Countries and the Scientific Community? *Law and Environment and Development Journal*. 6(3): 246. http://www.lead-journal.org/content/10246.pdf.

Mayan, M. J. (2009). *Essentials of Qualitative Inquiry*. New York, NY: Routledge Taylor & Francis Group.

McCombs, M. (2004). The Agenda Setting Approach. In: *Handbook of Political Communication*. ed. D. Nimmo, 121–40. Newbury Park, CA: Sage.

McDougall, C. L. (1995). *Intellectual Property Rights and the Biodiversity Convention: The Impact of GATT*. Friends of the Earth Trust.

Melon, C. C., M. Ray, S. Chakkalackal, M. Li, J. E. Cooper, J. Chadder, W. Ke, et al. (2009). A Survey of South-North Health Biotech Collaboration. *Nature Biotechnology*. 27(3): 229–32. http://dx.doi.org/10.1038/nbt0309-229.

Myers, N., R. A. Mittermeier, C. G. Mittermeier, G. A. B. da Fonseca, and J. Kent. (2000). Biodiversity Hotspots for Conservation Priorities. *Nature*. 403(6772): 853–58. http://dx.doi.org/10.1038/35002501.

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.

Ostrom, E. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action*. 1st ed. Cambridge University Press.

Ostrom, E. (1992). Crafting Institutions for Self-Governing Irrigation Systems. http://agris.fao.org/agris-search/search/display.do?f=2012/OV/OV201205717005717.xml;US19920103636.

Ostrom, E. (2005a). *Understanding Institutional Diversity*. Princeton, New Jersey: Princeton University Press.

Ostrom, E. (2005b). Understanding the Diversity of Structured Human Interactions. In: *Understanding Institutional Diversity*.3–31. Princeton University Press.

Ostrom, E., J. Burger, C. B. Field, R. B. Norgaard, and D. Policansky. (1999). Revisiting the Commons: Local Lessons, Global Challenges. *Science*. 284(5412): 278–82. http://dx.doi.org/10.1126/science.284.5412.278.

Ostrom, E., and R. Gardner. (1993). Coping with Asymmetries in the Commons: Self-Governing Irrigation Systems Can Work. *The Journal of Economic Perspectives*. 7(4): 93–112. http://dx.doi.org/10.2307/2138503.

Padma, T. V. (2005). India's Drug Tests. *Nature*. 436(7050): 485–485. http://dx.doi.org/10.1038/436485a.

Poteete, A. R., M. A. Janssen, and E. Ostrom. (2010). Small-N Case Studies: Putting the Commons under a Magnifying Glass. In: *Working Together: Collective Action, the Commons, and Multiple Methods in Practice*.3–30. Princeton University Press.

Ray, M., A. S. Daar, P. A. Singer, and H. Thorsteinsdóttir. (2009). Globetrotting Firms: Canada's Health Biotechnology Collaborations with Developing Countries. *Nature Biotechnology*. 27(9): 806–14. http://dx.doi.org/10.1038/nbt0909-806.

Ruttan, L. M. (2006). Sociocultural Heterogeneity and the Commons. *Current Anthropology*. 47(5): 843–53. http://dx.doi.org/10.1086/507185.

Ruttan, L. M. (2008). Economic Heterogeneity and the Commons: Effects on Collective Action and Collective Goods Provisioning. *World Development*. 36(5): 969–85. http://dx.doi.org/10.1016/j.worlddev.2007.05.005.

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.

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.

Thorsteinsdottir, H., U. Quach, A. S. Daar, and P. A. Singer. (2004). Conclusions: Promoting Biotechnology Innovation in Developing Countries. *Nature Biotechnology*. 22: DC48-DC52. http://dx.doi.org/10.1038/nbt1204supp-DC48.

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.

Chapter 2: Global inequities and sharing genetic resources for non-commercial research: A case-study of the DNA barcode commons

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. I 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. My 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. My data sources included key informant interviews, organizational documents, scientific outputs of the DNA barcoding community, and DNA barcode record submissions. My research suggested that the goal of creating a globally inclusive DNA barcode commons has not yet been fully achieved, and that the risks and benefits of participating in the commons are not equitably shared across heterogeneous global participants. DNA barcode organizations can mitigate the challenges caused by its global membership through ensuring its governance is representative and considers restrictions on use that may enhance participation in the commons.

Introduction

Life sciences research that uses genetic resources can be enhanced through collective action to create and manage the resources and their supportive research infrastructure. Research infrastructure includes databases to store sequence and associated metadata, and bio-repositories to store biological samples. Collectively, data, metadata and biological samples comprise a

research commons – a set of resources managed according to terms that encourage efficiency, equitable use, and sustainability. Collective action by stakeholders to create and use research commons 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.

Building on a substantial body of research on institutional arrangements that govern natural resource commons, a new framework has emerged to analyse knowledge commons, including those used for research (Hess and Ostrom 2011a; Frischmann et al. 2014; Berge and Laerhoven 2011). The framework is based on the work of Nobel Laureate Elinor Ostrom, who developed the Institutional and Analysis and Development (IAD) Framework to enable the systematic study of the institutions that govern natural resource commons. Her work was the empirical response to the widely perpetuated, yet unsubstantiated, claim that depletion of resources was inevitable without government intervention or private property rights (Hardin 1968). Ostrom identified the attributes that facilitate the sustainable governance of shared resources by communities (Ostrom 1990).

The adapted IAD framework for knowledge commons (KC-IAD) accounts for differences between natural resource and knowledge commons. While natural resource commons often have innate boundaries, such as where a river flows, knowledge commons are generally boundless, making the exclusion of potential users difficult. Unlike natural resources, knowledge resources are non-rivalrous, meaning that their use by an individual need not interfere with use by another. Governance of natural resource commons focuses on sustaining an existing resource, while many knowledge commons are built to solve a particular problem. Therefore, governance of knowledge commons must encourage both the use and the creation of the resource. However, the burden of creating knowledge commons and potential benefits derived from their utilization are not necessarily equitably shared among participants in the commons (Bubela et al. 2012; Strandburg et al. 2014).

Analyses of knowledge commons reported to date have focused on those situated in high income countries, whose activities were based on shared infrastructure (Bubela et al. 2012; Contreras 2014; Dedeurwaerdere 2010c). Little research has focused on the establishment and ongoing governance of knowledge commons that engage a heterogeneous community, including

institutional and individual participants from disparate regions of the world with historical power and economic imbalances. In addition to economic, language, and cultural differences, wide geographic distribution of participants makes the communication to facilitate collective action difficult.

I address this research gap by examining the attributes of a global knowledge commons exemplar: The DNA barcode commons. The purpose of the DNA barcode commons is to facilitate: large-scale documentation of life on earth, and identification of unknown specimens in situations where standard taxonomic identification is not possible. DNA barcodes are short, standardized gene regions that can be used to inexpensively identify unknown specimens. Identification compares the barcode sequence of the unknown specimen 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. Two organizations have led the coordination efforts: The Consortium for the Barcode of Life (CBOL, founded in 2003 at the Smithsonian Institution in Washington, DC) and the International Barcode of Life Project (iBOL founded in 2009 at the University of Guelph in Ontario).

With coordination efforts, the DNA barcoding community has built or adapted existing supporting infrastructure, comprised of databases and biorepositories. The infrastructure houses DNA barcode sequences and associated metadata and linked reference biological specimens, respectively. Scientists involved in DNA barcoding efforts have collectively produced millions of barcode records (Ratnasingham 2015) and published thousands of scientific papers (Bubela et al 2015), both key indicators of a successful DNA barcode commons. However, the global DNA barcoding environment is characterized by an inequitable distribution of risks and benefits in the utilization of the genetic resources that comprise the commons. There is a history of mistrust between potential participants from diverse regions.

My 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. To answer these questions, I employed a case-study approach guided by the KC-IAD framework. In the next section, I outline the history of the DNA barcode commons. I then introduce the case-study approach, describing the key factors examined

in the case study, the specific data sources, and the methods used for data collection and analysis. I present the results of the analyses and discussion of the implications in the context of the framework.

History of the DNA barcode commons

In January of 2003, Paul Hebert and colleagues published a paper proposing DNA barcodes as a standardized global species identification system (Hebert et al. 2003). Shortly thereafter, in March 2003, a group of taxonomists, molecular biologists, and bioinformaticians held an exploratory workshop to discuss the initiation of large-scale DNA barcoding efforts. The participants of the workshop discussed the potential benefits to society of adopting barcodes as a standardized identification system and its application beyond taxonomy and bioconservation (Stoeckle 2003). By December 2003, proponents had developed standards to build a DNA barcode database, including the linkage of barcode sequences to voucher specimens. Proponents worked to overcome opposition to using DNA barcodes as a tool for species identification, promote the system, and build a global network of DNA barcoders (Stoeckle 2003). The first formal organization created for DNA barcoding was the Consortium for the Barcode of Life (CBOL), founded in 2004 at the Smithsonian in Washington, DC. The first International Barcode of Life Conference was organized by CBOL and held at the Natural History Museum in London in 2005 (Consortium for the Barcode of Life 2015b). It later became a biennial event.

DNA barcoding quickly gained global momentum because it enabled a range of practical applications (Hebert et al. 2003). An open access, comprehensive database of DNA barcodes linked to metadata and a voucher specimen facilitated rapid identification of unknown specimens in situations where morphological identification was impossible. Such situations arise, *inter alia*, where traditional taxonomic expertise is unavailable or the specimen lacks distinguishing features, such as butchered meat or insect larvae. Proponents argued that a robust DNA barcode database would enable rapid identification of unknown species to meet the needs of regulatory agencies, including trafficking of endangered species and mislabelling or illegally importing fish or agricultural pests. The system that proponents envisaged included the shipment of unknown specimens to a laboratory equipped to produce low-cost DNA barcodes, which could then be

matched against known barcodes (Pennisi 2003). To generate DNA barcodes *in situ,* some barcoders called for the future development of a handheld device (Stoeckle and Hebert 2008).

Paul Hebert led an international initiative to build a comprehensive barcode reference database. Canadian funders supported the development of the necessary infrastructure, 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 and was funded through Genome Canada's International Consortium Initiative. iBOL included 26 nations as 'nodes' partnered through formal agreements (iBOL 2015d). The main mission of iBOL was to build an openly-accessible database including 5,000,000 barcodes, representing 500,000 species by 2015 (iBOL 2015i). Its activities were supported by five working groups: Barcode Database; Methods and Technologies; Informatics; Coordination; and GE^{3}LS (a term used by Genome Canada to refer to the ethical, economic, environmental, legal and social aspects of genomics research). All of the DNA barcode records produced through iBOL infrastructure at the University of Guelph are housed in BOLD, although any individual can use BOLD. The online system allows individuals who produced barcode records (either in their own labs or through central sequencing facilities) to work with the sequences on a private "workbench", and then later publish the sequence to the open access database (accessible through BOLD's public data portal). All records published on BOLD are copied to GenBank. The project has resulted in many "barcode of life" (BOL) sub-projects, focused on countries (MexBOL) (Escalante et al. 2010), topics (PolarBOL), or taxonomic groups (TreeBOL) (Jinbo et al. 2011).

During the 6th International Barcode of Life Conference in August 2015, barcode community members participated in a workshop to establish the International Society for the Barcode of Life (ISBOL). The Society would "coordinate completion of the [barcode] registry, to facilitate the development of barcode applications and to communicate with stakeholders at all levels" (Castle et al. 2015). Membership was automatically granted to all registrants of the conference, but left open to all interested parties. A governance council to initiate ISBOL, comprised of 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, was

proposed as an interim measure. The council is seeking feedback on proposed structure and governance from the broader DNA barcoding community.

Methods:

Case study approach

I used multiple methods in a case-study approach to analyse how the factors outlined in the IAD framework influence governance of the DNA barcode commons. Case study methodology is considered one of the best research methods available for explanatory answers to how questions about contemporary events (Yin 2009); it is frequently employed by scholars to study commons (Poteete et al. 2010c). Here I define a case study as an in-depth study of a relatively well-bounded phenomenon (Poteete et al. 2010c). One of the strengths of this approach is to enable an understanding of *how* factors in a commons might inhibit or promote effective governance and collective action. Such understanding facilitates recommendations of practical changes to commons governance to improve outcomes (Poteete et al. 2010c).

I used a variety of data sources and analytical methods as part of my case-study approach. The majority of the analyses focused on a document and literature search and key informant interviews. I used these data sources to identify facts about how the DNA barcode commons was established and is run, and gather diverse stakeholder perspectives. I supplemented this analysis with a bibliometric analysis of barcoding publications, and an analysis of barcode record submissions to BOLD. Bibliometric analysis provided insight into collaborations within the DNA barcode community, as illustrated through co-authorship, and approximated scientific output related to DNA barcode projects. Barcode record submission analysis provided insight into where DNA barcode record specimens were collected and stored, and which regions were contributing to the DNA barcode commons.

The KC-IAD framework

I used the KC-IAD framework to guide my analysis of the DNA barcode commons, as illustrated in Figure 1. This framework provided an *apriori* analytical frame that I used to identify relevant data in analysis. The framework demonstrates how the components (boxes) and

interactions (arrows) take place within the context of a background environment. Because knowledge commons are often purpose-built to solve a particular problem, understanding the history of the commons as well as its goals and objectives is central to an analysis of the social dilemmas faced by participating actors (such as whether or not to contribute to shared resources that others may use without contributing). Such understanding also enables an evaluation of the outcomes of the knowledge commons, which are related to the objectives of its actors and the dilemmas that need to be solved through collective action.

Central in the KC-IAD are action arenas, where "actors interact [e.g. , exchange goods and services, solve problems] as they are affected by exogenous variables" (Ostrom 2005b, 13). The exogenous variables are resource characteristics, attributes of the community, and governance. The Figure 1 arrows between exogenous variables demonstrate the complex relationship between them. The products of action arenas are patterns of interactions, which are outcomes themselves in a knowledge commons, as they are "inextricably linked with and determinative of the form and content of the knowledge or information output of the commons" (Frischmann et al. 2014, 19). For example, in the DNA barcode commons, the patterns of interactions can be observed in expected outputs of a research commons, such as co-authorship on barcode publications, and identification of the geographic origins of barcode records. The evaluative criteria of the patterns of interactions are based on the goals and objectives of the commons and are used to determine how the exogenous variables and behaviours in action arenas might eventually change and respond over time. This feedback is illustrated in Figure 1 by dashed arrows and indicates that both the experiences of actors in the action arenas and the patterns of interactions impact the exogenous variables of a knowledge commons.

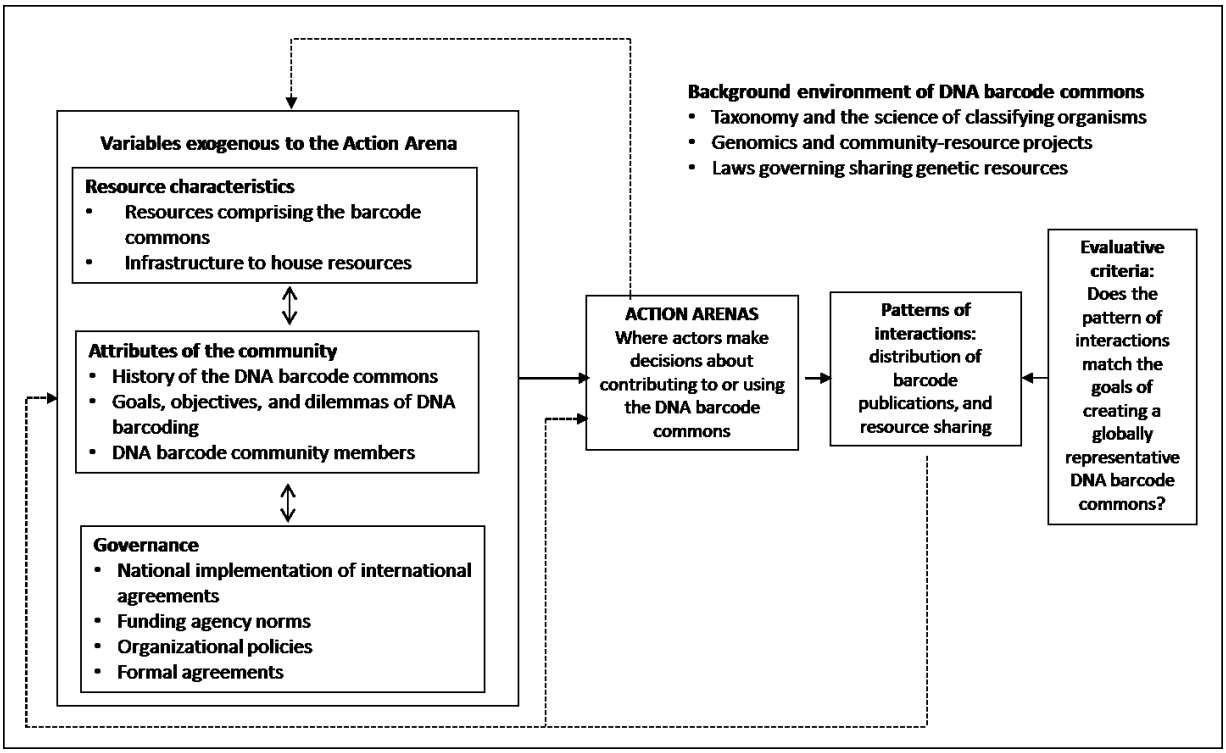

Figure 1. The DNA barcode commons described within the KC-IAD Framework (Frischmann et al. 2014).

Data collection and analysis

Document search and analysis

I 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 2015b), including: iBOL Node Memorandum of Understanding (MOU); iBOL Node MOU Appendix; Data Standards for Barcode Records; the Banbury Report on Taxonomy, DNA, and the Barcode of Life; and Guidelines to Authors of Barcode Data Release Papers. I obtained additional documents directly from iBOL project staff during a visit to the Canadian Centre for DNA Barcoding in 2012, including: the biological materials transfer agreement; iBOL data release policy; and microplate and data submission instructions. I reviewed key publications that detailed: the science of DNA barcoding (Hebert et al. 2003); controversies about the science (Moritz and Cicero 2004; Dupuis et al. 2012; Collins and Cruickshank 2013; Gregory 2005); the international efforts (Vernooy et al. 2010; Schindel et

al. 2008; Schindel 2010; Schindel et al. 2015); potential applications (Yancy et al. 2008; Gross 2012; Wong and Hanner 2008); organizational efforts of DNA barcoding proponents (Adamowicz 2015; Castle et al. 2015); and database-building efforts (Ratnasingham and Hebert 2007; Sonet et al. 2013).

Key Informant Interviews

I and two research assistants interviewed 50 key informants, including 35 individuals who participated in DNA barcoding projects, three policy makers involved in funding and oversight of DNA barcoding projects, and 12 individuals involved in genetic resource governance. This research received ethical approval from the University of Alberta Research Ethics Board – Health Panel (Appendix 7). From 14 countries, I selected interviewees who had extensive knowledge and expertise on the research question (Higginbottom 2004). I interviewed 25 of the 50 interviewees in person at two iBOL conferences, held in 2011 (Adelaide, Australia) and 2013 (Kunming, China). The remainder were interviewed over the phone (n=8) or in-person (n=17) by myself or a research assistant. I and research assistants used a semi-structured interview guide (Appendix 1) developed based on previous knowledge of the subject matter, with questions related to: research partnerships; sources of genetic resources; the impact of the *Convention on Biological Diversity*; and the feasibility and structure of genetic resource commons. Early interviewees informed minor adjustments to the interview guide. Interviews lasted between 30 and 90 minutes; the interviewer audio recorded each interview. I used a paid service to transcribe each interview verbatim to facilitate systematic analysis using NVivo 11 (NVivo qualitative data analysis Software; QSR International Pty Ltd. Version 11, 2014).

I analyzed the interviews using the KC-IAD framework to guide a content analysis, defined as "a systematic coding and categorizing approach used for exploring large amounts of textual information unobtrusively to determine trends and patterns of words used, their frequency, their relationships, and the structures and discourses of communication" (Vaismoradi et al. 2013, 400). First, I listened to each interview to verify the transcription was correct. Simultaneously, I made notes about central concepts that were shared by interviewees to iteratively inform subsequent data collection and analysis. I then coded the interview transcripts in NVivo 11 based on broad categories from the KC-IAD framework, including: governance structures; how the interviewee made decisions in different action arenas; and views on

participation in a global knowledge commons. Within these broad categories, I assigned detailed descriptive codes to each statement. I then grouped codes to form themes, and separately examined the themes within and between each group of interviewees. When reporting direct quotes, I edited the quotes to fix grammatical errors and improve clarity.

I grouped interviewees based on their main work affiliation at the time of the interview, as being from: Like-Minded Mega Diverse countries (LMMC) and non-LMMC. LMMC is a group of countries established in 2002 to promote their similar interests in protecting biodiversity (South Africa Department of Environmental Affairs, 2016). The LMMC group included Brazil, China, Colombia, Ghana, India, Indonesia, Kenya, Mexico, and South Africa, and the non-LMMC group included Australia, Canada, New Zealand, United Kingdom (UK), and the United States (US). Despite not being a member of the official LMMC group, I included Ghana in the LMMC group because its interests aligned with other African countries, such as Kenya and South Africa. In cases where identifying the country would risk identifying an individual participant, I have referenced the individual's region.

Bibliometric Analysis

I searched the Scopus database for peer-reviewed literature that referenced any of four seminal barcode papers (Hebert et al. 2003; Hollingsworth et al. 2009; Schoch et al. 2012; Stoeckle and Hebert 2008). The search, conducted in January 2015, yielded 3557 scientific journal publications from 2003-2014. I, and a research assistant, compiled a database of information about each article including: publication source; publication year; number of citations; and author names and institutional affiliations. A research assistant trained in programming used a customized computer program to combine synonymous names of single individuals and separate identical names of different individuals (the process is called author name disambiguation and I have detailed the methods in Appendix 2).

From this database I generated variables for presence of authors with institutional affiliations in high, middle, or low income countries (World Bank 2016); presence of authors from mega-biodiverse countries (Mittermeier 1997); and whether or not the paper was published in a highly ranked journal. The large sample size allowed me to use the four-category Gross National Income (GNI) per capita levels from the World Bank (upper income, >=$12,476; upper middle income, $4,036 to $12,475; lower middle income $1,026 to $4,035; and low income,

≤$1,025) rather than the dichotomous categorization I used for the qualitative analysis (World Bank 2016). I used InCites Journal Citation Reports (Thomson Reuters 2016) to identify the top 10 ranking journals in each field category relevant to DNA barcoding. I 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.

I, along with a research assistant, also created a geographic map of co-authorships. Using Gephi 0.8.2 Beta (Bastian et al. 2009), a research assistant familiar with Gephi geographically displayed the location of the primary affiliation of each author in the database, and created links between authors based on co-authorship on a single paper. I described the author sets (the set of authors on a single paper) based on the proportions of papers that span across different geographic regions.

Barcode record submissions to BOLD

I examined patterns of sharing barcode records and biological specimens across different country income levels (World Bank 2016) using two exemplars: (1) barcodes of medicinal plants and (2) 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).

I 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 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 what country specimens originate from. 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 (GPS (global positioning system) coordinates and/or country) and stored (institution name), and other metadata like time of collection. The public data portal also mines barcode gene region sequences from GenBank.

I downloaded plant records from the BOLD public data portal in January 2013 (150,220 records). Two research assistants assisted me in creating 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 1,300 known medicinal plant species names (obtained from http://www.ars-grin.gov/duke/ethnobot.html). I was unable to search each of the 1,300 plants in the taxonomy browser, so I did not estimate unpublished medicinal plant records. Of the identified public medicinal plant records, 5,788 had the latitude and longitude where the specimen was collected, and an additional 8,151 included specimen's country of origin. Of 3,956 records without any specimen collection information, 2,036 were mined from GenBank. I created a variable that indicates if the specimen was stored in the country where it was collected. I then used SPSS v.19 to tabulate the published medical plant records separately by the country income level, which allowed me to determine the proportion of materials that are stored outside of the country of origin for different country income levels.

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

Engagement with the DNA barcoding community

I actively sought input and feedback from the DNA barcoding community beyond their participation in formal interviews. I 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. I shared interview guides with barcode leaders and organizational administrators and invited feedback to ensure that questions were relevant to the barcoding

community. I, along with my supervisor, presented preliminary findings at three international DNA barcoding conferences (Bubela 2013; Bubela et al 2015; Geary et al. 2016; Geary and Bubela 2015) and invited feedback from conference attendees and interviewees from my study. I participated in a workshop in February 2013 that discussed medicinal plant barcoding and issues related to sharing genetic resources. The workshop resulted in a publication with leaders in the barcoding community (Schindel et al. 2015).

Results and Discussion

In this section, I present the results of my analyses using the KC-IAD framework (Figure 1). First, I detail the background environment including taxonomy, other influential genomics initiatives, and relevant laws. Next, I describe the resources and infrastructure that comprise the DNA barcode commons, and the attributes of the community, including the goals and dilemmas. I then describe governance structures, before moving on to detail interviewee experiences in three types of actions arenas in the commons: sharing genetic materials; generating and sharing DNA barcode records; and accessing data in the DNA barcode commons. Finally, I examine the patterns of interactions including evidence for 1) collaborations; 2) producing scientific outputs; and 3) sharing genetic resources.

Background environment of the DNA barcode commons

Taxonomy and the science of biological identifications through DNA barcodes

For most of its history, taxonomists (scientists who classify organisms) differentiated species based on morphological distinctions, which is slow and requires specialized expertise. Only 10% of an estimated 10-20 million species have been described over the last 250 years (Wilson 2003), and the discipline of taxonomy has been in decline as specialists have retired and have not been replaced due to shifting research priorities (Waterton et al. 2013). One group estimated that 15,000 trained taxonomists would be needed in perpetuity to effectively identify the world's biodiversity using traditional methods (Hebert et al. 2003).The field has been criticized for its focus on particular groups (such as vertebrates and flowering plants) over ecologically more important species (such as nematodes or mites) and a literature base that is difficult to access (Tautz et al. 2003).

Nevertheless, the documentation of global biodiversity is critical for the mitigation of anthropogenic and other threats, including climate change and habitat destruction (Hebert et al. 2003). As a result, in the early 2000s, the field of taxonomy experienced a resurgence with the advance of DNA sequencing technology and bioinformatics infrastructure (Waterton et al. 2013). Modernization of the discipline included digitizing and standardizing taxonomic knowledge to make it more accessible (Godfray 2002) and the adoption of molecular methods. Such DNA-based taxonomy is less dependent on human resources than traditional taxonomy, and offers the potential to accelerate and lower the cost of taxonomic identification (Hebert et al. 2003). However, proposals to expand DNA-methods for taxonomy were met with early resistance (Tautz et al. 2003), because the use of multiple non-standardized gene regions to differentiate species prevented automated analyses at the scale needed to document Earth's biodiversity (Moritz and Cicero 2004).

In 2003, Paul Hebert proposed DNA barcodes as a solution to the issues of scalability and standardization. DNA barcodes are short and ubiquitous gene sequences that (1) have high inter-species variability combined with low intra-species variability (to produce an identifiable barcode 'gap' between species), and (2) are sufficiently conserved to allow the use of universal sequencing primers within target taxonomic groups (Dentinger et al. 2011). The cytochrome c oxidase I (COI) mitochondrial gene was the first barcode proposed and has become the universal DNA barcode for animals (Hebert et al. 2003). Subsequent work has identified DNA barcodes for plants (Hollingsworth et al. 2009) and fungi (Schoch et al. 2012).

However, DNA barcoding was not accepted by taxonomists without controversy and significant scientific debate. As one researcher summarized in a publication, "some taxonomists view DNA barcoding as an Orwellian nightmare – In their eyes this would essentially kill the science of taxonomic research" (Smith 2005, 842). Opponents were concerned that focusing on fast identifications using standardized genes would divert funds from traditional taxonomy (Ebach and Holdrege 2005), overly simplify and incorrectly characterize the complex relationship between genes and species (Dupuis et al. 2012), and result in incorrect species identifications (Will and Rubinoff 2004). Despite these critiques, DNA barcoding has nevertheless gained prominence as a taxonomic tool, and one barcode database included 4,950,751 barcode records (published and unpublished) as of May 2016 (BOLD Systems 2015).

GenBank, the globally recognized open access repository for genetic sequences, contained 1,478,701 sequences labeled as "barcodes" as of August 2016.

Knowledge Commons for Genomics Data

The DNA barcoding initiative is similar to the Human Genome Project (HGP), which began in 1990 (Collins et al. 2003), in that it represents a coordinated effort among participants to use high-throughput methods to generate sequence databases as one aim. Such knowledge commons comprise "massive collections of [genomics] data stored in electronic databases across the world and made available through public networks" (Contreras 2014, 102).

The HGP was influential in the development of data sharing norms and rules within communities of genomics researchers. Starting in 1991, the International Human Genome Sequencing Consortium (NHGRI) published rapid data release standards for the HGP. The principles of rapid data release were reconfirmed in the 1996 Bermuda Accord (Bermuda Sequence Policies Archive 2016). In 2003, genomics community leaders convened in Fort Lauderdale to discuss updated standards (NHGRI 2003) based on the assumption that rapid and free release of genomics sequence data would promote the best interests of both science and the public. The principles were expanded to apply beyond major sequencing centres and funding agencies at a meeting in Toronto in 2009 (Toronto International Data Release Workshop Authors 2009). In general, the principles set out that sequencing data should be released immediately (within 24 hours in some instances) to an open access database. However, the principles also include the standard that researchers should refrain from disrespecting legitimate interests of the sequence producers in publications based on their data, for example, by suggesting a delay in publication of secondary analyses of the data.

Our case study points to a notable omission from current genomics data release principles, namely, an explicit 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 to date because few lower income countries have sequencing centres or access to the associated research and development funding (Helmy et al. 2016). For example, eighty percent of the 71 authors of the Toronto principles were

from just three countries (US, Canada, and the UK), with another 16% from European countries and 4% from Japan and China (Toronto International Data Release Workshop Authors 2009).

Furthermore, the genomics databases are located in high income countries. The Los Alamos National Laboratory in the US initiated local genomics data infrastructure in the late 1970s (Strasser 2008; Benson et al. 2013). The National Institutes of Health provided funding to create publicly available data infrastructure, and the Los Alamos database became GenBank in 1982. In the late 1980s, the National Center for Biotechnology Information (NCBI) was formed and absorbed the management of GenBank. NCBI is part of the International Nucleotide Sequence Database Collaboration (INSCD) 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).

Laws governing the sharing and utilization of genetic resources

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. Genetic resources refer to both biological samples and derived sequence data. However, several interviewees directly involved in various BOL activities were unaware of the legal instruments or did not believe in their applicability. Here I discuss these instruments at the international level, and I will focus on their national implementation further in this paper.

At the international level, the sharing of genetic resources 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: 1) conservation of biological diversity; 2) sustainable use of its components; and 3) 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 largely 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 to genetic resources. To protect legitimate interests in how they govern access to their

biodiversity, a group of 17 lower income countries with large amounts of biodiversity came together in 2002 to form the group of LMMC (LMMC 2002). This group submits shared statements to the *CBD* Working Group and Intergovernmental Committees to promote sovereign rights over genetic resources and express concerns regarding the limitations of the international agreements.

Despite the limited reach of the *Nagoya Protocol*, researchers in high income countries were concerned that it would negatively impact non-commercial biodiversity research in support of 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, in partnership with the German Research Foundation, hosted a workshop to address the challenges that an overly-restrictive agreement on access and benefits sharing of genetic resources would create for non-commercial research, such as BOL research. The group put forward a joint 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). The *Nagoya Protocol* was adopted at COP10, and included an article to require that parties "create conditions to promote and encourage research which contributes to the conservation and sustainable use of biological diversity, particularly in developing countries, including through simplified measures on access for non-commercial research purposes" (United Nations 2010).

Patterns of national implementation of the *CBD* and *Nagoya Protocol* pose additional challenges for research using genetic resources. 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* (Table 1). 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 would if the US was a party. Since entering into force in 2014, the *Nagoya Protocol* has 76 parties (August 2016), although most of the ratifications are by lower income countries; 63 of the 76 are countries that have a gross national income per capita of less than $12,736 (World Bank 2016). Because few

high income countries are parties to this agreement, it is difficult to enforce unauthorized utilization of genetic resources in many high income countries. Six of thirteen countries with formal participation in iBOL activities are not party to the *Nagoya Protocol* (Table 1) and do not have legislation that governs genetic resources or ABS.

Individuals and institutions interested in participating in genetic resource exchange across national borders contend with additional rules and oversight, often enshrined in materials and data transfer agreements. Such agreements stipulate conditions for how the shared resources may be used and in what cases, if any, the resources may be transferred to third parties. Often, individuals wishing to send genetic resources must show relevant national authorities proof of materials transfer agreements (MTAs) or data transfer agreements before receiving export permission.

Resources and infrastructure that comprise the DNA barcode commons

The DNA Barcode Production Pipeline

The process for creating a barcode record begins with a specimen whose taxonomic identification is known (Figure 2). Specimens can be derived from existing collections, or collected from the field and stored as a voucher specimen for the DNA barcode record (Figure 2 Processes A and E). Only a small sample of the specimen is needed to extract DNA (Figure 2, Process B), or the whole specimen may be used if the organism is small (e.g., an insect). The barcode gene region(s) is then amplified using polymerase chain reaction (PCR). The region amplified for animals is a region of the COI gene in the mitochondrial genome (Hebert et al. 2003); plants require at least two regions from the plastid genome including a sub-unit of ribulose 1, 5 bisphosphate carboxylase/oxygenas (rbcL) and maturase K (matK) (Hollingsworth et al. 2009); and the region amplified for fungal barcodes is the nuclear ribosomal internal transcribed spacer (ITS) (Schoch et al. 2012). PCR reactions require short primers, which are conserved sequences that flank the target barcode region, and these are based on the broad taxonomic group of the specimen (e.g., land plants or mammals) (Figure 2, Process C).

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. Individuals with access to PCR equipment can create the PCR product to ship to sequencing

Table 1: Characteristics of countries with formal iBOL participation

Country	ISSC^	CBD	Nagoya Protocol	Genetic Resources or associated TK^^ legislation (WIPO 2016)	ABS legislation (United Nations 2010)	# of Records in BOLD	% with species names	% mined from GenBank
High income (≥$12,476**)								
Canada	Yes	Party	Non-party	No	No	1,757,321	32.3	0.2
United States	Yes	Non-party	Non-party	No	No	349,017	30.1	10.7
Australia	No	Party	Non-party	No	No	204,066	32.2	7.1
Germany	Yes	Party	Party	Yes	Yes	137,062	36	4.7
Argentina	Yes	Party	Non-party	No	No	94,861	5.8	2.3
Norway	Yes	Party	Party	Yes	Yes	53,655	17.5	4.7
Finland	Yes	Party	Party	No	Yes	44,323	44.3	6.2
New Zealand	No	Party	Non-party	Yes	No	32,000	26	21.2
Russia	No	Party	Non-party	No	No	30,567	40.4	21.9
France	No	Party	Party	No	Yes	26,829	70.4	25.6
United Kingdom	Yes	Party	Party	No	Yes	16,401	83.8	30.4
Saudi Arabia	No	Party	Non-party	No	No	8,823	8.6	5
Portugal	No	Party	Non-party	Yes	No	7,926	85.1	56.2
Korea	No	Party	Non-party	No	No	7,765	84.7	77.4
Switzerland	Yes	Party	Party	Yes	Yes	5,104	68.6	35.6
Netherlands	No	Party	Party	No	Yes	4,514	83.7	27.7
Upper Middle income ($4,036 to $12,475)								
Costa Rica*	Yes	Party	Non-party	Yes	No	382,387	47.8	0.7
South Africa*	Yes	Party	Party	Yes	Yes	69,352	19.6	7.2
China*	Yes	Party	Non-party	Yes	No	69,278	47.3	33.3
Mexico	Yes	Party	Party	No	Yes	63,031	29.1	10.9
Brazil*	Yes	Party	Non-party	Yes	Yes	33,667	64.5	25.7
Panama	No	Party	Party	Yes	No	18,498	40	13.4

Country	ISSC^	CBD	Nagoya Protocol	Genetic Resources or associated TK^^ legislation (WIPO 2016)	ABS legislation (United Nations 2010)	# of Records in BOLD	% with species names	% mined from GenBank
Peru	No	Party	Party	Yes	Yes	17,543	25.5	20.4
Colombia*	No	Party	Non-party	No	No	5,210	63.2	32
Lower middle income ($1,026 to $4,035)								
Pakistan	No	Party	Party	No	No	39,050	12.1	1.2
Kenya*	No	Party	Party	No	Yes	26,915	11.3	5.9
Papua New Guinea	No	Party	Non-party	No	No	25,589	27.1	8.2
India*	No	Party	Party	Yes	Yes	15,519	88	53.6
Low income (≤$1,025)								
Madagascar*	No	Party	Party	No	No	48,805	35.2	5.1

^International Scientific Steering Committee of the International Barcode of Life project

^^Traditional Knowledge

* Member of the group of Like-Minded Mega-diverse countries

** Income levels from the World Bank Country classification of Gross National Income per capita (https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups)

facilities. Unlike whole specimens or extracted DNA, PCR products only contain the small, amplified, barcode region, not the entire genome of the specimen.

The PCR product is sequenced in a DNA sequencer (e.g., 3730xl DNA analyzer) (Figure 2, Process C) to produce the DNA barcode, which comprises a string of ordered nucleotides (adenine, cytosine, guanine, and thymine). The COI barcode, for example, is a string of 648 nucleotides, which can be represented simply as text with letters representing each nucleotide in the sequence (ACGCTGTAC…etc). Once the DNA barcode sequence is quality controlled (Figure 2, Process F) and linked to its metadata (information that describes the data), it becomes the barcode record (Figure 2, Process G). 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 additionally include photographs of the specimens. Barcode records enable scientific, curiosity-based, and regulatory uses by others (Figure 2, Process H).

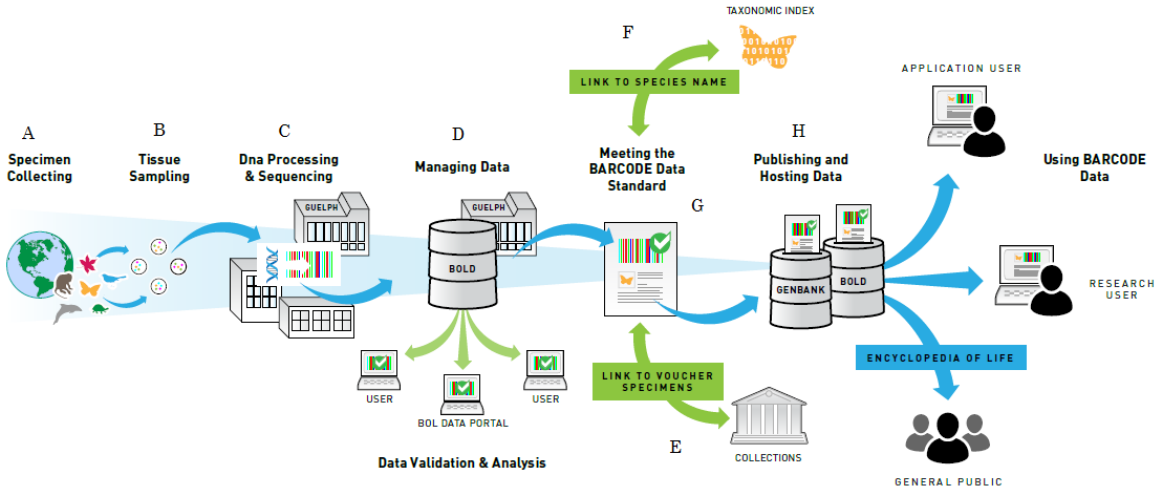

Figure 2. The DNA Barcode Pipeline (pipeline image reproduced from CBOL (Consortium for the Barcode of Life 2015b))

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

Infrastructure to house DNA barcode resources

The DNA barcode commons requires infrastructure to enable (a) large-scale production of barcode records, (b) storage of specimens and data, (c) access for use of the resources; and (d) value-added re-contribution of biomaterials, data and metadata to the commons.

Since 2003, barcoders have used existing data infrastructure, namely the International Nucleotide Sequence Database Collaboration (INSDC), 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. At that time, researchers estimated that the size of the barcode database would quickly surpass the approximately 52 million records on GenBank, necessitating the development of independent data infrastructure. In addition, the database would require specialized informatics to handle data curation throughout the analytical pathway and record storage (Ratnasingham and Hebert 2007). In 2007, the Canadian Centre for DNA Barcoding launched BOLD. While BOLD is housed within an Ontario Institute, it has grown from 102 users at its inception in 2005 to over 14,000 users from 94 countries in 2015 (Ratnasingham 2015).

BOLD is now established as the main curator of barcode records (5,248,177 as of Nov 2016, (BOLD Systems 2015)), and it includes open access and private data. Interviewees from non-LMMC and LMMC 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 also view raw sequence outputs and metadata. 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 directly with other platforms, and sequences published within BOLD are automatically submitted to GenBank. Barcode records, once published, are not subject to any restrictions on their use.

The use of standardized regions allows DNA barcode sequences to be generated via a cost-effective, high-throughput production pipeline. To facilitate the pipeline, the iBOL project promoted the establishment of core sequencing facilities (iBOL 2012), the largest of which, the CCBD, produces 600,000 barcode sequences each year (Centre for Biodiversity Genomics 2016).

In addition to barcode sequencing facilities, infrastructure is also needed to house at least one voucher specimen for each unique barcode record. Specimens must be taxonomically identified and 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 operating units to manage collections of the voucher specimens it receives.

Attributes of the DNA Barcoding Community

Goals and dilemmas of the DNA barcode commons

The goals of the DNA barcode commons are to: speed up the process of documenting global biodiversity; facilitate monitoring; and enable a broad array of applications based on an open access globally-representative DNA barcode record database (iBOL 2015i). 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 (knowledge spillover), and re-contribute value-added data (network effect) (Bubela et al. 2012; Schofield et al. 2009; Dedeurwaerdere 2010a). Individuals may stop contributing to the commons if they feel others are utilizing the resource without contributing to it ("free-riding") (Dedeurwaerdere 2010c). Similarly, scientists fear that others will receive first credit for analysis and publication of their data ("being scooped"). Concerns over publication priority are common in many scientific disciplines (Contreras 2010; Joly et al. 2012) and were frequently cited by interviewees. LMMC interviewees preferred that data release be delayed until after publication. One interviewee described the extent to which a colleague was concerned 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. In contrast, other knowledge commons for genomics research can feasibly be developed and maintained by non-LMMC participants alone (Dedeurwaerdere et al. 2013; Bubela et al. 2012; Collins et al. 2003). However, DNA barcode reference libraries cannot represent all life on earth without the contributions of global participants. This sentiment was expressed by a researcher from South Africa:

> *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 concerns LMMCs have about genetic resource misappropriation. DNA barcode records comprise data, metadata and biological specimens, all of which are genetic resources and therefore governed by the associated legal frameworks for their access and utilization. Those frameworks protect against misappropriation of DNA barcode records for commercial research, although some LMMC interviewees explained that lack of trust was still a significant barrier to shipping specimens 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 entirely:

> *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 also presents issues for non-commercial research, as 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, and

non-LMMC participants are more likely to have access to the financial resources and domestic infrastructure required to develop, maintain, and use the barcode commons. This distribution of resources also leads to a much broader free-rider dilemma, where LMMCs might feel that the entire barcoding effort is free-riding on their biodiversity: it requires LMMC genetic resources without necessarily providing corresponding benefits.

While scholars have demonstrated that participants in a commons can develop the necessary trust that promotes collective action, a key driver of this is face-to-face communication (Ostrom 2003), which is hampered by distance. Because of this, a significant dilemma remains as to how to govern this global knowledge commons to ensure that participants who contribute in different manners (i.e., research infrastructure vs biological specimens) are able to receive equitable benefits from the project and share the burden of risks equitably as well.

Community Members

The success of the DNA barcode commons relies on the active participation of a heterogeneous and global set of partners, some based in countries with limited financial resources for research and development. Here, I compile data on the groups of actors who contribute to and use the DNA barcode commons.

Community Leaders

Many interviewees spoke about key individuals who influenced the field of DNA barcoding and the development of the commons. Paul Hebert led initiatives to obtain funding for barcode infrastructure in Guelph (iBOL newsletter, Dec 2003). Scott Miller and David Schindel led the Consortium for the Barcode of Life, which shaped DNA barcoding policies to create global standards for barcode records (Consortium for the Barcode of Life 2015b).

Community leaders also had influence beyond the barcode commons. David Schindel advocated to the *CBD* during negotiations for the *Nagoya Protocol*, arguing for simplified measures for non-commercial research conducted for the purpose of bioconservation (Schindel 2010). He also promoted using standard agreements to facilitate ABS for non-commercial research to engage LMMC countries in barcoding efforts (Schindel et al. 2015; Vernooy et al. 2010). One of the most important roles of community leaders is to adopt a new technology, and prominent biologists like Dan Janzen and Winnie Hallwachs using DNA barcodes in 2003 had a

large impact on how the system became accepted throughout the scientific community (Janzen 2004; Burns et al. 2008).

DNA barcode record contributors

The majority of contributors to DNA barcode records are researchers (including taxonomists, ecologists, evolutionary biologists, systematists, and bio-informaticians), who work at universities or other research-intensive institutions, including museums and herbariums. As of 2015, BOLD had 14,000 users from 94 countries (Ratnasingham 2015). Contributions are broad and can include: adding data or specimens to the commons; developing quality-control measures; refining methods for producing or utilizing barcodes (Meusnier et al. 2008); and studying the utility of barcodes.

Lay participants can contribute to the commons. For example, these participants can suggest changes to taxonomic identifiers and highlight errors in the dataset. The LifeScanner program allows individuals to collect specimens for DNA barcoding (including whole specimens or tissue samples) and to receive information about the specimen (Biodiversity Institute of Ontario 2015). The resulting barcode records are then deposited into an open access database.

DNA barcode record users

In addition to researchers, users include other individuals who work for agencies reliant on species identification to identify unknown specimens, such as regulatory agencies (Yancy et al. 2008) or border control services (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 from both non-LMMC and LMMC emphasized that DNA barcode data and related taxonomic information should be openly available online at no cost to enhance public knowledge of biodiversity.

Database management

As the requirement to publish sequence data with scientific publications predates DNA barcoding efforts and BOLD, DNA barcoders initially deposited their sequence data into genomics repositories like GenBank. GenBank is run by the NCBI, which is part of INSDC

(with the European Nucleotide Archive (ENA) and the DNA Databank of Japan (DDBJ)). INSDC policy provides open access to all its records (Nakamura et al. 2013).

BOLD was developed to provide "an integrated bioinformatics platform that supports all phases of the analytical pathway from specimen collection to tightly validated barcode [database]" (Ratnasingham and Hebert 2007, 356). It is now the main curator of barcode records, both published and unpublished. BOLD provides a bioinformatics workbench for researchers to upload and work with their own privately held sequences in aggregation with the open access database. Bioinformaticians can easily download compilations of barcode records from the open access database in several formats enabling statistical comparisons (.xml, .tsv) and phylogenetic analyses (FASTA, TRACE).

The BOLD data policies initially stipulated that a complete barcode record would include GPS coordinates (Ratnasingham and Hebert 2007). However, interviewees felt that sharing specific GPS coordinates enabled unauthorized collection of specimens. The data standards for barcode records suggest sharing of GPS coordinates, but do not require it (Hanner 2009). The minimum standards for submitting barcode records to BOLD and receiving a barcode identifier on GenBank are: (1) voucher information (including unique identifiers and the institution storing the specimen); (2) the taxonomic phylum; and (3) the country where 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 the database of sequences for phylogenetic analyses.

Repository management

Specimen collections are stored in a range of facilities, including individual research collections in laboratories, institutional repositories, and national collections housed in herbariums and museums. Each repository determines processes for accessing the specimens that it stores, although individuals who choose to house their genetic resources in these repositories can opt to retain ownership and therefore control over who may subsequently access the material. This continued ownership means that while the facility can provide the infrastructure to store voucher specimens, staff of the facility may not use the specimens for unauthorized work and may not share the specimens with third parties.

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. In general, funds are distributed to two types of projects: large-scale resource-building initiatives (national or international) and smaller country-level projects that generate barcode data based on institutional or individual research targets.

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 2015b)). Canadian research funders have provided substantial funding to building the community. 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 2015f) and initiate the iBOL Project.

Many other funders supported barcoding projects that primarily focused on generating barcode records and expanding the taxonomic coverage of the reference database. iBOL lists 35 funders (from 15 countries) that provided more than $100,000 to support iBOL research and scientific objectives in participating nations (iBOL 2015f). The Canadian International Research and Development Centre (IDRC) provided $2.2 million to support developing countries in their barcoding efforts.

Governance of the DNA barcode commons

The DNA barcoding commons is regulated by formal laws over utilization of genetic resources in some jurisdictions as a result of national implementation of the *CBD* and *Nagoya Protocol*. These may be supplemented by institutional policies for research conduct and research with genetic resources. Researchers are also subject to the terms and conditions that are attached to funding. Finally, researchers are influenced by scientific norms in genomics, which may be enforced through written agreements between institutions. In this section, I describe these aspects of governance of the DNA barcoding commons.

National implementation of CBD and Nagoya Protocol

National laws and regulations that implement the *CBD* and *Nagoya Protocol*, if they exist (Table 1), inform how researchers should access and share genetic resources. Those laws and regulations have international reach in that they govern the export of genetic resources and their utilization in any country. They impose bureaucratic requirements for export permitting, place limits on utilization and generally impose a system of ABS. Therefore, all researchers importing genetic resources from countries with national ABS laws should be aware of their scope and both 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. Such policies may include requiring Prior Informed Consent between genetic resources users and local communities, and Mutually Agreed Terms that outline the intended use of the resource. Countries may also implement policies to encourage research that contributes to bioconservation, including simplified measures for accessing genetic resources. 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).

None of the non-LMMC interviewees were able to discuss implementation of the *CBD* or *Nagoya Protocol* in their own countries. In contrast, LMMC interviewees were more aware of the legal instruments and how national implementation impacted their own work. Many LMMC and non-LMMC interviewees spoke with frustration about implementation of policies by governments 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

Indirect governance by funding agencies

Funding agencies set rules for research conduct that recipients must follow. 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 benefits of research. Despite the wide range of funders, individuals participating in iBOL through the Canadian Centre for DNA Barcoding services were responsible for adhering to Genome Canada policies, and iBOL administrators reported its progress via a corporate board of three senior Genome Canada staff (iBOL 2015e).

In the early phases of iBOL, Genome Canada helped fund DNA barcode sequencing for any specimens sent to BIO, with the requirement that data produced was deposited in an open access database such as GenBank. Interviewees had a wide range of opinions on the appropriate delay prior to data release, although most supported a limited delay for contributor 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.*

iBOL was also guided by a Research Oversight Committee appointed by the board (iBOL 2015e). Perspectives from outside of this structure were represented by the 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 funding of over $250,000 for barcode research. 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 structure of policy-making contributed to decision-making inequities and a lack of representation from lower income countries. One researcher explained the impact of the centralized 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

Formal agreements to govern actions

iBOL and CBOL have both had influence on the legal instruments that govern the exchange of genetic resources within the barcode commons. iBOL developed a standard MTA for materials (specimens, tissue samples, PCR products) that were sent to the Canadian Centre for DNA Barcoding. The MTA was between the Canadian Centre and the institution of the individual providing the specimen, and included the default provision that the material was on permanent loan. The MTA also contained provisions that the provider deposit the data into open access databases, initially requiring only sequence data and high-level taxonomic information to be deposited, and later all associated information including species identifiers.

At the international level, members of CBOL have been actively involved in developing ABS agreements for non-commercial research. Such agreements establish how benefits and risks are shared between partners conducting research that accesses genetic resources, and provide reassurance to provider countries that there will be no un-approved commercial use of their genetic resources. Benefits to be shared 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. Several interviewees favoured using standard MTAs and ABS Agreements for convenience, for example, researcher interviewees did not want to contend with a large amount of paperwork and some preferred short agreements. A LMMC policy maker confirmed that standard agreements for ABS also 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

LMMC interviewees emphasized that MTAs were essential for ensuring that specimens were not used for commercial research or research beyond the original scope of the MTA without permission:

> *The person who is not using it as a commercial thing must sign an agreement to say that if it is at the stage he thinks that you want to use it for commercial purposes, you must come [back] to the institution for discussion* – Researcher, 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. For example, an LMMC researcher reported that his/her collaborator had used shared specimens for scientific studies outside of the original agreed-upon purpose detailed in the MTA. The recipient scientist had not been provided with the species names. S/he disclosed the unauthorized research when requesting the species names to present the findings at a conference. The providing institution refused to provide the species names, citing the breach of the MTA, effectively preventing the unauthorized work from being presented or ultimately published. The use of safeguards such as this in standardized agreements would facilitate other researchers from protecting their resources against misappropriation.

Main action arenas within the DNA barcode commons

The action arenas of DNA barcoding commons are any situation where participants interact and affect or produce outcomes. Here I detail the perspectives of interviewees regarding three key action arenas for DNA barcode commons: generating and sharing DNA barcode records; sharing genetic materials; and accessing the DNA barcodes commons databases.

Generating and sharing DNA barcode records

Individuals and institutions in countries with advanced scientific infrastructure and access to funding sources often favour the rapid and open sharing of genetic resources, particularly data (Field et al. 2009). This preference is reflected in policies that iBOL has developed, and was largely shared by interviewees in the DNA barcode commons. Interviewees felt that the benefits to science outweighed the risks to individual researchers and the best way to increase the coverage of the DNA barcode record database was to require release of barcode records. They also felt the requirement for data release was justified in the early stages of iBOL, when the sequencing was provided free of charge by CCBD.

While LMMC interviewees appreciated the history of data release generated by other large-scale genomics projects, they felt that the unique circumstances of biodiversity research warranted a different approach. Because of this, most did not approve of the standard for rapid, pre-publication release of data. Researcher interviewees pointed out that generating barcodes was labour intensive, and that too much emphasis 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, taking the samples, preserving, shipping. All that should not be underestimated". LMMC researchers also felt that they were disadvantaged by requirements to release data before publication. Indeed, publication timing was the most important factor for LMMC researchers in deciding whether to release data. A delay in release until after publication also served as a quality control through peer-review:

> *A good number of taxonomists [are] people that are not very confident with molecular work. And then they think that this is the first time they get this kind of data, and they are very reluctant to show this data to other people because this data, most of the times, have some [errors]. And that is the fear; really, that is the fear* – Researcher, Mexico

Interviewees who produced barcode records preferred the enhanced, barcode-specific capabilities of BOLD over GenBank. Features they highlighted were 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 vouchers, 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 that the central databases should have more involvement from international stakeholders to consider the needs of contributors and users from lower-resourced settings. Some researcher interviewees also cited lack of trust in North American and European research institutions as a reason to duplicate national-level data from BOLD on local servers.

Sharing biological materials to produce DNA barcodes

As described above, the *Nagoya* Protocol applies to sharing genetic resources other than just for commercial exploitation. Proponents of DNA barcoding were concerned about the ramifications of the *Nagoya Protocol* on biodiversity research, fearing that a restrictive agreement for accessing genetic resources would have the unintended consequence of also slowing biodiversity science. They advocated for simplified ABS terms for genetic resources used for non-commercial research (Schindel et al. 2008), and such terms were included in the adopted protocol.

Many interviewees spoke of the need for researchers to access genetic resources. They argued that because only the barcode would be sequenced and barcoding regions of the genome have no commercial value, the threat of misappropriation from DNA barcoding was overstated. However, many LMMC interviewees described nuanced challenges for the governance of how genetic resources for barcoding are accessed and shared.

Several researchers from non-LMMC expressed concerns about countries blocking important biodiversity research because of fears of misappropriation. As one UK iBOL project participant explained:

> *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. And that, I think, is the main driver for why many of those countries are resisting

This common perspective, however, failed to acknowledge mistrust stemming from a long history of misappropriation of genetic resources. The interviewed LMMC researchers and policy makers understood that genetic resources being shared for DNA barcoding projects were intended for biodiversity science. However, in contrast to a barcode region PCR product, when a specimen or tissue sample is shipped internationally, the whole genome is made available. 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 it". One interviewee mentioned that there was more protection for genetic resources when there was a potential for commercialization, because laws are clear on needing to prove the source of the materials, whereas scientific publications do not have the same requirement.

Some interviewees preferred to share genetic resources for barcoding only on the condition that the specimens and extracted DNA would be destroyed after the barcode sequence was generated. Other participants stated that storing specimens and DNA extracts was necessary to allow for quality control and future research, giving 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, we don't know all the potential purposes. For example genome sequencing will get cheaper, and 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 that specimens should be stored in the country of origin. For many interviewees from LMMCs, this would require the development of expensive storage infrastructure. To mitigate this, the iBOL model enabled countries without commons infrastructure to export genetic resources for barcoding to countries with existing infrastructure. However, LMMC

interviewees were frustrated that 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; that's where the work was published. As a result of that, if I as a researcher, an expert on my fauna [expected] to deliver on my fauna, to do the work, I've got to now spend a lot of my time and money visiting, extracting from those institutions scattered around the world at enormous difficulty; In other words, [these are impediments] to doing the work that I'm required to do to satisfy national interests*

A few LMMC researchers expressed the view 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.

Access to and use of the DNA barcode commons data

Databases used to store DNA barcode records, including BOLD and GenBank, were purposefully designed to allow open access to data and unrestricted use, under the assumption that this benefits the highest number of people (Ratnasingham and Hebert 2007; Nakamura et al. 2013). The requirement to be open access 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. In the view of one Mexican researcher, "[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". As one Canadian research explained: "I like the idea that somebody in India in a third-tier university 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 the use or distribution of the data (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. As one 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”. LMMC interviewees emphasized that different levels of access could be granted to certain types of users, such as: scientists; data contributors; data users without a record of misappropriation of genetic resources; and individuals with a good understanding of ABS principles.

Interviewees were divided on potential restrictions on use of data, 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 that individuals who used data should, at a minimum, acknowledge data producers. The acknowledgement of original data producers was considered most important for data users who did not typically collect specimens or generate data, but simply analysed it:

> *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, even for commercial applications. One interviewee explained that commercial applications were the main benefit of having the open database:

> *Once you get that [barcode record] database then, yeah, there are commercial applications that will be developed and there are academic applications that will be developed. I mean 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

Patterns of interactions and outcomes of the commons

In this section, I first describe how interviewees perceived partnerships and scientific collaborations within the DNA barcode commons. I then discuss (1) the pattern of collaborations as reflected in scientific publication outcomes, and (2) the pattern of specimen collection and data sharing as reflected in the outcome of deposited barcode records.

Collaborations in the commons

Collaborations within the DNA barcode commons define how the commons is built, maintained, and used by participants. They further facilitate the entry of new participants to the commons. Researcher interviewees identified reciprocity as a key factor they consider when determining which collaborations they enter into, but the definition of reciprocity varied. While many interviewees mentioned mutual scientific goals and complimentary 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. Researchers from non-LMMC and LMMC spoke about the importance of good personal relationships with collaborators. One South American researcher succinctly stated his "no assholes" rule. However, the reliance on personal relationships can result in the exacerbation of inequities. Such personal connections are often developed at conferences and meetings that researchers in lower income countries may not have funding to attend.

Some LMMC researchers and policy makers based decisions about undertaking collaborations with non-LMMC researchers on their own experience as well as historical experience, and preferred partnering with groups that shared their overall perspective:

> *We take funding only from those kinds of organizations which share our kind of a world view and which will not interrupt us in our policies - which will not tell us what to do. But once we put in a proposal, the only power that they have is to ensure that we're working according to what we have promised them that we will do in our proposal. So we look for liberal or international organizations which share our views and our perspectives* – Researcher, India

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

> *The people don't want [genetic resources] to be stolen by [the US and Canada] again. But every history is different [for] each person, no? In my case, for example, 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

DNA barcoding publications

Peer-reviewed publications are the primary outcome of basic research and are used as a metric to evaluate researchers for grants, awards, and promotions (Nelkin 1998). Publications indicate that authors are researching a specific topic and provide data on who is collaborating and at which institutions. The ability to produce publications using existing data is a key benefit for academic users of the DNA barcode commons. Many arguments for open access management structures for databases include the claim that researchers in lower income countries would benefit, as they would also be able to access data to produce their own research publications and enhance 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 actually go and get all that stuff for nothing, because someone else has paid for it* – iBOL project participant, United Kingdom

The number of publications referencing seminal DNA barcode papers rose sharply from 2003 to 2011, and leveled off to around 600 publications each year from 2012-2014 (Figure 3). This leveling off is expected as the field matured and references to original publications declined (Bouabid 2011; Barnett 1992). From 2003-2005, every article in my dataset had at least one author from a high income country (Figure 3), which suggests that barcoding activity during this time was conducted exclusively by high income country researchers or in partnership with them. While the proportion of articles with authors from middle or low income countries has risen, the majority of DNA barcode publications have been produced solely by authors in high income countries. A small proportion of papers include low income country authors, rising above two percent of total papers in only two years (2005 and 2011). These data suggest that the growing DNA barcoding community has not expanded to include low income country researchers at the same pace as middle income country researchers.

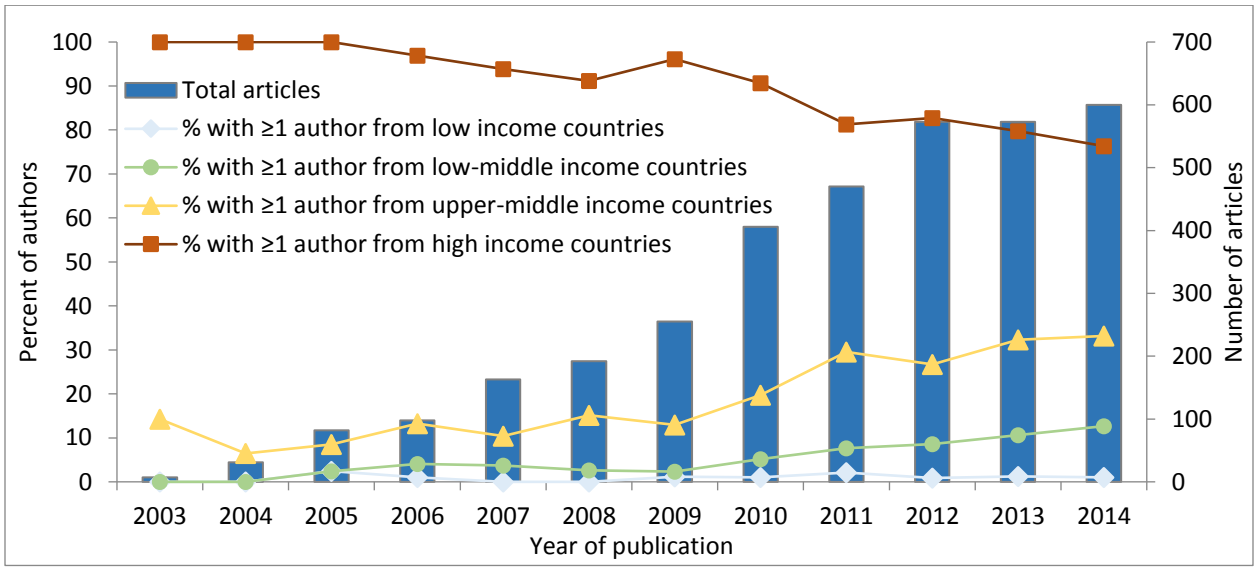

Figure 3. The number of articles citing four seminal barcode papers, published each year during 2003-2014, and percent of articles with at least one author from the specified income group of countries. Income levels are as defined by the World Bank Country and Lending Groups (World Bank 2016).

When examining the co-authorship patterns geographically, I found that co-authorship on barcoding publications was most frequent for researchers in high income "western" countries (defined by the United Nations as Canada, US, Western Europe, Australia and New Zealand (United Nations DGACM 2016)). Figure 4 shows the location of each author in the publication dataset, linked to their co-authors. The coloured lines represent co-authorships when at least one of the authors is not from the western country group, and displays how few co-authorship relationships exist between non-western country authors.

I counted the number of articles with authors from each region, and the number with author sets that spanned more than one region. Over half (54%) of the 3557 articles in my dataset had authors only from Western countries (Table 2). Because only 2% (80/3557) of the articles had author sets that spanned more than two regions, I did not include these co-authorships in Table 2. A full table of the distribution of author sets by region is included as Appendix 3. Regions rich in biodiversity, such as Africa and South America, had few author sets within or across their regions. For example, compared to the 54% of articles with author sets restricted to the Western country group, only 2.5% had author sets confined to Eastern Europe, 3.9% to Latin

America and the Caribbean, 0.8% to Africa, and 17% to Asia and the Pacific (Table 2). Articles with authors from more than one region that did not include western countries only made up 3.2% of the articles in my database.

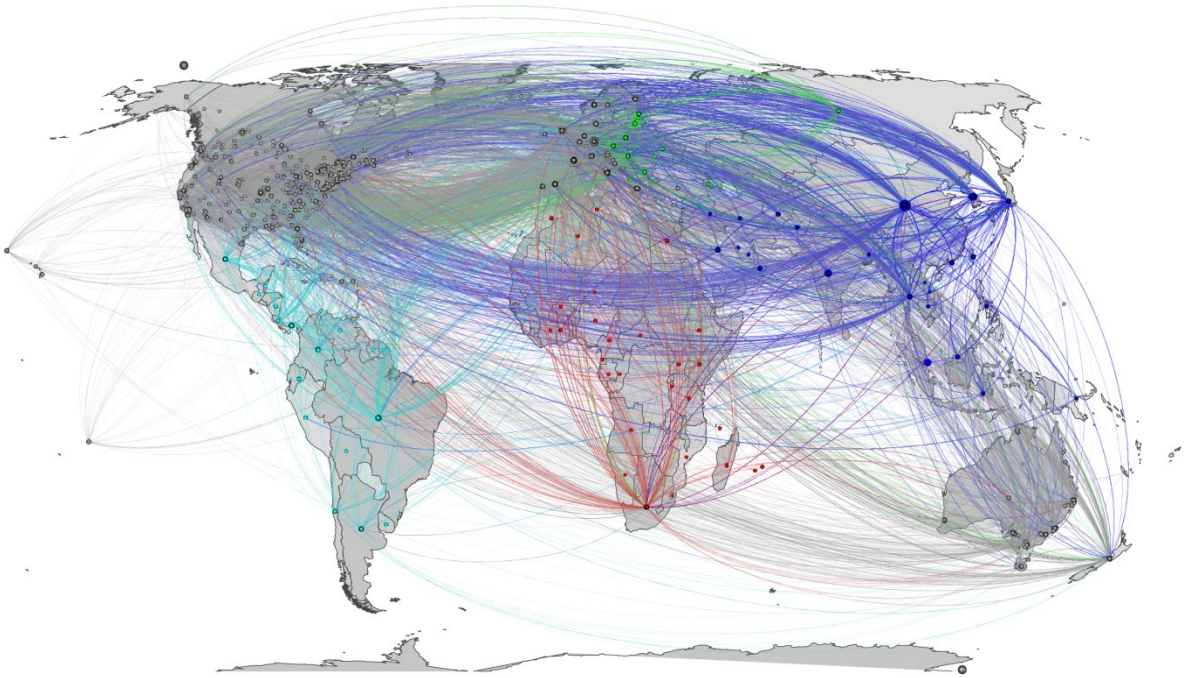

Figure 4. Co-authorship in the DNA barcoding publication database. Each node represents an author, and size of node indicates relative number of times the author has been mentioned in the database. Each line between nodes indicates that the authors co-authored a publication. Lines in grey indicate collaborations restricted to western countries. The coloured lines represent collaborations with other regions (United Nations DGACM 2016).

Researchers are evaluated in part by the prestige of the journals in which they publish their work (Callaham et al. 2002). Regression results (Table 3) showed articles with authors from middle or low income countries had lower odds of being published in highly ranked journals compared to articles with only high income country authors, and that this relationship strengthened somewhat after adjustment for biodiversity status. The estimated ORs adjusted for biodiversity status showed that articles with author sets from a mix of income levels had 65% of the odds of publication in highly ranked journals compared to articles with author sets restricted to high income countries and articles with author sets restricted to middle or low income countries had just 8% of these odds. After adjustment for income level, articles with authors from megabiodiverse countries had 50% greater odds of being published in highly ranked

journals than articles without authors from megabiodiverse countries. While there are a number of reasons that lower income country authors publish less frequently or in less prestigious journals (Salager-Meyer 2008), my data suggest that the academic benefits received by the growing DNA barcoding community are not readily available to middle and low income country researchers.

Table 2: Distribution (percent) of 3557 articles by geographic regions of residence of author sets (excluding 80 with author sets spanning more than 2 regions). Darker shading indicates a higher proportion of articles with some or all authors from that region. Bolded numbers indicate author sets restricted to that region.

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

^Canada, US, Western Europe, Australia, New Zealand
*Latin American and the Caribbean
+Asia and the Pacific

Table 3: Adjusted and unadjusted ORs for the association between characteristics of authors' country of residence and publication in a highly ranked journal, among 3557 articles referencing four seminal DNA barcoding papers.

Characteristic of authors' country of residence	N	Unadjusted OR	95% CI	Adjusted OR	95% CI
Income level					
Only high income country authors	2,386	1.0		1.0	
Mix of high and middle or low income country authors	615	0.76	0.54-1.1	0.65	0.45-0.92
Only middle or low	556	0.09	0.04-0.23	0.08	0.03-0.19
Biodiversity status					
Megabiodiverse country (vs not megabiodiverse)	1,939	1.1	0.82-1.4	1.5	1.2-2.0

OR, odds ratio. Models adjusted for income level and biodiversity status.

BOLD records for exemplar species: Medicinal plants and mosquito disease vectors

I chose two exemplars to examine the patterns of specimen collection and storage for DNA barcode records in BOLD: medicinal plants and mosquito disease vectors.

I identified 17,895 published medicinal plant records in BOLD as of February 2013, of which 11,685 specified specimen origin (Table 4). Fifty-four percent (6,297/11,685) of published medicinal plant records with origin data on BOLD were collected in high income countries, while only 0.4% (50/11,685) were collected in LMMCs. When information about where the voucher specimen was stored was available (9,477), I found only 3% (280/9,477) of voucher specimens were stored outside of the origin country. This may indicate the unwillingness of low income country researchers to share genetic resources with foreign collaborators, as was described by interviewees.

Table 4: 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 records with storage site)
Low income	50	3	47 (94%)	3 (100%)
Low-middle income	640	395	245 (38%)	33 (8%)
Upper-middle income	4,698	4,226	472 (10%)	142 (3%)
High income	6,297	4,853	1,444 (23%)	102 (2%)
Total	11,685	9,477	2,208 (19%)	280 (3%)

I identified 17,297 published barcode records for the genera *Anopheles*, *Aedes,* and *Culex* as of May 2016 (Table 5). Relative to medicinal plants, an even smaller number of records published in BOLD for the three mosquito species were linked to specimens originating from low income countries, with only one found in my dataset. Twenty-one percent (2,521/12,243) of mosquito voucher specimens were stored in collections outside of the origin country, considerably higher than the 3% of medicinal plant specimens stored outside of the origin country, indicating more willingness to export mosquitoes than medicinal plants.

As I was able to determine the number of unpublished records for the mosquito genera, I also compared the number of published and unpublished mosquito records on BOLD. I 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 were published and therefore 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.

Table 5: 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 records with storage site)
Low income	313	1	312 (99%)	1 (100%)
Low-middle income	2,817	1,577	1,240 (44%)	88 (6%)
Upper-middle income	3,312	1,251	2,061 (62%)	944 (76%)
High income	10,855	9,414	1,441 (13%)	1,488 (16%)
Total	17,297	12,243	5,054 (29%)	2,521 (21%)

These two exemplars suggest that individuals from high income countries are contributing more data and specimens to the DNA barcode commons, which is contrary to the goals of having a globally representative database. While many interviewees expressed the view that open access databases would provide the most benefits for potential users, and some emphasized the benefits for low income country researchers, my analysis of the patterns of interactions and outcomes demonstrates that global participation should not be assumed.

Conclusions

Limitations

My study has a number of limitations. I mainly interviewed individuals with direct involvement in BOL organizations and efforts, meaning my analysis did not represent perspectives of those who independently participate in DNA barcoding activities. Similarly, I did not analyse all barcoding publications, only those that referenced seminal papers. I also only examined a small subset of BOLD records relevant to my two exemplars; other exemplars may have revealed different patterns of use. My data interpretation was limited to my own perspectives, and it is possible that another individual might have different views. However, my use of an established theoretical framework reduced the reliance on my individual interpretation

of data. Finally, while my findings are not necessarily generalizable to other global knowledge commons, the rich description of case study provides the contextual details that are necessary to enable transferability to future studies of global knowledge commons governance.

Challenges and opportunities for governing the DNA barcode commons

My case study of the DNA barcode commons examined factors that impact efforts to create, maintain and govern a knowledge commons comprising heterogeneous global participants and resources. The DNA barcode commons includes a large and continually growing number of open access records, and its use results in hundreds of scientific papers every year. However, the goal of the DNA barcode commons is not only to create a comprehensive and accessible repository of barcode records, but also to ensure that the commons is *globally inclusive* with respect to content and *equitable* with respect to contribution, access and use. As knowledge commons should be evaluated against all their goals (Strandburg et al. 2014), my study suggests that the DNA commons has not yet fully accomplished its goals. Here, I summarize the challenges identified by my study, and potential mechanisms through which the DNA barcoding community can take its heterogeneous membership into account and improve governance to promote global participation.

My analysis of the DNA barcode commons action arenas identified how the factors outlined in the IAD framework influence governance of the DNA barcode commons (Figure 5). While the diverse interviewees agreed on the importance of improving taxonomic methods to increase the documentation of life on earth, the initiation of this work has been dominated by non-LMMC participants and their institutional contexts. Many of the policies related to DNA barcoding were influenced by other large-scale genomics projects (Collins et al. 2003), without consideration of how a global project that relies on flow of genetic resources across borders would impact all members of the community. While DNA barcoding leaders have been directly involved in negotiations for the *Nagoya Protocol* (Schindel 2010; Schindel et al. 2008), some members of the barcoding community, including one policy maker, were still unaware of the *CBD* and *Nagoya Protocol*, and their implications for barcoding policies.

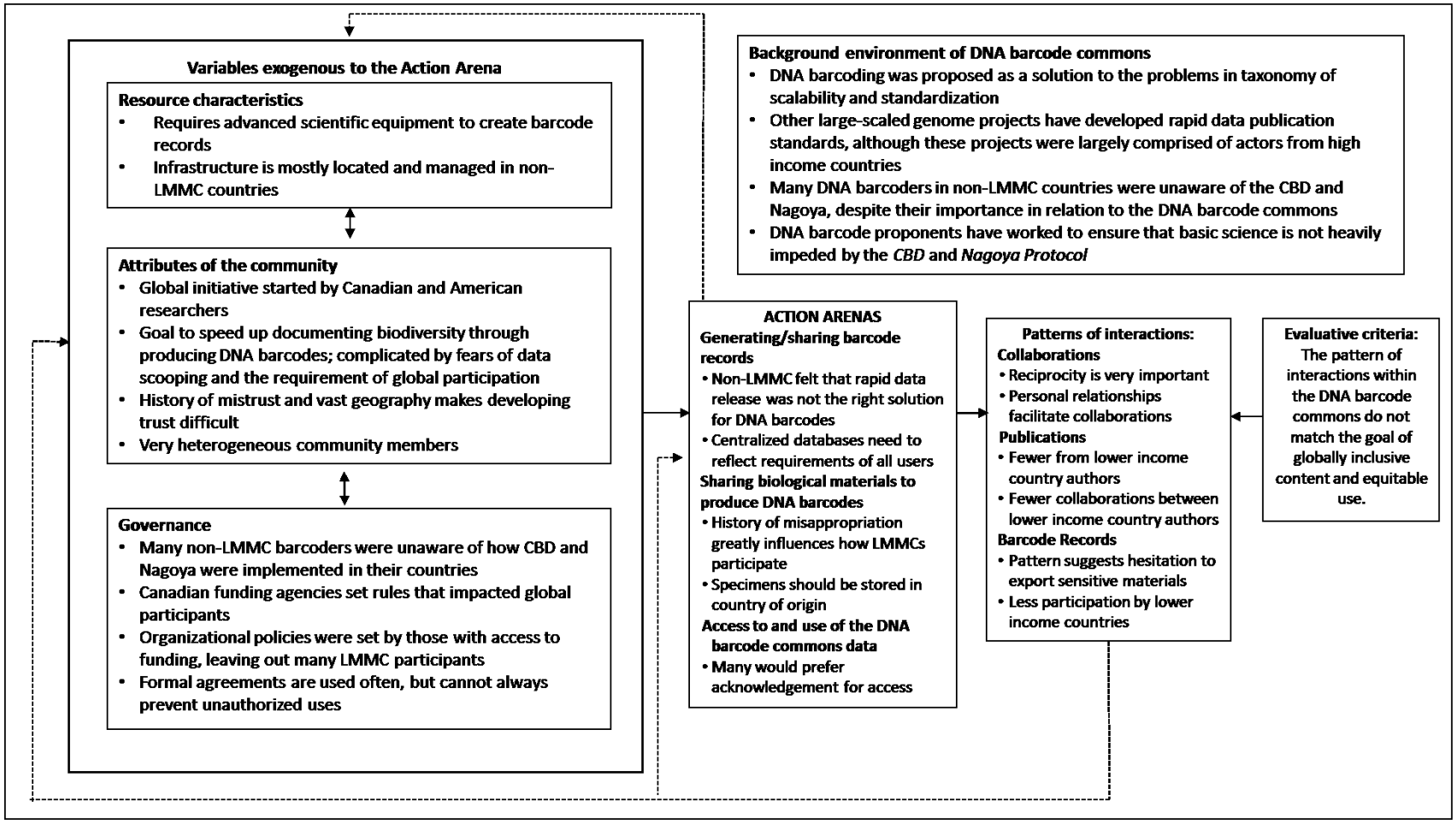

Figure 5. The results of my case-study situated within the KC-IAD Framework.

The background context provided by these international legal instruments is especially salient for the resources that comprise the DNA barcode commons. Generating barcode records requires advanced scientific capacity beyond that of many researchers and institutions in LMMCs. Interviewees explained that even though DNA barcodes contain only a small amount of genetic information, the generation of these records often requires sharing the specimen. That specimen contains the whole genome of the organism. The history of mistrust between countries with differential scientific capacity, and in particular, the ability to benefit from the utilization of genetic resources, has made many LMMC researchers unwilling to export their genetic materials for barcode projects.

The difficulty in establishing trust is also often found heterogeneous communities (Ruttan 2006). Representative governance structures can be effective in overcoming this mistrust (Poteete and Ostrom 2004), and previous research on commons has demonstrated that governance is more successful when actors participate in crafting rules (Ostrom 2003). However, I found that the governance in DNA barcoding has been dominated by the norms and standards from high income countries. The main project, iBOL, implemented rules established by Canadian funding agencies. While the project governance included means to receive feedback from international researchers, participation required substantial funding commitments, which by default excluded perspectives from researchers without access to funding but with interest in contributing barcode records. Indeed, many of the concerns about participating in the DNA barcoding commons shared with me by interviewees from lower income countries were not reflected in the policies of the community.

It is not surprising, therefore, that the heterogeneous interviewees in this study reported different behaviours in the action arenas for DNA barcoding. As behaviours in the action arenas are responsible for the pattern of interactions (Frischmann et al. 2014; Ostrom 2005b), the different behaviours interviewees described to me are likely the cause of the imbalanced pattern of participation in the DNA barcode commons that I observed. Notably, the proportion of DNA barcode papers with low income country authors has remained stagnant over the decade-long project. There are fewer barcode records from specimens collected in lower income countries, and there is evidence that data that are produced from these countries are not published into the barcode commons.

To promote wide-spread use and re-contribution of value-added data (Bubela et al. 2012; Schofield et al. 2009; Dedeurwaerdere 2010a), governing bodies for the DNA barcode commons have relied on promulgating community data-sharing norms of previous large-scale genomics projects (Field et al. 2009; Toronto International Data Release Workshop Authors 2009), and top-down rules requiring rapid data sharing (Genome Canada 2008; iBOL 2015b). These strategies are based on the assumption that open access and unrestricted use are the best ways to achieve the network effect. However, my study suggests that, in a global knowledge commons, open access and unrestricted use inhibit participation by global participants and therefore reduce the network effect. Interviewees in my study were hesitant to share data when they received limited benefits (e.g., scientific credit, increased capacity) in return. Setting restrictions on use, such as requiring citation, attribution, or an embargo period for first use of the data by the contributor, may enhance the release of resources into the commons and subsequently enhance its value. In addition, protection of sensitive information, such as geo-location data for endangered species, is essential.

In addition to helping increase participation, restrictions on the use of the DNA barcode commons are necessary to comply with 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. Barcode records stored in BOLD should be accompanied by terms that outline restrictions on use of the data for commercial applications and research.

Finally, strategies to manage whole specimens for barcoding could also be modified to encourage participation in the barcode commons. Destroying specimens and DNA extracts after the barcode sequence is produced may reassure potential barcode participants that their genetic resources would not be misappropriated, but would have to be balanced against the negative impacts on the integrity of the associated data. Supporting lower income countries in developing infrastructure to store and manage voucher specimens would reduce the need to store the specimens in foreign countries, and could be considered a potential benefit for in exchange for providing access to the country's genetic resources.

I have demonstrated that the goal of creating a globally inclusive DNA barcode commons has not yet been fully achieved. My research also suggests that the risks and benefits of participating in the commons are not equitably shared across heterogeneous global participants. The newly created ISBOL can mitigate the challenges caused by its global membership through ensuring its governance is representative and considers restrictions on use that may enhance participation in the commons.

References

Adamowicz, S. J. (2015). International Barcode of Life: Evolution of a Global Research Community. *Genome*. 58(5): 151–62. 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): D36–42. 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): 160–87. 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.

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.

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 27-31.

Bubela, T., J. Geary, M. Bieber, and D. S. Hik. (2015). Diffusing Barcoding: The Global Spread of a Good Idea. In: *Scientific Abstracts from the 6th International Barcode of Life Conference*.58:201–2. Genome. http://www.nrcresearchpress.com/doi/10.1139/gen-2015-0087.

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): 6350–55. 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. *JAMA*. 287(21): 2847–50. 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.

Centre for Biodiversity Genomics. (2016). Organization. http://biodiversitygenomics.net/about/organization/.

Collins, F. S., M. Morgan, and A. Patrinos. (2003). The Human Genome Project: Lessons from Large-Scale Biology. *Science*. 300(5617): 286–90. http://dx.doi.org/10.1126/science.1084564.

Collins, R. A., and R. H. Cruickshank. (2013). The Seven Deadly Sins of DNA Barcoding. *Molecular Ecology Resources*. 13(6): 969–75. http://dx.doi.org/10.1111/1755-0998.12046.

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): 1601–72. http://dx.doi.org/10.15779/Z38GW9F.

Contreras, J. (2014). Constructing the Genome Commons. In: *Governing Knowledge Commons*. eds. B. M. Frischmann, M. J. Madison, and K. J. Strandburg, 99–136. 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): 414–21. http://dx.doi.org/10.1016/j.resmic.2010.04.012.

Dedeurwaerdere, T., P. M. Stromberg, and U. Pascual. (2013). Social Motivations and Incentives in Ex Situ Conservation of Microbial Genetic Resources. In: *The Digital Public Domain: Foundations for an Open Culture*. eds. M. Dulong de Rosnay and J. C. De Martin, 111–124. Cambridge, UK: OpenBook Publishers.

Dentinger, B. T. M., M. Y. Didukh, and J.-M. Moncalvo. (2011). Comparing COI and ITS as DNA Barcode Markers for Mushrooms and Allies (Agaricomycotina). *PLoS ONE*. 6(9). http://dx.doi.org/10.1371/journal.pone.0025081.

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): 1905–16. 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): 4422–36. 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.

Escalante, P., R. Desalle, and S.-O. Kolokotronis. (2010). The MexBOL Initiative. *Mitochondrial DNA*. 21(sup1): 1–2. http://dx.doi.org/10.3109/19401736.2010.551715.

Field, D., S.-A. Sansone, A. Collis, T. Booth, P. Dukes, S. K. Gregurick, K. Kennedy, et al. (2009). 'Omics Data Sharing. *Science (New York, N.Y.)*. 326(5950): 234–36. 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): 219–20. 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*. 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.

Godfray, H. C. J. (2002). Challenges for Taxonomy. *Nature*. 417(6884): 17–19. http://dx.doi.org/10.1038/417017a.

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): R73–76. 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.

Hardin, G. (1968). The Tragedy of the Commons. *Science*. 162(3859): 1243–48. http://dx.doi.org/10.1126/science.162.3859.1243.

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): 313–21. 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*. 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 Knowlege Commons. In: *Understanding Knowledge as a Commons: From Theory to Practice*. eds. C. Hess and E. Ostrom, 3–26. MIT Press.

Higginbottom, G. M. A. (2004). Sampling Issues in Qualitative Research. *Nurse Researcher*. 12(1): 7–19. http://dx.doi.org/10.7748/nr2004.07.12.1.7.c5927.

Hollingsworth, P. M., L. L. Forrest, J. L. Spouge, M. Hajibabaei, S. Ratnasingham, M. van der Bank, M. W. Chase, et al. (2009). A DNA Barcode for Land Plants. *Proceedings of the National Academy of Sciences*. 106(31): 12794–97. 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): 731–32. http://dx.doi.org/10.1098/rstb.2003.1444.

Jinbo, U., T. Kato, and M. Ito. (2011). Current Progress in DNA Barcoding and Future Implications for Entomology. *Entomological Science*. 14(2): 107–24. http://dx.doi.org/10.1111/j.1479-8298.2011.00449.x.

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.-Z., 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.

LMMC. (2002). Cancun Declaration of Like-Minded Megadiversity Countries. Cancun, Mexico. https://www.environment.gov.za/sites/default/files/docs/cancun_lmmc_declaration1.pdf.

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.

Mittermeier, R. A. (1997). *Megadiversity: Earth's Biologically Wealthiest Nations*. CEMEX.

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.

Nakamura, Y., G. Cochrane, and I. Karsch-Mizrachi. (2013). The International Nucleotide Sequence Database Collaboration. *Nucleic Acids Research*. 41(D1): D21–24. 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): 893–94. 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. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action*. 1sted. Cambridge University Press.

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. Russell Sage Foundation Publications.

Ostrom, E. (2005). Understanding the Diversity of Structured Human Interactions. In: *Understanding Institutional Diversity*.3–31. Princeton University Press.

Pennisi, E. (2003). Modernizing the Tree of Life. *Science*. 300(5626): 1692–97. http://dx.doi.org/10.1126/science.300.5626.1692.

Poteete, A. R., M. A. Janssen, and E. Ostrom. (2010). Small-N Case Studies: Putting the Commons under a Magnifying Glass. In: *Working Together: Collective Action, the Commons, and Multiple Methods in Practice*.3–30. Princeton University Press.

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): 435–61. 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): 355–64. http://dx.doi.org/10.1111/j.1471-8286.2007.01678.x.

Ruttan, L. M. (2006). Sociocultural Heterogeneity and the Commons. *Current Anthropology*. 47(5): 843–53. http://dx.doi.org/10.1086/507185.

Salager-Meyer, F. (2008). Scientific Publishing in Developing Countries: Challenges for the Future. *Journal of English for Academic Purposes*., English for Research Publication Purposes, 7(2): 121–32. http://dx.doi.org/10.1016/j.jeap.2008.03.009.

Schindel, D. (2010). Biology without Borders. *Nature*. 467(7317): 779–81. 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, et al. (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): 6241–46. 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.

Smith, V. S. (2005). DNA Barcoding: Perspectives from a "Partnerships for Enhancing Expertise in Taxonomy" (PEET) Debate. *Systematic Biology*. 54(5): 841–44. http://dx.doi.org/10.1080/10635150500354894.

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*., no. 365(December): 307–28. http://dx.doi.org/10.3897/zookeys.365.6027.

Stoeckle, M. (2003). Taxonomy, DNA, and the Bar Code of Life. *BioScience*. 53(9): 796–97. http://dx.doi.org/10.1641/0006-3568(2003)053[0796:TDATBC]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.

Strasser, B. J. (2008). GenBank--Natural History in the 21st Century? *Science*. 322(5901): 537–38. http://dx.doi.org/10.1126/science.1163399.

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.

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): 168–70. 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.

Vaismoradi, M., H. Turunen, and T. Bondas. (2013). Content Analysis and Thematic Analysis: Implications for Conducting a Qualitative Descriptive Study. *Nursing & Health Sciences*. 15(3): 398–405. http://dx.doi.org/10.1111/nhs.12048.

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 Biol*. 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.

WIPO. (2016). World Intellectual Property Organization Intellectual Property Laws and Treaties Database (WIPO Lex). http://www.wipo.int/wipolex/en/.

Wong, E. H.-K., and R. H. Hanner. (2008). DNA Barcoding Detects Market Substitution in North American Seafood. *Food Research International*. 41(8): 828–37. 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, et al. (2008). Potential Use of DNA Barcodes in Regulatory Science: Applications of the Regulatory Fish Encyclopedia. *Journal of Food Protection*. 71(1): 210–17.

Yin, R. K. (2009). *Case Study Research: Design and Methods*. 4th ed. Vol. 5. Applied Social Research Methods. SAGE.

Chapter 3: The impact of heterogeneity in a global knowledge commons: Implications for governance of the DNA Barcode Commons.

Abstract

The extent of actor heterogeneity is known to influence the outcomes in natural resource commons, and scholars have recently begun addressed the impact of heterogeneity on knowledge commons creation and sustainability. There is increasing evidence to challenge the dominant theory that heterogeneity is uniformly disadvantageous, but little is known about heterogeneity in knowledge commons. Here, I analyse heterogeneity as it applies to rules for governing a 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. The dominant actors in the commons are researchers in diverse fields, and the global scope of barcoding means these researchers work in countries with varying levels of biodiversity, research infrastructure, and financial resources for scientific endeavours. This cultural and wealth heterogeneity among actors results in challenges for constructing and governing the commons, including its supporting infrastructure of databases and biorepositories. I interviewed participants in DNA barcoding, and collected organizational documents. I applied the grammar of institutions to identify institutional statements, and categorized each statement based on institutional logics theory. I found that institutional logics theory is an effective applied research tool to study heterogeneity in knowledge commons. My analysis also suggested that heterogeneity is a challenge to developing shared expectations in global knowledge commons, but participants can design institutional statements to bridge gaps in expectations.

Introduction

Commons vary in their degree of heterogeneity; with respect to commons, heterogeneity refers to the cultural and financial differences between actors (Vedeld 2000). Scholars have extensively studied the impact of heterogeneity on successful outcomes in natural resource commons but have only recently addressed the impact of heterogeneity on knowledge commons creation and sustainability (Mishra and Bubela 2014; Frischmann et al. 2014; Bubela et al. 2012;

Dedeurwaerdere 2010a). Empirical studies and theoretical models of natural resource commons demonstrate that heterogeneity can result in collective action challenges for commons governance (Ruttan 2006). However, there is increasing evidence to challenge the assumption that heterogeneity is uniformly disadvantageous (Agrawal and Gibson 2001). Different types of heterogeneity may have positive or negative impacts on commons governance. For example, while cultural heterogeneity is often a challenge to creating suitable rules for governing commons (Ruttan 2006; Hayo and Vollan 2012), wealth heterogeneity may have a positive impact on commons outcomes, as economically advantaged individuals may be inclined to contribute to sustaining a collective resource for their own benefit (Ruttan 2008). Here, I analyse heterogeneity as it applies to rules for governing a knowledge commons - the DNA barcode commons.

As a result of global promotion of DNA barcoding, the DNA barcode commons has evolved into a standard process for species identification, coordinated through two main organizations: the Consortium for the Barcode of Life (CBOL) and the International Barcode of Life Project (iBOL). In 2013, barcoding leaders proposed a new organization, the International Society for the Barcode of Life, to coordinate and promote the future activities of the DNA barcoding community (Castle et al. 2015).

The DNA barcode commons has widespread support because DNA barcoding has broad applications, including monitoring biodiversity, identifying food market substitutions, or reducing cross-border smuggling of endangered species (Barbuto et al. 2010; Gross 2012). All uses of barcoding rely on a robust, accessible reference database of verified DNA barcodes linked to voucher specimens and additional information (metadata) about the species and the specimen.

The dominant actors in the commons are researchers in fields that include taxonomy, ecology, conservation biology, genetics, and bioinformatics. The global scope of barcoding means these researchers work in countries with varying levels of biodiversity, research infrastructure, and financial resources for scientific endeavours. This cultural and wealth heterogeneity among actors results in challenges for constructing and governing the commons, including its supporting infrastructure of databases (for barcodes and metadata) and biorepositories (for voucher specimens and tissue samples). For example, countries with the most

biodiversity to contribute to the commons often have limited financial and scientific resources to use the commons, creating inequities. Actors and infrastructure in different countries also confront differing legal environments. National laws vary in the degree to which they regulate the export and import of biological specimens and sequence information, collectively termed "genetic resources".

The behaviours of individual actors who participate in DNA barcode commons, as in all commons, are determined by a mix of regulations, instructions, precepts, and principles (Black 1962), generally referred to collectively as 'rules' and referred to as "institutional statements" in the commons literature (Basurto et al. 2010). Rules established and disseminated by an authority are labeled "rules-in-form". For DNA barcode commons, these rules-in-form come from international treaties, such as the *Convention on Biological Diversity* that governs the use of genetic resources (including biological materials and sequence data), funding agency requirements, and organizations like CBOL and iBOL. However, the "rules-in-use" that individuals follow in their day-to-day behaviour are often different from rules-in-form, and institutional theorists have argued that understanding these rules-in-use is central to deeper institutional analysis of commons (Crawford and Ostrom 2005).

The grammar of institutions was introduced in 1995 to facilitate analysis of human action within institutional settings (Crawford and Ostrom 1995). Identification of this grammar is a tool for systematically differentiating between types of institutional statements (*Rules* with consequences, conditional *Norms*, and simplistic shared *Strategies*[1]),). Scholars have employed it extensively to examine the institutional statements that are 'in use' (Basurto et al. 2010; Siddiki et al. 2011). However, understanding the impact of heterogeneity on what institutional statements are used also requires an examination of what influences individuals' behaviours and their expectations of the behaviours of others. In addition to a tool for examining institutional statements, scholars can use other theoretical perspectives to expand their ability to study heterogeneity in relation to rules. To add to this toolkit of theoretical perspectives, I employ here

---

[1] I have used capitalization and italics to indicate when referring to the specific *Rules, Norms* and *Strategies* as defined in the syntax of the grammar of institutions. Otherwise, the term rules is referring generally to any statement that could be understood as rules-in-form or rules-in-use.

theory concerning institutional logics, that focuses on the importance of overarching institutional orders that shape how individuals behave and expect others to behave (Friedland and Alford 1991). Many empirical studies have drawn on the concept of institutional logics to examine and understand organizational behaviour (Thornton and Ocasio 2008; Thornton et al. 2012).

Using the DNA barcode commons as a case study, my goals are twofold: 1) to examine empirically how heterogeneity relates to institutional statements in a global knowledge commons, and 2) to contribute an expanded, theoretically grounded method for the study of heterogeneity and knowledge commons, using the concept of institutional logics. In the next sections, I outline the relevant theories and the conceptual framework used to accomplish my aims. I first describe how rules are studied within a commons using the grammar of institutions developed by Elinor Ostrom and colleagues (Crawford and Ostrom 2005). The grammar sets out a syntax that is frequently used to study systematically institutional statements that govern commons (Basurto et al. 2010; Siddiki et al. 2012; Siddiki et al. 2011). I then introduce the concept of institutional logics to expand the understanding of heterogeneity in relation to institutional statements. Institutional logics are a helpful analytic tool because they are useful in understanding the influence of multiple or unknown forms of heterogeneity. Finally, I describe my case study and go on to detail the analytical framework used to examine logics in the rules that guide actors who participate in the DNA barcode commons.

Literature and Background

Heterogeneity in the commons

Heterogeneity is an important aspect of a commons community. Broadly, it refers to the degree to which actors differ on characteristics. Communities can vary on any number of factors, making it difficult to identify which factors to empirically measure and study. Nonetheless, substantial research has focused on understanding the impact of heterogeneity on commons governance. The emerging theoretical consensus is that increased heterogeneity is a challenge for communities to overcome, rather than a necessary determinant of failure (Ostrom 2005). Evidence also suggests that institutional arrangements (defined as rules-in-use by a community to determine the nature of the use of and access to a resource, including how they will be

enforced (Ostrom 1987)) are more highly impactful on heterogeneous than homogenous communities (Brito et al. 1997), and the governance of commons is more successful when actors are involved in making their own rules (Varughese and Ostrom 2001).

Looking beyond the question of whether heterogeneity influences outcomes, scholars have increasingly focused on the impacts of specific sources of heterogeneity. Most studies are quantitative and assess just one narrowly defined variable as a source of heterogeneity, for example, knowledge (Lindahl 2012), dependence on the resource (Sakane et al. 2014), social caste (Shiferaw et al. 2012), contributions to the resource (Burlando and Guala 2005), social preferences (Fischbacher and Gaechter 2006), and productivity (Brito et al. 1997). A review of quantitative heterogeneity studies suggested that how heterogeneity impacts a commons depends on the source of heterogeneity that is examined as the independent variable, and whether the study measured collective action or the provision of a collective good as the dependant outcome variable (Ruttan 2008). For example, wealth heterogeneity has been associated with the long-term sustainability of a resource (Balooni et al. 2010), while socio-cultural heterogeneity has been associated with reduced collective action (Benin and Pender 2006). In expanding on this finding, Poteete and colleagues (2010a) demonstrated that social heterogeneity negatively impacted collective action because it was associated with lower levels of trust. In contrast, wealth heterogeneity positively affected provision of goods because wealthy actors were highly motivated to maintain provision of a good that provided them with benefits.

While Poteete et al. (2010a)'s theory remains a plausible explanation for outcomes in commons for rivalrous (i.e. depletable) resources in natural resource settings, few studies have explored the relationship between heterogeneity and knowledge commons governance. Knowledge commons require separate study and theory development because they differ from natural resource commons in several key ways (Frischmann et al. 2014). One of the most important differences is that knowledge commons typically comprise non-rivalrous resources. Therefore, the goal of successful knowledge commons governance is not preventing overuse and ensuring sustainability, but increasing use of the resource, promoting active re-contribution of value-added information, and fairly distributing the costs and benefits of the resource between users distributed across a potentially unlimited geography. Many institutional arrangements for these non-subtractable resources provide free and open access. In theory, open access makes use

more efficient, but in practice, it creates additional challenges. Open access leaves the resource vulnerable to free riders - those who take from the commons without contributing to it – and free riding reduces trust and therefore the ability to establish shared *Norms* within the community. Thus, heterogeneity in a knowledge commons may include many different dimensions compared to the previously studied natural resource commons communities. Qualitative approaches have the benefit of avoiding reliance on assumptions about selected variables (Jansen 2010). Instead, the qualitative approach I employ here enables an examination of the depth and diversity of heterogeneity in knowledge commons.

Grammar of Institutions

Institutional statements are the *Rules*, *Norms*, and *Strategies* that are either formally documented (rules-in-form) or employed by actors in their day-to-day behaviours (rules-in-use). The grammar of institutions was developed as a common syntax to aid the empirical study of institutional statements (Crawford and Ostrom 1995). The focus on rules is useful because they are modifiable, unlike other characteristics of a commons such as its biophysical characteristics or the attributes of the community (Crawford and Ostrom 2005).

The grammar of institutions includes five working parts summarized by the acronym "ADICO" (Crawford and Ostrom 1995). *Attributes* (A) refer to the subset of community members to which the particular statement applies, ranging from a single individual to the entire group. If not specified, readers can assume institutional statements apply to all actors. The *Deontic* (D) distinguishes prescriptive (must, forbidden) from non-prescriptive (should, may). The *Aim* (I) is the specific action or outcome intended by the statement. *Conditions* (C) set out the situations when an institutional statement is applied and are especially important in defining the situations in which a specific behaviour is permitted or forbidden. Unless specified, conditions include all situations. The *Or Else* (O) of the statement defines a consequence for the actor if s/he does not comply with the statement.

Researchers can use the ADICO syntax to classify institutional statements, either written in organizational documents or articulated by individuals, as *Rules, Norms,* or *Strategies*. *Strategies* (AIC) are the least constraining and represent a general guideline without conditions that allow or forbid the particular behaviour, and without consequences for non-compliance.

Norms (ADIC) are more specific and lay out the conditions under which a behaviour occurs, although *Norms* do not specify the consequences for non-performance of the behaviour. *Rules* (ADICO) are enforceable statements with specific conditions and consequences. Although useful as an empirical tool, the ADICO syntax can also be used as a guideline when creating institutional statements to ensure that are enforceable rules, rather than suggestions (Crawford and Ostrom 2005).

Institutional Logics

Institutional logics are a concept within institutional theory developed to help explain the influence of society on organizational behaviour (Friedland and Alford 1991). Thus, institutional logics are an attempt to account for society in institutional analysis by specifically considering the overarching organizational structures of Western society that shape how individuals and organizations behave within institutions. In the language of institutional logics theory, these organizational structures are called "institutional orders", and they include capitalist market, bureaucratic state, democracy, families, and religion; each has a central logic that guides behaviour. In short, logics allow actors to make sense of situations by providing "assumptions and values, usually implicit, about how to interpret organizational reality, what constitutes appropriate behaviour, and how to succeed" (Thornton 2004, 70).

Scholars have expanded the list of logics and theorized about how logics might interact with each other within an institution. The current framework includes seven logics: family, community, religion, state, market, profession, and corporation (Thornton et al. 2012). These logics often co-exist, although many early empirical studies focused solely on identifying a dominant logic (Goodrick 2002; Lounsbury 2002; Thornton 2002; Thornton and Ocasio 1999), largely ignoring any secondary logics. Scholars noted that the focus on dominant logics was a divergence from Friedland and Alford's (1991) original work, which stated organizational are typically subject to multiple logics. As a result, empirical work began to focus on how multiple logics jointly guide behaviour.

Initial work to explain how logics co-existed focused on mechanisms of competition for dominance (Hensmans 2003; Hoffman 1999), unresolved conflict requiring 'uneasy truces' between competing logics, (Reay and Hinings 2005), and various ways segmentation allowed for

multiple logics to co-exist without interacting (Reay and Hinings 2009; Lounsbury 2007; Thornton et al. 2005). However, none of these explanations allowed for the reality that multiple logics guide behaviour of single actors, an idea introduced by Goodrick and Reay (2011). Goodrick and Reay studied pharmacists, who faced pressure from four different logics: profession (obligation to professional standards), corporate (work within large organizations), market (sell product), and state (adhere to government regulations). Their analysis resulted in a rich description of how multiple societal level logics collectively influenced professional work (Goodrick and Reay 2011).

Recent research has suggested that specific contexts can act as filters that alter how logics inform behaviours within an organization (Jaskiewicz et al. 2015; Lee and Lounsbury 2015). These filters are attributes of the organization that impact how it draws on institutional logics, and can help explain organizational success (or failure) in managing competing logics. For example, Jaskiewicz et al (2015) found that while family and market logics informed family businesses, how these logics informed behaviour depended on the filters of family culture and leadership style within each business.

Ideal types as an empirical tool

Empirical study of logics was grounded on the concept of ideal types (Thornton 2004; Thornton and Ocasio 2008), which represent how behaviour would be organized if each logic was the only influence on the behaviour (Thornton et al. 2012). As an analytical tool, defining the ideal types of logics allows researchers to systematically cluster behaviours into categories to facilitate comparison (Thornton et al. 2012). Prior to empirical analysis, researchers must identify the relevant logics for the study, and define how behaviour would be organized if guided solely by each logic (Reay and Jones 2016). Next, I describe four logics from the current framework of institutional logics that I determined were relevant to the DNA barcode commons: profession, state, market, and corporation.

The profession logic is characterized by the specialized knowledge gained through professional education (Freidson 2001). In DNA barcoding, the main professions are related to taxonomy and biology, as these provide the training and expertise to create DNA barcodes and use them. Professionals alone determine DNA barcoding goals and standards, which are enforced through professional associations, such as academic journals requiring sequence data release as a

stipulation on publishing an article. While state actors can legislate to provide authority to professional associations, they are not involved in determining the standards (such as how specimens must be prepared) that professional associations enforce under these laws.

In contrast, the ideal type state logic does involve direct government control through either legislation or state actors. State actors, including public funding agencies and regulatory bodies, determine the DNA barcode projects that receive resources and the standards and protocols to which DNA barcoders must adhere. The strategy of the government is to use DNA barcoding to increase public good.

In the ideal type market logic, individuals compete in an open system free of any regulation, whether produced through a state agency or professional association. Any individual would be able to produce DNA barcodes, and user preferences determine what standards are acceptable quality. The strategy of the market logic is to increase efficiency and personal gains, either reputational or profit-based. Even in a market logic-dominated DNA barcode commons, professionals (biologists) would be the main actors producing DNA barcodes, resulting in some overlap between personal and professional reputational benefits. However, a biologist is driven by a market logic rather than a profession logic when the reputational benefits are derived from personal gains independent of professional associations, and when professional standards are disregarded in favour of standards which provide other personal benefits. While specialized knowledge may be valued, it is available to anyone and not necessarily obtained through formal education (Goodrick and Reay 2011).

Corporate logics place emphasis on organizational hierarchy and administrative control of actions. The relevant organization in DNA barcoding is the iBOL project, which established rules and project targets for participants. Administration controls actions to standardize production, aiming for predictability and efficiency. The strategy associated with a corporate logic is to increase its size and diversification.

Summary of theoretical background

The above approach developed to study institutional logics as the basis for behaviour in organizations can be adapted to help understand the mechanisms through which heterogeneity affects rules governing behaviour in knowledge commons. The grammar of institutions provides

syntax to deduce institutional statements from descriptions of how individuals behave in their daily practices, and the concept of institutional logics provides a mechanism to identify the bases for these actions, and examine possible sources of conflict that might arise from a lack of shared logics due to heterogeneity across actors. I examine a single case study to provide an in-depth analysis of heterogeneity in a knowledge commons: the DNA barcode commons.

Case Study: DNA barcode commons

In this section, I describe the salient aspects of how the DNA barcode commons comprises heterogeneous actors. First, however, I briefly explain what DNA barcoding is and how it became a global scientific endeavor.

The process and promise of DNA barcoding

In 2003, a team led by Paul Hebert at the University of Guelph in Canada proposed DNA barcoding as a tool to accelerate documenting life on earth – a pre-requisite for the study of anthropogenic and other impacts on biodiversity (Hebert et al. 2003). The proposal demonstrated the potential to differentiate species by way of sequencing a small region of an organism's DNA. Previous efforts at classifying organisms were inefficient (requiring highly specialized taxonomists), expensive (requiring extensive DNA sequencing), and not scalable due to lack of standardization (Tautz et al. 2003). Hebert et al proposed using short, highly conserved, and relatively ubiquitous DNA sequences. These sequence characteristics make high-throughput analyses possible, enabling a barcoding pipeline with attendant economies of scale.

While DNA barcoding is promoted as a tool to help identify life on earth, the process to create a DNA barcode begins with taxonomically classified specimens (Figure 6). Biological materials, therefore, are an integral component of the DNA barcode commons. Specimens are collected from the field or sampled from existing collections. These specimens can be shipped to locations capable of performing other tasks associated with creating barcode records, including isolating DNA, sequencing the relevant barcode, and storing the voucher specimen. The result of the barcoding process is an open access, comprehensive database of DNA barcodes linked to metadata and a reference specimen.

The effort to build the DNA barcoding database quickly gained momentum not only because of its potential contributions to the field of taxonomy but also its enablement of a range of practical applications (Hebert et al. 2003). Access to such a database facilitates rapid identification of unknown specimens in situations where morphological identification is impossible. Such situations arise in many contexts, for example, where the necessary expertise to make identifications is unavailable, or the specimen needing identification is indistinguishable from other similar species, such as butchered meat or insect larvae (Costa and Carvalho 2007). Identifications are made by matching the DNA barcode from an unknown specimen to the known barcode record linked to the voucher specimen. DNA barcodes from unknown specimens may be generated by sending (1) the entire specimen (2) the specimen's whole DNA extract, or (3) only the DNA barcode extract from the specimen to a laboratory equipped to produce DNA barcodes.

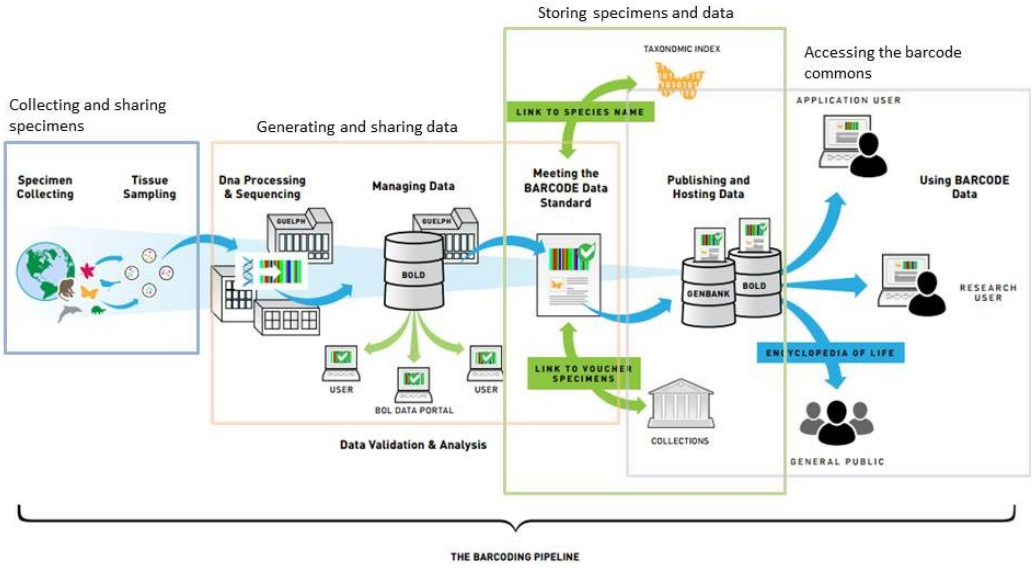

Figure 6. The DNA barcoding pipeline. Adapted from CBOL website (Consortium for the Barcode of Life 2016)

Barcoding as an organized global effort

Creating a comprehensive database of DNA barcode records was necessarily a global endeavor, and scientists quickly organized to generate international participation through two main organizations: The Consortium for the Barcode of Life (CBOL) and the International

Barcode of Life Project (IBOL). CBOL was founded in 2004 at the Smithsonian Institution with support from the Alfred P. Sloan Foundation, and focused on promoting the DNA barcode system and developing global standards. In 2009, Paul Hebert led an international initiative to build a barcode reference library for global biodiversity; this initiative resulted in the launch of the International Barcode of Life (iBOL) Project in 2010. Funded through Genome Canada's International Consortium Initiative, iBOL included 26 nations as iBOL 'nodes' partnering through formal agreements (iBOL 2015d). Genome Canada originally committed $25 million from until 2015, along with 35 international sponsors from 15 countries (iBOL 2015f). The main mission of iBOL was to build a publicly accessible database including 5,000,000 barcodes representing 500,000 species by 2015 (iBOL 2015i). The Barcode of Life Datasystem (BOLD), housed at the informatics unit of the Biodiversity Institute of Ontario (described in more detail in the next section), grew from 102 users in 2005 to over 14,000 users from 94 countries in 2015 (Ratnasingham 2015).

The formal governance structures of CBOL and iBOL established standards for participating in the DNA barcoding effort, but other overarching polices and norms influenced both the development of these standards and the behaviours of participants. At the international level, sharing genetic resources is governed 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 most direct influence on the organizational rules of iBOL derived from Genome Canada; as with all of Genome Canada projects, iBOL was appointed a Board from Genome Canada to oversee and direct how the funding was used. The Board subsequently appointed a Research Oversight Committee that is independent of the iBOL but provides strategic directions on the research to ensure it achieves its Board-approved objectives and milestones (iBOL 2015e).

In 2013, leaders within the DNA barcoding community proposed a new international coordinating organization to sustain the functions of CBOL and iBOL (Li et al. 2013). The organization, named the International Society for the Barcode of Life (ISBOL) was launched in 2015 at the 6th International Barcode of Life Conference in Guelph, Ontario. The authors of the Kunming Declaration and representatives from key regions and organizations formed an Interim

Governance Council to develop a governance structure for the new organization and establish its practices.

The heterogeneous DNA barcoding community

The DNA barcode commons comprises actors from a set of countries that differ in their ability to support equal participation. To participate independently in barcoding, a researcher must have access to specimen collections, DNA extraction and sequencing equipment, and infrastructure to store voucher specimens and to process and share data. Many of the countries that possess the greatest amount of biodiversity within their borders (Australian Government 2008) have few resources for biodiversity research, including The Congo, Madagascar, India, Colombia, and Venezuela.

To address disparities among participating countries in scientific infrastructure, iBOL established regional and central nodes. The aim of regional and central nodes was to provide scientific infrastructure and support to national nodes that lacked the necessary resources (iBOL 2015d). Canadian national and provincial funders supported additional infrastructure - the Canadian Center for DNA Barcoding (CCDB) housed within the Centre for Biodiversity Genomics (CBG) at BIO in Guelph. CCBD is the largest contributor of DNA barcode records, and CBG includes operating units to manage biological collections, production of genomic data, informatics, international development, and education and outreach (Centre for Biodiversity Genomics 2016).

Despite available resources, however, national laws and institutional policies may limit the extent to which researchers can share the genetic resources required to produce barcodes, including both the biological specimens and the DNA barcode sequences. Many biodiverse countries have implemented the *Nagoya Protocol* through national legislation on access and benefit sharing (ABS) that protects sovereignty over genetic resources (United Nations 2010), with the intent of preventing misappropriation of resources. In contrast, high income countries like Canada, the United States (US), and Australia have not. Researchers from countries with ABS laws may not be allowed to send specimens or biological samples out of country or deposit DNA barcodes into databases.

Actors also vary in their ability to influence BOL efforts. The two main BOL organizations that create standards and protocols for creating and sharing DNA barcodes, iBOL and CBOL, are based in Canada and the US. While iBOL included an International Scientific Steering Committee, participation was limited to individuals who had secured large amounts of local funding for BOL projects (iBOL 2015e).

Finally, the ability to benefit from an open-access barcode database varies substantially. Researchers and regulators in countries with better scientific infrastructure are more likely to be able to utilize a barcode database (Yancy et al. 2008). Professional benefits are also more likely to accrue to researchers with training in bioinformatics, access to computers and broadband internet connections, English language competence for publications and presentations, and funding to travel to scientific venues for developing collaborative networks.

Summary

Actors in the DNA barcode commons, an exemplar global knowledge commons, are highly heterogeneous. Commons resources comprise multiple forms of data, metadata and biological materials, the sharing and utilization of which are governed by a complex web of rules. Both iBOL and CBOL established formal institutional arrangements for participating in the DNA barcode commons, however, these groups represent a small subset of the community largely centralized in North America. Such institutional arrangements have the potential to encourage collective action to create and maintain the commons, but collective action may be impacted by socio-cultural and wealth heterogeneity among actors. Current empirical approaches to study the impacts of heterogeneity fail to consider multiple forms of heterogeneity, or provide insight on which institutional arrangements might best overcome the problems arising from heterogeneity, especially that which arises from a global context. My empirical analysis addresses the study of heterogeneity through combined use of the grammar of institutions to identify rules-in-use, and institutional logics necessary to understand the relationship between heterogeneity and collective action within the global DNA barcode commons.

Methods

My empirical work aimed to broaden research on heterogeneity and rules-in-use beyond the impact of single variables. I employed a mixed-methods approach to combine theoretical frameworks and data analysis methods to analyse a single case study, using multiple data sources (Poteete et al. 2010b).

Data sources

My data derived from two sources: key informant interviews and organizational documents. This research received ethical approval from the University of Alberta Research Ethics Board – Health Panel (Appendix 7). I, along with the assistance of a research assistant, interviewed 35 individuals DNA barcoders, iBOL and CBOL administrators, and key individuals in the international barcoding community who were engaged in policy discussions. I selected the number of interviewees based on a combination of resources available to attend events to recruit participants, ability to recruit appropriate interviewees, and an ongoing review of data to determine if additional interviews were needed to explore emerging themes. The country affiliations of the interviewees were: Australia, Canada, China, Colombia, Ghana, India, Indonesia, Kenya, Mexico, New Zealand, South Africa, United Kingdom, and the US. I conducted most interviews in person at International Barcode of Life conferences in 2011 (Adelaide, Australia), 2013 (Kunming, China) and 2015 (Guelph, Ontario), with the rest occurring over the phone or in person at other locations by myself or a research assistant. We followed a semi-structured interview guide that queried about research collaborations, sources of genetic resources, views on the national and international frameworks governing ABS for genetic resources, and participation in DNA barcode commons (Appendix 4). I used a paid service to transcribe each interview verbatim and verified each transcription against the original interview recording.

I analyzed organizational documents and statements developed by the two main BOL organizations: IBOL and CBOL. I collected documents and policy statements that described their purpose or the governance approaches, or provided instructions to project participants. The formal documents were: the biological materials transfer agreements; iBOL data release policy;

microplate and data submission instructions; iBOL Node Memorandum of Understanding; iBOL Node MOU Appendix; Data Standards for Barcode Records; the Banbury Report on Taxonomy, DNA, and the Barcode of Life; and Guidelines to Authors of Barcode Data Release Papers. I obtained these documents either by downloading them from the BOL websites or directly from the BOL project administrators. I included one document not produced directly by BOL proponents or actors, *The Fort Lauderdale Report on Sharing Data from Large-scale Biological Research Project,* because other BOL documents referenced it as a guiding principle for data sharing. In addition, I accessed informal project descriptions from iBOL.org of global iBOL-supported barcode projects, DNA barcoding, and the iBOL governance structure (iBOL 2015i).

Data analysis

I developed a framework to guide analysis of each interview transcript and document. This framework integrated qualitative and quantitative approaches for systematic identification and characterization of institutional statements about DNA barcoding activities in the interviews and documents. I used NVivo qualitative analysis Software (QSR International Pty Ltd. Version 10, 2012) to facilitate data organization and analysis.

I developed the list of DNA barcoding behaviours based on project descriptions on the iBOL website (iBOL 2015i), and the workflow required to produce and share barcodes (Figure 7). The actions to participate in barcoding include: setting the scope of barcoding activities; collaborating with others to conduct a barcode project; collecting, sharing, and storing biological specimens (including whole specimens, tissues, and purified genetic material); generating, sharing, and storing data associated with biological specimens (including barcode sequences and meta-data such as sample locations and voucher specimen pictures); and accessing the barcode commons. I read each transcript and document and marked each instance of these behaviours (Figure 7a).

Using NVivo software, I focused my analysis on the text surrounding the identified behaviours. Individuals do not usually express institutional statements explicitly (Crawford and Ostrom 2005), thus researchers must examine texts to assemble the statements from impartial references contained in larger blocks of interview transcripts (Figure 7b). Each behaviour I identified was a potential *aim* in the grammar of institutions, and I examined the surrounding text

to identify other components of the syntax. I then re-organized the statement to produce a simplified version (Figure 7c), and coded them based on what type of institutional statement they were (*Strategy*, *Norm*, or *Rule*).

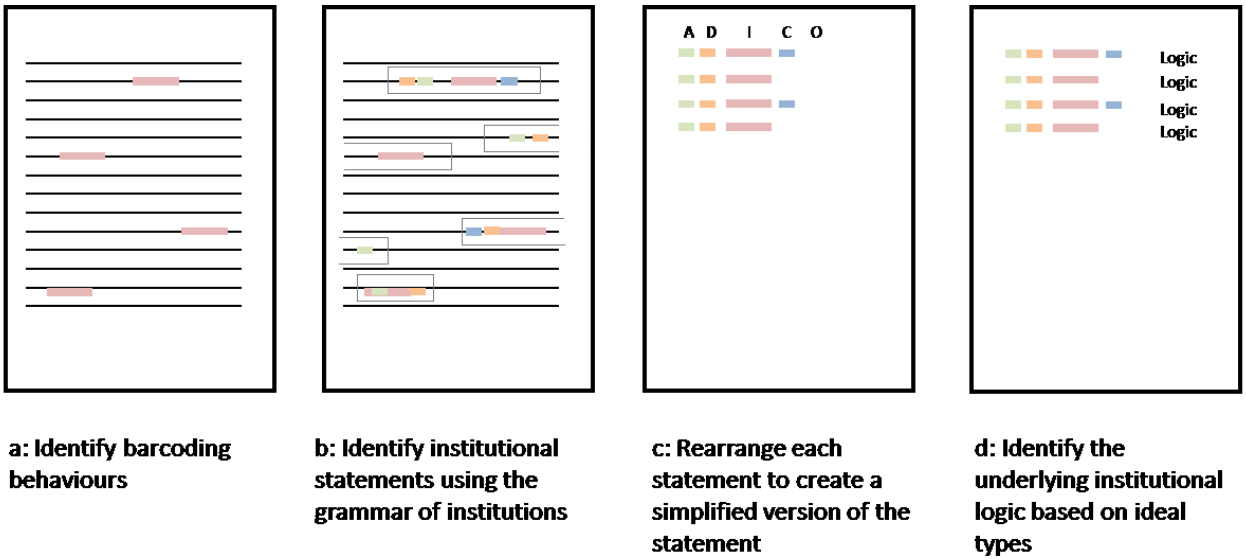

Figure 7. Analytical workflow for identifying institutional statements and the logics that influence each statement.

I and a research assistant then coded each institutional statement based on which of the logics appeared to influence it (Figure 7d), determined by comparing the statement to descriptions of ideal types (Table 6) and the general characteristics of the logics. Multiple logics could influence a single statement, and any statements that appeared to be influenced by logics in addition to or other than the four I defined were coded as "other".

I compared the coding by each independent coder. We agreed on 67% of our initial coding. We discussed each disagreement to come to a consensus about which logics were influencing each statement. Although we did not blind ourselves to the participants' country affiliation, we did not link institutional statements to participant attributes and we organized the statements randomly when coding for logics. This reduced the likelihood of systematic errors based on the country of participants. Table 7 gives an example of statements coded as each behaviour and logic type.

Table 6: Descriptions of DNA barcoding behaviours as determined by each of the four identified ideal logic types

	Corporation	**Market**	**Profession**	**State**
Setting the scope of barcoding efforts	BOL administration assigns commons participants species to barcode	Any individual can decide what species they will barcode	Researchers decide what species they will barcode	State agencies determine what species are barcoded
Collaboration Setting (who is included and excluded from collaborations)	BOL administration determines who will collaborate and how collaborations function	Any individual can decide who they will collaborate with and how collaborations function	Researchers decide who they will collaborate with and how collaborations function	State agencies determine who will collaborate and how collaborations function
Standard setting (creating commons) Collect specimens	BOL mandates credentials and standards for specimen collection	Any individual can set their own standards for specimen collection	Researchers decide on credentials and standards for specimen collection	State agencies mandate credentials and standards for specimen collection
Store specimens	BOL mandates standards for specimen storage	Any individual can set their own standards for specimen storage	Researchers decide standards for specimen storage	State agencies mandate standards for specimen storage
Share specimens	BOL determines who can share specimens when	Any individual can decide who they will share specimens with and when	Researchers decide who can share specimens when	State agencies determine who can share specimens when
Generate data	BOL administration mandates data quality standards	Any individual can set their own data quality standards	Researchers set data quality standards	State agencies mandate data quality standards
Store data	BOL administration mandates data storage requirements	Any individual can determine their own data storage requirements	Researchers determine data storage requirements	State agencies mandate data storage requirements
Share data	BOL administration determines who releases data when	Any individual can decide when they release their data	Researchers decide who releases data when	State agencies determine who releases data when

	Corporation	Market	Profession	State
Access data (use commons)	BOL administration determines who can access the data	Any individual can decide who can access their data	Researchers decide who can access the data	State agencies determine who can access the data

Table 7: Examples of institutional statements shared by interviewees coded for each ideal logic type and barcoding behaviour.

	Corporation	Market	Profession	State	Other
Set scope of project	iBOL Nodes should have a steering committee to guide the scope of the project	Individuals contribute data based on their own agenda	Researchers make barcode databases broad enough for their own projects	Researchers collect the number of specimens that is required by their funding agency	No examples
Start collaborations	iBOL Project Leaders should emphasize the legal and ethical issues faced by some countries participating as Nodes	Individuals form contractual one-off arrangements to avoid ongoing responsibilities	Researchers collaborate when projects will have impact or are novel	State agencies should not allow participation in international projects when researchers are required to house their specimens in another country	Developing country researchers negotiate collaborations when they seek to receive benefits from sharing specimens
Collect specimens	Individuals may use museum collections when they are *bona fide* researchers	Any individual is permitted to collect a specimen for barcoding	Researchers find work-arounds when they cannot fulfill specimen-collection permit requirements	Individuals should not ignore state laws and process in the utilization of genetic resources for commercial or non-commercial purposes	Researchers may collect specimens from foreign institutions that house their national specimens, if they have the time and resources to do so

	Corporation	Market	Profession	State	Other
Share Specimens	Repositories should backup duplicates of materials by sharing with other repositories	Individuals share samples with project collaborators for the purpose of joint publication[1]	Researchers share specimens when they meet at conferences	Researchers share specimens in accordance with State laws and process, including export permitting.	Researchers from countries concerned about sharing whole specimens may instead send the amplified PCR product of the barcode region for sequencing
Store specimens	Repositories store DNA extracts for quality control purposes only, when the materials were sent to them for barcoding	Customers paying for DNA barcoding services may choose to not store their specimens	Researchers store DNA materials when the quality is good.	Institutions should not store specimens outside the country when it is in their national interest to house them domestically	Developed nation institutions should not be the default location to store specimens for developing nations
Generate data	Database managers enforce data generation standards through pre-submission data validation	Any individual may contribute data to barcoding efforts	Researchers should contribute data to barcoding efforts and perform quality control	Researchers may generate barcode sequences from their specimens when the have the proper permits to access genetic information	Researchers who are not allowed to export materials can access training programs to develop in-country barcodes[2]

[1] Coded as market and profession

[2] Coded as state and other

	Corporation	Market	Profession	State	Other
Share data	Researchers must agree to release sequence data to GenBank when they are part of the iBOL project	Individuals use open access databases when they want to allow scrutiny	Researchers release data after publication in a scientific journal	Researchers release data to the public domain when government funders require it	Researchers do not release data when the species is medicinal
Store data	Barcode sequencers put barcodes into BOLD	Individuals should store their data where it can be retrieved easily at any time by anyone	Researchers store their own data and materials to have ownership over their research direction	Countries develop mirror sites to store all data domestically[3]	Researchers create a mirror site in their own country when people do not want their data on foreign servers only[4]
Access data	Database managers should not release specimen location coordinates	Individuals accessing public databases should use discretion when the information may not be validated	Professionals should have access to data	Everyone should have access to data	Database managers should provide levels of access to data, highest levels going to those who respect the rules

[3] Coded as state and other

[4] Coded as state and other

I classified interviewees based on the location of their main work affiliation at the time of the interview, and placed them into two groups: Like-Minded Mega Diverse countries (LMMC) and non-LMMC. LMMC is a group of countries established in 2002 to promote their similar interests in protecting biodiversity (South Africa Department of Environmental Affairs, 2016). The LMMC group included China, Colombia, Ghana, India, Indonesia, Kenya, Mexico, and South Africa, and the non-LMMC group included Australia, Canada, New Zealand, United Kingdom, and the US. Despite not being a member of the official LMMC group, I included Ghana in the LMMC group because, as a lower income country in Africa, it more closely aligns with countries like Kenya and South Africa than Australia and the USA.

I used SPSS v. 19 to calculate odds ratios (ORs) and 95% confidence intervals (CI) as measures of association between the institutional statements that non-LMMC interviewees stated and LMMC interviewees. I described the number of logics influencing BOL organization documents using proportion estimates.

I also conducted an in-depth thematic analysis of statements influenced by "other" logics. After selecting each section of transcript that we coded as "other", I identified explicit themes in the text. I linked the text with a description of the theme, and grouped similar themes together. I then separated the text into interviewee categories (LMMC and non-LMMC) to identify dominant themes within each category.

Results

My analysis revealed the type and distribution of institutional logics that influenced the institutional statements. Overall, the types of logics represented in statements varied depending on the source of the statement, and the DNA barcoding behaviour to which the statement referred (Table 8). Here, I detail the institutional statements and logics, describing organizational documents first followed by interview transcripts.

Institutional Statements in BOL organization documents

I identified 64 institutional statements in the twelve BOL documents (Table 8). Most statements were *Norms* (44) or *Strategies* (19). I identified only one *Rule* statement in the documents. The statements focused on behaviours related to building the commons (Figure 8),

including 21 statements about storing data, 15 statements about sharing data, and 12 statements about sharing specimens.

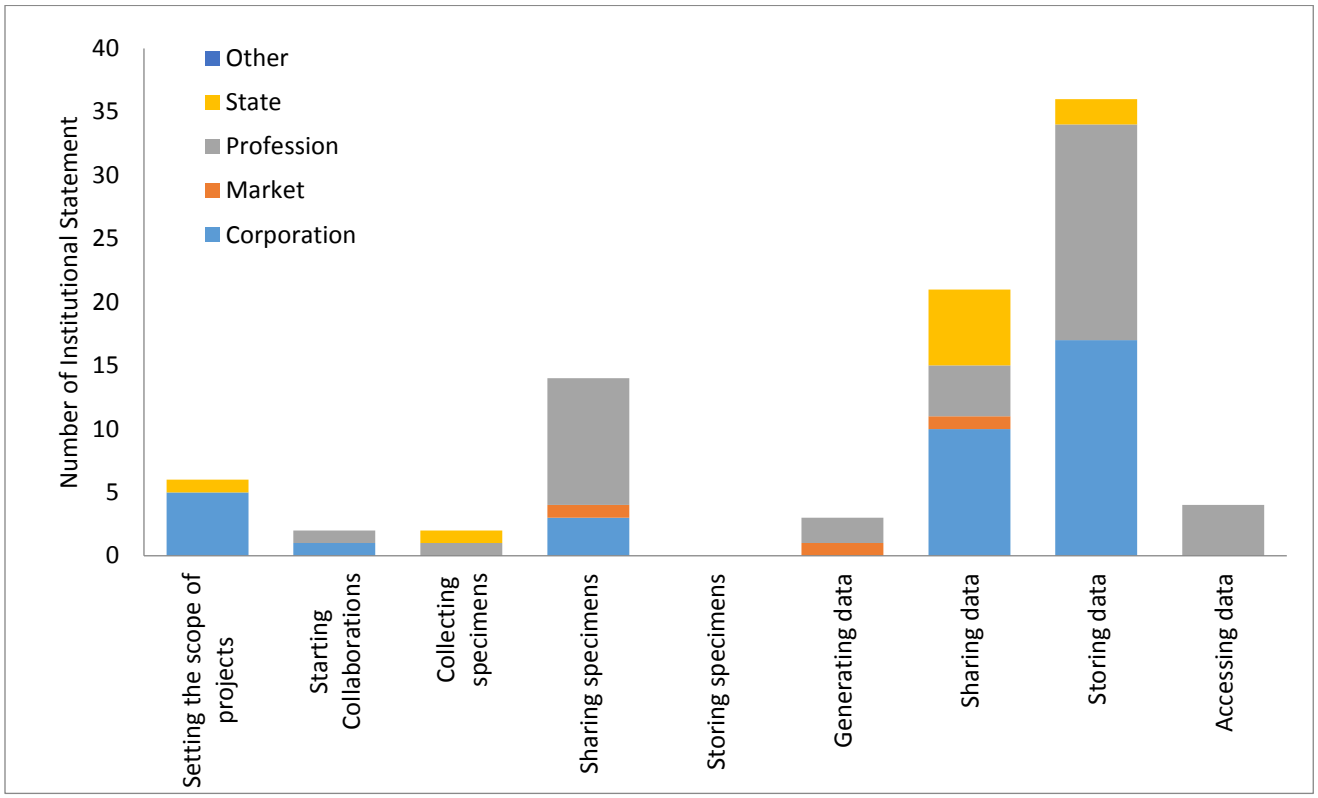

Figure 8. Logic types represented in institutional statements for each barcoding behaviour. Numbers indicate the total number of statements. Note that more than one logic type could influence each statement. As a result, the sum of logic types across behaviour type could be more than the number of total statements.

The most common logics that influenced institutional statements were profession (n=40) and corporation (n=37). The profession logic was a prominent influence on statements about sharing specimens, and generating and storing data (Table 8). The corporation logic most prominently influenced statements about the scope of projects, and data generating and sharing. The state logic had little influence on the institutional statements in organizational documents, except for specimen collection.

Table 8: Frequency of institutional logics influencing institutional statements by barcoding behaviour and statement source.

Statement influenced by:			Corporation		Market		Profession		State		Other	
	Statement source*	n	n	%	n	%	n	%	n	%	n	%
Setting the scope of projects	non-LMMC	6	4	67%	0	0%	0	0%	2	33%	0	0%
	LMMC	8	1	13%	1	13%	7	88%	2	25%	0	0%
	Documents	5	5	100%	0	0%	0	0%	1	20%	0	0%
Starting collaborations	non-LMMC	18	1	6%	9	50%	16	89%	3	17%	2	11%
	LMMC	21	2	10%	4	19%	18	86%	1	5%	4	19%
	Documents	2	1	50%	0	0%	1	50%	0	0%	0	0%
Collecting specimens	non-LMMC	12	1	8%	2	17%	5	42%	5	42%	1	8%
	LMMC	17	2	12%	0	0%	4	24%	15	88%	1	6%
	Documents	2	0	0%	0	0%	1	50%	1	50%	0	0%
Sharing specimens	non-LMMC	21	10	48%	1	5%	9	43%	5	24%	2	10%
	LMMC	31	7	23%	1	3%	11	35%	17	55%	2	6%
	Documents	12	3	25%	1	8%	10	83%	0	0%	0	0%
Storing specimens	non-LMMC	14	3	21%	4	29%	12	86%	3	21%	0	0%
	LMMC	9	2	22%	0	0%	3	33%	3	33%	1	11%
	Documents	0	0	-	0	-	0	-	0	-	0	-
Generating data	non-LMMC	11	4	36%	2	18%	6	55%	1	9%	1	9%
	LMMC	5	1	20%	1	20%	2	40%	1	20%	0	0%
	Documents	3	1	33%	1	33%	3	100%	0	0%	0	0%
Sharing data	non-LMMC	35	9	26%	3	9%	14	40%	9	26%	2	6%
	LMMC	33	2	6%	2	6%	19	58%	11	33%	5	15%
	Documents	15	10	67%	1	7%	4	27%	6	40%	0	0%
Storing data	non-LMMC	6	2	33%	0	0%	5	83%	0	0%	0	0%
	LMMC	9	1	11%	0	0%	7	78%	2	22%	2	22%
	Documents	21	17	81%	0	0%	17	81%	2	10%	0	0%
Accessing data in the commons	non-LMMC	10	0	0%	5	50%	0	0%	6	60%	1	10%
	LMMC	8	0	0%	1	13%	0	0%	6	75%	1	13%
	Documents	4	0	0%	0	0%	4	100%	0	0%	0	0%

*Statement sources include: interview transcripts from non-likeminded megadiverse country interviewees (non-LMMC), interview transcripts from likeminded megadiverse country interviewees (LMMC), and BOL organization documents (documents).

Institutional statements articulated by interviewees

The 35 interviewees expressed 274 institutional statements about DNA barcoding behaviours. Only five statements were *Rules*; most were *Norms* (127) or *Strategies* (142).

LMMC and non-LMMC interviewees expressed a similar number of institutional statements (141 and 133, respectively), distributed similarly across behaviours (Figure 9, Table 9). However, there was variation in the pattern of logics that influenced the institutional statements given by each interviewee type (Figure 10). Compared to LMMC interviewee statements, non-LMMC statements had 2.3 and 3.0 times the odds of being influenced by corporation and market logics, respectively, and 53% of the odds of being influenced by state logics (Table 9).

Table 9: Relative odds of statements types comparing interviewees from non-likeminded megadiverse countries to the reference group of interviewees from likeminded megadiverse countries.

	OR	95% CI
Behaviour		
Setting the scope of projects	0.79	0.27 - 2.3
Starting Collaborations	0.89	0.45 - 1.8
Collecting specimens	0.72	0.33 - 1.6
Sharing specimens	0.67	0.36 - 1.2
Storing specimens	1.7	0.72 - 4.1
Generating data	2.5	0.83 - 7.3
Sharing data	1.2	0.68 - 2.0
Storing data	0.69	0.24 - 2.0
Accessing data	1.4	0.52 - 3.5
Logics		
Corporation	2.3	1.3 - 4.4
Market	3.0	1.4 - 6.6
Profession	0.97	0.60 - 1.6
State	0.53	0.32 - 0.88
Other	0.61	0.26 - 1.4

OR, odds ratio

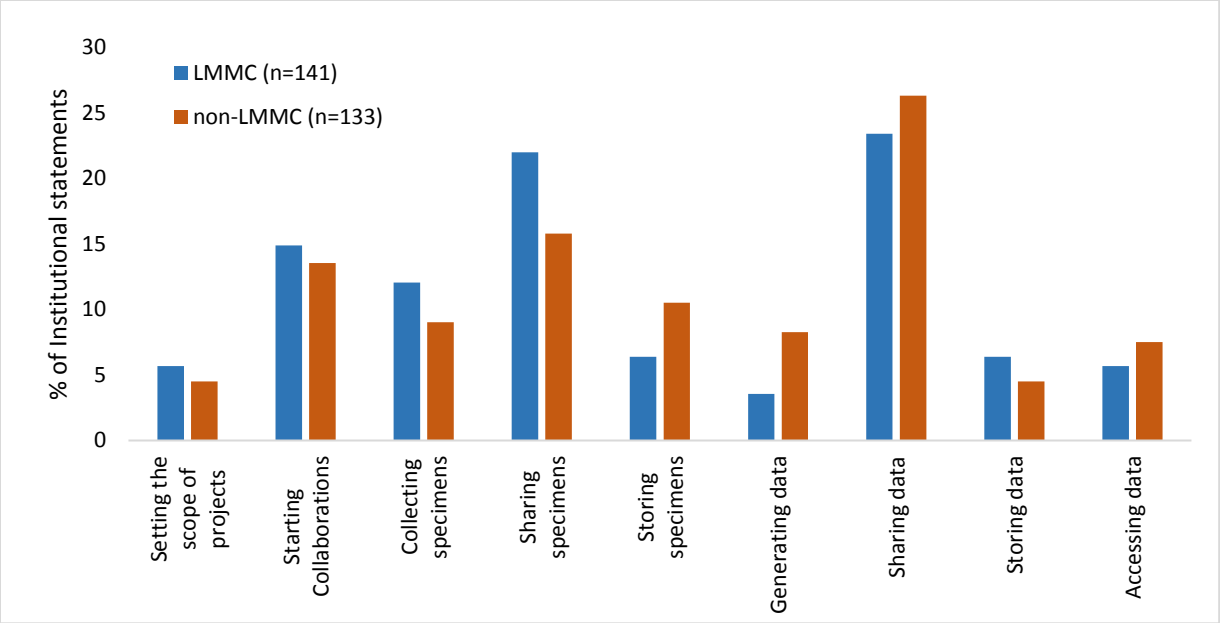

Figure 9. The distribution of DNA barcoding behaviour types across articulated institutional statements, by interviewee status.

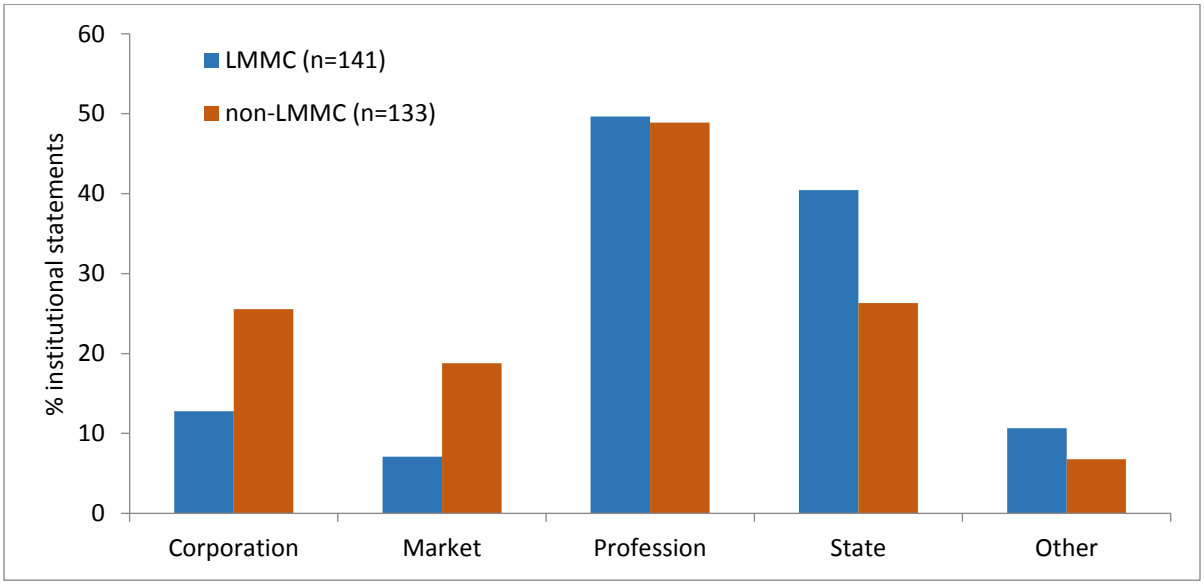

Figure 10. The distribution of logic types influencing institutional statements by interviewee status.

Next, I examined each of the barcoding behaviours separately to provide an in-depth depiction of the logics underlying the institutional statements.

Setting the scope of DNA barcoding activities

Only 14 statements were about scope setting behaviours. Different logics influenced the 8 LMMC statements and the 6 non-LMMC interviewee statements. Corporation (4) and state (2) logics influenced non-LMMC, expressed as fulfilling iBOL project targets for barcoding and undertaking projects that create public good. The profession logic most frequently (7) influenced LMMC statements about scope-setting statements, citing considerations such as academic interests and project feasibility.

Collaborating for DNA barcode projects

LMMC and non-LMMC statements about collaborations were largely influenced by the profession logic (21 and 18 statements, respectively), with frequent referrals to the professional benefits that could be gained through collaborations with some overlap with non-professional personal considerations like relationships, personalities, and trust (Figure 11). Non-LMMC interviewee statements had 3.4 times the odds (Table 10) of being influenced by the market logic (e.g. Researchers collaborate when they like the person).

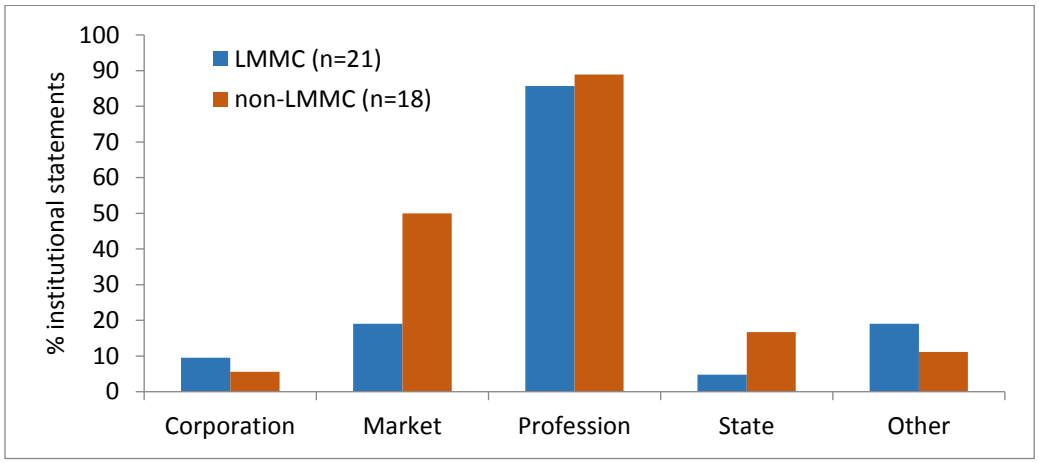

Figure 11. The distribution of logic types influencing institutional statements about collaboration behaviours for barcode projects, by interviewee status.

Collecting, sharing, and storing specimens

Different logics influenced LMMC and non-LMMC statements on collecting, sharing, and storing specimens. More statements focused on specimen sharing (52) than collection (29) or storage (23).

Collecting specimens

Both interviewee groups expressed statements about collecting specimens that were largely influenced by the state logic (Figure 12), although non-LMMC statements had 13% of the odds of being influenced by the state compared to LMMC statements (Table 10). While both groups focused on permitting and state policies, LMMC interviewees were more specific about the requirement to adhere to state policies.

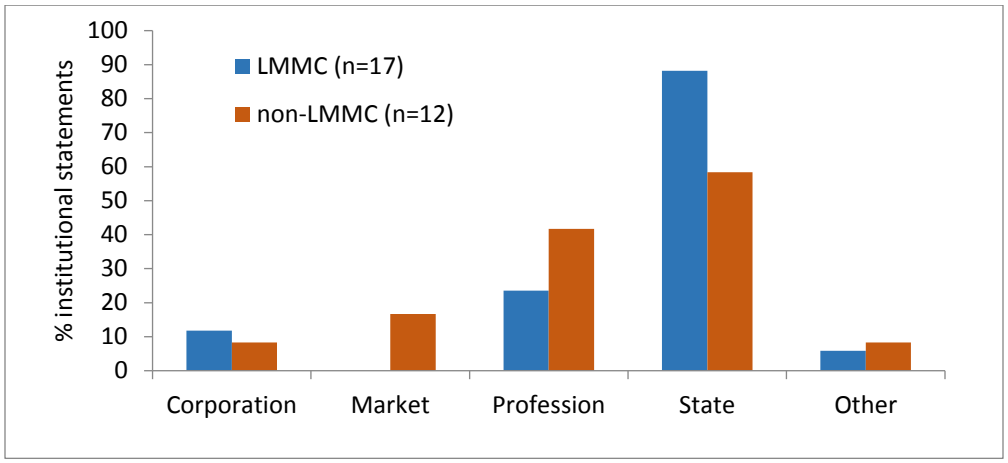

Figure 12. The distribution of logic types influencing institutional statements about specimen collection behaviours, by interviewee status.

Table 10: Relative odds of statements pertaining to behavior types within categories of logics type comparing interviewees from non-likeminded megadiverse countries to the reference group of interviewees from likeminded megadiverse countries.

	Corporation		Market		Profession		State		Other	
Behaviour	**OR**	95% CI	**OR**	95% CI	**OR**	95% CI	**OR**	95% CI	**OR**	95% CI
Setting the scope of projects	**14**	0.94, 208			**0.13**	0.02, 0.78	**1.5**	0.15, 15	**-**	
Starting Collaborations	**0.56**	0.05, 6.7	**3.4**	0.81, 14	**0.58**	0.11, 3.0	**4.0**	0.38, 42	**0.75**	0.11, 5.1
Collecting specimens	**0.68**	0.05, 8.5	**-**		**2.3**	0.47, 12	**0.13**	0.02, 0.86	**1.5**	0.08, 26
Sharing specimens	**3.1**	0.94, 10	**1.5**	0.01, 25	**1.1**	0.36, 3.5	**0.29**	0.09, 1.0	**1.5**	0.20, 12
Storing specimens	**0.95**	0.13, 7.2	**-**		**7.3**	1.1, 48	**0.55**	0.08, 3.6	**-**	
Generating data	**2.3**	0.19, 28	**0.89**	0.06, 13	**1.8**	0.21, 15	**0.40**	0.02, 8.1	**-**	
Sharing data	**5.4**	1.1, 27	**1.5**	0.23, 9.3	**0.70**	0.27, 1.8	**0.69**	0.24, 2.0	**0.34**	0.06, 1.9
Storing data	**4.0**	0.27, 59	**-**		**1.4**	0.10, 20	**-**		**-**	
Accessing data	**-**		**7.0**	0.61, 80	**-**		**0.50**	0.07, 3.8	**0.78**	0.04, 15

"-" indicates that ORs could not be calculated due to low numbers

Sharing specimens

LMMC and non-LMMC statements about sharing specimens had similar odds of being influenced by the profession logic (Figure 13). However, non-LMMC statements had 3.1 times the odds of being influenced by the corporation logic, and 29% of the odds of being influenced by the state logic (Table 10).

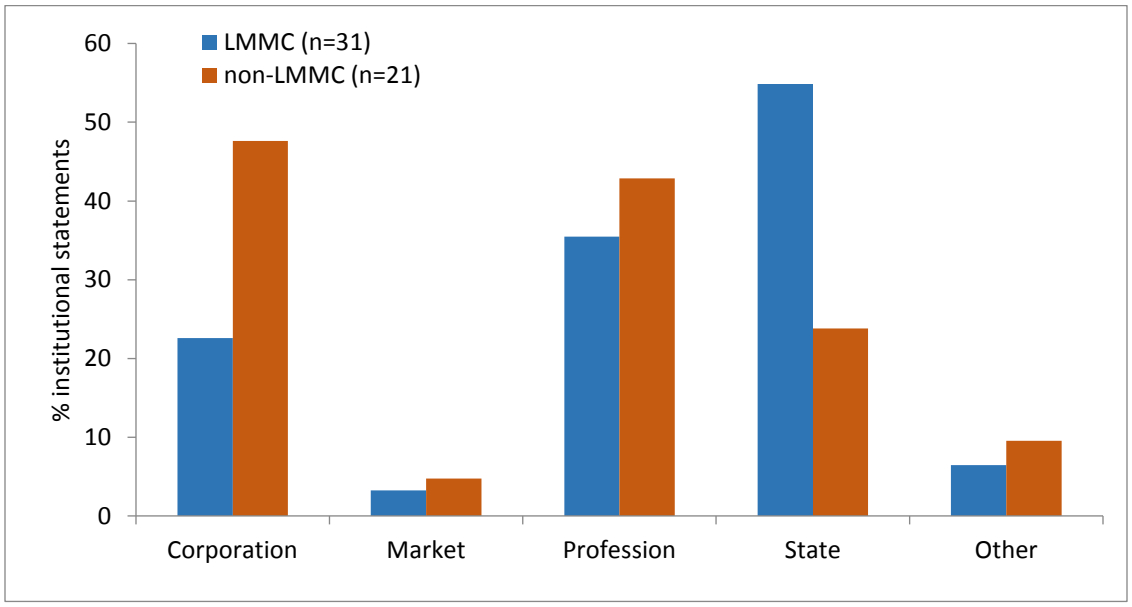

Figure 13. The distribution of logic types influencing institutional statements about specimen sharing behaviours, by interviewee status.

Storing specimens

Non-LMMC statements about storing specimens were often influenced by the profession logic (Table 8). The nine LMMC statements did not have any dominant influence from any logics. Statements from both LMMC and non-LMMC interviewees that were informed by the profession logic were typically about the need to store duplicate specimens in multiple locations for quality control and back-up. LMMC statements influenced by the state logic focused on storing specimens domestically to protect national sovereignty over genetic resources.

Generating, sharing, and storing data

A variety of logics also influenced LMMC and non-LMMC statements about generating, sharing, and storing. Many more statements focused on data sharing (68) than generation (16) or storage (15).

Generating data

Interviewees expressed few statements (16) that described behaviours related to generating data. There were no dominantly influential logics on LMMC statements (Table 8). There were 11 statements about data generation from non-LMMC interviewees, most influenced by profession (6) and corporation (4) logics. These statements focused on maintaining and improving data quality.

Sharing data

Institutional statements about sharing data were the most frequently identified, with 68 statements (Table 8). The state logic influenced data sharing statements at a similar frequency for both LMMC and non-LMMC interviewees (Figure 14). The profession logic was the most common influence on LMMC statements, with interviewees emphasizing the need to make their own judgements about what data to share based on their needs, including the ability to make professional contributions (e.g., publications). Non-LMMC statements had 5.4 times the odds of being influenced by the corporation logic (Table 10). Interviewees often expressed this as the requirement to release data when using iBOL sequencing facilities to produce DNA barcodes.

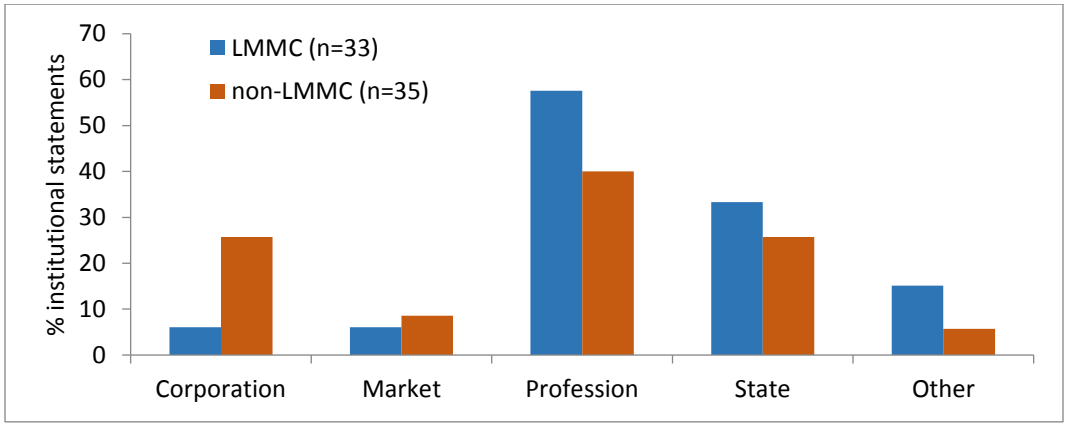

Figure 14. The distribution of logic types influencing institutional statements about data sharing behaviours, by interviewee status.

Storing data

Both LMMC and non-LMMC statements (9 and 6, respectively) regarding storing data were most influenced by the profession logic (Table 8), emphasizing the professional requirement to keep backups and ensure data quality.

Accessing data in the commons

Eight LMMC statements and ten non-LMMC statements about accessing data behaviours were influenced by the state logic (Table 8). LMMC statements emphasized that specific types of data should be accessible for the public good, such as information that would help the public understand biodiversity. Although the estimate is imprecise, non-LMMC interviewees had 7.0 times the odds of being influenced by market logics (Table 10).

Global inequities as a filtering effect on logics

I identified 38 institutional statements that had apparent influences beyond the four logics I defined for my study (27 from LMMC interviewees and 11 from non-LMMC interviewees). Many of these statements related to historical tensions between developed and developing countries, and persistent inequities that have resulted. LMMC statements often related to an imbalance between rich and poor countries and institutions, and the historical imbalances that continued to affect the ability of developed countries to participate in research endeavors as equal partners. Interviewees mentioned: problems with accessing materials housed in foreign institutions, the desire to store project resources domestically (both data and specimens), an imbalance between capacity to participate in barcode projects, the need to protect sensitive data from misappropriation (especially traditional knowledge associated with genetic resources), and the need for all DNA barcoders to be aware of rules for access and benefits sharing of genetic resources. One interviewee stated:

> *I think [they will not store Mexican specimens in a foreign repository] because the people firstly will think to develop locally. You see, they are angry because the server with the BOLD system is in Canada - that is why we are going to create a mirror in Mexico because they don't want even the information to be in a server that is outside the country – Researcher, Mexico*

Most of the non-LMMC statements influenced by "other" logics also acknowledged imbalances, although much more vaguely. These statements acknowledged the need for capacity building in lower income countries, suggested options to ship PCR products or store specimens in the country of origin when concerns about shipping materials exist, and favoured negotiated agreements to ensure developing country participants benefitted from specimen sharing. As one

interviewee stated, “Institutions in the [global] north are going to have to get used to being more open and more careful about documenting where they got specimens from, who they lend specimens to, all that has to be rethought and made much more transparent and accountable” – Researcher, USA.

I also noted that interviewees expressed logics differently depending on whether they were from LMMC or non-LMMC. The focus of state logics in particular varied depending on whether LMMC or non-LMMC expressed the statement. When LMMC interviewees discussed sharing specimens, they emphasized sovereignty over genetic resources, whereas non-LMMC interviewees focused on exporting and permitting. Non-LMMC interviewees emphasized the importance of open-access data to produce public benefits, whereas LMMC interviewees emphasized the need for state control of sensitive information.

Discussion

The DNA barcoding community is a global and heterogeneous network of individuals and organizations in different countries that work collectively to build and maintain infrastructure and resources for rapid species identification in support of biodiversity research and other applications. To date, organizations based in Canada and the US have led the international coordination of the DNA barcode commons. Various *Rules*, *Norms* and *Strategies* influence expectations for how individuals participate in the commons. These range from undocumented and unenforceable scientific community standards to organizational and funder policies to national and international laws. However, individuals participating in DNA barcoding do not necessarily share expectations, nor do these expectations necessarily align with the commons and its supporting *Rules*, *Norms* and *Strategies*. As efforts continue to institutionalize the commons with the establishment of the ISBOL, it is timely to consider which expectations participants share and do not share. Such understanding can inform the development of institutional governance, its policies, and appropriate incentives for participation. Such incentives should encourage contributions of resources to the commons, and distribute the risks and benefits in a manner that builds trust among the heterogeneous participants.

Here I discuss the implications of heterogeneity in the DNA barcode commons; current gaps between rules-in-use of participants and the rules-in-form; which behaviours are

underpinned by shared expectations; and what types of rules could be implemented to bridge the gap in expectations of participants.

Implications of heterogeneity for the DNA Barcoding Community

The relationship I identified between logics and rules-in-use suggests that the heterogeneity of the DNA barcoding community contributes to governance challenges. The challenge arises because, like previous work has suggested (Ruttan 2006; Varughese and Ostrom 2001; Hayo and Vollan 2012), this heterogeneous community is made of individuals with different expectations for what constitutes appropriate behaviour, and different ideas about which rules should govern behaviour. While appropriate institutional arrangements can mitigate a lack of shared expectations among participants (Varughese and Ostrom 2001), I also found that formal institutional arrangements in the community were insufficient. These statements identified in organizational documents were more influenced by corporation logics than the LMMC interviewee statements, and failed to provide consequences for incorrect actions.

Because of its global nature, the DNA barcode commons manifests multiple forms of heterogeneity, including cultural, historical, geographical, technical, and financial differences. While studies have identified challenges resulting from many of these forms of heterogeneity (Lindahl 2012; Ruttan 2006; Vedeld 2000; Poteete and Ostrom 2004), wealth is one relevant type of heterogeneity that has been shown to be favourable in specific contexts like sustaining natural resource commons (Rattan 2005). Rattan hypothesized this was possible because wealthy actors with a stake in the resource could use their wealth to ensure the continued availability of the resource. Researchers categorize this sustainability as a positive outcome, regardless of how equitably actors used or contributed to the resource (Ruttan 2006). However, the characteristics of a successful knowledge commons are very different from natural resource commons, and are more dependent on collective action to establish the resource (Ostrom 2005a; Frischmann et al. 2014).

A robust DNA barcode commons relies on collective action to coordinate distribution of the resources needed to build the commons, which in turn is dependent on participants trusting that they will be able to derive equitable benefits from the created resource. Relying on collective action makes a knowledge commons more susceptible to problems associated with establishing

shared expectations. While participants from high income countries may provide the financial resources to develop a knowledge commons, they are unable to obligate participations from people in lower income countries. Stimulating this collective action requires effective rules; otherwise, potential contributors of DNA barcodes from lower income countries may not be able (or willing) to add their resources to the commons.

In cases such as DNA barcoding in which it is difficult for participants to engage actively in creating effective governance (Ostrom 2005), hierarchy and "legitimate authority" can have a positive impact on establishing shared norms (Schweik and English 2012; Fleischman et al. 2014). For the DNA barcoding community, the impending development of the International Society for the Barcode of Life (Castle et al. 2015) offers an opportunity to establish legitimate authority that represents the diverse interests of the community, and ensures that LMMC perspectives are also included in the development of an enforceable governance structure.

Identifying appropriate institutional statements for the DNA Barcoding Community

LMMC and non-LMMC interviewees and BOL organizational documents had a similar distribution of statements across behaviour types, suggesting a shared perspective about which behaviours require institutional arrangements. However, the logics influencing those statements varied across interviewees, BOL documents, and behaviour types. Institutional statements that reflect the expectations of commons participants are more likely to be viewed as legitimate, and therefore, followed (Tyler 1990). I observed the most conflicts between logics that influenced controversial behaviours: collecting specimens for research, sharing specimens, and sharing data. Establishing enforceable institutional arrangements for these behaviours should be a priority for the ISBOL. Based on conflicts between logics influencing BOL documents and the interviewees, I suggest guidelines for institutional statements for DNA barcoding in Table 11.

Table 11: Suggested guidelines for institutional statements for DNA barcoding based on the grammar of institutions *ADICO* syntax

Behaviour	Suggestion	Justification
Setting the scope of activities	Establish *Aims* that relate to professional targets as well as corporation goals	LMMC interviewees were not influenced by corporation logic, despite participation in BOL organizations
Collaborating	Set out *Conditions* to limit how participants extract personal benefits from collaborations	Non-LMMC interviewees were more likely to be influenced by market logic, indicating a potential need to mitigate the impact of seeking personal gain within barcoding partnerships
Collecting specimens	*Conditions* should be used to indicate *Aims* for collecting specimens are only permitted when abiding by state policies that apply to where the specimen is being collected from, or when individuals receive training on laws that govern collection	Interviewees did not share expectations for collecting specimens. Given sensitivities with respect to specimen collection, BOL organizations should specify that actors are expected to abide by state policies, and should develop an educational course for potential DNA barcoders to learn about relevant laws and ABS standards
Sharing specimens	*Conditions* should be used to indicate that *Aims* for sharing specimens are only permitted when abiding by appropriate state policies	Interviewees did not share expectations. Given the sensitivity toward sharing specimens, organizations need to specify that actors are expected to abide by state policies, and facilitate enforcing compliance
Storing specimens	*Aims* should focus on storing collections in the country of origin, possibly including *Conditions* under which storage in foreign repositories is permitted	Interviewees shared expectations about storage for data quality, while the different expectations of where to locate storage facilities could create conflicts
Generating data	*Aims* should follow guidelines of professional organizations for barcode data standardization	Interviewees were most influenced by profession logics and would respect standardization

Behaviour	Suggestion	Justification
Sharing data	*Aims* should emphasize the professional goals of increasing visibility and use of data, and *Conditions* should indicate which kinds of sensitive data can be omitted from barcode record requirements	The large number of BOL organization rules about data sharing did not influence participants' behaviours. Many individuals indicated they would feel more comfortable sharing less sensitive information.
Storing data	*Aims* should emphasize the professional requirement to maintain data quality and backups	Interviewees shared expectations about storing data to maintain quality, but existing organizational documents were based on leaders of the organization enforcing their expectations on individuals who participate.
Accessing data in the commons	*Aims* should consider public access and efficiency, but with *Conditions* that specify when permissions to access the commons may be revoked	Interviewees did not share expectations about accessing data; more specific institutional statements could facilitate sharing data

Overall, new institutional arrangements for the DNA barcode commons could calibrate expectations across actors. This will require explicit recognition of the lasting effects of historic inequities, and the implications of this history for current international partnerships. There were no institutional statements that took global inequities into account in any BOL documents. As long as participants remain influenced by inequities across countries, institutional arrangements need to reflect the inequities to avoid the risk that potential contributors will view the organization as illegitimate. Even though many aspects of a DNA barcode commons are non-rivalrous, many of the inputs, such as research funding, are rivalrous. The 'leveling of the playing field' is not about making the outputs freely available, but about equalizing the cost at which contributors get equivalent benefits (Frischmann et al. 2014).

Conclusions

Contributions to methods

Research on heterogeneity in communities that engage in knowledge commons has focused on selecting a particular source of heterogeneity, and quantifying the impact that the selected heterogeneous variable has on a measured outcome. Outcomes in studies of the impact of heterogeneity on knowledge commons are usually selected measures of the quality of the resource or the level of collective action (Ruttan 2008). While these methods have identified that heterogeneity has a measurable impact on important outcomes, they fail to illustrate *how* heterogeneity produces the outcome, or suggest *how* to mitigate undesired outcomes. The analytical framework I developed that combines the grammar of institutions and institutional logics helps identify where participant expectations and actions conflict with each other and with formal policies. This approach identified which behaviours would benefit from direct institutional arrangements without requiring a priori knowledge of or guesswork about which type of heterogeneity was salient in the knowledge commons of interest.

While my framework does not quantify or model the impacts of heterogeneity, it is useful for understanding the mechanism through which heterogeneity affects a commons. Previous work has suggested that heterogeneity impacts collective action because of a lack of shared expectations and common ground (Ruttan 2006). Characterizing institutional statements using

logics helps pinpoint expectations of behaviours that participants do and do not share. This understanding can improve: 1) quantitative evaluations by providing more sophisticated inputs for quantitative models and 2) a community's ability to overcome heterogeneity challenges by identifying behaviours that lack shared expectations and may generate conflicts. In the case of DNA barcoding, my approach identified which behaviours will likely require more precise institutional statements based on a lack of shared expectations. Individuals can create institutional statements to establish shared expectations that might otherwise naturally exist in a more homogeneous group.

Limitations

This study has a number of limitations. I interviewed participants over several years, and, therefore, responses may have been influenced by factors that changed over time. However, the overall structure of DNA barcoding governance did not change during the course of data collection. Individual behaviour might also depend on experience and length of time a participant has engaged in barcoding efforts, and I did not account for these potential differences between interviewees. I also was the only researcher who interpreted the data regarding the "other" logics. While it is possible that another researcher may have drawn different conclusions, my use of qualitative software to track my analytical decisions enables confirmability of my findings.

Contributions to theory

Knowledge commons

Scholars suggest that all basic attributes of a knowledge commons, including resources, community, goals, and objectives, are heavily dependent on the history of the commons (Frischmann et al. 2014). My data supported this perspective, given actors' explicit references to the negative history of genetic resource misappropriation by the biodiversity research community as influences on their institutional statements.

My data also demonstrate that institutional statements reflect the heterogeneity of participants. While many studies of heterogeneity have focused on measuring specific impacts on outcome variables, my study demonstrates the utility of examining how heterogeneity broadly influences decisions that participants make in a commons. Studying institutional arrangements is appealing because, unlike many characteristics of commons, institutional statements can be

modified (Crawford and Ostrom 1995). While the grammar of institutions allows researchers to demonstrate where institutional statements are too vague or weak to be enforceable and therefore effective, logics analysis enables recommendations about how institutional statements could mitigate the negative effects of heterogeneity.

Institutional logics

My study demonstrated the practical utility of using ideal types of logics as an analytical tool, and further supports the conceptualization that multiple logics simultaneously influence the behaviours of individuals within organizations (Goodrick and Reay 2011). My study also makes one main contribution to institutional logics theory: it demonstrates how filters modify the impact of institutional logics, suggesting that global inequities are an important filter in organizations that span both high and low income countries.

Research examining institutional logics has failed to develop robust theories for how overarching societal orders might change when considered in an international context. Although I found evidence that institutional logics are universal and found across countries and cultures, often showing quite similar patterns, I also found evidence that global inequities may act as a filtering mechanism through which logics exert influence. I identified global inequities much less frequently than each logic I included in my analytical framework, and global inequities never appeared to be a dominant influence on any behaviours. Nonetheless, I identified the theme of global inequities consistently throughout statements pertaining to each behaviour. I recommend that researchers explicitly consider global inequities as a filter mechanism in subsequent studies where it may have an impact, especially when researchers expect state logics to be an important influence.

References

Agrawal, A., and C. C. Gibson. (2001). Introduction: The Role of Community in Natural Resource Conservation. In: *Communities and the Environment: Ethnicity, Gender, and the State in Community-Based Conservation*. eds. A. Agrawal and C. C. Gibson, 1–31. New Brunswick, New Jersey, and London: Rutgers University Press.

Australian Government. (2008). Biodiversity Hotspots. Text. *Department of the Environment and Energy*. September 2. http://www.environment.gov.au/biodiversity/conservation/hotspots.

Balooni, K., J. F. Lund, C. Kumar, and M. Inoue. (2010). Curse or Blessing? Local Elites in Joint Forest Management in India's Shiwaliks. *International Journal of the Commons*. 4(2): 707–28. http://dx.doi.org/10.18352/ijc.217.

Barbuto, M., A. Galimberti, E. Ferri, M. Labra, R. Malandra, P. Galli, and M. Casiraghi. (2010). DNA Barcoding Reveals Fraudulent Substitutions in Shark Seafood Products: The Italian Case of "palombo" (Mustelus Spp.). *Food Research International*. 43(1): 376–81. http://dx.doi.org/10.1016/j.foodres.2009.10.009.

Basurto, X., G. Kingsley, K. McQueen, M. Smith, and C. M. Weible. (2010). A Systematic Approach to Institutional Analysis: Applying Crawford and Ostrom's Grammar. *Political Research Quarterly*. 63(3): 523–37. http://dx.doi.org/10.1177/1065912909334430.

Benin, S., and J. Pender. (2006). Collective Action in Community Management of Grazing Lands: The Case of the Highlands of Northern Ethiopia. *Environment and Development Economics*. 11(1): 127–49. http://dx.doi.org/10.1017/S1355770X05002688.

Black, M. (1962). *Models and Metaphors: Studies in Language and Philosophy*. 1st Edition edition. Cornell University Press.

Brito, D. L., M. D. Intriligator, and E. Sheshinski. (1997). Privatization and the Distribution of Income in the Commons. *Journal of Public Economics*. 64(2): 181–205. http://dx.doi.org/10.1016/S0047-2727(96)01611-8.

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.

Burlando, R. M., and F. Guala. (2005). Heterogeneous Agents in Public Goods Experiments. *Experimental Economics*. 8(1): 35–54. http://dx.doi.org/10.1007/s10683-005-0436-4.

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.

Centre for Biodiversity Genomics. (2016). Organization. http://biodiversitygenomics.net/about/organization/.

Consortium for the Barcode of Life. (2016). What Is DNA Barcoding? http://www.barcodeoflife.org/content/about/what-dna-barcoding.

Costa, F. O., and G. R. Carvalho. (2007). The Barcode of Life Initiative: Synopsis and Prospective Societal Impacts of DNA Barcoding of Fish. *Life Sciences Society and Policy*. 3(2): 29. http://dx.doi.org/10.1186/1746-5354-3-2-29.

Crawford, S., and E. Ostrom. (1995). A Grammar of Institutions. *American Political Science Review*. 89(3): 582–600. http://dx.doi.org/10.2307/2082975.

Crawford, S., and E. Ostrom. (2005). A Grammar of Institutions. In: *Understanding Institutional Diversity*. ed. E. Ostrom, 137–74. Princeton University Press.

Dedeurwaerdere, T. (2010). 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.

Fischbacher, U., and S. Gaechter. (2006). Heterogeneous Social Preferences and the Dynamics of Free Riding in Public Goods. *SSRN*., March. http://papers.ssrn.com/abstract=874928.

Fleischman, F. D., N. C. Ban, L. S. Evans, N. C. Ban, L. S. Evans, G. Epstein, S. Villamayor-Tomas, G. Garcia-Lopez, and S. Villamayor-Tomas. (2014). Governing Large-Scale Social-Ecological Systems: Lessons from Five Cases. *International Journal of the Commons*. 8(2): 428–56. http://dx.doi.org/10.18352/ijc.416.

Freidson, E. (2001). *Professionalism, the Third Logic*. Chicago: The University of Chicago Press.

Friedland, R., and R. Alford. (1991). Bringing Society Back In: Symbols, Practices and Institutional Contradictions. In: *The New Institutionalism in Organizational Analysis*. eds. W. Powell and P. Dimaggio, 232–63. University Of Chicago Press.

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.

Goodrick, E. (2002). From Management as a Vocation to Management as a Scientific Activity: An Institutional Account of a Paradigm Shift. *Journal of Management*. 28(5): 649–68. http://dx.doi.org/10.1177/014920630202800505.

Goodrick, E., and T. Reay. (2011). Constellations of Institutional Logics Changes in the Professional Work of Pharmacists. *Work and Occupations*. 38(3): 372–416. http://wox.sagepub.com/content/38/3/372.short.

Gross, M. (2012). Barcoding Biodiversity. *Current Biology*. 22(3): R73–76. http://dx.doi.org/10.1016/j.cub.2012.01.036.

Hayo, B., and B. Vollan. (2012). Group Interaction, Heterogeneity, Rules, and Co-Operative Behaviour: Evidence from a Common-Pool Resource Experiment in South Africa and Namibia. *Journal of Economic Behavior and Organization*. 81(1): 9–28. http://dx.doi.org/10.1016/j.jebo.2011.09.002.

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): 313–21. http://dx.doi.org/10.1098/rspb.2002.2218.

Hensmans, M. (2003). Social Movement Organizations: A Metaphor for Strategic Actors in Institutional Fields. *Organization Studies*. 24(3): 355–81. http://dx.doi.org/10.1177/0170840603024003908.

Hoffman, A. J. (1999). Institutional Evolution and Change: Environmentalism and the U.S. Chemical Industry. *The Academy of Management Journal*. 42(4): 351–71. http://dx.doi.org/10.2307/257008.

iBOL. (2015a). Partner Nations. *International Barcode of Life*. http://ibol.org/about-us/partner-nations/.

iBOL. (2015b). Project Governance. The Organization of the iBOL Team. *International Barcode of Life*. http://ibol.org/about-us/project-governance/#.

iBOL. (2015c). Sponsors | iBOL. *International Barcode of Life*. http://ibol.org/about-us/sponsors/.

iBOL. (2015d). What Is iBOL | iBOL. *International Barcode of Life*. http://ibol.org/about-us/what-is-ibol/.

Jansen, H. (2010). The Logic of Qualitative Survey Research and Its Position in the Field of Social Research Methods. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*. 11(2). http://www.qualitative-research.net/index.php/fqs/article/view/1450.

Jaskiewicz, P., K. Heinrichs, S. B. Rau, and T. Reay. (2015). To Be or Not to Be: How Family Firms Manage Family and Commercial Logics in Succession. *Entrepreneurship Theory and Practice*. http://dx.doi.org/10.1111/etap.12146.

Lee, M.-D. P., and M. Lounsbury. (2015). Filtering Institutional Logics: Community Logic Variation and Differential Responses to the Institutional Complexity of Toxic Waste. *Organization Science*, February. http://dx.doi.org/10.1287/orsc.2014.0959.

Li, D.-Z., 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.

Lindahl, T. (2012). Coordination Problems and Resource Collapse in the Commons - Exploring the Role of Knowledge Heterogeneity. *Ecological Economics*. 79: 52–59. http://dx.doi.org/10.1016/j.ecolecon.2012.04.016.

Lounsbury, M. (2002). Institutional Transformation and Status Mobility: The Professionalization of the Field of Finance. *Academy of Management Journal*. 45(1): 255–66. http://dx.doi.org/10.2307/3069295.

Lounsbury, M. (2007). A Tale of Two Cities: Competing Logics and Practice Variation in the Professionalizing of Mutual Funds. *Academy of Management Journal*. 50(2): 289–307. http://dx.doi.org/10.5465/AMJ.2007.24634436.

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): 254–73. http://dx.doi.org/10.1089/omi.2013.0158.

Ostrom, E. (1987). Institutional Arrangements for Resolving the Commons Dilemma: Some Contending Approaches. *The Question of the Commons. The Culture and Ecology of Communal Resources*. 250–265. http://dlc.dlib.indiana.edu/dlc/handle/10535/2274.

Ostrom, E. (2005). *Understanding Institutional Diversity*. Princeton, New Jersey: Princeton University Press.

Poteete, A. R., M. A. Janssen, and E. Ostrom. (2010a). Meta-Analysis: Getting the Big Picture through Synthesis. In: *Working Together: Collective Action, the Commons, and Multiple Methods in Practice*.89–112. Princeton University Press.

Poteete, A. R., M. A. Janssen, and E. Ostrom. (2010b). Overcoming Methodological Challenges. In: *Working Together: Collective Action, the Commons, and Multiple Methods in Practice*.3–30. Princeton University Press.

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): 435–61. 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.

Reay, T., and C. R. Hinings. (2009). Managing the Rivalry of Competing Institutional Logics. *Organization Studies*. 30(6): 629–52. http://dx.doi.org/10.1177/0170840609104803.

Reay, T., and C. R. (Bob) Hinings. (2005). The Recomposition of an Organizational Field: Health Care in Alberta. *Organization Studies*. 26(3): 351–84. http://dx.doi.org/10.1177/0170840605050872.

Reay, T., and C. Jones. (2016). Qualitatively Capturing Institutional Logics. *Strategic Organization*. 14(4): 441–54. http://dx.doi.org/10.1177/1476127015589981.

Ruttan, L. M. (2006). Sociocultural Heterogeneity and the Commons. *Current Anthropology*. 47(5): 843–53. http://dx.doi.org/10.1086/507185.

Ruttan, L. M. (2008). Economic Heterogeneity and the Commons: Effects on Collective Action and Collective Goods Provisioning. *World Development*. 36(5): 969–85. http://dx.doi.org/10.1016/j.worlddev.2007.05.005.

Sakane, N., M. T. van Wijk, M. Langensiepen, and M. Becker. (2014). A Quantitative Model for Understanding and Exploring Land Use Decisions by Smallholder Agrowetland Households in Rural Areas of East Africa. *Agriculture, Ecosystems and Environment*. 197: 159–73. http://dx.doi.org/10.1016/j.agee.2014.07.011.

Schweik, C., and R. English. (2012). *Internet Success: A Study of Open-Source Software Commons*. Cambridge, Massachusetts; London, England: MIT Press. https://mitpress.mit.edu/books/internet-success.

Shiferaw, B., T. A. Kebede, and V. R. Reddy. (2012). Community Watershed Management in Semiarid India: The State of Collective Action and Its Effects on Natural Resources and Rural Livelihoods. In: *Collective Action and Property Rights for Poverty Reduction: Insights from Africa and Asia*.153–88. University of Pennsylvania Press.

Siddiki, S., X. Basurto, and C. M. Weible. (2012). Using the Institutional Grammar Tool to Understand Regulatory Compliance: The Case of Colorado Aquaculture. *Regulation & Governance*. 6(2): 167–88. http://dx.doi.org/10.1111/j.1748-5991.2012.01132.x.

Siddiki, S., C. M. Weible, X. Basurto, and J. Calanni. (2011). Dissecting Policy Designs: An Application of the Institutional Grammar Tool. *Policy Studies Journal*. 39(1): 79–103. http://dx.doi.org/10.1111/j.1541-0072.2010.00397.x.

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.

Thornton, P. H. (2002). The Rise of the Corporation in a Craft Industry: Conflict and Conformity in Institutional Logics. *The Academy of Management Journal*. 45(1): 81–101. http://dx.doi.org/10.2307/3069286.

Thornton, P. H. (2004). *Markets from Culture: Institutional Logics and Organizational Decisions in Higher Education Publishing*. Stanford University Press.

Thornton, P. H., C. Jones, and K. Kury. (2005). Institutional Logics and Institutional Change in Organizations: Transformation in Accounting, Architecture, and Publishing. *Research in the Sociology of Organizations*. 23: 125–170.

Thornton, P. H., and W. Ocasio. (1999). Institutional Logics and the Historical Contingency of Power in Organizations: Executive Succession in the Higher Education Publishing Industry, 1958- 1990. *American Journal of Sociology*. 105(3): 801–43. http://dx.doi.org/10.1086/210361.

Thornton, P. H., and W. Ocasio. (2008). Institutional Logics. In: *The SAGE Handbook of Organizational Institutionalism*. eds. R. Greenwood, C. Oliver, R. Suddaby, and K. Sahlin-Andersson, . Sage Publications Ltd.

Thornton, P. H., W. Ocasio, and M. Lounsbury. (2012). *The Institutional Logics Perspective: A New Approach to Culture, Structure and Process*. Oxford University Press, USA.

Tyler, T. R. (1990). *Why People Obey the Law: Procedural Justice, Legitimacy, and Compliance*. New Haven, CT: Yale University Press.

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.

Varughese, G., and E. Ostrom. (2001). The Contested Role of Heterogeneity in Collective Action: Some Evidence from Community Forestry in Nepal. *World Development*. 29(5): 747–65. http://dx.doi.org/10.1016/S0305-750X(01)00012-2.

Vedeld, T. (2000). Village Politics: Heterogeneity, Leadership and Collective Action. *Journal of Development Studies*. 36(5): 105–34. http://dx.doi.org/10.1080/00220380008422648.

Yancy, H. F., T. S. Zemlak, J. A. Mason, J. D. Washington, B. J. Tenge, N.-L. T. Nguyen, J. D. Barnett, et al. (2008). Potential Use of DNA Barcodes in Regulatory Science: Applications of the Regulatory Fish Encyclopedia. *Journal of Food Protection*. 71(1): 210–17.

Chapter 4: DNA barcoding in the media: Does coverage of cool science reflect its social context?

Abstract

Paul Hebert and colleagues first described DNA barcoding in 2003, which led to international efforts to promote and coordinate its use. Since its inception, DNA barcoding has generated considerable media coverage. We analysed whether this coverage reflected both the scientific and social mandates of international barcoding organizations.

We searched newspaper databases to identify 900 English-language articles from 2003-2013. Coverage of the science of DNA barcoding was highly positive but lacked context for key topics. Coverage omissions pose challenges for public understanding of the science and applications of DNA barcoding; these included coverage of governance structures and issues related to the sharing of genetic resources across national borders.

Our analysis provided insight into how barcoding communications efforts have translated into media coverage; more targeted communication efforts may focus media attention on previously omitted, but important topics. Our analysis is timely as the DNA barcoding community works to establish the International Society for the Barcode of Life.

Introduction

DNA barcoding is a taxonomic system that uses short, conserved genetic sequences of animals (Hebert et al. 2003), plants (Hollingsworth et al. 2009), or fungi (Schoch et al. 2012) as a standardized marker for identifying species. Paul Hebert and colleagues first formally described barcoding as a large-scale system in 2003 (Hebert et al. 2003). The Consortium for the Barcode of Life (CBOL) was subsequently launched in 2004 to develop DNA barcoding as a global standard (Consortium for the Barcode of Life 2015b), and the International Barcode of Life (iBOL) Project launched in 2010 to extend the coverage of the barcode reference library (Stoeckle and Hebert 2008). Both organizations bring together multi-sectoral, multi-national stakeholders to promote the uptake and use of barcoding by scientific, regulatory, and lay communities. Regulatory uses for barcoding include product testing for consumer protection and rapid identification for illegal trafficking of endangered species. IBOL, in partnership with

CBOL, also actively promotes the use of the system by citizen scientists, including school children, to understand and contribute to broader knowledge of local biodiversity (e.g.: the School Malaise Trap Program (Biodiversity Institute of Ontario 2013)). The DNA barcoding community's mission therefore includes public outreach and communication through education programs and media (iBOL 2015c). As a result, both DNA barcoding and Barcoding of Life (BOL) initiatives have received considerable media coverage.

While outreach communications raise the profile of the science of barcoding, such communications also enable public discussion of biodiversity policy, climate change, human impacts on biodiversity, and other societal implications of barcoding (Costa and Carvalho 2007). An important component of biodiversity policy is the sharing of barcoding data and the specimen material from which barcodes are generated; such sharing is fundamental to BOL initiatives. Collectively, data and materials are termed genetic resources, and sharing these resources is governed at the international level by the *Convention on Biological Diversity* (*CBD*) and the *Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization* (*Nagoya Protocol*). These international laws are specifically implemented at the national level, especially in the biodiversity-rich countries of the Southern Hemisphere and Asia. But the sharing of genetic resources across national borders is controversial internationally, because of the fear of biopiracy, and national laws may be highly restrictive (Bubela and Gold 2012; Schindel et al. 2015). Biopiracy references the concerns of many biodiversity-rich, but research-infrastructure poor, countries that their genetic resources will be exploited for commercial gain without appropriate return of benefits, monetary or otherwise, to the country of origin.

We analysed English-language print media coverage of DNA barcoding and BOL initiatives and organizations. In addition to fulfilling a key objective of the iBOL project (iBOL 2015h), characterizing DNA barcoding media coverage is timely as the community plans the establishment of an International Society for the Barcode of Life (Castle et al. 2015). An analysis of media coverage provides scientists and administrators with insight into how their communications efforts are portrayed. We analysed general coverage of DNA barcoding and, specifically, whether media coverage reflected the breadth of the scientific and social mandates of BOL organizations– what is covered well, and what is omitted?

Methods

We searched media databases LexisNexis, Factiva, and Canadian Newsstand for English language articles and limited our search to print and online newspaper coverage. These language and print medium constraints represent limitations of the study. However, newspapers remain a key tool for public outreach related to policy efforts even though non-traditional media sources on the Internet or television content are also important sources of news (Nisbet et al. 2002). Indeed, newspaper content is often online, as well as in print sources, and newspapers have an agenda-setting role over Internet and television content, with content mirrored in later, derivative media (Caulfield 2004).

We searched publications in iBOL partner nations (iBOL 2015d) from January 2003-Dec 2013; 2003 was the publication year of the first paper formally proposing DNA barcoding as a large-scale, standardized system for specimen identification (Hebert et al. 2003). We designed our search algorithms to capture articles with iterations of "Barcode of Life" or "DNA barcoding" found anywhere in the article (Appendix 5). We manually excluded duplicated (identical publications obtained from more than one database) but not syndicated articles (similar/identical articles published in more than one newspaper under one management or derived from a paid service subscription, such as a newswire service), and non-print articles (for example, newswires that were not published). We also excluded a small number of articles in which DNA barcoding referred to activities other than the differentiation of species using a standardized short DNA sequence. For example, we excluded articles in which "DNA barcoding" described methods used to track an individual animal from farm to market.

We developed *a priori* codes (Appendix 6) for content analysis of each article based on descriptions of the iBOL Project and DNA barcoding from the iBOL website and other publications (iBOL 2015g; Hebert et al. 2003; Stoeckle and Hebert 2008), including: article characteristics, iBOL missions and working groups, barcoding science, and genetic resource sharing. A single researcher (EC) read each article and assigned it codes. We assessed the reliability of this data collection method by having a second researcher (JG) code a subset of the articles (n=183) to estimate inter-rater agreement. Our acceptable level of agreement ranged from 70-99% for each category used in our analysis, and the average agreement was 87% (Stemler 2004).

Additionally, for reasons of sample size, we only included articles in our comparative analyses that were published in the top five countries/regions by frequency. We examined associations between media coverage and country/regions using Pearson's Chi Square (SPSSv19). The greater coverage in Canada allowed a more detailed examination of media coverage in that country over time. We grouped Canadian media coverage into three time periods: 2003-2007, during the emergence and early diffusion of DNA barcoding and prior to iBOL; 2008-2009, when the iBOL project was securing funding and community support to launch internationally; and 2010-2013, when the system and iBOL were well established (2010-2013). We examined associations between media coverage and the three time periods using Pearson's Chi Square (SPSSv19).

Results and Discussion

General Content of Newspaper Articles on DNA Barcoding

Like other genetic technologies (Caulfield and Bubela 2005; Bubela et al. 2009), DNA barcoding has received considerable media attention, especially in Canada where the system and iBOL originated. Our search identified 900 unique English language articles, the majority published in five countries/regions: Canada (304), USA (159), United Kingdom (145), India (89), and Australia/New Zealand (81). The remaining 122 articles were in publications from China, South Africa, France, Kenya, Germany, Netherlands, and Pakistan.

Indicative of the prominence of newspaper articles is length of the article, the type of article (e.g., general news item or editorial), position in the newspaper (e.g., front page), and the key information sources – other publications or interviewees (Bubela and Caulfield 2004). The articles in our set were short; they averaged 580 words (median 491) (Table 12), which is similar to other estimates (Pellechia 1997). The majority of articles (86%) were current news articles, and only 8% were in-depth investigative pieces. Some articles featured prominently within the publication, with 24% in the front page/general news section. The DNA barcoding system itself was the main focus of 16% of the articles, and was mentioned only in passing in 15% (Table 12). The total number of articles increased over time, with only Canada showing a distinct peak in number of articles in 2008/2009 (Figure 15). Studies of science journalism show that peaks in

media coverage tend to coincide with important events (Nisbet and Lewenstein 2002). In Canada, this peak coincided with the funding of iBOL by Genome Canada Funding in 2008-2009. Similar peaks in media coverage did not occur in other countries, despite the prominence of other organizations like CBOL and the international reach of iBOL (Table 13).

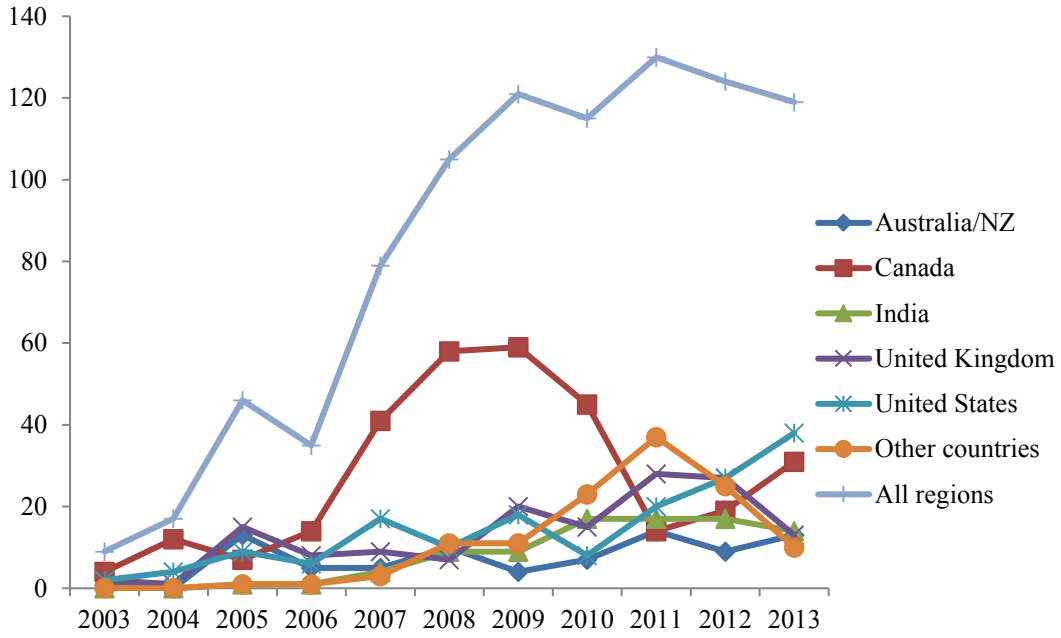

Figure 15. Number of newspaper articles per year for each country/region and all regions combined.

Table 12: Prevalence of variables representing the general content of newspaper articles on DNA barcoding, and comparison between major regions

	All articles (n=900)		Australia/NZ (n=81)		Canada (n=304)		India (n=89)		United Kingdom (n=145)		USA (n=159)		Other regions (n=122)		Chi-square* (p value)
Average page length (words)	580		499		582		436		580		785		470		n/a
	n	**%**	n	**%**	n	**%**	n	**%**	n	**%**	n	**%**	n	**%**	
Front Page/general news section	215	**24**	15	**19**	109	**36**	1	**1**	38	**26**	30	**19**	22	**18**	51.53 (0.000)
Article Format															
Latest news	775	**86**	68	**84**	267	**88**	86	**97**	116	**80**	125	**79**	113	**93**	19.44 (0.001)
In-depth investigation	68	**8**	8	**10**	15	**5**	1	**1**	15	**10**	24	**15**	5	**4**	21.69 (0.000)
Opinion piece or commentary	34	**4**	3	**4**	13	**4**	0	**0**	11	**8**	3	**2**	4	**3**	-
Article Focus on DNA Barcoding															
Barcoding is main focus	143	**16**	19	**24**	46	**15**	9	**10**	33	**23**	25	**16**	11	**9**	9.74 (0.045)
Barcoding mentioned in passing	133	**15**	18	**22**	24	**8**	25	**28**	30	**21**	25	**16**	11	**9**	29.68 (0.000)
Research publication mentioned	205	**23**	10	**12**	68	**22**	7	**8**	18	**12**	42	**26**	60	**49**	21.76 (0.000)
Any regulatory agencies and government departments	116	**13**	7	**9**	63	**21**	4	**5**	13	**9**	27	**17**	2	**2**	23.08 (0.000)
Food and drug agency†	93	**80**	3	**43**	58	**92**	1	**25**	10	**77**	20	**74**	1	**1**	-
Environmental protection or agricultural agency†	33	**28**	2	**29**	14	**22**	3	**75**	6	**46**	8	**30**	0	**0**	-
Sources cited in article															
Scientists	629	**70**	58	**72**	254	**84**	44	**49**	104	**72**	115	**72**	54	**44**	43.09 (0.000)
Government representative	60	**7**	8	**10**	24	**8**	7	**8**	13	**9**	5	**3**	3	**3**	5.63 (0.229)
Industry	38	**4**	2	**3**	15	**5**	3	**3**	3	**2**	14	**9**	1	**1**	9.40 (0.052)
Environmental group	21	**2**	0	**0**	9	**3**	0	**0**	2	**1**	9	**6**	1	**1**	-

*Chi-square is the result of comparing the five major regions: Australia/New Zealand, Canada, India, United Kingdom, and United States of America

† Proportions of types of regulatory agencies and government departments are calculated out of articles with any mentioned

Table 13: The frequency of barcoding and BOL categories in Canadian newspaper coverage over three time periods.

	2003-2007 (n=78)		2008-2009 (n=117)		2010-2013 (n=109)		Chi-square (p value)
	n	**%**	n	**%**	n	**%**	
Building the reference library	59	**76**	34	**29**	45	**41**	42.13 (0.000)
Methods or evaluating barcoding	5	**6**	9	**8**	6	**6**	0.44 (0.801)
Frame							
Science description or celebration	74	**95**	80	**68**	73	**67**	22.70 (0.000)
Public accountability	0	**0**	12	**10**	14	**13**	10.30 (0.006)
Scientific reference	27	**35**	23	**20**	18	**17**	9.38 (0.009)
Barcoding not explained	54	**69**	95	**81**	79	**73**	4.20 (0.379)
Main theme is basic science	61	**78**	42	**36**	38	**35**	42.75 (0.000)
Main theme is barcoding applications	1	**1**	49	**42**	31	**28**	39.75 (0.000)
Barcoding theme in body of article							
Basic Science	75	**96**	57	**49**	57	**52**	51.83 (0.000)
Monitoring invasive species	14	**18**	20	**17**	18	**17**	0.07 (0.968)
Food systems	13	**17**	73	**62**	50	**46**	39.67 (0.000)
Bioconservation	28	**36**	33	**28**	31	**28**	1.58 (0.454)
Forensics	3	**4**	16	**14**	7	**6**	6.77 (0.034)
Import/Export Monitoring	5	**6**	17	**15**	18	**17**	4.38 (0.112)
Medically relevant species monitoring	11	**14**	12	**10**	14	**13**	0.72 (0.698)
Citizen engagement	0	**0**	29	**25**	4	**4**	38.78 (0.000)
Any application of barcoding*	35	**45**	98	**84**	79	**73**	34.14 (0.000)
Regulatory agencies and government bodies	9	**12**	35	**30**	19	**17**	10.74 (0.005)
Type of agency (n=63)							
Food and Drug†	8	**89**	33	**94**	17	**90**	0.54 (0.765)
Agriculture or Environment†	2	**22**	10	**29**	2	**11**	2.32 (0.313)
Sources cited in article							
Scientists	63	**81**	104	**89**	87	**80**	3.97 (0.137)
Government representative	2	**3**	10	**9**	12	**11**	4.57 (0.102)
Industry	0	**0**	4	**3**	11	**10**	-
Environmental group	0	**0**	7	**6**	2	**2**	-
Partnerships mentioned	13	**17**	49	**42**	33	**30**	13.92 (0.001)
Vague reference to teams (n=95)	9	**69**	40	**82**	14	**42**	13.63 (0.001)
BOL organizations	2	**3**	18	**15**	23	**21**	13.11 (0.001)
Project funding sources mentioned	18	**23**	29	**25**	23	**21**	0.43 (0.806)

*Themes combined to create these variables include monitoring invasive species, food systems, bioconservation, forensics, import/export monitoring, and medically relevant species.
† Proportions of types of agencies are calculated out of articles with any agency mentioned

In keeping with the focus of newspaper articles on the science of barcoding, scientists were cited as sources or directly quoted in most articles (up to 84% in Canada, but as low as 49% in India) (Table 12). Media articles on science often reference research publications (Caulfield

and Bubela 2005), and we observed this most commonly in the USA (26%) and Canada (22%), possibly because researchers in these countries published many of the early publications that promoted and defended barcoding (Hebert et al. 2003; Stoeckle and Hebert 2008; Schindel and Miller 2005; Hebert and Gregory 2005; Miller 2007). Regulatory agencies and government departments were mentioned most frequently in Canada (21%) and the USA (17%) (Table 12). Of the articles mentioning regulatory agencies or government departments, 80% referenced food and drug regulators. Despite bioconservation being an important goal of DNA barcoding efforts, environmental protection or agricultural agencies were referred to much less frequently than food agencies, representing only 28% of agencies mentioned. Few articles (7%) included quotes or source information directly from government representatives, and even fewer quoted or sourced anyone from industry (4%) or environmental groups (2%) (Table 12).

Themes covered in newspaper articles: Basic science versus application?

Each article may have more than one theme, but the "main" theme or topic corresponds to that encountered in the first paragraphs of the article where reader attention is concentrated (Tewksbury and Althaus 2000). Not surprisingly, given the search strategy, the main theme in 64% of articles was explicitly related to DNA barcoding (Table 14). Canadian articles showcased DNA barcoding as the main theme the most frequently, which was also not surprising because of the prominence of Canadian researchers in the DNA barcoding community. The most common DNA barcoding themes were descriptions of the basic science underpinning the barcoding system and its taxonomic applications (Table 14), followed by other applications of DNA barcoding.

Applications of DNA barcoding co-occurred in most articles as a theme with the basic science of DNA barcoding, although the latter was the *main* theme in more articles (Table 14). The most common applications appearing as the main theme were related to food systems (Table 14). While not a common main theme, bioconservation applications were frequently mentioned, most commonly in mega-diverse Australia and highly endemic New Zealand (Table 14).

Canadian coverage of the basic science of DNA barcoding decreased over time, from the main theme in 78% of articles in 2003-2007 to 35-36% in 2008-2013 (Table 13). Mentions of basic science also decreased from 96% of articles in 2003-2007 to only 49-52% in later time

periods (Table 13), while mentions of barcoding applications rose from 45% of articles to 73-84% (Table 13). Food system applications were found frequently in later years as well, included in 62% of articles in 2008-2009. These patterns mirror trends in scientific reporting at major barcoding conferences, notably an increase in presentations on the socio-economic applications of DNA barcoding in later years (Adamowicz 2015). Other barcoding application themes such as bioconservation, invasive species monitoring, and forensic applications did not change over time (Table 13). The rise in coverage of food system applications is likely due to the number of scientific publications on the utility of DNA barcoding for monitoring seafood labeling fraud (Barbuto et al. 2010; Wong and Hanner 2008), illegal bushmeat sales (Eaton et al. 2009; Dalton and Kotze 2011), and the general appeal of such projects to the public.

The theme of citizen engagement was one of the least common themes, represented in only five percent of total articles. Most of these articles were from Canada and the USA, with an

Table 14: Proportions of articles with various themes in newspaper articles

	All articles (n=900)		Australia/NZ (n=81)		Canada (n=304)		India (n=89)		United Kingdom (n=145)		USA (n=159)		Other regions (n=122)		Chi-square* (p value)
	n	**%**	n	**%**	n	**%**	n	%	**n**	%	**n**	%	n	**%**	
Main theme anything related to barcoding	573	**64**	48	**59**	227	**75**	38	**43**	81	**56**	88	**55**	91	**75**	40.34 (0.000)
Basic science	370	**41**	31	**38**	141	**46**	29	**33**	51	**35**	54	**34**	64	**53**	11.04 (0.026)
BOL-project	17	**2**	2	**3**	5	**2**	1	**1**	3	**2**	1	**1**	5	**4**	1.72 (0.787)
Monitoring invasive species	9	**1**	3	**4**	2	**1**	0	**0**	0	**0**	0	**0**	4	**3**	-
Food systems	114	**13**	7	**9**	65	**21**	1	**1**	15	**10**	22	**14**	4	**3**	29.30 (0.000)
Bioconservation	13	**1**	2	**3**	3	**1**	1	**1**	2	**1**	2	**1**	3	**3**	-
Forensics	9	**1**	2	**3**	1	**0**	0	**0**	2	**1**	1	**1**	3	**3**	-
Import/Export monitoring	3	**0**	0	**0**	0	**0**	0	**0**	0	**0**	1	**1**	2	**2**	-
Medically relevant species	31	**3**	1	**1**	9	**3**	1	**1**	8	**6**	6	**4**	6	**5**	-
Citizen engagement	16	**2**	0	**0**	8	**3**	0	**0**	3	**2**	5	**3**	0	**0**	-
Any application of barcoding†	186	**21**	15	**19**	81	**27**	8	**9**	27	**19**	33	**21**	22	**18**	14.34 (0.006)
Themes mentioned anywhere in article‡															
Basic science	590	**66**	60	**74**	189	**62**	64	**72**	96	**66**	98	**62**	83	**68**	6.76 (0.149)
BOL-project	95	**11**	11	**14**	43	**14**	3	**3**	13	**9**	5	**3**	20	**16**	20.32 (0.000)
Monitoring invasive species	84	**9**	6	**7**	52	**17**	4	**5**	8	**6**	4	**3**	10	**8**	35.12 (0.000)
Food systems	264	**29**	22	**27**	136	**45**	9	**10**	34	**23**	53	**33**	10	**8**	47.48 (0.000)
Bioconservation	301	**33**	40	**49**	92	**30**	27	**30**	51	**35**	54	**34**	37	**30**	11.06 (0.026)
Forensics	93	**10**	12	**15**	26	**9**	4	**5**	24	**17**	9	**6**	18	**15**	16.45 (0.002)
Import/Export monitoring	115	**13**	13	**16**	40	**13**	6	**7**	20	**14**	17	**11**	19	**16**	4.48 (0.345)
Medically relevant species	130	**14**	12	**15**	37	**12**	10	**11**	24	**17**	23	**15**	24	**20**	2.23 (0.694)
Citizen engagement	49	**5**	0	**0**	33	**11**	0	**0**	3	**2**	13	**8**	0	**0**	27.47 (0.000)
Any application of barcoding†	588	**65**	66	**82**	212	**70**	43	**48**	92	**63**	112	**70**	63	**52**	24.55 (0.000)

*Chi-square is the result of comparing the five major regions: Australia/New Zealand, Canada, India, United Kingdom, and United States of America. A dash indicates that expected values are too low to calculate.

†Themes combined to create these variables include monitoring invasive species, food systems, bioconservation, forensics, import/export monitoring, and medically relevant species.

‡Coding is not mutually exclusive, and one article may have contained multiple DNA-barcoding related themes

additional two articles from the UK, and most of the Canadian articles were published 2008-2009.

The pattern of coverage in Canada (Table 13) could be a result of specific outreach efforts by iBOL project leaders. Earlier efforts likely focused on promoting the science of DNA barcoding at a time when the proponents were attempting to elicit support for the science. Once the scientific premise was more generally accepted (Frézal and Leblois 2008), media outreach shifted to promoting real-world applications of DNA barcoding in an effort to attract the attention of potential users of the system and more diversified sources of funding.

Framing

Media analyses focus not only on the quantity, article type, and prominence/placement of coverage, but also on the framing of that coverage (Gamson and Modigliani 1989). Frames can influence how issues are perceived and understood (Kahlor and Stout 2009). They represent interpretive schema to help make sense of and discuss an issue and are used by journalists to condense complex content into appealing news reports (Scheufele 1999; Nisbet and Scheufele 2007). Frames examined by other media studies of genetic technologies include: a general description of a scientific topic, celebration of progress in the context of human research/ingenuity; controversy; risks; economic prospects; public accountability; and human interest (Bubela and Caulfield 2004; Nisbet 2009a). Genetic research, especially that with benign or medical applications, is commonly framed as a celebration of progress (Bubela and Caulfield 2004). Similarly for DNA barcoding, we found that the majority of articles (87%) were either neutral descriptions of the science or framed as a celebration of scientific progress (Table 15). This result highlights that the majority of articles were news items rather than investigative journalism, and focused on the science of barcoding.

Table 15: Prevalence of frames used in newspaper coverage of DNA barcoding

	All articles (n=900)		Australia/NZ (n=81)		Canada (n=304)		India (n=89)		United Kingdom (n=145)		USA (n=159)		Other regions (n=122)		Chi-square* (p value)
	n	**%**	n	**%**	n	**%**	n	%	**n**	%	**n**	%	n	**%**	
Description of science or celebration of progress	781	**87**	76	**94**	227	**75**	87	**98**	136	**94**	139	**87**	116	**95**	52.51 (0.000)
Controversy	5	**1**	0	**0**	4	**1**	0	**0**	0	**0**	1	**1**	0	**0**	-
Risks	64	**7**	5	**6**	40	**13**	1	**1**	3	**2**	13	**8**	2	**2**	24.10 (0.000)
Public Accountability	30	**3**	0	**0**	26	**9**	0	**0**	3	**2**	1	**1**	0	**0**	30.62 (0.000)

*Chi-square is the result of comparing the five major regions: Australia/New Zealand, Canada, India, United Kingdom, and United States of America. A dash indicates that expected values are too low to calculate.

In Canada, early articles framed barcoding as a positive scientific advance (95%) (Table 13). This framing largely continued into later periods, but 12-14% of articles used the "public accountability" frame from 2008-2013. This frame represented calls for public control or participation, for example, coverage of a study on food mis-labeling combined with an explanation of how DNA barcoding could be used as a regulatory/consumer protection tool to address the problem. This shift in Canadian media frames coincided with an increase of coverage of regulatory applications for barcoding, which followed an evaluation of DNA barcoding by the United States Food and Drug Administration (FDA). The evaluation determined that barcodes were a suitable marker with broad practical applications. Applications included the FDA Regulatory Fish Encyclopedia used to identify species substitution; substitution can result in adverse health outcomes and economic fraud (Yancy et al. 2008).

In the political arena, the most common framing by media is controversy (Nisbet 2009a). In the realm of science, this framing is apparent when science and politics converge, for example in the coverage of climate change (Nisbet and Mooney 2007; Nisbet 2009b), agricultural biotechnology (Nisbet and Huge 2006), and human embryonic stem cell research (Nisbet and Goidel 2007; Nisbet et al. 2003). Given that barcoding raises elements of social controversy in the context of utilization of genetic resources, it is surprising that only five articles were so framed. While controversial issues over the utilization of genetic resources are far less prevalent

in developed countries, and lack of such framing reflects the prevalence of English-language, developed country articles in our sample, none of the 206 articles from middle income or developing countries referenced either the *CBD* or the *Nagoya Protocol*.

Controversy in science coverage may also arise in the context of controversy over the science itself. For example, media coverage of climate science is often criticized for over-representing a small amount of dissent in the scientific community (Boykoff and Boykoff 2007). In contrast, the media has so far ignored debates about DNA barcoding in the scientific literature (e.g., (Ebach and Holdrege 2005; Dupuis et al. 2012; Will and Rubinoff 2004)). The cytochrome C oxidase I mitochondrial gene region was the first barcode region described, chosen as a standardized marker for animals because of increased phylogenetic signal relative to other mitochondrial genes, reduced problems of frame-shifting as caused by insertions and deletions in ribosomal genes, and the ability to use robust universal primers (making high-throughput analyses feasible) (Hebert et al. 2003). However, opposing taxonomists were concerned that focusing on fast identifications using a single mitochondrial gene would divert funds from traditional taxonomy (Ebach and Holdrege 2005), overly simplify and incorrectly characterize the complex relationship between genes and species (Dupuis et al. 2012), and result in incorrect species identifications (Will and Rubinoff 2004). These debates were not reflected in the media coverage beyond a few opinion articles and letters-to-the-editor submitted by the opposing scientists themselves (Table 12).

What the media miss: Errors of omission versus commission

Studies on media reporting of genetics research report few errors of commission (Bubela et al. 2009) – where facts are reported, they are generally accurate, especially when articles report on a scientific publication. However, errors of omission – what is *not* reported– are common (Caulfield and Bubela 2005). The public engagement mandate of iBOL provided the organisation with funding and infrastructure for media outreach. The main goal of these outreach and engagement activities was to inform about the science of DNA barcoding, facets of which were discussed in most articles. Articles from Australia/New Zealand and India were most likely to include the theme of barcoding science (74% and 72%, respectively, Table 10). Conversely, very few articles (7%) explained basics of the process of barcoding, information that is needed to

understand the legal and social issues relevant to barcoding (Table 16). This is surprising given the high percentage of articles that cited or quoted barcoding researchers.

Another issue largely omitted from coverage was the core mission of iBOL to add 500,000 species to a publicly accessible reference database by 2015. Only 53% of articles mentioned adding barcoding data to the reference database, and only 12% of articles detailed the specifics of the goal (Table 16). Only 45% of Canadian articles mentioned adding barcodes to the reference database (Table 16), even though the database, Barcode of Life Data Systems BOLD (Ratnasingham and Hebert 2007), is housed at the University of Guelph (Biodiversity Institute of Ontario 2012).

Table 16: Prevalence of important DNA barcoding topics that were largely omitted from media coverage

	All articles (n=900)		Australia/NZ (n=81)		Canada (n=304)		India (n=89)		United Kingdom (n=145)		USA (n=159)		Other regions (n=122)		Chi-square* (p value)
	n	%	n	%	n	%	n	%	n	%	n	%	n	%	
Scientific process of DNA barcoding															23.46 (0.003)
Not described at all	691	**77**	58	**72**	228	**75**	77	**87**	124	**86**	112	**70**	92	**75**	
Explained briefly	148	**16**	17	**21**	55	**18**	11	**12**	9	**6**	35	**22**	21	**17**	
Explained fully†	61	**7**	6	**7**	21	**7**	1	**1**	12	**8**	12	**8**	9	**7**	
Adding barcodes to reference library															
Specifics of iBOL core mission included	105	**12**	8	**10**	36	**12**	9	**10**	25	**17**	16	**10**	11	**9**	4.96 (0.292)
Non-specific mention of adding barcodes	478	**53**	54	**67**	138	**45**	56	**63**	94	**65**	68	**43**	68	**56**	31.43 (0.000)
Any funding sources mentioned	152	**17**	13	**16**	70	**23**	21	**24**	12	**8**	26	**16**	10	**8**	16.67 (0.002)
Governments‡	50	**33**	2	**15**	29	**41**	12	**57**	4	**33**	3	**12**	0	**0**	14.26 (0.007)
National Science agencies‡	16	**11**	0	**0**	11	**16**	1	**5**	1	**8**	1	**4**	2	**0**	5.83 (0.213)
Private foundations‡	26	**17**	1	**8**	8	**11**	1	**5**	2	**17**	14	**54**	0	**0**	27.75 (0.000)
Partnerships mentioned	263	**29**	20	**25**	95	**31**	17	**19**	51	**35**	43	**27**	37	**30**	8.68 (0.70)
Vague reference to teams§	160	**61**	12	**60**	63	**66**	9	**53**	29	**57**	26	**60**	21	**57**	1.96 (0.743)
Specific partnerships between developed countries§	59	**22**	4	**20**	22	**23**	2	**12**	19	**37**	9	**21**	3	**8**	6.52 (0.163)
Specific partnerships between developing and developed countries§	13	**5**	0	**0**	0	**0**	2	**12**	3	**6**	1	**2**	7	**19**	-
Partnerships between or led by developing countries	17	**7**	0	**0**	0	**0**	4	**24**	2	**4**	1	**2**	10	**27**	-
Sample locations different from media source locations‖	169	**36**	11	**22**	61	**39**	6	**17**	28	**41**	34	**40**	29	**38**	11.71 (0.020)
Commercial utilization of genetic resources	3	**0**	0	**0**	1	**0**	0	**0**	0	**0**	0	**0**	2	**2**	-
CBD or Nagoya mentioned	1	**0**	0	**0**	0	**0**	0	**0**	0	**0**	1	**0**	0	**0**	-

*Chi-square is the result of comparing the five major regions: Australia/New Zealand, Canada, India, United Kingdom, and United States of America. A dash indicates that expected values are too low to calculate.

†Articles were coded as “explained barcoding fully” if they included that barcodes were short, standardized gene regions that must be compared to a reference database of known species
‡Proportions of types of funders are calculated out of articles with any funder mentioned
§Proportions of types of partnerships are calculated out of articles with any partnerships mentioned
|| Calculated from articles with both locations of sample collections and media sources (n=471)

Large-scale, international research initiatives, such as CBOL and iBOL, require equally ambitious funding commitments. However, a minority of articles mentioned or discussed funding sources (Table 16). Of the articles that included funding sources, the most frequent reference was to unspecified government sources, common in Canada and India. National science funding agencies, such as Genome Canada, were rarely mentioned despite their substantial financial contributions. Funding from private organizations (e.g., Alfred P. Sloan Foundation) was the most common type of funding mentioned in the USA. Omission of funding sources is common across media reporting of science, and is frequently recommended to improve coverage in light of assessment of credibility of research and potential conflicts of interest (Cook et al. 2007; McComas and Simone 2003).

Omissions also occurred in newspaper coverage of the social mandates of barcoding. Despite the emphasis of the iBOL website on its partnerships (iBOL 2015a), these were only mentioned in 29% of articles (263 articles). Of these 263 articles, 61% made vague references to teams, 22% mentioned specific partnerships between partners in developed countries in the Northern Hemisphere, and only 5% mentioned partnerships between researchers or research institutions in developed and developing countries (Table 16). A further 7% mentioned partnerships solely between researchers in developing countries, or led by a developing country partner. The omission is surprising because research institutions in Canada, the USA, and the UK are intended to function within iBOL as central nodes, providing support to other countries with less scientific capacity (iBOL 2015d).

Canadian funders, in particular, require partnerships with biodiverse, developing countries to meet the species targets for the reference library, although none of the Canadian articles in our dataset described such partnerships (Table 16). In Canada, partnerships and teams were more frequently mentioned in years when barcoding proponents were applying to Genome Canada as part of the International Consortium Initiative, for which iBOL was funded in 2009 (Table 13). However, partnerships between developed countries were only detailed in 23% of Canadian articles that mentioned partnerships (Table 16).

Major DNA barcoding initiatives like CBOL, iBOL, and BOL sub-projects (such as PolarBOL), were the main theme in only 2% of articles (Table 14). These initiatives were also mentioned in a minority of articles (Table 14) – 14% in Canada and Australia/NZ, 9% in the UK

(9%), and 3% in both USA and India. Canadian articles had the most coverage of BOL projects, with the highest frequency of coverage coinciding with the official launch of iBOL in 2010 (21%). Although iBOL itself is a Canadian initiative, other major BOL projects are based in other countries. CBOL is based in the USA and led many of the early DNA barcoding initiatives (Consortium for the Barcode of Life 2015b). Yet, only one of the USA articles had a BOL project as a main theme, and BOL projects were mentioned in only five additional articles (Table 14).

Political/Social Context of DNA Barcoding and Genetic Resources

DNA barcoding, developed a decade after the *CBD* entered into force, has both the potential to help achieve the objectives of the *CBD* and to run counter to its premise of national sovereignty over genetic resources. The core objective of the *CBD* is the conservation of biological diversity, and DNA barcoding is a useful system to document and monitor biodiversity (Hebert et al. 2003). However, the process of generating the barcode library often requires genetic resources to traverse national borders, because many countries lack the infrastructure to generate DNA barcodes. The large-scale barcoding initiatives require barcode sequences to be placed in the public domain, along with meta-data that describes the sample (including its collection site) and the storage of a reference specimen. While the DNA barcode itself represents a short sequence with minimal utility beyond species identification, the reference specimen and/or material from which the DNA barcode is extracted represents the entire genome of the specimen. This raises concerns over the potential for the misappropriation (use other than for barcoding) of genetic resources, in other words, biopiracy (Schindel et al. 2015). The issue is even more contentious when the specimens originate in areas inhabited by indigenous peoples, even when the resources are accessed by domestic researchers (Beas-Rodrigues 2012). Thus, like other forms of non-commercial research, DNA barcoding initiatives have been slowed or stopped entirely because of national laws and regulations over access to genetic resources, utilization of genetic resources, and the return of benefits to the country of origin (Vernooy et al. 2010). These legal frameworks are collectively referred to as Access and Benefit Sharing (ABS).

As part of CBOL's main functions, the organization has advocated to the *CBD* for reduced restrictions on non-commercial science, such as DNA barcoding (Consortium for the Barcode of Life 2015a). Barcoding proponents have argued reduced restrictions are essential when the primary goal of that science promotes the main goal of the *CBD* – the protection of global biodiversity (Schindel 2010; Schindel et al. 2015). Indeed a core working group objective for iBOL was to develop equitable use of genetic resources in the context of DNA barcoding (iBOL 2015c). Despite the international prominence of debates over ABS, and the participation of DNA barcoding researchers and practitioners in those debates, the topic was virtually absent from media coverage (Table 16). Only one Canadian article mentioned commercial utilization of genetic resources, and only one USA article mentioned the *CBD*. No articles mentioned the *Nagoya Protocol*. However, our analysis did not examine coverage in non-English speaking, biodiversity-rich countries, where interest in the protection of national biodiversity may be higher.

The lack of reporting on ABS issues appears to be limited to the sharing of benefits and not access to genetic resources. Many articles (169) described barcoding projects or research using internationally sourced genetic materials (Table 16) but did not go on to discuss the surrounding policy context or debate. The majority of articles in Canada and USA described projects that used genetic materials not collected in those countries (Figure 16). Only in Australia and India did a majority of articles mention genetic materials that were collected in those countries. The majority of Canadian articles mentioned foreign genetic resources, but later coverage included more domestic resources (Figure 17). The lack of coverage of genetic resources and ABS issues therefore represents a major gap in the coverage of the social mandate of DNA barcoding and its supporting organizations.

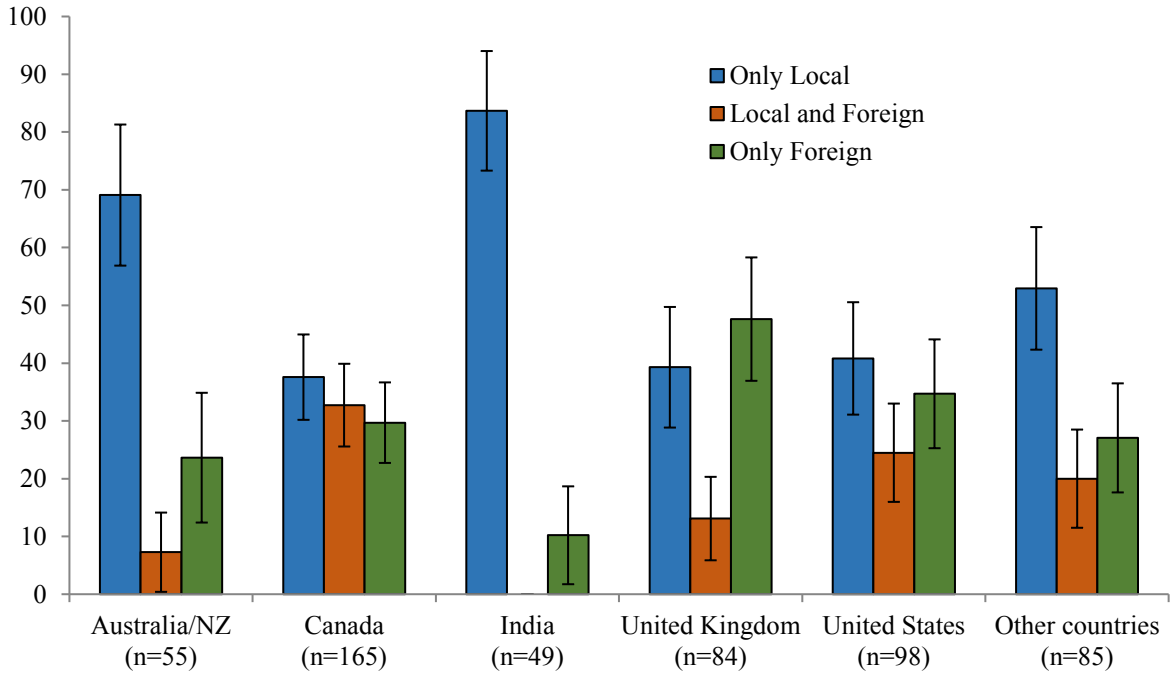

Figure 16. Percentage of articles that mentioned each sampling location, from articles that did so (n=536, 60% of 900 articles). Error bars indicate 95% confidence intervals.

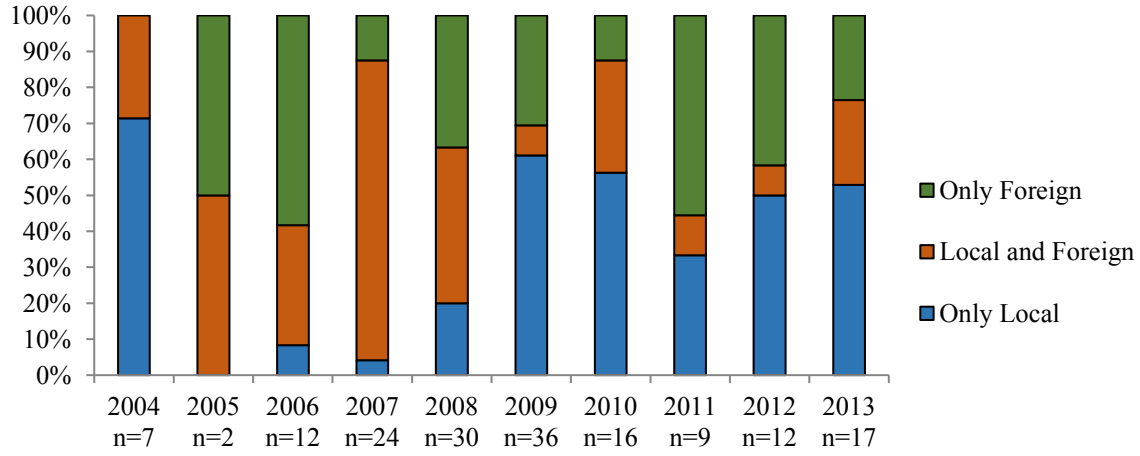

Figure 17. Percentage of articles that mentioned each sampling locations, from articles that did so (n=165, 54% of 304 Canadian articles). No articles mentioned sampling locations in 2003.

Conclusions

The iBOL project included a core objective of media research to understand how its science and technology is perceived in the context of avoiding an "undernourished environment for wider public discussion of the project's results" (iBOL 2015c). BOL initiatives and DNA barcoding have received considerable coverage in print media, and coverage continues to increase across regions. This continued interest in DNA barcoding and its applications offers an opportunity to enrich public understanding and debate over the applications of the science and its governance, as well as to contribute to broader understanding of issues related to biological diversity.

The major gap we identified between media coverage and the mandates of BOL organizations was the sharing of genetic resources within international partnerships. Since such sharing is regulated by national legal frameworks that implement the now-in-force *Nagoya Protocol*, publics need to be made aware of the existence of these frameworks. Information is also needed on other regulations that restrict the collection of specimens for barcoding, including permitting and collections within National Parks. Such awareness is particularly important in the context of projects, like LifeScanner (http://www.scientificamerican.com/citizen-science/lifescanner/), that aim to engage the public in biodiversity monitoring through specimen collection and barcoding. Scientists would similarly benefit from enhanced discussion and understanding of ABS issues, because partnership-based research, led by institutions and researchers in developed countries, can still be colonial in nature when partners exist in developing regions (Waterton et al. 2013). A lack of understanding of the history of exploitation, and related concerns about biopiracy, leaves international research projects vulnerable to a lack of cooperative goal-setting, which can hinder the development of trust (Bagshaw et al. 2007). Trust is essential to the establishment and maintenance of international research partnerships and/or consortia.

The lack of media coverage on social and political implications of barcoding is one indication that discussions on these topics are not prominent within the barcoding community and its outreach efforts. The development of the International Society for the Barcode of Life as a governance organization for future DNA barcoding efforts provides the opportunity to redress the balance in public outreach: to reflect both barcoding's scientific and its social mandates.

Future research can support this effort through examining the effectiveness of specific media outreach methods, such as press releases.

Acknowledgements

JG received doctoral awards from Canadian Institutes of Health Research, Alberta Innovates-Health Solutions, and the Killam Trust at the University of Alberta. The research was funded through grants to TB for the iBOL project (PI: Paul Hebert) from *Genome Canada and the Ontario Genomics Institute*.

References

Bubela, T., J. Geary, M. Bieber, W. Luth, and D. Hik. (2015). Diffusing Barcoding: The Global Spread of a Good Idea. *Genome*., Scientific abstracts from the 6th International Barcode of Life Conference, 58(May): 201–2. http://dx.doi.org/10.1139/gen-2015-0087.

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.

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.

Contreras, J. (2014). Constructing the Genome Commons. In: *Governing Knowledge Commons*. eds. B. M. Frischmann, M. J. Madison, and K. J. Strandburg, 99–136. Oxford University Press.

Crawford, S., and E. Ostrom. (1995). A Grammar of Institutions. *American Political Science Review*. 89(3): 582–600. http://dx.doi.org/10.2307/2082975.

Dedeurwaerdere, T. (2006). The Institutional Economics of Sharing Biological Information. *International Social Science Journal*. 58(188): 351–68. http://dx.doi.org/10.1111/j.1468-2451.2006.00623.x.

Dedeurwaerdere, T. (2010). Institutionalizing Global Genetic Resource Commons: Towards Aternative Models for Facilitating Access in the Global Biodiversity Regime. *SSRN*. May. http://dx.doi.org/10.2139/ssrn.1611549.

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, Part 1(January): 1–10. http://dx.doi.org/10.1016/j.envsci.2015.08.006.

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): 219–20. http://dx.doi.org/10.1139/gen-2015-0087.

Geary, J., E. Camicioli, and T. Bubela. (2013). The Public Face of Barcoding: How iBOL and DNA Barcoding Are Represented in Newspapers from around the World. Presented at the 5th International Barcode of Life Conference. Kunming, China, Oct 27-31.

Goodrick, E., and T. Reay. (2011). Constellations of Institutional Logics Changes in the Professional Work of Pharmacists. *Work and Occupations*. 38(3): 372–416. http://wox.sagepub.com/content/38/3/372.short.

Li, D.-Z., 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.

Madison, M. J. (2014). Commons at the Intersection of Peer Production, Citizen Science, and Big Data. In: *Governing Knowledge Commons*. eds. B. M. Frischmann, M. J. Madison, and K. J. Strandburg, 209–54. Oxford University Press.

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): 254–73. http://dx.doi.org/10.1089/omi.2013.0158.

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.

Chapter 5: Conclusions

Summary of empirical findings

I sought to contribute to theory and practice of governing knowledge commons through a case-study of the DNA barcode commons. While knowledge commons scholarship has increasingly accounted for the context of international research environments (Dedeurwaerdere 2006; Dedeurwaerdere 2010b; Mishra and Bubela 2014), my research adds a perspective on how such commons function in a heterogeneous *global* research environment.

In addressing my first research question, "How do factors identified by the Institutional Analysis and Development (IAD) framework influence effective governance of a global knowledge commons?", I found that the factors relevant to the DNA barcoding commons created challenges for the collective action required for effective governance. While the DNA barcode commons has functioned to generate and share barcode records, the strategy to openly share data and materials has not yet produced a globally representative barcode resource. To achieve this goal, governance structures should promote an equitable distribution of burdens and benefits for contribution, access, and use.

These findings relate closely to my second research question: "How does heterogeneity inform the rules used by actors to govern their behaviours?" I found that the rules in use were based on a variety of logics, with a pattern that suggested a lack of shared understanding for crafting rules between participants from the like-minded megadiverse countries (LMMC) and non-LMMC stakeholders. National laws that govern access and benefits sharing in LMMCs were a prominent factor in determining how researchers in LMMCs participated in DNA barcoding, yet these laws were sometimes disregarded by researchers and policy makers in high income countries, like Canada and the US, who have driven barcoding efforts. Nevertheless, my research points to mechanisms that could develop enforceable rules for participation in a global knowledge commons based on shared expectations in contexts where the heterogeneous participants might otherwise choose rules that drive conflicting behaviours.

My findings from my final research question ("How are topics and issues relevant to governing the global DNA barcode commons presented in newspaper coverage") provided insight into how the factors that inhibit collective action in the DNA barcode commons are

represented in public discourse, namely the print media. I concluded that critical issues, such as fair and equitable access and benefits sharing, were omitted from public barcoding discourse in countries where influential policies and guidelines are being developed. The considerable media coverage focused on positive aspects of barcoding science. There exists an opportunity, therefore, for leaders of the barcoding community to generate more awareness of the social and policy context of DNA barcoding activities and their conservation/regulatory goals.

Limitations

My research has a number of limitations. I primarily interviewed individuals with direct involvement in BOL organizations and efforts, and, therefore, I likely missed the perspectives of those who participate in DNA barcoding activities independently. I also interviewed participants over several years, and, therefore, responses may have been influenced by factors that changed over time. However, the overall structure of DNA barcoding governance did not change during the course of data collection. Individual behaviour might also depend on experience and length of time a participant has engaged in barcoding efforts, and I did not account for these potential differences between interviewees. My data interpretation was limited to my own perspectives, and it is possible that another individual might have different views. However, my use of an established theoretical framework reduced the reliance on my individual interpretation of data.

In my bibliometric analysis (Chapter 2), I only examined DNA barcoding publications that referenced seminal papers. This method was selected for simplicity and efficiency, because an initial attempt to identify barcoding articles through topic-based searches yielded an unacceptable percentage of irrelevant articles. However, references to the seminal papers started leveling off in 2012. This suggests that I may have missed some later publications if they did not reference early work, yet were still about DNA barcoding. Also, I did not differentiate between articles that were either critical or positive towards DNA barcoding.

In the second exemplar in Chapter 2, I selected a very small subset of BOLD records that were representative of specific topics of interest: those that had potential public health applications (disease-carrying mosquitos) and those that were more likely to have restrictions on specimen exporting (plant-based medicines). The pattern of the records I chose are likely not representative of other groups of DNA barcode records (such as endangered species or those

with limited applications beyond taxonomy and bioconservation). Finally, I had limited access to information about unpublished records. A larger analysis that includes more information about unpublished records might reveal different patterns of data and specimen sharing within the DNA barcode community.

My research also has limitations inherent from my case-study approach, as I only examined a single global knowledge commons. I cannot estimate how generalizable the findings are to other global knowledge commons. However, my analytical approach and interpretations of my data were grounded in a well-established framework. The use of this framework adds transferability, as the key contextual factors that define cases within this framework are determined not by my work, but by the previous work of others. Other examples of knowledge commons that share the characteristics outlined in the IAD framework will likely find the results from my study applicable.

Theoretical contributions

Despite its limitations, my research has demonstrated the importance of distinguishing between international (i.e. participation simply spans multiple countries) and global (i.e. outcomes depend on global participation) knowledge commons when describing principles for effective governance. The key difference between such commons is whether or not achieving the goals of the commons requires participation from actors in countries that differ on relevant factors. Previous work on knowledge commons for research has not explicitly accounted for this distinction (Contreras 2014; Bubela et al. 2012; Dedeurwaerdere 2010b), although work on other knowledge commons has emphasized the importance of being attentive to the social organization of participants (Madison 2014). There are some similarities between global and international knowledge commons, which may have made distinguishing between them less apparent in previous work. For example, other work has demonstrated that researcher motivations to participate in commons are dominated by reputation and social identity influences in the scientific community (Dedeurwaerdere et al. 2016). Similar motivations that I identified were tempered by the negative history (colonialism and resource misappropriation, resulting in wealth and power inequities) of the commons, a factor that previous scholars have suggested is important when studying all types of knowledge commons (Frischmann et al. 2014).

While many previous studies of heterogeneity among actors in knowledge commons have focused on measuring specific impacts on quantifiable outcome variables, my research demonstrated the utility of examining how heterogeneity broadly influences participant behaviours in a commons. Studying institutional arrangements is appealing because, unlike many characteristics of commons, institutional statements can be modified (Crawford and Ostrom 1995). While the grammar of institutions allows researchers to demonstrate where institutional statements are too vague or weak to be effective, logics analysis enables recommendations about how institutional statements could mitigate the negative effects of heterogeneity without needing to precisely define or measure it.

My research made three contributions to institutional logics. First, it demonstrated the practicality of using ideal types of logics as an analytical tool. Second, it supported previous evidence that multiple logics simultaneously influence the behaviours of individuals within organizations (Goodrick and Reay 2011). Finally, it demonstrated how filters modify the impact of institutional logics and suggested that global inequities are an important filter in organizations that span high and low income countries.

Overall, my research demonstrated the value of using multiple and mixed methods and theoretical perspectives to approach studying knowledge commons. My in-depth interviews provided rich data about individual experiences and behaviours in the commons. My quantitative analysis of the outcomes of the commons, however, provided an important context for understanding the perspectives of interviewees. Using the IAD framework provided a theoretically grounded foundation for understanding how the factors of a global knowledge commons influenced governance. However, previous research on heterogeneous commons governance relied on precise characterizations and modeling the effect of a single aspect of heterogeneity on specific quantifiable outcome variables. Incorporating institutional logics allowed me to study heterogeneity in my case study as a broad concept, without having to identify in advance the salient factors to measure.

Summary of Recommendations for the DNA Barcoding Community

During the 6th International Barcode of Life Conference in August 2015, barcode community members participated in a workshop to establish the International Society for the

Barcode of Life (ISBOL). The Society would "coordinate completion of the [barcode] registry, to facilitate the development of barcode applications and to communicate with stakeholders at all levels" (Castle et al. 2015). Membership was automatically granted to all registrants of the conference, but is open to all interested parties. A governance council to initiate ISBOL, comprised of 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, was proposed as an interim measure. The council will seek feedback on proposed structure and governance from the broader DNA barcoding community.

The findings of the empirical research presented in this thesis point to several recommendations for the new ISBOL and other BOL organizations, funding agencies that support DNA barcoding efforts, and individuals participating in the DNA barcoding community. Although there are limitations to my data and analysis, for the sake of clarity I have written my recommendations as definitive statements for the barcoding community to consider. Here, I outline those recommendations and their justifications:

ISBOL and other BOL organizations:

1) The Interim Council for the International Society for the Barcode of Life (ISBOL) should ensure that (a) the governing body of ISBOL is representative of the diversity of the barcoding community; and (b) mechanisms are put in place to solicit input for policy development that is also representative of the diversity of the barcoding community.

    Justification 1: Perspectives that represent the diversity of the DNA barcoding community have not been adequately accounted for in the governance of BOL organisations, an inadequacy that impedes the development of effective rules, norms and strategies.

2) The Interim Council for ISBOL should not require barcoding project funding as a prerequisite for individuals to participate in DNA barcoding governance.

    Justification 2: Requiring funding commitments precludes the participation of individuals who may only be able to contribute non-financial but nevertheless valuable resources (e.g., collected samples) to barcoding efforts.

3) 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 *Convention on Biological Diversity* and *the Nagoya Protocol*; and b) promulgate educational materials for the community on the legal and policy context of DNA barcoding activities to enhance understanding and compliance.

    Justification 3: Current barcoding governance documents have not adequately referenced international legal instruments or national laws, nor have the rules expressed in these documents reflected the intent of these laws. Furthermore, the diversity among barcoding participants in the application of laws that govern genetic resources 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 concerns of LMMC participants, in particular, and facilitate their participation in barcoding efforts. Enhanced education about the legal framework would help researchers understand why it is necessary to govern barcoding activities in a manner that respects LMMC concerns about access to and utilization of genetic resources.

4) BOL organizations should implement barcode database access and use requirements that a) encourage the community to contribute to and use barcoding databases while respecting community concerns, such as restricted access to sensitive data, a reasonable embargo period to enable publication, and acknowledging the originators of the data; and b) restrict data access to non-commercial use.

    Justification 4: Researchers are concerned with unrestricted access to and use of their data. At a minimum, members of the barcoding community expressed a desire for restrictions that protect sensitive data, a reasonable embargo period to enable first publication, and acknowledgement of the contributor of the data. It is common for open access databases to impose such conditions of use to encourage community participation. Additionally, unrestricted use allows data to be used in ways that may preclude barcoders from accessing genetic resources under

"simplified measures for non-commercial research" under the *CBD,* and *Nagoya Protocol,* and national laws.

5) BOL organizations should develop advisory guides for the development of Materials Transfer Agreements (MTAs) that address community concerns about sharing materials (such as requiring specimen destruction after the generation of the barcode or restriction of the use of the specimen to generating a barcode sequence).

   Justification 5: Researchers are concerned that materials shared for the generation of DNA barcodes may be used by recipients for purposes they would not wish to support. In addition, technology transfer offices that mediate exchanges of MTAs may not have capacity to draft MTAs that are compliant with the legal framework for genetic resources. Capacity may be lacking in LMMCs with respect to funding for and staffing of technology transfer offices, and capacity in non-LMMC institutions may be lacking due to lack of knowledge or understanding of this legal framework. Guidance on the drafting of MTAs that respect international and national laws for the sharing and use of genetic resources is therefore needed.

6) To fulfill their goals related to effective public engagement, BOL organizations should create media communication strategies on how to report social, cultural and political issues relevant to DNA barcoding activities.

   Justification 6: To date, media coverage of DNA barcoding has largely omitted topics relevant to the social, cultural and political mandates of BOL organizations. BOL organizations can facilitate inclusion of these topics in media coverage by providing guidance on key media messages.

Funding Agencies

7) Funding agencies interested in increasing barcode activities should support infrastructure for storing biological materials (including specimens and DNA extracts) in LMMC countries.

   Justification 7: LMMC researchers may be more hesitant to participate if they are required to store specimens outside of their country.

8) Funding agencies should work with each other and BOL organizations to 1) coordinate use-of-funds policies for international barcoding projects; and 2) ensure adequate funding is provided for long-term sustainability of barcode infrastructure.

Justification 8: If funding agencies do not coordinate funding policies, they risk implementing policies that are in conflict with each other. Additionally, the specialized infrastructure needed for the DNA barcode commons will be unsustainable without provisions to ensure continued funding (for example, through cost-recovery fees or ongoing funding commitments).

DNA Barcode Researchers

9) Researchers should seek learning opportunities regarding the *CBD* and *Nagoya Protocol*, and, in particular, invite their LMMC partners to provide training on these topics.

Justification 9: Implementing effective governance will be difficult in a community where some members are unaware of these legal instruments. Inviting LMMC partners to train non-LMMC researchers and staff would increase the equity of partnerships that include LMMC and non-LMMC members by providing LMMC partners with professional benefits.

10) Researchers should learn and comply with the legal requirements of their academic institutions and other relevant institutions in the countries where they work.

Justification 10: Not all researchers are compliant with legal requirements, and they must be for the DNA barcode commons to function successfully.

11) Researchers who use the resources of the DNA barcode commons should contribute back to it, and they should acknowledge the contributions of resource providers.

Justification 11: Acknowledgements and re-contribution will help encourage other researchers to participate, and increase the value of the DNA barcode commons.

Knowledge translation

Knowledge translation for this project includes engagement with the DNA barcoding community, dissemination of preliminary findings, and the collaborative production of reports or other materials for knowledge users.

I actively sought input and feedback from the DNA barcoding community beyond their participation in formal interviews. In May 2012, I visited the Biodiversity Institute of Ontario (BIO) to learn about the facility and its workflow for producing barcode records. I traveled to three International Barcode of Life conferences, in 2011, 2013, and 2015, where I attended many sessions on diverse barcode topics and interacted closely with barcode participants. I presented preliminary findings at two international DNA barcoding conferences (Geary and Bubela 2015; Geary et al. 2013; Bubela et al 2015) and invited feedback from conference attendees and interviewees from our study. I participated in a workshop in February 2013 that included discussions of medicinal plant barcoding and issues related to sharing genetic resources. The workshop resulted in a publication with leaders in the barcoding community (Schindel et al. 2015). I published the findings from my media analysis in a special DNA Barcoding edition of the journal *Genome* to reach an audience of DNA barcode researchers and decision makers.

I have built relationships with key DNA barcoding leaders, and I am working with them to disseminate findings from Chapters 2 and 3. I have been discussing my findings with the Chair of the IBOL Governance and Knowledge Mobilization Working Group to ensure my dissemination strategies and communications are appropriate for the DNA barcoding community. The 7th annual International Barcode of Life conference is in South Africa in late 2017, and while I may not attend, I will work with DNA barcoding leaders to develop key messages to be incorporated into other presentations at the conference. I will develop summaries appropriate for the IBOL Newsletter (Barcode Bulletin) and the iBOL website. In addition to empirical papers that I aim to publish in the International Journal of the Commons, I will also seek to publish a summary of my findings in a journal that is read by DNA barcode researchers and other stakeholders. I will also prepare other dissemination materials if my discussions with the DNA barcoding community reveal additional communication strategies.

Future research

Based on the findings of my research, there are several areas of future research I suggest. Overall, more research is needed to understand how governance of *global* (i.e. outcomes depend on global participation) and *international* (i.e. participation simply spans multiple countries) knowledge commons are different. Subsequent research should explicitly differentiate between these scenarios.

Our understanding of the DNA barcode commons would also benefit from additional research. Currently, we know little about individuals who complete DNA barcoding projects but are not interested in contributing materials or data to the barcode commons. Future research that examines why these individuals choose to not participate may provide more insight for improving governance. Further in-depth bibliometric analysis of barcoding publications would improve our understanding of how DNA barcoding has diffused across disciplines and regions. Future research should also examine how the DNA barcode commons shifts after its transition to governance by ISBOL. Researchers should examine how the characteristics of the new organization might impact media coverage, participation in barcoding projects, contribution of biological materials and data to the barcode commons, and co-authorships/partnerships within the DNA barcoding community.

Conclusion

My empirical research aimed to inform best practices for the governance of global knowledge commons. I have outlined several strategies based on the exemplar DNA barcode commons. Through facilitating representative governance, collective rules-making, and consideration of the complex social and policy context, global knowledge commons can enhance partnerships between researchers in higher and lower income countries with benefits for all involved.

References

Bubela, T., J. Geary, M. Bieber, W. Luth, and D. Hik. (2015). Diffusing Barcoding: The Global Spread of a Good Idea. *Genome*., Scientific abstracts from the 6th International Barcode of Life Conference, 58(May): 201–2. http://dx.doi.org/10.1139/gen-2015-0087.

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.

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.

Contreras, J. (2014). Constructing the Genome Commons. In: *Governing Knowledge Commons*. eds. B. M. Frischmann, M. J. Madison, and K. J. Strandburg, 99–136. Oxford University Press.

Crawford, S., and E. Ostrom. (1995). A Grammar of Institutions. *American Political Science Review*. 89(3): 582–600. http://dx.doi.org/10.2307/2082975.

Dedeurwaerdere, T. (2006). The Institutional Economics of Sharing Biological Information. *International Social Science Journal*. 58(188): 351–68. http://dx.doi.org/10.1111/j.1468-2451.2006.00623.x.

Dedeurwaerdere, T. (2010). Institutionalizing Global Genetic Resource Commons: Towards Aternative Models for Facilitating Access in the Global Biodiversity Regime. *SSRN*. May. http://dx.doi.org/10.2139/ssrn.1611549.

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, Part 1(January): 1–10. http://dx.doi.org/10.1016/j.envsci.2015.08.006.

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): 219–20. http://dx.doi.org/10.1139/gen-2015-0087.

Geary, J., E. Camicioli, and T. Bubela. (2013). The Public Face of Barcoding: How iBOL and DNA Barcoding Are Represented in Newspapers from around the World. Presented at the 5th International Barcode of Life Conference. Kunming, China, Oct 27-31.

Goodrick, E., and T. Reay. (2011). Constellations of Institutional Logics Changes in the Professional Work of Pharmacists. *Work and Occupations*. 38(3): 372–416. http://wox.sagepub.com/content/38/3/372.short.

Li, D.-Z., 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.

Madison, M. J. (2014). Commons at the Intersection of Peer Production, Citizen Science, and Big Data. In: *Governing Knowledge Commons*. eds. B. M. Frischmann, M. J. Madison, and K. J. Strandburg, 209–54. Oxford University Press.

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): 254–73. http://dx.doi.org/10.1089/omi.2013.0158.

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.

Works Cited

Adamowicz, S. J. (2015). International Barcode of Life: Evolution of a Global Research Community. *Genome*. 58(5): 151–62. http://dx.doi.org/10.1139/gen-2015-0094.

Agrawal, A., and C. C. Gibson. (2001). Introduction: The Role of Community in Natural Resource Conservation. In: *Communities and the Environment: Ethnicity, Gender, and the State in Community-Based Conservation*. eds. A. Agrawal and C. C. Gibson, 1–31. New Brunswick, New Jersey, and London: Rutgers University Press.

Andersson, K. (2006). Understanding Decentralized Forest Governance: An Application of the Institutional Analysis and Development Framework. *Sustainability: Science, Practice, & Policy*. 2(1): 25–35. http://search.proquest.com/openview/50b2eef538348d33635cad951e68368c/1?pq-origsite=gscholar&cbl=136137.

Australian Government. (2008). Biodiversity Hotspots. Text. *Department of the Environment and Energy*. September 2. http://www.environment.gov.au/biodiversity/conservation/hotspots.

Bagshaw, D., M. Lepp, and C. R. Zorn. (2007). International Research Collaboration: Building Teams and Managing Conflicts. *Conflict Resolution Quarterly*. 24(4): 433–46. http://dx.doi.org/10.1002/crq.183.

Balooni, K., J. F. Lund, C. Kumar, and M. Inoue. (2010). Curse or Blessing? Local Elites in Joint Forest Management in India's Shiwaliks. *International Journal of the Commons*. 4(2): 707–28. http://dx.doi.org/10.18352/ijc.217.

Barbuto, M., A. Galimberti, E. Ferri, M. Labra, R. Malandra, P. Galli, and M. Casiraghi. (2010). DNA Barcoding Reveals Fraudulent Substitutions in Shark Seafood Products: The Italian Case of "palombo" (Mustelus Spp.). *Food Research International*. 43(1): 376–81. http://dx.doi.org/10.1016/j.foodres.2009.10.009.

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. In: . https://gephi.org/users/publications/.

Basurto, X., G. Kingsley, K. McQueen, M. Smith, and C. M. Weible. (2010). A Systematic Approach to Institutional Analysis: Applying Crawford and Ostrom's Grammar. *Political Research Quarterly*. 63(3): 523–37. http://dx.doi.org/10.1177/1065912909334430.

Beas-Rodrigues, E. (2012). Property Rights, Biocultural Resources and Two Tragedies: Some Lessons from Brazil. In: *Genetic Resources and Traditional Knowledge: Case Studies and Conflicting Interests*. eds. T. Bubela and E. Gold, . Edward Elgar.

Behnke, R. (1994). Natural Resource Management in Pastoral Africa. *Development Policy Review*. 12(1): 5–28. http://dx.doi.org/10.1111/j.1467-7679.1994.tb00053.x.

Benin, S., and J. Pender. (2006). Collective Action in Community Management of Grazing Lands: The Case of the Highlands of Northern Ethiopia. *Environment and Development Economics*. 11(1): 127–49. http://dx.doi.org/10.1017/S1355770X05002688.

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): D36–42. 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): 160–87. 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.

Biodiversity Institute of Ontario. (2012). BOLD Systems. *Barcode of Life Data Systems*. http://www.boldsystems.org/views/login.php.

Biodiversity Institute of Ontario. (2013). School Malaise Trap Program. http://malaiseprogram.ca.

Biodiversity Institute of Ontario. (2015). Lifescanner. *Lifescanner*. http://lifescanner.net/.

Black, M. (1962). *Models and Metaphors: Studies in Language and Philosophy*. 1st Edition edition. Cornell University Press.

BOLD Systems. (2015). Public Data Portal - Record List. *BOLD Systems*. 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.

Boykoff, M. T., and J. M. Boykoff. (2007). Climate Change and Journalistic Norms: A Case-Study of US Mass-Media Coverage. *Geoforum*., Theme Issue: Geographies of Generosity, 38(6): 1190–1204. http://dx.doi.org/10.1016/j.geoforum.2007.01.008.

Brito, D. L., M. D. Intriligator, and E. Sheshinski. (1997). Privatization and the Distribution of Income in the Commons. *Journal of Public Economics*. 64(2): 181–205. http://dx.doi.org/10.1016/S0047-2727(96)01611-8.

Bromley, D. W. (1992). The Commons, Common Property, and Environmental Policy. *Environmental and Resource Economics*. 2(1): 1–17. http://dx.doi.org/10.1007/BF00324686.

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., and T. Caulfield. (2004). Do the Print Media "hype" Genetic Research? A Comparison of Newspaper Stories and Peer-Reviewed Research Papers. *Canadian Medical Association Journal*. 170(9): 1399–1407. http://dx.doi.org/10.1503/cmaj.1030762.

Bubela, T., J. Geary, M. Bieber, W. Luth, and D. Hik. (2015). Diffusing Barcoding: The Global Spread of a Good Idea. *Genome*., Scientific abstracts from the 6th International Barcode of Life Conference, 58(May): 201–2. http://dx.doi.org/10.1139/gen-2015-0087.

Bubela, T., and R. E. Gold, eds. (2012). *Genetic Resources and Traditional Knowledge: Case Studies and Conflicting Interests*. Cheltenham, UK and Northampton, MA: Edward Elgar.

Bubela, T., M. C. Nisbet, R. Borchelt, F. Brunger, C. Critchley, E. Einsiedel, G. Geller, et al. (2009). Science Communication Reconsidered. *Nature Biotechnology*. 27(6): 514–518. http://dx.doi.org/doi:10.1038/nbt0609-514.

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.

Burlando, R. M., and F. Guala. (2005). Heterogeneous Agents in Public Goods Experiments. *Experimental Economics*. 8(1): 35–54. http://dx.doi.org/10.1007/s10683-005-0436-4.

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): 6350–55. 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. *JAMA*. 287(21): 2847–50. 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.

Caulfield, T. (2004). Popular Media, Biotechnology, and the Cycle of Hype. *Houston Journal of Health Law & Policy*. 5: 213–32. http://heinonline.org/HOL/Page?handle=hein.journals/hhpol5&id=225&div=&collection=journals.

Caulfield, T., and T. Bubela. (2005). Media Representations of Genetic Research. In: *Crossing Over: Genomics in the Public Arena*. eds. E. F. Einsiedel and F. Timmermans, 17–30. Calgary, AB: University of Calgary Press.

Centre for Biodiversity Genomics. (2016). Organization. http://biodiversitygenomics.net/about/organization/.

Collins, F. S., M. Morgan, and A. Patrinos. (2003). The Human Genome Project: Lessons from Large-Scale Biology. *Science*. 300(5617): 286–90. http://dx.doi.org/10.1126/science.1084564.

Collins, R. A., and R. H. Cruickshank. (2013). The Seven Deadly Sins of DNA Barcoding. *Molecular Ecology Resources*. 13(6): 969–75. http://dx.doi.org/10.1111/1755-0998.12046.

Consortium for the Barcode of Life. (2015a). Mission Statement. *Barcode of Life*. http://www.barcodeoflife.org/content/about/mission-statement.

Consortium for the Barcode of Life. (2015b). What Is CBOL? *Barcode of Life*. http://www.barcodeoflife.org/content/about/what-cbol.

Consortium for the Barcode of Life. (2016). What Is DNA Barcoding? http://www.barcodeoflife.org/content/about/what-dna-barcoding.

Contreras, J. (2010). Data Sharing, Latency Variables, and Science Commons. *Berkeley Technology Law Journal*. 25(4): 1601–72. http://dx.doi.org/10.15779/Z38GW9F.

Contreras, J. (2014). Constructing the Genome Commons. In: *Governing Knowledge Commons*. eds. B. M. Frischmann, M. J. Madison, and K. J. Strandburg, 99–136. Oxford University Press.

Convention on Biological Diversity. 2012. List of Parties. *Convention on Biological Diversity*. Accessed November 25. http://www.cbd.int/convention/parties/list/.

Cook, D. M., E. A. Boyd, C. Grossmann, and L. A. Bero. (2007). Reporting Science and Conflicts of Interest in the Lay Press. *PLoS ONE*. 2(12): e1266. http://dx.doi.org/10.1371/journal.pone.0001266.

Costa, F. O., and G. R. Carvalho. (2007). The Barcode of Life Initiative: Synopsis and Prospective Societal Impacts of DNA Barcoding of Fish. *Life Sciences Society and Policy*. 3(2): 29. http://dx.doi.org/10.1186/1746-5354-3-2-29.

Crawford, S., and E. Ostrom. (1995). A Grammar of Institutions. *American Political Science Review*. 89(3): 582–600. http://dx.doi.org/10.2307/2082975.

Crawford, S., and E. Ostrom. (2005). A Grammar of Institutions. In: *Understanding Institutional Diversity*. ed. E. Ostrom, 137–74. Princeton University Press.

Dalton, D. L., and A. Kotze. (2011). DNA Barcoding as a Tool for Species Identification in Three Forensic Wildlife Cases in South Africa. *Forensic Science International*. 207(1–3): e51–54. http://dx.doi.org/10.1016/j.forsciint.2010.12.017.

Dedeurwaerdere, T. (2006). The Institutional Economics of Sharing Biological Information. *International Social Science Journal*. 58(188): 351–68. http://dx.doi.org/10.1111/j.1468-2451.2006.00623.x.

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). Institutionalizing Global Genetic Resource Commons: Towards Aternative Models for Facilitating Access in the Global Biodiversity Regime. *SSRN*., May. http://dx.doi.org/10.2139/ssrn.1611549.

Dedeurwaerdere, T. (2010c). Global Microbial Commons: Institutional Challenges for the Global Exchange and Distribution of Microorganisms in the Life Sciences. *Research in Microbiology*. 161(6): 414–21. 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, Part 1(January): 1–10. http://dx.doi.org/10.1016/j.envsci.2015.08.006.

Dedeurwaerdere, T., P. M. Stromberg, and U. Pascual. (2013). Social Motivations and Incentives in Ex Situ Conservation of Microbial Genetic Resources. In: *The Digital Public Domain: Foundations for an Open Culture*. eds. M. Dulong de Rosnay and J. C. De Martin, 111–124. Cambridge, UK: OpenBook Publishers.

Dentinger, B. T. M., M. Y. Didukh, and J.-M. Moncalvo. (2011). Comparing COI and ITS as DNA Barcode Markers for Mushrooms and Allies (Agaricomycotina). *PLoS ONE*. 6(9). http://dx.doi.org/10.1371/journal.pone.0025081.

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): 1905–16. 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): 4422–36. http://dx.doi.org/10.1111/j.1365-294X.2012.05642.x.

Eaton, M. J., G. L. Meyers, S.-O. Kolokotronis, M. S. Leslie, A. P. Martin, and G. Amato. (2009). Barcoding Bushmeat: Molecular Identification of Central African and South American Harvested Vertebrates. *Conservation Genetics*. 11(4): 1389–1404. http://dx.doi.org/10.1007/s10592-009-9967-0.

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.

Escalante, P., R. Desalle, and S.-O. Kolokotronis. (2010). The MexBOL Initiative. *Mitochondrial DNA*. 21(sup1): 1–2. http://dx.doi.org/10.3109/19401736.2010.551715.

Feeny, D., F. Berkes, B. J. McCay, and J. M. Acheson. (1990). The Tragedy of the Commons: Twenty-Two Years Later. *Human Ecology*. 18(1): 1–19. http://dx.doi.org/10.1007/BF00889070.

Field, D., S.-A. Sansone, A. Collis, T. Booth, P. Dukes, S. K. Gregurick, K. Kennedy, et al. (2009). 'Omics Data Sharing. *Science (New York, N.Y.)*. 326(5950): 234–36. http://dx.doi.org/10.1126/science.1180598.

Fischbacher, U., and S. Gaechter. (2006). Heterogeneous Social Preferences and the Dynamics of Free Riding in Public Goods. *SSRN*. March. http://papers.ssrn.com/abstract=874928.

Fleischman, F. D., N. C. Ban, L. S. Evans, N. C. Ban, L. S. Evans, G. Epstein, S. Villamayor-Tomas, G. Garcia-Lopez, and S. Villamayor-Tomas. (2014). Governing Large-Scale Social-Ecological Systems: Lessons from Five Cases. *International Journal of the Commons*. 8(2): 428–56. http://dx.doi.org/10.18352/ijc.416.

Freeman, R. B., and W. Huang. (2014). Collaboration: Strength in Diversity. *Nature*. 513(7518): 305. http://dx.doi.org/10.1038/513305a.

Freidson, E. (2001). *Professionalism, the Third Logic*. Chicago: The University of Chicago Press.

Frézal, L., and R. Leblois. (2008). Four Years of DNA Barcoding: Current Advances and Prospects. *Infection, Genetics and Evolution*. 8(5): 727–36. http://dx.doi.org/10.1016/j.meegid.2008.05.005.

Friedland, R., and R. Alford. (1991). Bringing Society Back In: Symbols, Practices and Institutional Contradictions. In: *The New Institutionalism in Organizational Analysis*. eds. W. Powell and P. Dimaggio, 232–63. University Of Chicago Press.

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.

Gamson, W. A., and A. Modigliani. (1989). Media Discourse and Public Opinion on Nuclear Power: A Constructionist Approach. *American Journal of Sociology*. 95(1): 1–37. http://www.jstor.org/stable/2780405.

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): 219–20. http://dx.doi.org/10.1139/gen-2015-0087.

Geary, J., E. Camicioli, and T. Bubela. (2013). The Public Face of Barcoding: How iBOL and DNA Barcoding Are Represented in Newspapers from around the World. Presented at the 5th International Barcode of Life Conference. Kunming, China, Oct 27-31.

Geary, J., E. Camicioli, and T. Bubela. (2016). DNA Barcoding in the Media: Does Coverage of Cool Science Reflect Its Social Context. *Genome*. 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.

Godfray, H. C. J. (2002). Challenges for Taxonomy. *Nature*. 417(6884): 17–19. http://dx.doi.org/10.1038/417017a.

Goodrick, E. (2002). From Management as a Vocation to Management as a Scientific Activity: An Institutional Account of a Paradigm Shift. *Journal of Management*. 28(5): 649–68. http://dx.doi.org/10.1177/014920630202800505.

Goodrick, E., and T. Reay. (2011). Constellations of Institutional Logics Changes in the Professional Work of Pharmacists. *Work and Occupations*. 38(3): 372–416. http://wox.sagepub.com/content/38/3/372.short.

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): R73–76. 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.

Hardin, G. (1968). The Tragedy of the Commons. *Science*. 162(3859): 1243–48. http://dx.doi.org/10.1126/science.162.3859.1243.

Hayo, B., and B. Vollan. (2012). Group Interaction, Heterogeneity, Rules, and Co-Operative Behaviour: Evidence from a Common-Pool Resource Experiment in South Africa and Namibia. *Journal of Economic Behavior and Organization*. 81(1): 9–28. http://dx.doi.org/10.1016/j.jebo.2011.09.002.

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): 313–21. http://dx.doi.org/10.1098/rspb.2002.2218.

Hebert, P. D. N., and T. R. Gregory. (2005). The Promise of DNA Barcoding for Taxonomy. *Systematic Biology*. 54(5): 852–59. http://dx.doi.org/10.1080/10635150500354886.

Helmy, M., M. Awad, and K. A. Mosa. 2016. Limited Resources of Genome Sequencing in Developing Countries: Challenges and Solutions. *Applied & Translational Genomics*. Accessed June 2. http://dx.doi.org/10.1016/j.atg.2016.03.003.

Hensmans, M. (2003). Social Movement Organizations: A Metaphor for Strategic Actors in Institutional Fields. *Organization Studies*. 24(3): 355–81. http://dx.doi.org/10.1177/0170840603024003908.

Hess, C. (2012). The Unfolding of the Knowledge Commons. *St Antony's International Review*. 8(1): 13–24. http://www.ingentaconnect.com/content/stair/stair/2012/00000008/00000001/art00003.

Hess, C., and E. Ostrom. (2006). A Framework for Analysing the Microbiological Commons. *International Social Science Journal*. 58(188): 335–349. http://dx.doi.org/10.1111/j.1468-2451.2006.00622.x.

Hess, C., and E. Ostrom. (2011a). Introduction: An Overview of the Knowlege Commons. In: *Understanding Knowledge as a Commons: From Theory to Practice*. eds. C. Hess and E. Ostrom, 3–26. MIT Press.

Hess, C., and E. Ostrom. (2011b). *Understanding Knowledge as a Commons: From Theory to Practice*. MIT Press.

Higginbottom, G. M. A. (2004). Sampling Issues in Qualitative Research. *Nurse Researcher*. 12(1): 7–19. http://dx.doi.org/10.7748/nr2004.07.12.1.7.c5927.

Hoffman, A. J. (1999). Institutional Evolution and Change: Environmentalism and the U.S. Chemical Industry. *The Academy of Management Journal*. 42(4): 351–71. http://dx.doi.org/10.2307/257008.

Hollingsworth, P. M., L. L. Forrest, J. L. Spouge, M. Hajibabaei, S. Ratnasingham, M. van der Bank, M. W. Chase, et al. (2009). A DNA Barcode for Land Plants. *Proceedings of the National Academy of Sciences*. 106(31): 12794–97. http://dx.doi.org/10.1073/pnas.0905845106.

iBOL. (2012). iBOL. *iBOL*. http://ibol.org/.

iBOL. (2015a). Campaigns. *International Barcode of Life*. http://ibol.org/about-us/campaigns/.

iBOL. (2015b). Data & Resource Sharing Policies. *International Barcode of Life*. http://ibol.org/resources/data-release-policy/.

iBOL. (2015c). International Barcode of Life. *International Barcode of Life*. http://ibol.org/working-groups/.

iBOL. (2015d). Partner Nations. *International Barcode of Life*. http://ibol.org/about-us/partner-nations/.

iBOL. (2015e). Project Governance. The Organization of the iBOL Team. *International Barcode of Life*. http://ibol.org/about-us/project-governance/#.

iBOL. (2015f). Sponsors | iBOL. *International Barcode of Life*. http://ibol.org/about-us/sponsors/.

iBOL. (2015g). The International Barcode of Life Project. *Genome Canada*. http://www.genomecanada.ca/genomicsOnTheHill/images2007/iBol/iBOLhandout.pdf.

iBOL. (2015h). WG 5.4 - Education Initiatives for Schools and Media. *International Barcode of Life*. http://ibol.org/wg-5-4-education-initiatives-for-schools-and-media/.

iBOL. (2015i). What Is iBOL | iBOL. *International Barcode of Life*. http://ibol.org/about-us/what-is-ibol/.

Imperial, M. T., and T. Yandle. (2005). Taking Institutions Seriously: Using the IAD Framework to Analyze Fisheries Policy. *Society & Natural Resources*. 18(6): 493–509. http://dx.doi.org/10.1080/08941920590947922.

Jansen, H. (2010). The Logic of Qualitative Survey Research and Its Position in the Field of Social Research Methods. *Forum Qualitative Sozialforschung / Forum: Qualitative*

Social Research. 11(2). http://www.qualitative-research.net/index.php/fqs/article/view/1450.

Janzen, D. H. (2004). Now Is the Time. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 359(1444): 731–32. http://dx.doi.org/10.1098/rstb.2003.1444.

Jaskiewicz, P., K. Heinrichs, S. B. Rau, and T. Reay. (2015). To Be or Not to Be: How Family Firms Manage Family and Commercial Logics in Succession. *Entrepreneurship Theory and Practice*, February. http://dx.doi.org/10.1111/etap.12146.

Jinbo, U., T. Kato, and M. Ito. (2011). Current Progress in DNA Barcoding and Future Implications for Entomology. *Entomological Science*. 14(2): 107–24. http://dx.doi.org/10.1111/j.1479-8298.2011.00449.x.

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.

Kahan, D. M. (2002). Reciprocity, Collective Action, and Community Policing. *California Law Review*. 90(5): 1513–39. http://dx.doi.org/10.2307/3481363.

Kahlor, L., and P. Stout, eds. (2009). *Communicating Science: New Agendas in Communication*. 1 edition. Routledge.

Kamau, E. C., B. Fedder, and G. Winter. (2010). The Nagoya Protocol on Access to Genetic Resources and Benefit Sharing: What Is New and What Are the Implications for Provider and User Countries and the Scientific Community? *Law and Environment and Development Journal*. 6(3): 246. http://www.lead-journal.org/content/10246.pdf.

Lee, M.-D. P., and M. Lounsbury. (2015). Filtering Institutional Logics: Community Logic Variation and Differential Responses to the Institutional Complexity of Toxic Waste. *Organization Science*. http://dx.doi.org/10.1287/orsc.2014.0959.

Li, D.-Z., 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.

Lindahl, T. (2012). Coordination Problems and Resource Collapse in the Commons - Exploring the Role of Knowledge Heterogeneity. *Ecological Economics*. 79: 52–59. http://dx.doi.org/10.1016/j.ecolecon.2012.04.016.

LMMC. (2002). Cancun Declaration of Like-Minded Megadiversity Countries. Cancun, Mexico. https://www.environment.gov.za/sites/default/files/docs/cancun_lmmc_declaration1.pdf.

Lounsbury, M. (2002). Institutional Transformation and Status Mobility: The Professionalization of the Field of Finance. *Academy of Management Journal*. 45(1): 255–66. http://dx.doi.org/10.2307/3069295.

Lounsbury, M. (2007). A Tale of Two Cities: Competing Logics and Practice Variation in the Professionalizing of Mutual Funds. *Academy of Management Journal*. 50(2): 289–307. http://dx.doi.org/10.5465/AMJ.2007.24634436.

Madison, M. J. (2014). Commons at the Intersection of Peer Production, Citizen Science, and Big Data. In: *Governing Knowledge Commons*. eds. B. M. Frischmann, M. J. Madison, and K. J. Strandburg, 209–54. Oxford University Press.

McComas, K. A., and L. M. Simone. (2003). Media Coverage of Conflicts of Interest in Science. *Science Communication*. 24(4): 395–419. http://dx.doi.org/10.1177/1075547003024004001.

McCombs, M. (2004). The Agenda Setting Approach. In: *Handbook of Political Communication*. ed. D. Nimmo, 121–40. Newbury Park, CA: Sage.

McDougall, C. L. (1995). *Intellectual Property Rights and the Biodiversity Convention: The Impact of GATT*. Friends of the Earth Trust.

McGinnis, M. D. (2011). An Introduction to IAD and the Language of the Ostrom Workshop: A Simple Guide to a Complex Framework. *Policy Studies Journal*. 39(1): 169–83. http://dx.doi.org/10.1111/j.1541-0072.2010.00401.x.

Melon, C. C., M. Ray, S. Chakkalackal, M. Li, J. E. Cooper, J. Chadder, W. Ke, et al. (2009). A Survey of South-North Health Biotech Collaboration. *Nature Biotechnology*. 27(3): 229–32. http://dx.doi.org/10.1038/nbt0309-229.

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.

Miller, S. E. (2007). DNA Barcoding and the Renaissance of Taxonomy. *Proceedings of the National Academy of Sciences*. 104(12): 4775–76. http://dx.doi.org/10.1073/pnas.0700466104.

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): 254–73. http://dx.doi.org/10.1089/omi.2013.0158.

Mittermeier, R. A. (1997). *Megadiversity: Earth's Biologically Wealthiest Nations*. CEMEX.

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.

Myers, N., R. A. Mittermeier, C. G. Mittermeier, G. A. B. da Fonseca, and J. Kent. (2000). Biodiversity Hotspots for Conservation Priorities. *Nature*. 403(6772): 853–58. http://dx.doi.org/10.1038/35002501.

Nakamura, Y., G. Cochrane, and I. Karsch-Mizrachi. (2013). The International Nucleotide Sequence Database Collaboration. *Nucleic Acids Research*. 41(D1): D21–24. http://dx.doi.org/10.1093/nar/gks1084.

Nature Education. (2014). Polymerase Chain Reaction / PCR | Learn Science at Scitable. http://www.nature.com/scitable/definition/polymerase-chain-reaction-pcr-110.

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): 893–94. 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/.

Nisbet, M. C. (2009a). Framing Science: A New Paradigm in Public Engagement. In: *Communicating Science: New Agendas in Communication*. eds. L. Kahlor and Stout, Patricia, , New York40–67. Taylor and Francis.

Nisbet, M. C. (2009b). Communicating Climate Change: Why Frames Matter for Public Engagement. *Environment: Science and Policy for Sustainable Development*. 51(2): 12–23. http://dx.doi.org/10.3200/ENVT.51.2.12-23.

Nisbet, M. C., D. Brossard, and A. Kroepsch. (2003). Framing Science The Stem Cell Controversy in an Age of Press/Politics. *The International Journal of Press/Politics*. 8(2): 36–70. http://dx.doi.org/10.1177/1081180X02251047.

Nisbet, M. C., and R. K. Goidel. (2007). Understanding Citizen Perceptions of Science Controversy: Bridging the Ethnographic—survey Research Divide. *Public Understanding of Science*. 16(4): 421–40. http://dx.doi.org/10.1177/0963662506065558.

Nisbet, M. C., and M. Huge. (2006). Attention Cycles and Frames in the Plant Biotechnology Debate Managing Power and Participation through the Press/Policy Connection. *The Harvard International Journal of Press/Politics*. 11(2): 3–40. http://hij.sagepub.com/content/11/2/3.

Nisbet, M. C., and B. V. Lewenstein. (2002). Biotechnology and the American Media The Policy Process and the Elite Press, 1970 to 1999. *Science Communication*. 23(4): 359–91. http://dx.doi.org/10.1177/107554700202300401.

Nisbet, M. C., and C. Mooney. (2007). Framing Science. *Science*. 316(5821): 56–56. http://dx.doi.org/10.1126/science.1142030.

Nisbet, M. C., and D. A. Scheufele. (2007). The Future of Public Engagement | The Scientist Magazine®. *The Scientist*. http://www.the-scientist.com/?articles.view/articleNo/25447/title/The-Future-of-Public-Engagement/.

Nisbet, M. C., D. A. Scheufele, J. Shanahan, P. Moy, D. Brossard, and B. V. Lewenstein. (2002). Knowledge, Reservations, or Promise? A Media Effects Model for Public Perceptions of Science and Technology. *Communication Research*. 29(5): 584–608. http://dx.doi.org/10.1177/009365002236196.

Ostrom, E. (1987). Institutional Arrangements for Resolving the Commons Dilemma: Some Contending Approaches. *The Question of the Commons. The Culture and Ecology of Communal Resources.* 250–265. http://dlc.dlib.indiana.edu/dlc/handle/10535/2274.

Ostrom, E. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action*. 1sted. Cambridge University Press.

Ostrom, E. (1992). Crafting Institutions for Self-Governing Irrigation Systems. http://agris.fao.org/agris-search/search/display.do?f=2012/OV/OV201205717005717.xml;US19920103636.

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. Russell Sage Foundation Publications.

Ostrom, E. (2005a). *Understanding Institutional Diversity*. Princeton, New Jersey: Princeton University Press.

Ostrom, E. (2005b). Understanding the Diversity of Structured Human Interactions. In: *Understanding Institutional Diversity*.3–31. Princeton University Press.

Ostrom, E., J. Burger, C. B. Field, R. B. Norgaard, and D. Policansky. (1999). Revisiting the Commons: Local Lessons, Global Challenges. *Science*. 284(5412): 278–82. http://dx.doi.org/10.1126/science.284.5412.278.

Ostrom, E., and R. Gardner. (1993). Coping with Asymmetries in the Commons: Self-Governing Irrigation Systems Can Work. *The Journal of Economic Perspectives*. 7(4): 93–112. http://dx.doi.org/10.2307/2138503.

Padma, T. V. (2005). India's Drug Tests. *Nature*. 436(7050): 485–485. http://dx.doi.org/10.1038/436485a.

Pellechia, M. G. (1997). Trends in Science Coverage: A Content Analysis of Three US Newspapers. *Public Understanding of Science*. 6(1): 49–68. http://dx.doi.org/10.1088/0963-6625/6/1/004.

Pennisi, E. (2003). Modernizing the Tree of Life. *Science*. 300(5626): 1692–97. http://dx.doi.org/10.1126/science.300.5626.1692.

Poteete, A. R., M. A. Janssen, and E. Ostrom. (2010a). Meta-Analysis: Getting the Big Picture through Synthesis. In: *Working Together: Collective Action, the Commons, and Multiple Methods in Practice*.89–112. Princeton University Press.

Poteete, A. R., M. A. Janssen, and E. Ostrom. (2010b). Overcoming Methodological Challenges. In: *Working Together: Collective Action, the Commons, and Multiple Methods in Practice*.3–30. Princeton University Press.

Poteete, A. R., M. A. Janssen, and E. Ostrom. (2010c). Small-N Case Studies: Putting the Commons under a Magnifying Glass. In: *Working Together: Collective Action, the Commons, and Multiple Methods in Practice*.3–30. Princeton University Press.

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): 435–61. 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): 355–64. http://dx.doi.org/10.1111/j.1471-8286.2007.01678.x.

Ray, M., A. S. Daar, P. A. Singer, and H. Thorsteinsdóttir. (2009). Globetrotting Firms: Canada's Health Biotechnology Collaborations with Developing Countries. *Nature Biotechnology*. 27(9): 806–14. http://dx.doi.org/10.1038/nbt0909-806.

Reay, T., and C. R. Hinings. (2009). Managing the Rivalry of Competing Institutional Logics. *Organization Studies*. 30(6): 629–52. http://dx.doi.org/10.1177/0170840609104803.

Reay, T., and C. R. (Bob) Hinings. (2005). The Recomposition of an Organizational Field: Health Care in Alberta. *Organization Studies*. 26(3): 351–84. http://dx.doi.org/10.1177/0170840605050872.

Reay, T., and C. Jones. (2016). Qualitatively Capturing Institutional Logics. *Strategic Organization*. 14(4): 441–54. http://dx.doi.org/10.1177/1476127015589981.

Ruttan, L. M. (2006). Sociocultural Heterogeneity and the Commons. *Current Anthropology*. 47(5): 843–53. http://dx.doi.org/10.1086/507185.

Ruttan, L. M. (2008). Economic Heterogeneity and the Commons: Effects on Collective Action and Collective Goods Provisioning. *World Development*. 36(5): 969–85. http://dx.doi.org/10.1016/j.worlddev.2007.05.005.

Sakane, N., M. T. van Wijk, M. Langensiepen, and M. Becker. (2014). A Quantitative Model for Understanding and Exploring Land Use Decisions by Smallholder Agrowetland

Households in Rural Areas of East Africa. *Agriculture, Ecosystems and Environment*. 197: 159–73. http://dx.doi.org/10.1016/j.agee.2014.07.011.

Salager-Meyer, F. (2008). Scientific Publishing in Developing Countries: Challenges for the Future. *Journal of English for Academic Purposes*, English for Research Publication Purposes, 7(2): 121–32. http://dx.doi.org/10.1016/j.jeap.2008.03.009.

Scheufele, D. (1999). Framing as a Theory of Media Effects. *Journal of Communication*. 49(1): 103–22. http://dx.doi.org/10.1111/j.1460-2466.1999.tb02784.x.

Schindel, D. (2010). Biology without Borders. *Nature*. 467(7317): 779–81. 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. Bonn, Germany. https://69.90.183.227/abs/doc/cbol-dfg-bonn-abs-article-en.pdf.

Schindel, D., and S. Miller. (2005). DNA Barcoding a Useful Tool for Taxonomists. *Nature*. 435(7038): 17–17. http://dx.doi.org/10.1038/435017b.

Schoch, C. L., K. A. Seifert, S. Huhndorf, V. Robert, J. L. Spouge, C. A. Levesque, W. Chen, et al. (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): 6241–46. 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.

Schweik, C., and R. English. (2012). *Internet Success: A Study of Open-Source Software Commons*. Cambridge, Massachusetts; London, England: MIT Press. https://mitpress.mit.edu/books/internet-success.

Shiferaw, B., T. A. Kebede, and V. R. Reddy. (2012). Community Watershed Management in Semiarid India: The State of Collective Action and Its Effects on Natural Resources and Rural Livelihoods. In: *Collective Action and Property Rights for Poverty Reduction: Insights from Africa and Asia*.153–88. University of Pennsylvania Press.

Siddiki, S., X. Basurto, and C. M. Weible. (2012). Using the Institutional Grammar Tool to Understand Regulatory Compliance: The Case of Colorado Aquaculture. *Regulation & Governance*. 6(2): 167–88. http://dx.doi.org/10.1111/j.1748-5991.2012.01132.x.

Siddiki, S., C. M. Weible, X. Basurto, and J. Calanni. (2011). Dissecting Policy Designs: An Application of the Institutional Grammar Tool. *Policy Studies Journal*. 39(1): 79–103. http://dx.doi.org/10.1111/j.1541-0072.2010.00397.x.

Smith, V. S. (2005). DNA Barcoding: Perspectives from a "Partnerships for Enhancing Expertise in Taxonomy" (PEET) Debate. *Systematic Biology*. 54(5): 841–44. http://dx.doi.org/10.1080/10635150500354894.

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*., no. 365(December): 307–28. http://dx.doi.org/10.3897/zookeys.365.6027.

Stemler, S. E. (2004). A Comparison of Consensus, Consistency, and Measurement Approaches to Estimating Interrater Reliability. *Practical Assessment, Research & Evaluation*. 9(4). http://PAREonline.net/getvn.asp?v=9&n=4.

Stoeckle, M. (2003). Taxonomy, DNA, and the Bar Code of Life. *BioScience*. 53(9): 796–97. http://dx.doi.org/10.1641/0006-3568(2003)053[0796:TDATBC]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.

Strasser, B. J. (2008). GenBank--Natural History in the 21st Century? *Science*. 322(5901): 537–38. http://dx.doi.org/10.1126/science.1163399.

Suber, P. (2011). Introduction: An Overview of the Knowlege Commons. In: *Understanding Knowledge as a Commons: From Theory to Practice*. eds. C. Hess and E. Ostrom, 171–208. MIT 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.

Tewksbury, D., and S. L. Althaus. (2000). Differences in Knowledge Acquisition among Readers of the Paper and Online Versions of a National Newspaper. *Journalism & Mass Communication Quarterly*. 77(3): 457–79. http://dx.doi.org/10.1177/107769900007700301.

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.

Thornton, P. H. (2002). The Rise of the Corporation in a Craft Industry: Conflict and Conformity in Institutional Logics. *The Academy of Management Journal*. 45(1): 81–101. http://dx.doi.org/10.2307/3069286.

Thornton, P. H. (2004). *Markets from Culture: Institutional Logics and Organizational Decisions in Higher Education Publishing*. Stanford University Press.

Thornton, P. H., C. Jones, and K. Kury. (2005). Institutional Logics and Institutional Change in Organizations: Transformation in Accounting, Architecture, and Publishing. *Research in the Sociology of Organizations*. 23: 125–170.

Thornton, P. H., and W. Ocasio. (1999). Institutional Logics and the Historical Contingency of Power in Organizations: Executive Succession in the Higher Education Publishing Industry, 1958- 1990. *American Journal of Sociology*. 105(3): 801–43. http://dx.doi.org/10.1086/210361.

Thornton, P. H. (2008). Institutional Logics. In: *The SAGE Handbook of Organizational Institutionalism*. eds. R. Greenwood, C. Oliver, R. Suddaby, and K. Sahlin-Andersson, . Sage Publications Ltd.

Thornton, P. H., W. Ocasio, and M. Lounsbury. (2012). *The Institutional Logics Perspective: A New Approach to Culture, Structure and Process*. Oxford University Press, USA.

Thorsteinsdottir, H., U. Quach, A. S. Daar, and P. A. Singer. (2004). Conclusions: Promoting Biotechnology Innovation in Developing Countries. *Nature Biotechnology*. 22: DC48-DC52. http://dx.doi.org/10.1038/nbt1204supp-DC48.

Toronto International Data Release Workshop Authors. (2009). Prepublication Data Sharing. *Nature*. 461(7261): 168–70. http://dx.doi.org/10.1038/461168a.

Tyler, T. R. (1990). *Why People Obey the Law: Procedural Justice, Legitimacy, and Compliance*. New Haven, CT: Yale University Press.

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.

Vaismoradi, M., H. Turunen, and T. Bondas. (2013). Content Analysis and Thematic Analysis: Implications for Conducting a Qualitative Descriptive Study. *Nursing & Health Sciences*. 15(3): 398–405. http://dx.doi.org/10.1111/nhs.12048.

Varughese, G., and E. Ostrom. (2001). The Contested Role of Heterogeneity in Collective Action: Some Evidence from Community Forestry in Nepal. *World Development*. 29(5): 747–65. http://dx.doi.org/10.1016/S0305-750X(01)00012-2.

Vedeld, T. (2000). Village Politics: Heterogeneity, Leadership and Collective Action. *Journal of Development Studies*. 36(5): 105–34. http://dx.doi.org/10.1080/00220380008422648.

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 Biol*. 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.

WIPO. (2016). World Intellectual Property Organization Intellectual Property Laws and Treaties Database (WIPO Lex). http://www.wipo.int/wipolex/en/.

Wong, E. H.-K., and R. H. Hanner. (2008). DNA Barcoding Detects Market Substitution in North American Seafood. *Food Research International*. 41(8): 828–37. 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, et al. (2008). Potential Use of DNA Barcodes in Regulatory Science: Applications of the Regulatory Fish Encyclopedia. *Journal of Food Protection*. 71(1): 210–17.

Yin, R. K. (2009). *Case Study Research: Design and Methods*. 4th ed. Vol. 5. Applied Social Research Methods. SAGE.

Appendices

Appendix 1: Interview guide for chapter 2

Researcher

I. Biography / Attributes / Conditions

1. Can you please tell me about your research programs? [PROMPT: how funded (public research grant, industry, other?, in kind contribution)? Scale and scope]
2. Can you please describe your past and current research activities in the field of [INSERT PERSONALISED FIELD OF EXPERTISE]?

II. Research Partnership/Collaborations

3. Do you have any current and/or future research collaborations?
    i. If yes:
        a. Please describe your current collaborative projects (PROMPT: with whom you collaborate; institutions, individuals, industry)
        b. Where are they located?
        c. What is the nature of collaboration?
        d. What factors did you consider when you enter into the collaboration? (PROMP: trust, reciprocity, previous collaboration relationship, social learning)
    ii. If no:
        e. Is there any specific reason for not participating in any collaborative project?
4. Why do you collaborate with other organizations and individuals?
5. What are the costs and benefits of participating in international collaborative projects?
6. What do you expect when you enter into a collaborative research project?
7. What did you learn from your past collaborative work in terms of the management of research project and knowledge sharing?

III. Source of Genetic Resources and Equity

8. Where do you get genetic resources from? (PROMPT: Who collects and how? How the material reaches you, and in what form?)
9. What legal agreements such as Material Transfer Agreements do you enter into to access genetic resources?
10. If yes, what have your experiences been with Material Transfer Agreements (PROMPT: Positive or Negative?)
11. Do you have any framework and mechanism to share royalties with communities or organizations that provide you with genetic resources?
12. Do you have any research partnership with the communities that provide you genetic material? How do you see them and their knowledge about genetic material?

IV. IP and Marketing Strategies

13. Have you ever sought to commercialize any of your research? Please explain.

[How did you do it? Did you go through your Technology Transfer Office? Were you successful?]

14. Have you ever been blocked in any of your research endeavors by intellectual property? Please explain.
15. Do you share genetic material with other researchers in your collaborative team? (local, national, foreign?) If yes, on what conditions? If no, why? Please explain.

V. Genetic Resources and the *Convention on Biological Diversity*

16. How do you perceive and understand regulations that govern access to genetic resources and benefit-sharing under the Convention on Biological Diversity (*CBD*)?
17. What do you think of the national implementation of *CBD*, and how do you compare it with the pre-*CBD* era and its implications for your R&D activities? Please explain.

VI. The Genetic Resource Commons: Efficiency, Sustainability and Equity

18. What are your perceptions and understanding of biomaterials repositories and databases? Please explain.
19. Do you think developing large-scale, international biomaterials repositories and databases is desirable or feasible in your field of research? If yes, how? If no, why?
20. What do you think the overarching objective of a large-scale, international biomaterials repository or database should be?
21. What are the costs and benefits of large-scale, international repositories and databases to the participants? Please explain. [Prompt on whether the repositories and databases would benefit all participants equally]
22. What do you think are the main non-science challenges in setting up large-scale, international repositories and databases? Please explain (PROMPT: regulations, sustainability of funding, legal barriers, access to material, cultural barriers in terms of willingness to contribute materials or data)
23. What incentives and safeguards would you require to deposit your materials and data in large-scale international repositories and databases?
24. What kind of regulatory framework should be put in place for the effective and sustainable functioning of a biomedical repository?
    i. Who should and should not be allowed to participate? Or, what are the criteria of participation?
    ii. What kind of administrative structure should be developed?
    iii. What kind of institutions and infrastructure should be developed, and who should perform what actions for the establishment of a biomedical repository? (PROMPT: management structure, common ground/ rules, common language)
    iv. What mechanism should be developed to decide on what kind of information must, may, or must not be shared, and how?
    v. What kind of incentive and benefit-sharing model would work effectively for a biomedical repository?
    vi. What do you think are the best methods to handle differences (institutional, socio-cultural, geopolitical) among various partners?
25. Is there anything else you would like to add?

Policy Maker

Interview Guide (Policy Makers)

VII. Biography / Attributes / Conditions

1. Can you please tell me about your department/institution and its activities? [PROMPT: scale and scope of your organizational activities]
2. Can you please describe your role within your organization and your past and current activities in the field of [INSERT PERSONALISED FIELD OF EXPERTISE]?

VIII. The Genetic Resource Commons: Efficiency, Sustainability and Equity

3. What are your perceptions and understanding of biomaterials repositories and databases? Please explain.
4. What biomedical repositories you support? Why do you support?
5. What are the basic criteria to fund a biomedical repository?
6. What kind of measures you think would sustain biomedical repositories? What do you think are the major barriers to the sustainability of biomedical repositories?
7. What do you expect from biomedical repositories? Who do you think are the beneficiaries of biomedical repositories?
8. Do you think there should be some policies to govern data and biomaterial sharing? If yes, please explain. If not, why not? Please explain.
9. What kind of institutional mechanism you think should be developed to allow stakeholders to provide their input for policy formulation, implementation and evaluation?
10. How do you perceive and understand regulations that govern access to genetic resources and benefit-sharing under the Convention on Biological Diversity (*CBD*)?
11. What do you think of the national implementation of *CBD*, and how do you compare it with the pre-*CBD* era and its implications for R&D activities? Please explain.
12. Do you think developing large-scale, international biomaterials repositories and databases is desirable or feasible in your field of research? If yes, how? If no, why?
13. What do you think the overarching objective of a large-scale, international biomaterials repository or database should be?
14. What are the costs and benefits of large-scale, international repositories and databases to the participants? Please explain. [Prompt on whether the repositories and databases would benefit all participants equally]
15. What do you think are the main non-science challenges in setting up large-scale, international repositories and databases? Please explain (PROMPT: regulations, sustainability of funding, legal barriers, access to material, cultural barriers in terms of willingness to contribute materials or data)
16. What incentives and safeguards would you require to deposit your materials and data in large-scale international repositories and databases?
17. What kind of regulatory framework should be put in place for the effective and sustainable functioning of a biomedical repository?
    a. Who should and should not be allowed to participate? Or, what are the criteria of participation?
    b. What kind of administrative structure should be developed?

c. What kind of institutions and infrastructure should be developed, and who should perform what actions for the establishment of a biomedical repository? (PROMPT: management structure, common ground/ rules, common language)
d. What mechanism should be developed to decide on what kind of information must, may, or must not be shared, and how?
e. What kind of incentive and benefit-sharing model would work effectively for a biomedical repository?
f. What do you think are the best methods to handle differences (institutional, socio-cultural, geopolitical) among various partners?

18. In your opinion, what is the mission of universities and public research centres? How do you understand the public and private partnership in research and development?
19. How do you understand the co-management of resources by community, industry and researchers? Are there any best practices of co-management? Please explain.
20. Do you suggest any other mechanism or process that can strike the middle ground between commercialization and public access?

Appendix 2: Author name disambiguation methods

Authors: Luth W, Geary J, Bieber M.

We identified unique author names in our database of DNA barcoding papers using a multi-step unsupervised author clustering system. Current supervised cluster methods are more precise, however, they require trained human annotation of author records (Ferreira et al., 2010). In datasets as large as ours, supervised cluster methods are unfeasible (Ferreira et al., 2010). Therefore, we developed an unsupervised author clustering system that uses existing unique author identification numbers in Scopus and an existing disambiguation tool.

We addressed the lack of metadata available for authors by doing keyword searches on the web for departments and institutions. We then used a geolocation algorithm to generate country affiliations for all putative authors. For putative authors in Canada, the United States, and Australia we also collected state, city, and postal code data.

We organized authors by last name, first name, middle name (where applicable). We created author clusters based on matching last names and first initials (Ferreira, Gonçalves & Laender, 2012), which is the minimal name information in citations (Han et al. 2004).

We then disambiguated the ambiguous combinations of last name and first initial. The first step of this process was to use Scopus IDs. If author names had the same Scopus ID, they were considered the same, as the Scopus Author ID has been shown to reliably disambiguate authors (Kawashima & Tomizawa, 2015). If author names did not have unique Scopus IDs, we used the Harvard Disambiguation to disambiguate author clusters based on matching PubMed IDs.

We used metadata linked to each author name, including institution, country, and co-author data, to identify matching authors in the remaining clusters. We considered the author names remaining after this step to be unique. Using Reijnhoudt et al.'s (2014) calculation of the accuracy of Scopus Author IDs, our author disambiguation process has a recall of 95.5% and precision of 87.2%. With Kwashima and Tomizawa's (2015) Scopus Author ID calculations, the recall and precision of our author disambiguation process are 97.5% and 98.8% respectively.

Works Cited

Torvik, V. I. & Smalheiser, N. R. (2009). Author Name Disambiguation in MEDLINE. *ACM Trans Knowl Discov Data* 3(3).

Strotmann, A. & Zhao, D. (2012). Author name disambiguation: What difference does it make in author-based citation analysis? *J Am Soc Inf Sci Tec* 63(9).

Smalheiser, N. R. & Torvik, V. I. (2009). Author name disambiguation. *Ann Rev Info Sci Tech* 43(1).

Reijnhoudt, L., Costas, R., Noyons, E., Börner, K. & Scharnhorst A. (2014). 'Seed + expand': a general methodology for detecting publication oeuvres of individual researchers. *Scientometrics* 101(2).

Kwashima, H. & Tomizawa, H. (2015). Accuracy evaluation of Scopus Author ID based on the largest funding database in Japan. *Scientometrics* 103(3).
http://link.springer.com/article/10.1007/s11192-015-1580-z?no-access=true

Ferreira, A. A., Veloso, A., Gonçalves, M. A. & Laender, A. H. F. (2010). Effective Self-training Author Name Disambiguation in Scholarly Digital Libraries. Proceedings of the 10th Annual Joint Conference on Digital Libraries. New York: ACM.

Ferreira, A. A., Gonçalves, M. A. & Laender, A. H. F. (2012) A brief survey of automatic methods for author name disambiguation. *SIGMOD Rec* 41(2)

Han, H., Giles, L., Zha, H., Li, C. & Kostas Tsioutsiouliklis. (2004). Two Supervised Learning Approaches for Name Disambiguation in Author Citations. In *JDCL '04: Proceedings of the 4th ACM/IEEE Joint Conference on Digital Libraries.* New York: ACM.

Appendix 3: The number of articles with authors from each of the five world regions, and combinations of the regions.

# of Regions	Region	Number of Articles	% Total Articles
1	West Europe and Others (W)	1,917	53.9
1	Asia Pacific (AP)	616	17.3
1	Latin American and the Caribbean (L)	141	4.0
1	Eastern Europe (E)	83	2.3
1	Africa (AF)	32	0.9
2	W, AP	266	7.5
2	W, L	188	5.3
2	W, E	125	3.5
2	W, AF	91	2.6
2	AP, E	8	0.2
2	AP, AF	5	0.1
2	AP, L	3	0.1
2	E, AF	1	0.0
2	L, AF	1	0.0
3	W, AP, E	22	0.6
3	W, AP, AF	13	0.4
3	W, AP, L	11	0.3
3	W, L, E	11	0.3
3	W, L, AF	7	0.2
3	W, E, AF	6	0.2
4	W, AP, L, AF	3	0.1
4	W, AP, L, E	2	0.1
4	W, AP, E, AF	1	0.0
5	W, AP, L, E, AF	4	0.1

Appendix 4: Interview Guide for Chapter 3

I. Biography / Attributes / Conditions

1. Give a brief history of your career up until now, leading in to how you first got involved in IBOL and DNA barcoding efforts?

   PROMPT: how work is funded, scale and scope,

II. Source of Genetic Resources and Equity

2. Where do you get genetic resources (data, specimens, DNA) from? Describe how you use genetic resources from the process of collecting samples to creating and analyzing data.
3. Do you have any type of agreements (legal?) such as Material Transfer Agreements to guide accessing and using GR?

   PROMPT: If yes, describe the process of entering into an agreement

4. How do you perceive and understand regulations that govern access to genetic resources and benefit-sharing under the Convention on Biological Diversity (CBD)?

   PROMPT: Does the CBD apply to your work? Why or why not?

5. Do you have any framework and mechanism to share benefits with communities or organizations that provide you with genetic resources?

   PROMPT: What type of benefits would you be able to share? What type of benefits should be shared?

   IF: How do you think the community you partner with understands GR and benefits?

III. Research Partnership/Collaborations

6. Do you have any research partners that you share GR with?
   a. If yes:
   a. Please describe your current collaborative projects (PROMPT: with whom you collaborate; institutions, individuals, industry)
   b. Where are they located?
   c. What is the nature of collaboration?
   d. What factors did you consider when you enter into the collaboration? (PROMP: trust, reciprocity, previous collaboration relationship, social learning)
   b. If no:
   e. Is there any specific reason for not participating in any collaborative project?
7. Why do you share GR with other organizations and individuals?
8. What are the costs and benefits of sharing GR in collaborative projects?

IV. The Genetic Resource Commons: Efficiency, Sustainability and Equity

9. What is your main motivation for contributing to a resource like BOLD?
10. What do you think the goals of iBOL are? Do you think these are feasible? If yes, how? If no, why? What should the goals be?
11. What incentives and safeguards would you require to deposit your GR in large-scale international repositories and databases?
12. What kind of regulatory framework should be put in place for the effective and sustainable functioning of BOLD and material repositories related to iBOL?
    a. Who should and should not be allowed to participate? Or, what are the criteria of participation?
       PROMPT: management structure, common ground/ rules, common language)
    b. What mechanism should be developed to decide on what kind of information must, may, or must not be shared, and how?
    c. What kind of incentive and benefit-sharing model would work effectively for BOLD and material repositories related to iBOL?
    d. What do you think are the best methods to handle differences (institutional, socio-cultural, geopolitical) among various partners?
13. Is there anything else you would like to add?

Appendix 5: Search algorithms used to identify articles on DNA barcoding or Barcode of Life projects

LexisNexis

(dna W/5 (barcod! OR bar code!)) OR (gene w/5 (barcod! OR bar code!)) OR (genes w/5 (barcod! OR bar code!)) OR (genetic w/5 (barcod! OR bar code!)) OR (life W/3 (barcod! OR bar code!))

Factiva

(dna near5 barcod*) OR (dna near5 "bar code") OR (dna near5 "bar codes") OR (dna near5 "bar coding") OR (gene near5 barcod*) OR (gene near5 "bar code") OR (gene near5 "bar codes") OR (gene near5 "bar coding") OR (genes near5 barcod*) OR (genes near5 "bar code") OR (genes near5 "bar codes") OR (genes near5 "bar coding") OR (genetic near5 barcod*) OR (genetic near5 "bar code") OR (genetic near5 "bar codes") OR (genetic near5 "bar coding") OR (barcod* near3 life) OR ("bar code" near3 life) OR ("bar codes" near3 life) OR ("bar coding" near3 life)

Canadian Newsstand

(dna N/5 barcod*) OR (dna N/5 "bar code") OR (dna N/5 "bar codes") OR (dna N/5 "bar coding") OR (gene N/5 barcod*) OR (gene N/5 "bar code") OR (gene N/5 "bar codes") OR (gene N/5 "bar coding") OR (genes N/5 barcod*) OR (genes N/5 "bar code") OR (genes N/5 "bar codes") OR (genes N/5 "bar coding") OR (genetic N/5 barcod*) OR (genetic N/5 "bar code") OR (genetic N/5 "bar codes") OR (genetic N/5 "bar coding") OR (barcod* N/3 life) OR ("bar code" N/3 life) OR ("bar codes" N/3 life) OR ("bar coding" N/3 life)

We limited results from each search to only include the targeted date range, countries, and language.

Appendix 6: Categories and options used to code newspaper articles.

Newspaper section of article
Front page/general news
International News
National News
Local News
Commentary/Focus
Debate
Consumer
Editorial Page
Culture
Business
Science & Technology
Health & Medicine
Environment
Letters to the Editor
Entertainment
Sport
Lifestyle
Other, please indicate
Unknown
News Format
Article with latest News
Investigation, reportage, background
Interview (mainly)
Opinion Piece
Editorial (from the paper's editor)
Commentary from other people
Letters to the editor
Review of books, films etc.
Other
How prominently does DNA barcoding feature in the article?
Mentioned only in passing
Small and unimportant component of article
Large and integral component of overall story that includes other topics
Main focus of article
Is a scientific publication referenced?
No
Yes (specify** reference)
Are any national laws or regulatory agencies referenced?
No
Yes (specify** laws or regulatory agencies)

Whose voices are represented in the article (either directly quoted or mentioned as sources)
Researcher
General Public
Government
Industry
Environmental Group
No Sources are specified
Other (specify others)
Frame of article
Descriptive (a purely descriptive account of basic scientific research)
Progress (a celebration of new development)
Economic prospect (economic potential; prospects for investment and profits)
Ethical (call upon ethical principles)
Risks before the Event (call for restraint in the face of unknown risk)
Risks after the Event (fatalism after the innovation)
Public accountability (call for public control; participation; public involvement; regulatory mechanisms)
Globalization (call for global perspective; national competitiveness within an economy or isolationism)
Profile/Human Interest Story
Controversy
Main theme of article* (topic prominently placed first in article)
Monitoring invasive species
Verifying food labelling or monitoring food systems
Bioconservation efforts
Forensic investigations
Barcode of Life project or collaboration (iBOL, FISHBOL, etc)
Monitoring species that impact human health (disease vectors, traditional or plant medicines)
Science (taxonomy, biology, genetics, etc)
Other, not related to DNA barcoding or iBOL
Is the process of DNA barcoding described?
No, it is mentioned but not described
Partially, there is some description of the process of using DNA barcoding as a tool, but it's incomplete
Yes, DNA barcoding is described as using short regions of DNA to identify species through comparison to a reference database
Type of DNA barcoding projects mentioned
Contribute records to the barcode database
Use records from the barcode database
Develop methods or evaluate barcoding as a tool

Are funding sources mentioned?
No
Yes (specify** funding sources)
Are partnerships mentioned?
No
Yes (specify** all partnerships mentioned)
Does the article present the removal of genetic resources for commercial development in another country?
No
Yes, positively as bioprospecting
Yes, negatively as biopiracy
Whose ownership rights are emphasized?
Are any international treaties mentioned?
No
Yes, CITES
Yes, CBD
Yes, Nagoya Protocol
Yes, other (specify other)
Is the location where genetic resources used for barcoding were collected included?
No
Yes (specify** all locations mentioned)

*Main theme was coded as "other" unless the topic was specifically related to BOL organizations or DNA barcoding. The same topics were also recorded if they were present anywhere each article, reported separately.
** Author JG coded questions that required adding more specific information after the initial data collection, and determined relevant categories based on the data collected. We estimated percent agreement based on the *a priori* coding categories presented in this table.

Appendix 7: Ethics approval

Approval Form

Date:	October 6, 2010
Principal Investigator:	Tania Bubela
Study ID:	Pro00016720
Study Title:	Building Robust Research Commons to Support Large-Scale International Research Projects in Genomics and Life Sciences
Approval Expiry Date:	October 5, 2011
Sponsor/Funding Agency:	Genome Canada GC

Thank you for submitting the above study to the Health Research Ethics Board - Health Panel . Your application, including revisions received today, has been reviewed and approved on behalf of the committee.

A renewal report must be submitted next year prior to the expiry of this approval if your study still requires ethics approval. If you do not renew on or before the renewal expiry date, you will have to re-submit an ethics application.

Approval by the Health Research Ethics Board does not encompass authorization to access the patients, staff or resources of Alberta Health Services or other local health care institutions for the purposes of the research. Enquiries regarding Alberta Health Services administrative approval, and operational approval for areas impacted by the research, should be directed to the Alberta Health Services Regional Research Administration office, #1800 College Plaza, phone (780) 407-6041.

Sincerely,

Doug Gross, Ph.D.
Associate Chair, Health Research Ethics Board - Health Panel

Note: This correspondence includes an electronic signature (validation and approval via an online system).