

Tropical Plant Collections and ‘Big Data’

Review of presentations and discussion by the chairperson of this session, Stephen P. Hubbell, University of California, Los Angeles, USA.

The world is now engulfed by the perfect biodiversity storm: the concatenation of rapid global climate change, large-scale habitat loss, and accelerating extinction rates. Evidence from many sources indicates that we are in the midst of an anthropogenic mass extinction event, possibly the sixth great mass extinction in the history of life on earth. This perfect storm also comes at an inopportune time when a large fraction of the world’s biodiversity still remains undiscovered and undescribed, much less its distribution and abundance known and its ecological function understood.

Plant collections have a vital role to play in confronting the global biodiversity crisis, not only as repositories of types and reference specimens, but also as sources of new knowledge and syntheses of patterns of plant biodiversity, evolution, ecology, and biogeography across the globe. In the last two decades, enormous strides have been made in digitizing collections and making data available through advancements in electronic data capture, storage, retrieval, presentation, analysis, synthesis and dissemination. But the sheer volume of the data and the rapidity at which new information is becoming available online has created a series of major computational and analytic challenges now widely referred to in diverse fields as ‘big data’. In this section of the symposium volume, we presented papers from experts on

both the challenges and opportunities that ‘big data’ present in plant biodiversity studies that to varying degrees are collection-based. Many exciting ‘big data’ opportunities are now unfolding not only to expand our understanding of the evolutionary origins and ecological maintenance of plant biodiversity across the globe, but also to generate better, more data-informed strategies for conserving and managing this biodiversity.

Kenneth Feeley discussed diligently how collections-based research and fieldwork together can reveal historical rates of migration of plant species in response to climate change. Simon Queenborough gave us a fascinating look at how functional traits and their distribution across taxa can be studied from collections. This was followed by an outstanding synthetic paper by Alexandre Antonelli, integrating big data from phylogenetics, geography, modern studies of functional traits, and fossils to understand regional and global patterns of plant biodiversity and their evolutionary origins. Jorge Soberón finally took up the question of how big data can inform biogeographic inference and help in conservation planning, but not without thoughtful use of the data building on biological theories.

While listening to the presentations by the speakers in the ‘big data’ session of the symposium, several

questions came to mind that caused me to reflect on the prospects for using 'big data' to help with global conservation and ecosystem management efforts over the next several decades. I take the liberty of sharing these questions with the readers of these proceedings: (1) Who are the target audiences or 'consumers' of 'big data' on global plant biodiversity? (2) What are the most pressing specific questions about plant biodiversity that need answers for conservation or land management decision-making, and can these questions be answered by collections-based 'big data'? (3) In what ways can access to 'big data' increase the perceived value of the collections from which they are derived? (4) Conversely, what is the risk that the increasing availability of 'big data' will reduce the perceived value of the collections themselves? (5) Many sources of 'big data' used in the examples presented by Feeley, Queenborough and Antonelli are not collections-based, as, for example, remote-sensing data. As these additional sources are added, can we say that

this strengthen the long-term value of collections-derived data?

A concluding comment: As a tropical plant community ecologist, I am largely an outsider to the plant collections scientific community who attended this meeting. Nevertheless, I was honoured to be invited because my research program and that of my colleagues across the world who study tropical (and now also temperate) forest dynamics in a global network of large, permanent forest inventory plots (a consortium known as the Center for Tropical Forest Science or the Smithsonian Global Earth Observatory) would be impossible without the herbaria, the collections they house all over the world, and the botanical expertise they maintain. Indeed, collections are the essential foundation for all of plant ecology and plant conservation, and must not only survive, but must be strengthened, to successfully confront today's perfect biodiversity storm.