



BIOSCAN: ILLUMINATING BIODIVERSITY AND SUPPORTING SUSTAINABILITY

by DONALD HOBERN
Executive Secretary, International Barcode of Life Consortium

The International Barcode of Life Consortium (iBOL) launched its new research program BIOSCAN in June 2019, to scale up its efforts to inventory life on Earth at a time when an ecological crisis is threatening the planet.

Recent reports from the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) and the Intergovernmental Panel on Climate Change (IPCC) have highlighted the scale of the pressures that threaten the environment and that are triggering a massive extinction event. Public awareness of these issues is growing and there are increasing demands for policymakers to work to support the environment and to focus on sustainable solutions.

Large-scale datasets are key to empowering societies and politicians to make these changes. Such data are available for some global systems, such as climate and land cover, and national scale datasets are often

available for agriculture, human population, and land use. However, at present, biodiversity is not represented at the level of detail or at the scale and frequency required to support decision-making.

iBOL has been acquiring growing volumes of data on species and their distributions since 2010 with their first research program BARCODE 500K. By 2015, the program had delivered DNA barcodes representing 500,000 species via its online database called the Barcode of Life Data System (BOLD). These standardized reference sequences have offered researchers everywhere a transformational tool for rapid species identification as well as a range of applications across taxonomy, biogeography, ecology, biosecurity, and conservation. The benefits to researchers, policymakers, and the wider public are likely to be even greater through widespread adoption of metabarcoding as a survey tool. Metabarcoding uses DNA barcodes for cheap and efficient assessment of which species are

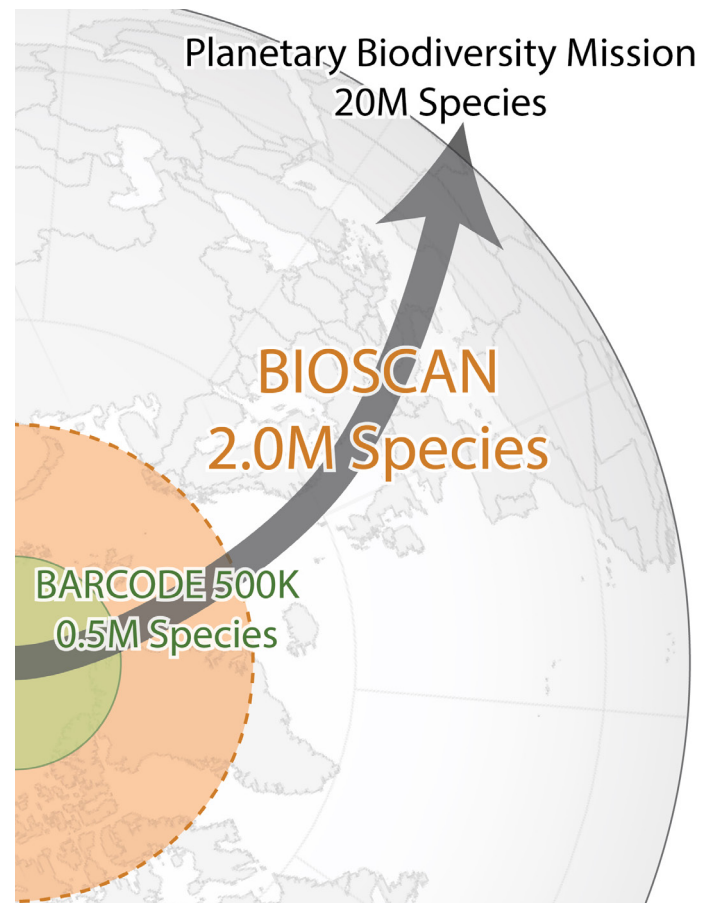
found in a bulk sample or have left residual traces of their DNA in water, soil, and other substrates (“environmental DNA” or eDNA).

Species identification has always been a central challenge for biological research, a task that has relied on the skill-base of the international taxonomic community and the deep and complex foundation of a quarter millennium of work naming and describing species. The importance and difficulty of being able to assign a name to any arbitrary organism of interest and the shortage of trained taxonomists and curators to do this work has become known as the taxonomic impediment and is recognized as an international problem. DNA barcoding has already revolutionized approaches and expectations around detection and diagnosis of species of interest. These changes have been most significant in contexts where morphological taxonomy has been most difficult, such as separation of cryptic species, identification of fragments or products derived from organisms, and recognition of species from poorly-characterized life stages.

BIOSCAN is accelerating support for reviewing and describing the millions of species still lacking scientific names. The Barcode Index Number (BIN) system offered by BOLD simplifies analysis and presentation of well-defined sets of specimens as diagnosable units of biodiversity. Each BIN represents a cluster of individuals that show minimal variation in the standard barcode markers and, in many cases, these clusters will correspond to different species that live and reproduce separately in the environment.

Since organisms can be assigned to a BIN even when no scientific name is available and even when the exact taxonomic significance of the BIN is unclear, the expanded collecting and sequencing effort planned for BIOSCAN can both assist taxonomists to work more rapidly and efficiently and can offer an interim framework for categorizing and mapping taxonomic units pending full taxonomic review. The significance of such a framework cannot be underestimated. Without a proper and timely catalogue of the units of biodiversity, we cannot fully study or understand the species with which we share the planet and with which our own future is intertwined.

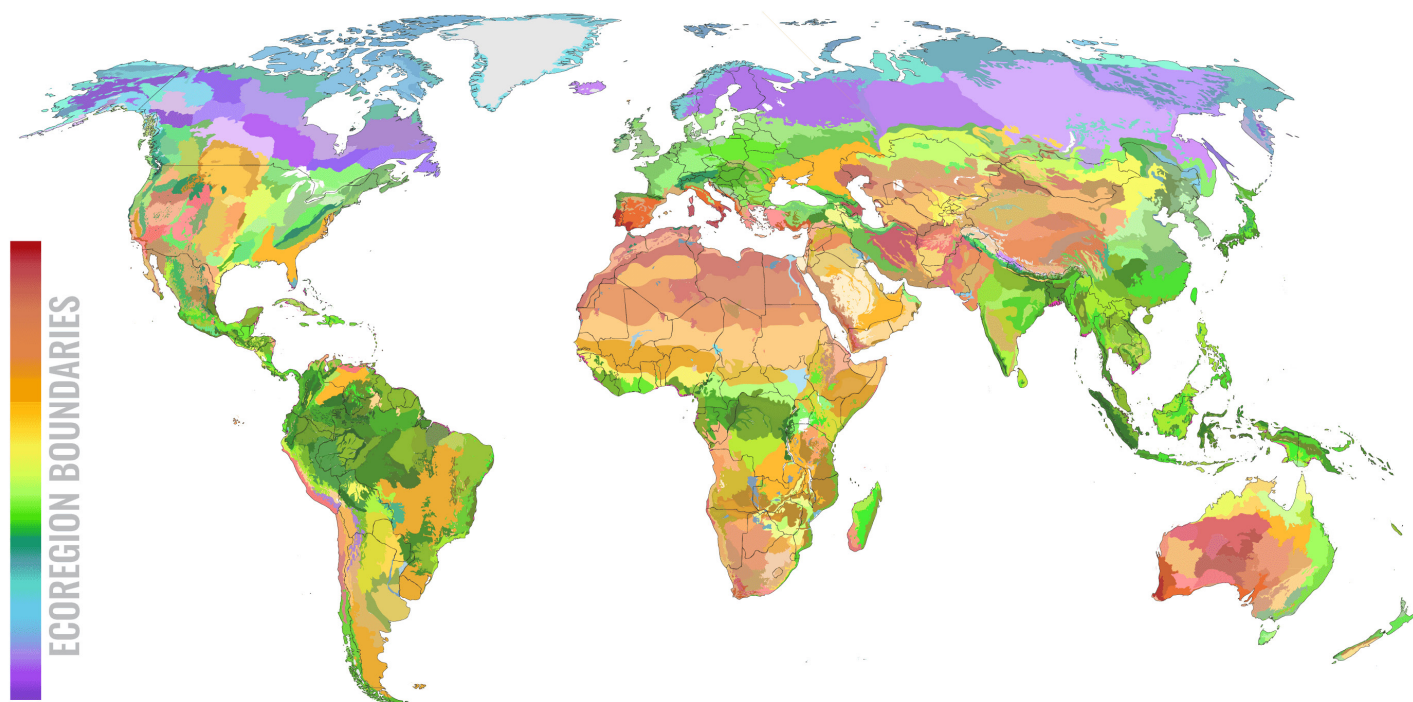
As a result of delivering an efficient tool for identifying and classifying any organism, we gain the ability to explore and track the patterns of communities and ecosystems through time and space. This is especially important for understanding hyperdiverse groups and



The BARCODE 500K research program established the sequencing facilities, analytical protocols, informatics platforms, and international collaborations needed to build the DNA barcode reference library. Building on this success, BIOSCAN launched in June 2019 to scan life and codify species interactions while expanding the reference library and demonstrating its utility. BIOSCAN will be the foundation for the Planetary Biodiversity Mission, a mission to save our living planet.

megadiverse regions. Detailed community analysis is unachievable, or at least unscalable when it depends on sorting and identifying thousands of cryptic organisms, which is the situation for most insects, fungi or marine organisms. As sequencing technologies and bioinformatics capabilities continue to advance, these same difficult groups can be routinely and regularly sampled and described. This offers whole new windows into the structure, ecology, and dynamics of each ecosystem, opening up unprecedented opportunities to understand and respond to biological systems. Perhaps most importantly of all, high-bandwidth DNA-based monitoring of biodiversity can support intelligent approaches to landscape-level conservation, agriculture and pest management, and response to climate change.

BIOSCAN comes at a time when technological advances are combining with the rich data held in BOLD to increase the cost-effectiveness of barcoding and



BIOSCAN will lay the foundation for an earth observation system. It will examine biological communities from at least half the world's ecoregions to begin the task of compiling comprehensive biodiversity baselines.

metabarcoding. The iBOL community internationally, and particularly the Centre for Biodiversity Genomics (CBG) at Guelph, are at the forefront in exploiting next-generation sequencing. iBOL's approach is to use the power and scale of these platforms to focus on a narrow subset of each species' genome as the tool that cheaply permits the broadest possible detection and identification of any species.

Going even further, the sensitivity of these platforms is unlocking the often hidden relationships between species, allowing us to document these interactions and clarify their role in structuring biological communities. Every organism interacts with representatives of other species as hosts or food and itself supports or contains a universe of parasites and microbes. These relationships have complex effects on the role that each species plays in each ecosystem. In the past, these associated species have often been detected as a source of potential confusion while deriving reference barcodes from specimens. Increased sensitivity from sequencing platforms will allow BIOSCAN to start treating these intermingled sequences not as noise but as a tool to document the set of species associated with a specimen, the organism's symbiome.

iBOL's new program will use these advances to build on the foundations of BARCODE 500K and

deliver the reference data, tools, and processes that will allow the world to survey and monitor all life. BIOSCAN's three main research themes aim to (1) increase the coverage of the barcode reference library to at least two million species, (2) exploit the power of new sequencing platforms to survey species communities at thousands of sites across different ecoregions and (3) to probe the biotic associations of millions of individual organisms. The CBG team has invested not only in upgrading sequencing hardware to support the scale and complexity of BIOSCAN but also in the informatics capability required to support it, now available as the Multiplex Barcode Research and Visualization Environment (mBRAVE). iBOL will also use this program to address outstanding issues around marker genes and sequencing protocols for challenging taxonomic groups and to standardize approaches to sampling taxa in different environments and ecosystems.

The efficiency of barcoding as a tool for identifying species or for validating other identifications also positions BIOSCAN as an essential activity in support of other genomics activities. The Earth Biogenome Project (EBP) and a suite of taxon-specific genomics networks aim to sequence full genomes or significant portions of the genome for many or all the world's

species. A significant challenge for these major projects will be to locate high-quality genetic material to represent each of these species. By building the reference library of DNA barcodes, each accompanied by vouchered specimens and extracted DNA, BIOSCAN's collecting activities can also enable these projects to proceed rapidly and with high confidence. The deliverables of BIOSCAN are fully complementary to those of EBP and similar efforts. BIOSCAN will deliver the reliable look-up mechanisms that verify the identifications associated with more extensive sequencing and will also deliver the biogeographic information to understand the distribution and variation for each species, along with their interactions. Complete-genome efforts will complement this with extensive additional data from examples of each species, enabling us to explore how species function and how evolution has shaped them.

By deep sequencing tens of millions of DNA extracts from single specimens and metabarcoding more than 100 million specimens from 2,000 sites spanning half the world's ecoregions, BIOSCAN will expose countless undescribed species and reveal their distri-

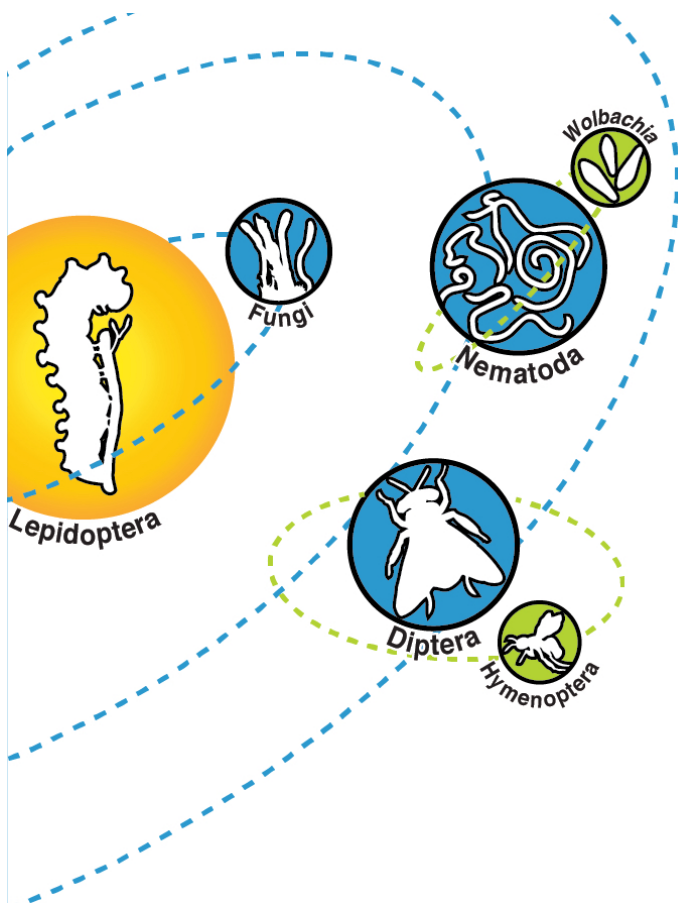
butions, dynamics and hidden interactions. Although BIOSCAN will not register all species or fully reveal their dynamics and interactions, it will be the foundation for a 20-year mission that will achieve these goals. Along the way, the aim is to develop the network to include practitioners and projects in all regions.

Participation is sought from researchers in all countries to expand iBOL's coalition and explore multi-cellular diversity throughout the world's ecosystems. iBOL welcomes comments and online discussion on the [draft Strategic Plan](#) for BIOSCAN.

We share our planet with more diversity than we yet recognise. This diversity drives the systems that keep the planet habitable for our species and those on which we depend. Now is the time to understand and monitor biodiversity everywhere. BIOSCAN is a key opportunity to make this happen.

Online:

<https://ibol.org/barcodebulletin/features/bioscan-illuminating-biodiversity-and-supporting-sustainability/>



BIOSCAN will use taxonomically targeted primer sets on the DNA extract from single specimens to reveal their commensals, mutualists, parasites and parasitoids – the symbiome.