Title: Cyberinfrastructure for an integrated botanical information network to investigate the ecological impacts of global climate change on plant biodiversity

Short title: Botanical Information and Ecology Network (BIEN)

Project Leaders

Brian J. Enquist, Associate Professor, University of Arizona, Tucson, AZ, 85721, USA; benquist@email.arizona.edu; (505) 626-3336 (main contact). Research covers the influence of plant traits, phylogeny, and diversity on larger scale ecological, ecosystem, and evolutionary patterns. Teaching interests cover general plant functional biology, macroecology, diversity, and community ecology. Outreach includes serving data from ecological forest plots via the SALVIAS network.

Richard Condit, Chief Scientist, Center for Tropical Forest Science, Global Forest Observatory Network, Smithsonian Tropical Research Institute, Unit 9100 Box 0948, DPO AA 34002; conditr@gmail.com. Research interests include population biology and community models. Teaching interests are in quantitative ecology and forest modeling, with an emphasis internationally.

Robert K. Peet, Professor, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-3280, USA; uniola@email.unc.edu. Research focuses on patterns of plant diversity in North America, the vegetation of the Southeastern United States, and bioinformatics as applied to ecology and systematics. Outreach includes serving vegetation data on the VEGBANK network, developing international exchange standards for ecological and taxonomic data, and work on copyright policy in the digital age.

Mark Schildhauer, Director of Computing, National Center for Ecological Analysis and Synthesis (NCEAS), 735 State St., Suite 300, Santa Barbara, CA 93202-3351 USA; schild@nceas.ucsb.edu. Research on software tools for storing, integrating, and manipulating ecological data, through involvement with several large ecoinformatics efforts, including the Knowledge Network for Biocomplexity (KNB), and the Science Environment for Ecological Knowledge (SEEK). Outreach includes efforts to make published datasets widely available, using advanced knowledge representation techniques to facilitate data discovery and integration, and teaching ecologists the importance of informatics in their science.

Barbara M. Thiers, Director, William and Lynda Steere Herbarium at the New York Botanical Garden, 2900 Southern Blvd., Bronx, NY 10458-5126 USA; bthiers@nybg.org. Research in the application of information technology to herbarium management, and outreach efforts to increase access to specimen-based data for the scientific community.

Core Team Members

Sandy Andelman (Conservation International, Vice President TEAM initiative, vegetation plots), Brad Boyle (University of Arizona, biodiversity informatics), Jeannine Cavender-Barres (University of Minnesota, plant traits and phylogenies), Steve Dolins (Bradley University, computer science), Stephanie Hampton (Deputy Director NCEAS, outreach and training), Jesse Kennedy (Napier University, Scotland, Computer visualization, Taxonomic Concept Schema), Brian McGill (University of Arizona, ecological informatics, macroecology), Hans ter Steege (University of Utrecht, Netherlands, collections and plot networks), Jens Christian Svenning (Aarhus University, Denmark, biogeography), Nathan Swenson (Michigan State University, plant traits and macroecology), Oliver Phillips (Leeds University, UK, RAINFOR network, global change and vegetation plots), Peter Jørgensen (Missouri Botanical Garden, collections and C3Tropicos), Dave Vieglais (University of Kansas Biodiversity Research Center, Darwin Core/DiGIR), Corine Vriesendorp (Field Museum, collections and outreach); Susan Wiser (Landcare Institute, New Zealand, plant ecology, IAVS vegetation plot exchange schema).

Project Summary

To answer many of the major questions in comparative botany, ecology, and global change biology it is necessary to extrapolate across enormous geographic, temporal and taxonomic scales. Yet much ecological knowledge is still based on observations conducted within a local area or even a few hundred square meters. Understanding ecological patterns and how plants respond to global warming and human alteration of landscapes and ecosystems necessitates a holistic approach. Such an approach must be conducted at a scale that is commensurate with the breadth of the questions being asked. Further, it requires identification, retrieval, and integration of diverse data from a global confederation of collaborating scientists across a broad range of disciplines. We propose to network core databases and data networks to create a novel resource for quantitative plant biodiversity science. The grand challenge is to assemble and share the world's rapidly accumulating botanical information from plots and collections to create a distributed, web-accessible, readily analyzable data resource. With such a resource, we will answer major questions of direct relevance to plant ecology, plant evolution, plant geography, conservation, global change biology, and protection of biodiversity and ecosystem services. In particular, *how does climate influence the distribution and abundance of plant species, how does the phylogenetic diversity of plants vary across broad environmental and climatic gradients, and how are plants assembled into ecological communities?* While these and associated questions are at the core of many research endeavors in comparative botany and ecology, our past collective inability to integrate data on a large scale has significantly limited our ability to address these questions head on. This proposed Grand Challenge team will create a data resource of unprecedented size and scope together with the tools for its use, thereby empowering botanists and the general public to better address fundamental issues in plant ecology and global change biology. Although we will focus on plants of the New World, the infrastructure and protocols developed will be scalable to all geographic regions and all types of organisms. Future steps will enable cross-cutting linkages to emerging networks on plant genomics, physiology, and phylogeny, allowing us to address fundamental genetic and evolutionary questions at unprecedented spatial and temporal scales.

I. – The grand challenge: *In a changing world, what grows where and why?*

I.a - Introduction - Understanding what controls the abundance and distribution of botanical diversity is fundamental to much research that underlies ecology, evolution, comparative plant biology, and global change biology [1-4]. For example, the geographic distributions of plant species reflect physiological tolerances [5, 6], evolutionary and climatic history, and offer insights into the traits that underlie adaptation [7-12] and the mechanisms involved in population divergence [13]. Abundance is a measure of ecological dominance and ecosystem services and often reflects fitness [3, 8, 14]. Indeed, together, information on abundance and distribution provide the ability to bridge the plant sciences by linking many central questions in plant biology [13, 14] to the great botanical diversity in nature. In addition, climate-induced shifts in the distribution and abundance of plant taxa can impact the diversity and function of local communities [15] and thereby alter ecosystem attributes [16]. In a changing world taxon abundances and geographic ranges will likely rapidly expand or contract [17-19] and some species will become extinct [20, 21] but we as a scientific community are not yet prepared to anticipate those changes [18].

Our ability to predict species' abundances and ranges, let alone how they will change, remains limited [18, 22]. In order for biologists to predict how individual taxa and entire communities will respond to a changing world requires understanding why plant taxa grow where they do and what limits their ranges. However, distributional and community shifts are broad, spanning large geographic gradients and sometimes continents. Further, range size abundance, and phylogenetic/taxonomic information have rarely been addressed in many parts of the globe, especially in the tropics due to the lack of integration of the many plot samples where abundances have been calculated. A full understanding of present and future patterns of biodiversity necessitates examination of processes and taxa across geographic and environmental gradients.

I.b - The Barriers - The lack of a global source of integrated and standardized biodiversity observation records is a fundamental impediment to advancing the plant sciences. As a result, the development of a global perspective on variation in basic floristic and ecological attributes has been limited. These problems are especially acute in the tropics where biodiversity is concentrated but poorly known [23-25].

Most datasets originate from individual researchers and span a few square kilometers [26, 27], recording varied kinds of data using idiosyncratic protocols and published (if at all) in various formats [28]. Further, even if we could integrate these original sources, we would not be able to place much confidence in the resulting list of taxa because: (*i*) there is no standardized global list to assess the validity of the names or circumscriptions of plant taxa[1]; (*ii*) there is no global standard for naming variants contained within taxa[2]; (*iii*) there are numerous technical and data quality issues with merging and serving data from disparate sources; and (*iv*) *there is no standardized process by which data on distribution and abundance of plant taxa can be combined with information from plant physiology, genomics, and phylogeny*. The last point is especially important because understanding why species are limited in where they grow requires genetic and physiological knowledge of traits that affect how an individual responds to the environment [29]. We can describe observed vegetation shifts across continents as temperatures rise, for example, but will not be able to predict future shifts until we know how individuals and taxa change across environmental gradients and how they react to changes in temperature and precipitation.

I. c - The Solution - iPlant offers a unique opportunity to create a confederated plant data network to allow scientists to address "what grows where and why". This network will also provide feedback and training to data providers as well as novel educational opportunities. The solution will require developing a transformative cyberinfrastructure, based on proven informatics approaches coupled with cutting edge software tools to: (*i*) make easily accessible the many disparate and individually limited datasets; *(ii)* standardize these data streams, *(iii)* integrate data streams adhering to divergent taxonomic standards; *(iv)* create a constantly updated but perfectly archived data resource that biologists and the public can query seamlessly; *(v)* provide feedback to the main data providers so that they can then better serve and standardize their data; and (*vi*) integrate the plant data streams with environmental data from a range of sources (e.g., climate, land cover change, etc.). By collecting and combining vegetation censuses, botanical surveys, and specimen records from herbaria, we will achieve comprehensive, cross-taxa

1 http://www.wakehurst.org/science/directory/projects/Target1GSPCGlobCheck.html

2 http://www.tdwg.org/about-tdwg/

coverage and provide a global perspective on variation in basic floristic and ecological attributes [25, 30-32]. Ultimately, we aim to create a "Data Discovery Environment" that will serve and process basic information for academics, conservationists, and the public. We call our integrated data resource BIEN, or the Botanical Information and Ecology Network[2].

I.d - The Grand Challenge Team - We propose to address the Grand Challenge by creating an integrated global network of botany data providers and users. What the BIEN team seeks for plant science is not simply a new data network and cyberinfrastructure, but a new paradigm with respect to how data are recorded and integrated. The need for this paradigm switch is well documented [33-35], but making it happen has proven difficult. With support from the National Center for Ecological Analysis and Synthesis (NCEAS[3]), the BIEN grand challenge team held a planning meeting in December 2008. The BIEN team and additional collaborators (Table 1) broadly represents the community of plant biodiversity sciences and includes members of many major botanical institutions, data networks, and biodiversity initiatives, as well as informatics experts. Together, we identified key data sources from around the globe, the kinds of data to be extracted from each, and identified the requirements for merging disparate data into an integrated framework. In the process of initiating this botanical network we identified short- and long-term cyberinfrastructure needs, including existing technologies and protocols, programming challenges, end-user tools, and web services. Drawing from the conclusions of our initial meeting, we here detail the barriers and specific solutions needed to integrate and provide access to botanical biodiversity data.

II. – Proposed scientific activities

Design and creation of the BIEN network will be guided by the goal of addressing the Grand Challenge question – *in a changing world, what grows where and why?* The Grand Challenge Team has identified three fundamental and tangible sub-questions that will enable us to directly address aspects of the grand challenge and demonstrate the power and utility of the cyberinfrastructure we propose to design and build. Each question addresses key problems in plant ecology, comparative plant biology, and biodiversity conservation.

Q 1: *How does climate influence the relative distribution of narrow and widespread species? Do these relationships vary in tropical and temperate environments?*

Q 2: *How are abundance and size of geographic range of taxa related? For example, do plants with small ranges tend to be rare relative to widespread species?*

Q 3: *What are the physiological, demographic, environmental, and phylogenetic correlates of rarity (small ranges, low local population size) and commonness (large range, high local population size) across environmental gradients at scales ranging from local to continental? Can these correlates be used to predict vulnerability or resistance to extinction for species and communities under differing scenarios of habitat loss and climate change?*

The first two questions are fundamental to ecology [1]. The third question has strong practical implications for land management decisions related to conservation hotspots and preservation efforts [23]. All three questions lead to predictions about the vulnerability of species and communities to extinction, and all will allow us to compare temperate and tropical floras precisely [8, 36, 37]. Range size abundance, and phylogenetic/taxonomic information have rarely been addressed in many parts of the globe, especially in the tropics, and we will be able to address the questions at a global scale. Most importantly, the project will result in an integrated data resource on the distribution of plant species that will be a baseline against which to gauge responses to global warming and global change. As we discuss below, questions relating range size and abundance to physiological or genetic traits rely on other initiatives that are currently underway, some under the auspices of iPlant.

[2] http://www.nceas.ucsb.edu/projects/12290
[3] http://www.nceas.ucsb.edu/

III. – Data and computational activities to address the grand challenge

III. a – The goal - Addressing the Grand Challenge will require a comprehensive, integrated and standardized data network of biodiversity observation records from across the globe. Data sources must range from the tiny datasets collected by individual scientists to data streams from large and long-lasting programs, established groups, institutions, and herbaria. We propose to develop a data integration network where plant biologists from many different disciplines, with many different research goals, upload, standardize, merge, and share data. This network will also serve as a permanent repository for legacy data. The end product will be the ability to address questions at spatial and temporal scales far exceeding the reach of any individual research program.

III. b –Biodiversity observation data - There is an enormous amount of existing data on plant distribution and abundance as well as several emerging sources of additional botanical information. Many are tiny datasets collected by a single scientist, whereas others represent much larger and long-lasting efforts. We identify two main sources of data crucial to questions about geographic range and abundance:

(1) Collection records. We estimate based on *Index Herbariorum*[4] that the world's museums hold ~300 million plant specimens, of which perhaps 15 million have been digitized and are potentially accessible for our project. Others are being steadily digitized. GBIF[5], the Global Biodiversity Information Facility, has begun the task of assembling digitized museum collection records, offering already 65 million individual species occurrence records from the Americas (including both plants and animals). There is the potential for a total of 95 million records from US herbaria alone [38]. GBIF records are available for our demonstration project. In addition, we are working directly with several US data sources, including the Missouri[6] and New York[7] Botanical Gardens, and have preliminary agreements to access other major collections in the US (Table 3).

(2) Vegetation plot records. Plant ecologists routinely delimit precise areas and assess plant species abundance, often by recording either individual trees by size or all plant taxa by percent cover. These 'plots' allow precise estimates of abundance of each species. Plot data are linked by accurate geocoordinates to specific site conditions. The grand challenge team identified vegetation plots already digitized and available to the BIEN data confederation: 1350 tropical forest plots in Central and South America and 325,000 North American vegetation plots, both forest and non-forest (Table 2). Most of these plots hold 10-100 plant species, so the total number of species occurrences is on the order of 15 million. At the first BIEN meeting we also examined the total number of digitized plots that could be integrated in a future network. We identified several hundred thousand additional North American plots potentially available, and there are over a million European vegetation plots that have been digitized in TurboVeg format alone [39] (Hennekens *pers. comm.*). Large digital plot archives such as those for New Zealand and South Africa are also potentially available. We estimate that within five years we can have networked data from in excess of 400,000 North American plots, 2000 Central and South American plots, and potentially another million plots from outside the Americas.

Based upon our first order approximations, *there is the potential for at least ~500 million taxonomic occurrence records*, where a single record is an observation of a plant characterized by a latitude and longitude coupled with a taxonomic determination of that plant. As we discuss below, accessing and integrating these two fundamental data units across multiple botanic data sources entails significant challenges in informatics. But successful integration will provide a powerful new cyberinfrastructure [38] to answer fundamental questions in ecology, evolution and global change research.

III. c Secondary Data Sources: Traits, Phylogeny, and everything -omics – Confederation of the data sources listed above will, for the first time, provide the botanical community with a baseline for the study of abundance and distribution on a global scale. It is these data that will be the core of the BIEN initiatives. However, in order to address the processes and mechanisms that influence and ultimately determine many aspects of abundance and distribution, the BIEN initiative must also merge this

[4] http://sciweb.nybg.org/science2/IndexHerbariorum.asp
[5] http://www.gbif.org/
[6] http://www.mobot.org/
[7] http://www.nybg.org/

framework with several additional sources of data being organized by groups outside of BIEN. The BIEN data network will become even more valuable when linked to these other plant informatics efforts. Indeed, there are exciting potentials for synergistic activities with emerging groups. These groups include: (*i*) functional traits; (*ii*) genomic data; and (*iii*) phylogenies. These attributes are, in turn, best linked through botanical nomenclature (a focus of BIEN) or indirectly through phylogenetic resources such as the GC "Tree of Life" team.

Functional traits are phenotypic attributes of the organism and are defined as quantifiable morpho- physio- and phenological attributes that impact fitness indirectly via their effects on growth, reproduction and survival, the three components of individual performance [40, 41]. There are several efforts underway to actively compile and network global information on several key [42, 43] plant traits. The key trait networking efforts include GLOPNET[11], the NSF RCN funded TraitNet[12], the National Phenology Network[13], and the global TRY[14] network. Variation in functional traits often influence the performance in plants in differing environments [44]. Thus, the ability to merge plant distribution and abundance data with information on plant functional traits will allow for the mechanistic linkage between abundance and distribution with variation in phenotypes.

Ultimately, in a changing world, both genes and environments are important in determining how plants respond to the environment and ultimately where they grow. Through iPlant the BIEN team has the potential to integrate with other GC teams such as the GC team "Cyberinfrastructural Support for Genetic and Ecophysiological Studies of Plant Phenological Control in Complex and Changing Environments") utilizing new molecular- to field-level models, databases, and techniques relating to the phenology of crops and plants in natural ecosystems. These new techniques, such as gene-based, ecophysiological approaches [12], can be used to chart how changes the abiotic environment can influence the life cycle across the geographic range of a plant species, and presumably across different species whose ranges occupy differing environments.

As we describe below, the cyberinfrastructure proposed for BIEN is key to integrating all these other GC efforts in that it provides the linkage through which all these data networks can be integrated by providing the critical semantic mediation of the many-to-many relationships between taxon names and concepts. This linkage will be central and absolutely necessary in order to merge genomic, trait, phylogenetic, distribution and abundance data.

III. d – 'Priming the Pump': Short-term needs of a global BIEN - In order to quickly address our core science questions, cyberinfrastructure needs, and obstacles to data integration, the BIEN core team is currently compiling and analyzing several data sources. Using the sources in Tables 2 and 3 we are generating the foundations of an integrated plant diversity database. During this process we have identified two critical problems that continue to limit attempts at broad-scale integration of biodiversity observation data. It is these problems that will first require the support of the iPlant collective.

- *The lack of taxonomic standardization is the most important informatics impediment in the plant sciences*. The plant sciences do not yet have the cyberinfrastructure needed for taxonomic standardization—the matching of taxon names and concepts in different data sources, and as a consequence plant scientists and the cyberinfrastructure they employ, are not prepared to provide high resolution identification of the taxa reported in literature and various botanical databases (such as occurrences, gene sequences, and traits). Tools are needed that will allow investigators to efficiently navigate a data landscape where one taxon might have many names (synonyms) and one name might refer to many taxa (taxon concepts). In addition, the system must encourage a continual update of taxonomic relationships by the taxonomic community. Members of the BIEN team have previously designed solutions to this problem using set theory mapping of the relationships among taxonomic concepts (a name as used by a specific authority) [45, 46]. Examples include design of the Taxon Concept Schema[15] recently adopted by TDWG as an international standard, and implementation of the

[11] http://forestecology.cfans.umn.edu/glopnet.html
[12] http://www.columbia.edu/cu/traitnet/
[13] http://www.usanpn.org/
[14] http://www.try-db.org/
[15] http://www.tdwg.org/standards/117/

core components in the VegBank[1] plot archive and in the SE Floristic Atlas[2]. What is missing is a cyberinfrastructure and component data that employs the critical taxon concept approach for taxon documentation and for integration of data from mixed sources that follow different taxonomic perspectives.

We propose three short-term demonstration projects aimed at solving the taxonomic barrier. First, we will create an approximation of an authoritative list of New World plants by combining a series of regional checklists, such as the USDA Plants list for North Americana north of the Mexican boundary, the Caribbean list of Acevedo (new version due in May 2009), the Missouri Botanical Garden's catalog of plants of Ecuador, Nicaragua, and Peru, Funk's catalog of plants of northeastern South America, and the Smithsonian Tropical Research Institute's Neotropical Tree Species Checklist[3]. We will attempt to place all the New World taxa reported in Tropicos either in the list or in the synonymy. Second, we will develop and deploy a taxonomic "scrubbing" tool for high-throughput detection and correction of taxonomic spelling and nomenclatural errors. This application will build on existing taxon name matching algorithms such as TaxaMatch[4] and TaxonScrubber[5]) and will use TROPICOS[6] and the IPNI Plant Names Database[7] as authoritative name references. Third, we will conduct a comprehensive demonstration of the power of taxon concept mapping using the approximately 80,000 taxon concepts relationships documented for the flora of the Southeastern US in the the SE Floristic Atlas[13], expanded to allow selection of alternative taxonomic perspectives. The SE Floristic Atlas (SEFA) is the only large-scale online use of taxon concept relationships to integrate diverse occurrence data from many sources including museum collections, literature references and plot data. This regional-scale project is a working example of the sort of tools we propose to put at the service of the entire ecology and biodiversity community.

- A second major infrastructural problem arises from interaction and networking among data sources, which leads to serious challenges with regards to data quality and data provenance. For example, experts may detect and correct errors in the raw data for their own use, but this secondary improvement often does not flow back to the original data source. As a result efforts are wasted, original data sources remain uncorrected, and it is often challenging to determine what constitutes the best, or even, unique, set of information. A cyberinfrastructure is needed that allows seamless feedback between data providers and data users in a process of data annotation and correction. This feedback would include revision of data, real-time addition of new data, perfect archiving so that data available at a given date can be easily viewed for reanalysis, and feedback to the data sources with respect to suggested revisions.

We will process approximately 1,350 tropical forest plots from Central and South America, 400,000 North American vegetation plots, and perhaps 400,0000 museum collections to validate names and geocoordinates. We will implement a workflow where all suggested changes are piped back to the source for validation and we will subsequently track the level of response achieved to guide us in design of more efficient and user-friendly tools.

In order to demonstrate the feasibility of an enormous, cross-continent, taxon-occurrence data network, the BIEN team will immediately use the data networks created through these projects to begin addressing our guiding scientific questions. A table will be created for all plot data that contains geocoordinates, survey date, and abundance for each species; a parallel table giving geocoordinates, date, and species name will be assembled from the specimen data. The end result will be, for the first time, the creation of data resources containing standardized and error-checked geographic occurrence and abundance records of several tens of thousands of plant species in the Americas.

III.e –Long-term feasibility of a global BIEN: Overview – With iPlant, we aspire to develop a global data integration network where plant biologists from many different disciplines can upload, standardize, merge, and share data. This network will also serve as a permanent repository for legacy data and provide a

[1] http://www.vegbank.org/
[2] http://www.herbarium.unc.edu/seflora/firstviewer.htm
[3] http://ctfs.si.edu/neotropicaltree/
[4] http://www.cmar.csiro.au/datacentre/irmng/
[5] http://www.salvias.net/pages/taxonscrubber.html
[6] http://www.tropicos.org/
[7] http://www.ipni.org/ipni/plantnamesearchpage.do
[13] http://www.herbarium.unc.edu/seflora/firstviewer.htm

benchmark against which to quantify the impact of climate change. The end product of such a network will be the ability to address questions at spatial and temporal scales far exceeding the reach of any individual research program.

Data integration at this scale will be an immense challenge, one that we believe will require an innovative hybrid solution with features of both a data warehouse and a data network. Proximately, as a data warehouse, the solution must allow for import and standardization of diverse data according to a common schema and vocabulary. Ultimately, as a distributed data network, the solution must empower the community to contribute, manage and update their own data. Engaging the community will be essential for long-term sustainability in the face of frequent updates.

Figure 1 provides an overview of the central activities of the BIEN network. Producing an integrated data network will require several main steps (labeled in the figure) including:

1) ***Coordination of core botanical data streams**: Provide tools for the creation, coordination and uploading (push) or harvesting (pull) of disparate data sources into the confederated iPlant Cyberinfrastructure, using common data communication protocols and exchange schema. The two main data streams are ecological plots/surveys and specimen records. In addition, we intend to also interact with outside groups representing the trait, phlogenetic and genomic communities.
2) ***Data integration and quality control:*** Here we will enable integration of disparate data according to a consistent model, while coupling concept-based taxonomic authority information with the raw data. This system identifies and iteratively refines taxonomic ambiguities and errors, while providing comprehensive feedback to original data provider.
3) ***Global scale, extensible, confederated database:*** Here we will create a web-based framework with data communication protocols and exchange schema that are compatible with a) broader ecological and environmental networks (e.g. NSF OCI/DataNet efforts and NSF OCI/Interop efforts in ecological and earth sciences, NEON), and b) other data frameworks (phylogenetic, traits, genomics)
4) ***Web-accessible end-user resource:*** Next, we propose to create a flexible and logical interface for powerful data querying, discovery and analysis, with download formats readily usable by all field of plant biology research
5) ***A Data-Discovery Environment (DDE)***: As discussed below, the main user interface of the coordinated BIEN network will be the "Data Discovery Environment" or DDE. The DDE is the end result of data integration, standardization, and confederation, accessed through a flexible querying interface that allows for deep exploration and analysis of the confederated database.
6) ***Iterative feedback to the community and data providers***: The BIEN project will not be static but instead will be a dynamic network. The process of #1-5 will be the creation of cyberinfrastructure (tools for data standardization, scrubbing, and exploration) that will allow for iterative feedback to the original data providers who can then modify their original data sources. Thus, over time, botanical diversity data are increasingly corrected and improved.

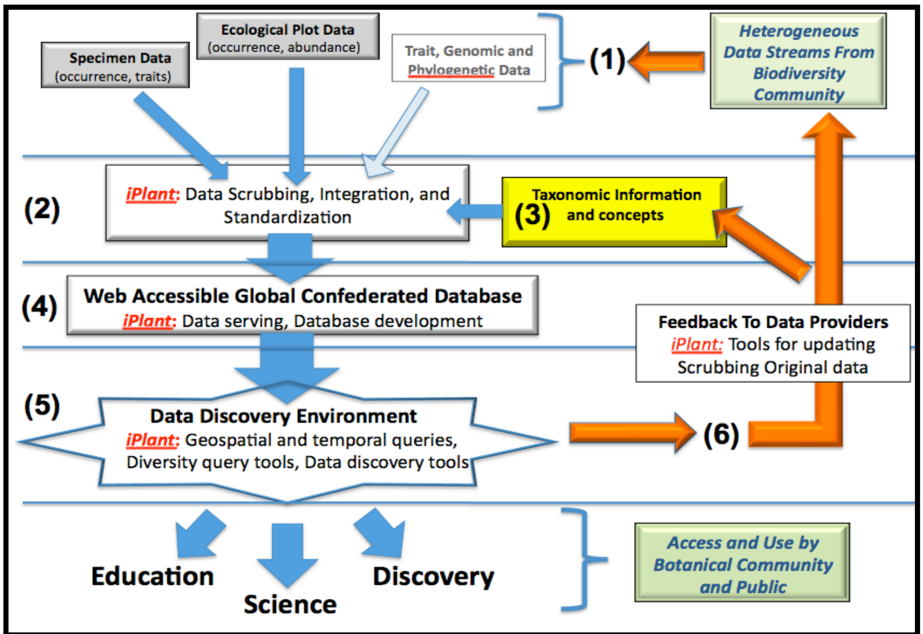

Figure 1. Overview of the proposed logical flow and important cyberinformatics steps in the creation of a Botanical Information and Ecology Network (BIEN). As described in the text, this global network is an iterative dynamic network that provides feedback to data providers as well as, over time, and increasingly improved source of standardized biodiversity data.

III. e - Long-term feasibility of a global BIEN: Data Discovery Environment - The main user interface of the coordinated BIEN network will be the "Data Discovery Environment" or DDE. The DDE will be the access point to the global confederated database for both academics and the general public. The DDE will allow for: (*i*) multiple opportunities for outreach and education (see section V. below); (*ii*) allow the user to view the data as originally collected or under alternative taxonomies and phylogenies; (*iii*) to visualize the distribution and density of data points; (*iv*) to create species-lists with linked species attributes; and (*v*) to pool different data sources (biological observations, traits, physiology, climate) at differing temporal and spatial scales. Appendix B provides a more detailed outline of the steps involved in the data integration and discovery process. Below we detail the key steps in the creation of BIEN.

IV. Proposed long-term cyberinfrastructure: Tools and Web Services

Addressing the Grand Challenge question will require not only the compilation of data but also the *maintenance of a comprehensive, integrated and standardized global data network*. Creation and maintenance of this resource will require a cyberinfrastructure composed of numerous tools and services. We define cyberinfrastructure as highly extensible, broadly compatible, highly useful information resources that are compatible with but extend the existing technology solutions on which the community currently relies. Figure 1 provides an overview of five central sets of activities the BIEN team will need to support by identifying, modifying or creating tools and services.

Modern botanical science is being dramatically altered by access to expanding quantities of data. Our grand challenge represents but one of many such topics on organizing and serving of such large data quantities. Of special interest here is that the core components of each of the six central sets of activities identified in Fig. 1 will be of critical value to scientists addressing other questions and will be of much broader value beyond our proposed project. Importantly, the cyberinfrastructure we are proposing can be generalized to observations of all types or organisms at all spatial scales.

Proposed cyberinfrastructure to create tools to empower and motivate sharing of biodiversity data

- **Taxon concept mapping tools**
- **Taxon concept resolution services**
- **Tools for taxonomists mapping concepts**
- **Tools for mass import**
- **Tools for aggregators mapping concepts**
- **Tools for the community to contribute**
- **Tools for data integration**
- **Tools for taxon mapping and prediction**

IV.a. The creation, coordination, and digestion of multiple core botanical data streams, each conforming to a specified exchange schema - To assure efficient and accurate access to biodiversity data, those data must be provided via community-sanctioned protocols that are supported by a suite of tools for efficient data export, discover, revision, and import. The protocols and tools are needed in part to empower and motivate the community to share biodiversity, and in part to provide assured direct and efficient access to large quantities of quality-controlled, standardized data. Moreover, providers of data, while often willing to share data, generally do not have the resources to develop idiosyncratic exports for individual users, but need to employ a single data export mechanism. The plant biodiversity and ecological communities have made major progress in establishing necessary protocols, and several implementations of these have significantly improved access to specimen data in particular over the past decade (e.g. Darwin Core). Yet much work remains to be done as not all types of observation data are yet supported, current implementations are often spotty in their data holdings, and errors and redundancies in the data need to be resolved.

- Creation of standardized biodiversity data-exchange schemas - Accurate and efficient data migration and ingestion requires broad international acceptance and compliance with established data exchange standards. Only with acceptance and widespread application of data exchange standards will we be able to absorb data from a broad array of sources. We propose working with existing confederation schemas, or where necessary, developing novel schemas, that preserve the richness of information contributed by heterogeneous data providers, and implementing these in production-ready systems available to researchers around the world. This effort will resolve the complex challenge of integrating vegetation, collection and observation data collected over vast spatiotemporal scales, using numerous collection methodologies, and currently stored in disparate data systems.

- Tools to export, search and insert schema data - To efficiently employ biodiversity data streams conforming to a common schema, tools will be needed to generate those data streams, view and edit them, and absorb them into the greater cyber infrastructure In some cases standard schema are already supported, such as in the transmission of collection records conforming to DarwinCore. However, to achieve the critical buy-in from the community we will need to provide interface tools for the other data systems, such as the Specify[1]and EMu[2] systems for collection management and the TEAM[3] (Conservation International), VegBank, and TurboVeg systems for vegetation plot data. In addition protocols and scripts should be provided for incorporation in new data systems to ease communication.

A single, widely used data exchange standard—Darwin Core[4]— exists for biological collections data, although there are several common variants that are not completely compatible. Nevertheless, the widespread use of this standard via the DiGIR[5] and TAPIR[6] data exchange protocols will simplify extraction of data from specimen databases. Darwin Core, however, only provides a description of a subset of the ecologically relevant information that might be available for a specimen record, and falls far short of the content needed for a broader range of taxon occurrence records such as those from vegetation plots where, for example, ecologically important measurements of association and relative abundance can vastly enhance our ability to understand the mechanisms influencing distribution and co-existence of plant species. Overall, a much more complex exchange schema is needed to bring together the numerous formats in which vegetation data are captured, as well as to harmonize vegetation, taxon occurrence and specimen data. Fortunately, development of a confederated vegetation data exchange schema, provisionally titled VegX, is nearing completion and will soon be available as an official TWDG[7] and IAVS standard[8]. This schema will be deployed as the common template for the import of specimen and vegetation data into the core BIEN database (see *Data import*, below).

Exchange standards are required for several types of data. Fortunately, BIEN team members play leading roles in the on-going development of many of these standards including those for collection data, vegetation plot data, general taxon observation data, and taxon concept data, BIEN members closely involved in standards development would work closely with the iPlant technologists to assure that emerging international standards for these data types are appropriately incorporated into the iPlant products.

IV. b. Quality control and standardization of the content of the data streams - Access to large quantities of biodiversity observation data does not assure the data consistency, accuracy and integration needed to address the grand challenge question. Quality control of the data stream must be assured through standard services and workflows. We here identify several key components for which tools and workflows will be required.

- Tools for taxonomic scrubbing - Names of organisms are infamously prone to spelling errors and orthographic variants. The only solution is to match names against standardized lists. The two primary lists available are W3Tropicos[9] and IPNI[10], both of which have administrators who have expressed interest in cooperating with us, and work on a centralized nomenclature is underway at GBIF/EoL. Applications are needed to perform exact and fuzzy matching of names, correct spelling errors, map synonymies and quantify taxonomic uncertainty. Existing applications such as TaxaMatch[11] and our own Taxon Scrubber[12] point the way to more sophisticated solutions (see 15, 18]). In addition. Consistency will be greatly enhanced by adoption of a set of universal identifiers (GUIDS) for taxon names as advocated by TDWG[13] and already implemented by IPNI[14] in the form

1 http://www.specifysoftware.org/Specify/specify/
2 http://www.kesoftware.com/content/view/512/356/lang,en/
3 http://www.teamnetwork.org/en/
4 http://wiki.tdwg.org/twiki/bin/view/DarwinCore/WebHome
5 http://digir.sourceforge.net/
6 http://wiki.tdwg.org/TAPIR
7 http://www.tdwg.org/
8 http://www.bio.unc.edu/Faculty/Peet/vegdata/standards.htm
9 http://www.tropicos.org/
10 http://www.ipni.org/
11 http://www.cmar.csiro.au/datacentre/irmng/
12 http://www.salvias.net/pages/taxonscrubber.html
13 http://www.tdwg.org/
14 http://www.ipni.org/

of LSIDs[16]. We will consistently apply taxon name GUIDS as available and work to unify such systems across name providers.

- Georeferencing Tools - Although latitude and longitude for the collection site are today standard standard components of a plant observation and collection records, most older specimens lack these data. At The New York Botanical Garden, only an estimated 30% of specimens include geocoordinates. In order for herbarium specimens to be used in vegetation analyses, they must be georeferenced. Adding coordinates to specimens retroactively is time-consuming, sometimes requiring as much time as digitization of all other specimen data. A cyberinfrastructure tool is needed to help supply these missing data, and to check existing geocoordinates against locality descriptions. Although several georeferencing applications are already available (e.g., Biogeomancer[17], SpeciesLink[18]), none are capable of the high-throughput georeferencing required by BIEN. One of our primary goals is to engage the community and build upon existing expertise wherever possible with the goal of, improving and assisting in the deployment of these applications as web services.

- Tool for the Detection of duplicates. Standard plant specimen techniques include collection in multiple sets and distributed to multiple herbaria. Though this practice is beneficial for users of individual herbaria, it can be an insidious source of error and pseudo-replication in analyses based on georeferenced specimen data. A cyberinfrastructure tool is needed to identify duplicates and remove all but one instance from a given analysis. Some work toward this end is already under way in the form of the Filtered-Push network[19] (see also http://mantis.cs.umb.edu/wiki/index.php/Main_Page)

- Tools to Compile Names of Collectors, Determiners, and Taxonomists. Authoritative lists of persons involved in collection and determination (plus linked GUIDS) are necessary to maximize data quality and consistency within the biodiversity community. This information will serve many purposes, the most important of which is inferring identification quality—an issue distinct from taxonomy. We will build upon existing resources within the community by accessing existing authoritative data sources such as the Harvard Names Database[20], TROPICOS[21], and the New York Botanical Garden's internal database of 150,000 plant collectors and taxonomists

IV. c. Taxonomic integration - Taxonomy is the common language for describing biodiversity, but as with any language, taxonomic names change over time as discoveries lead to new interpretations of how biological entities should be classified. Millions of observations and collections of plant specimens exist that are "identified" through these taxonomic names, but since the meanings of these names can change, with the consequence that the interpretation and definitive understanding of "what occurred where" is compromised, yet there remains no clear-cut mechanism for updating determinations through time. Clarifying the meanings and relationships of names applied to organisms by different researchers at different times is a fundamental challenge when integrating biological data.

Ambiguity can arise due to spelling errors, variant spellings, nomenclatural synonymy, but also due to taxonomic revisions where splitting and lumping change the circumscription of specimens associated with names [19], *making taxonomic standardization not only a major challenge for BIEN, but also a major impediment to merging data within any area of the biological sciences where the taxon is the linking variable*. Methods and services for taxonomic standardization must be provided that successful navigate a data landscape where one taxon might have many names and one name might refer to many taxa. Moreover, solutions must support a world where different institutions have different preferred taxonomies, and the accepted solutions are constantly changing as new information becomes available, yet must be perfectly achieved to capture the content at any time in the past.

16 http://en.wikipedia.org/wiki/LSID
17 http://www.biogeomancer.org/
18 http://splink.cria.org.br/tools?criaLANG=en
19 http://www.tdwg.org/proceedings/rt/metadata/351/0
20 http://asaweb.huh.harvard.edu:8080/databases/botanist_index.html
21 http://www.tropicos.org/PersonSearch.aspx

Resolution of the many-to-many relationship between names and taxon concepts is to be found in application of taxon concept relationships as first described by Berendsohn 1997 [47] and subsequently articulated by BIEN personnel in the recently adopted TDWG taxon concept schema and in related publications (e.g. [34, 35]), Variants of this schema have already been embedded in the VegBank plot data archive and in the SE Flora NatureServe Biotics and the Euro+Med flora biodiversity systems. Planning documents for a future release of USDA PLANTS also include taxon concepts. However, the major impediment to adoption of the approach is the need to create, update and resolve the relationships among taxon concepts for which specific tools and services are needed.

- Taxon concept mapping tools. Relationships among taxon concepts are routinely asserted by taxonomists when studying specific groups but are not captured in the data in any standardized format. In addition, aggregators of biodiversity information also make these decisions on a regular basis. We need to provide software tools to facilitate the capture and documentation of this process so that the information can be used for future data integration. As part of the SEEK project and to support the SE Atlas project, we developed a prototype taxon mapping tool, ConceptMapper[1]. ConceptMapper is a desktop tool to assist taxonomists to relate taxonomic concepts from one classification to another and to manage taxonomic concept metadata that precisely define taxonomic concepts. Concept data stored in the system can be retrieved, visualized and changed through the ConceptMapper user interface. Main functions include importing, exporting, querying and viewing concept data, adding and editing relationships, concepts, references and specimens.

- Taxon concept service. A service containing a central cache of taxon concepts and their relationships is needed so that users and data systems that wish to integrate data can find related concepts and make accurate matches. Preliminary design specifications and a prototype were developed as part of the recent SEEK project[2].

- Taxon concept matching tools. Users wishing to integrate datasets need to discover taxon concept relationships and be guided through the required integration decisions. In particular, we need a tool that increases our ability to import sets of concepts and find matches; suggesting when the ecological and biodiversity information can be merged and how to merge ambiguous matches. This is a complex task commonly required in meta-analysis.

IV.d. Data warehouse and network - Mechanisms and infrastructure must be provided to allow efficient access to and analysis of the available data and incorporation of other data types linked by commonality of taxon or location or both. Data integration at this scale will be an immense challenge, one that we believe will require an innovative hybrid solution with features of both a data warehouse and a data network. As a data warehouse, the solution must allow for import and standardization of diverse data according to a common schema and vocabulary and support efficient querying and data exploration. As a network, the solution must empower the community to contribute, manage and update their own data. Engaging the community will be essential for long-term sustainability in the face of frequent updates. We anticipate that the greater BIEN data network will be distributed across many institutions, and that this is unavoidable because of issues of ownership, confidentiality and funding. However, we anticipate drawing on the many sources to create a centralized data warehouse, optimized for user access and efficiency. This application will interact with all of the preceding applications and will manage mapping of source data to a common exchange schema, transfer to temporary staging tables, correction and standardization within the staging tables of taxonomy, locality information and geocoordinates, and final import and normalization to the BIEN core database. A user interface will allow for supervised execution of these processes, and will enable users to generate error reports that can be used to correct source data, thus providing feedback to the community and improving data quality during future imports.

IV.e. Data discovery and analysis environment - The data discovery environment will consist of a rich set of tools for user-driven data exploration, amalgamation and extraction. These tools will assist in the querying and compilation of candidate data sets and subsequent analysis. A common language of geospatial and phylogenetic queries will enable the user to build linkages between normally disparate

[1] http://cvs.ecoinformatics.org/cvs/cvsweb.cgi/~checkout~/seek/projects/taxon/conceptmapper/
[2] http://seek.ecoinformatics.org/Wiki.jsp?page=SEEKTaxonCommunity

datasets such as ecological inventories, specimen occurrences, trait databases and environmental observations.

- Tools for the implementation of workflows - Many types of data manipulation and analysis will be complex, yet will need to be repeated by numerous users. Examples include efficient phylogenetically structured queries, queries linking either children or parents, queries oriented around taxon concept resolution, flexible geospatial queries and data mapping, and integration of genomic, physiological or trait data. Often these may need to be established in advance to be run from a simple web interface. . We anticipate building on the Kepler[1] workflow environment and system developed for the Ecological and Geoscience communities. Such workflow infrastructure tools will allow automated analyses capability to be readily available to domain-level scientists and even the general public through web portals.
- Tools for data query and discovery – Two types of tools are necessary for performing data discovery: (1) A data environment allowing the extraction of any subset of data, and (2) a suite of analytical tools. Obviously useful subsets of the data would be single species or locations, but more complex queries, such as phylogenetic or trait-based, would also be useful. Analytical tools include mapping of many features, range prediction (thus rely on links to climatic data), and comparisons accounting to alternative taxonomic concepts.
- Tools for taxon/clade mapping, exploration, and analytical range modeling - Features of the data discovery environment must also include: flexible geospatial querying and mapping of data, efficient, phylogenetically-structured queries, and the ability of the user to retrieve all records for child taxa linked to a particular parent taxon/clade that occurs within a user specified geographic region. The BIEN team is interested in also working with iPlant to also provide simultaneous access to public domain climate, geophysical and satellite data. In addition, we will incorporate new research [48-51] into how to best 'predict' the geographic range of a given taxon/calde. Range predictions based on widely-used ecological niche modeling techniques (see [48-51] for a review) will be constructed by combining taxon observations from plots and specimens with public domain climate, geophysical and satellite data. These predictions will be cached and available for searching in the same manner as actual taxon observations, enabling users to predict species occurrences and richness in unexplored regions. The current University of Arizona Biodiversity Informatics Initiative (BDII[2]) provides an example of the potential power of such applications.

Proposed Cyberinfrastructure for Data Discovery Environment

- **Flexible geospatial querying and mapping of data.**
- **Efficient, phylogenetically-structured queries, enabling the user to retrieve all records for child taxa linked to a particular parent taxon.**
- **Support of alternative taxonomies and phylogenetic perspectives**
- **Integration with phylogenetic, genomic, physiological and trait datasets**
- **Simultaneous access to public domain climate, geophysical and satellite data**
- **Analytical tools for range modeling**

IV. f. Update and checking tools - - Data are constantly changing. Any comprehensive and robust biodiversity cyberinfrastructure must efficiently handle the dynamic nature of data where new data and corrections to previously available data are seamlessly incorporated and integrated on a continuous basis. This will require either that sources push corrections to the system are queried for updates on a routine basis. In addition, if data are to be cited and available for reanalysis, it will be necessary to have perfect versioning so that the data source can be viewed as it was at any time in the past, a capability already incorporated into some data sources, such as VegBank. Finally, as errors are identified and as value is added to records, this information should be returned to the source. Indeed, many of the tools we will be developing will provide useful feedback to collections as they continue to digitize their specimens.

V. – Education and outreach

Cyberinfrastructure tools created for BIEN will facilitate herbarium management, plant identification and training. These tools, developed for the Data Discovery Environment (DDE), will range from the

[1] http://kepler-project.org/Wiki.jsp?page=KeplerProject
[2] http://loco.biosci.arizona.edu/bdii/

management of specimens (in both herbaria as well as for ecologists in the field) to the ability to quickly identify plants. The ability to automatically automate taxonomy updates from experts (instead of searching through records manually) and to be able to rapidly identify duplicate specimens would greatly facilitate the collection and management of biodiversity data. This last example is akin to the Library of Congress where when you are registering a book, you can look up the ISBN. Something similar to an ISBN for herbarium specimens, where herbaria could access the full record for a specimen that has already been entered, would reduce data entry errors and accelerate specimen data entry. Additionally, the BIEN network could greatly increase the ability to identify plants in the field (aid in the creation of keys, regional guides etc.

Researchers armed with access to specimen data with well-defined taxonomy will be a tremendous improvement to the current system of access to specimens. The ability to generate geographic range maps from the collection of specimen locations could have a significant impact as a training tool and will likely be key for conservation scientists and professionals, both in terms of discovering the botanical diversity of protected and unprotected areas as well as setting priorities for inventories of areas with conservation potential. Smaller-scale herbaria, such as those in the developing world, are crippled by poorly-curated collections, limited resources to identify specimens, and almost no access to floristic treatments. When floristic treatments (monographs, catalogues, checklists) are available, there are no associated images and drawings. Moreover, maintaining up-to-date taxonomy is nearly as slow as it was a century ago; taxonomic changes still need to be evaluated name-by-name, and updated by hand, specimen-by specimen. This creates a tremendous obstacle for specimen identification and training, as well as hampering local scientists' abilities to conduct ecological and botanical research.

Cyberinfrastructure for Outreach and Education

Specimen management –
- Batch taxonomy updates.
- Specimen management - Identifying duplicate specimens.

Identification
- On-line taxon identification tools
- Range maps/Plant distribution maps.

A globally confederated database, with well-defined taxonomy and images of scanned specimens, lays the foundation for addressing these issues. If there is a central source for taxonomy and images, herbaria managers, botanists, ecologists, and students will have access to the tools they need to identify plants in the field and existing specimens. Further they will be better able train the next generation of botanists and ecologists, conduct stronger research, and do more informed conservation planning.

VI. – Computer and information science needs and available expertise

VI.a. - Technology partnerships: Community. The plant biodiversity and vegetation ecology communities collectively have long recognized the value of merging their data across institutional boundaries, and have made considerable strides in this regard over the past decade. The formation of GBIF; the development of exchange standards like Darwin Core, ABCD, TCS, and the emerging VegX; the implementation of distributed querying frameworks via DiGIR and TAPIR protocols—each of these technologically-informed standards and their implementation required skilled technologists working alongside plant scientists, assure the design of appropriate cyberinfrastructure solutions that enhance research opportunities, encourage community participation, and provide appropriate accreditation to the original data providers. A major problem, however, with developing cyberinfrastructure for an entire research discipline is mounting and sustaining a critical mass of engineering effort. The institutions and individuals involved with this proposal have all been involved in developing technology solutions for the plant biodiversity and ecological sciences. Most of these groups support their technology efforts through relatively modest, short-term research awards, where there are strong pressures to advance and publish about technological innovation, and not simply focus on creating and maintaining cyberinfrastructure.

VI.b. - Technology partnerships: iPlant - We envision iPlant providing a highly skilled software engineering workforce (managed by experienced information architects) and working alongside the BIEN domain scientists, to construct the next generation of cyberinfrastructure. iPlant can provide the technological expertise leading to efficient implementation and deployment, in compliance with existing and emerging standards. There should, however, be some committed engagement with related cyberinfrastructure efforts in the ecological sciences. In particular, an NSF/OCI INTEROP project, “Virtual Data Centers”, is attempting to confederate general ecological data, much of which contains taxonomic occurrence information. In addition, NSF/OCI’s DataNet program has the explicit mission of facilitating construction of long-term archives for scientific data, and the DataNet Observation Network for Earth, or DataNet ONE, is currently pending approval as a very broad effort to bring together data in the observational earth sciences. Assuring compatible authentication and access control mechanisms, and common data exchange and communication protocols with these and other relevant cyberinfrastructure projects will not only minimize redundancy of effort in technology development, but also potentially lead to greater compatibility of data resources. It would, for example, be highly advantageous for the BIEN data resources to be discoverable and accessible by ecologists and other researchers from outside of BIEN *per se*, even if with only a subset of features we are proposing here for the Data Discovery Environment.

We expect that iPlant staff will take principal responsibility for defining and developing the cyberinfrastructural products proposed here. We envision that iPlant engineers will be embedded within each of the proposed BIEN Working Groups (sectn VI). Their role will be to help clarify and define, through an iterative process, how the iPlant project will develop and implement the technologies proposed herein. While iPlant engineers undertake the actual process of constructing the cyberinfrastructure, these Working Group participants are committed to making themselves available for ad-hoc, as well as regularly scheduled consultation to help clarify details for the iPlant workforce. The participants represented here are also committed to facilitating iPlant understanding of the various software infrastructures that already exist. This will help prevent the undesirable outcome of iPlant reinventing existing solutions, or pursuing ends that ultimately are not compatible with the relevant data sources, or do not provide the relevant research features with which the listed participants are broadly familiar.

Aside from a dedicated and talented iPlant engineering workforce to help pursue these cyberinfrastructural goals, we feel that it is critical to the success of this mission to recognize and assist the invited participants in contributing their efforts. Accordingly, we recommend that the iPlant project consider how and whether the individual Working Groups might be assigned dedicated staff who might not only reside at the iPlant lead institutions, but also spend time at various of the participating institutions, as needed, to interact with the lead technology and plant researchers at those institutions. Another useful option would be iPlant’s assignment of some dedicated technical liaison staff at the PI institutions, to assure continued focus and closely coordinated communication with the main engineering workforce at iPlant.

VII. Project Management:

Much of the CI development will be done with assistance from Working Groups containing expertise in plant species, representative user groups, education and outreach audiences, evaluation teams, etc. We propose five Core Activities Working Groups to define and motivate the primary science and cyberinfrastructural goals of this project:

A. **Science Working Group:** will oversee and set the agenda for BIEN so that the development of cyberinfrastructure is guided by Science needs. This group will also initiate and conduct, with collaboration of with BIEN team members and collaborators, the research underlying our core research questions.
B. **Specimens, Plots, and Occurrences**: to bring together the significant data resource providers on this proposal, identify potential future providers, and clarify domain-related challenges to integrating these disparate data sources
C. **Confederation Data Model and Exchange Schema**: working in close conjunction with Group A, to develop a formal model that can integrate the relevant data resources, and provide a well-specified set of protocols to allow for future community participation
D. **Taxonomy**: to focus on specifying and resolving the taxonomic names issues with a robust exchange schema
E. **Georeferencing**: to focus on developing a standards-compliant approach to resolving and harmonizing any of the plant data having a georeferenced context.

F. **Data System Features and usability**: work closely with Working Group A to help define the end-user requirements for this framework, to assure their relevance in meeting the most critical needs of the targeted research audience especially in the Data Discovery Environment.

We also propose two Synergistic Activities Working Groups, to acknowledge the importance of linkages with these other areas of concern, and to potentially feed into the core cyberinfrastructure development activity:

A. **Phylogenetic Issues**: with much closely aligned work focused on phylogenetic analyses of plant communities, this working group will help identify and potentially coordinate participation of technology efforts in phylogenetics with BIEN.
B. **Trait/Genomic Integration**: many efforts are underway to more effectively enabling querying on functional traits of vegetation rather than taxonomic names, and this working group will identify and potentially coordinate the participation of those efforts with BIEN.
C. **Education and Outreach**: This group will design digital training tools (*i*) for global distribution of plant distributions and specimen management tools for herbaria and their associates; (*ii*) for conservation biology professionals; (*iii*) for instructors/educators to use the tool. In addition, this group will help design geographic discovery tools for predicting species in a polygon that can be used locally anywhere in the Americas, and then a usability assessment in several settings, Linkages to information to aid in identification

The associated domain scientists as well as key community personnel associated with each working group are detailed in the Appendix C. Once work groups are formed, a joint workshop will be held with other GC Teams with common CI interests and iPlant staff. We envision that iPlant could sponsor at least 1-2 meetings per year for each of the Working Groups., Several of these Working Groups might meet in close succession or simultaneously, to enable program-wide coordination of their efforts. (see suggested relative time line on Appendix D).

VI. b. Project progress monitoring and evaluation plan - The BIEN core team will establish an external evaluation team comprised of plant biology community. This evaluation team will be responsible for overseeing that the central goals as finally agreed upon between BIEN and iPlant are effectively and efficiently implemented, but also that the needs of the representative botanical communities are best met. Further, external evaluators will be responsible for assessing Cyberinfrastructure as well as educational and outreach.

VIII – Impact of successful infrastructural development on the broader field

Taxonomy underlies all biology and will be central to the development of synthetic biodiversity science at iPlant. We propose to address the taxonomy challenge head-on and provide tools that assure the necessary matching of names and concepts. Solving the 'taxonomy problem' is the key enabler and would be widely useful to botanists and zoologists alike. Indeed, we see taxonomy tools as central to integration within all of iPlant. We will work closely with other groups within iPlant and elsewhere (TRAITNET, TRY, TEAM, Tree of Life, Bar Code Consortium) to assure that taxonomic tools we create are widely used. We emphasize that it is the *integration* of taxonomic information with the significant information resources being brought together under this proposal that will significantly improve access and use of biodiversity data within the plant sciences. Indeed we see taxonomy tools as central to integration within all of iPlant, and a key requirement for creating the massive, global-scale confederation of plant biodiversity data that we propose here. The PI's and associated collaborators represent the range of institutions, authority, and competencies to provide the iPlant engineers with top-notch expertise in plant ecology and biodiversity cyberinfrastructural needs, along with close awareness of the relevant existing technological implementations underway within these areas. This combination of expertise is what we believe is necessary to create a relevant and transformative information resource for plant biology.

The BIEN network will provide biologists, ecologists, and conservation biologists easy access to widely dispersed data that everyone knows about but few can ever work with. The BIEN tools will be valuable to virtually every branch of ecology, even if simply to provide background information on a species of interest. Finally, by working with iPlant, we hope to remain near the forefront of bioinformatics, developing software that helps to assure the integrity, accuracy, and timeliness of widely used data. At least some of these tools may be relevant in other aspects of informatics.

Appendix A - Tables

Table 1. BIEN core team and collaborators outside the team of PIs. Core Team members responsibilities involve serving as a central domain scientist in one or more of the core working groups as well as overseeing the development of cyberinfrastructure. Collaborators may serve in BIEN working groups and/or advice the development of tools. All have pledged data or collaboration in the development of software tools and to represent the botanical communities that they represent.

Name	Institution	Core Team or Collaborator	Attended BIEN Workshop	Contribution
Sandy Andelman	Conservation International	Core	x	Data, tools, outreach, science
Jeannine Cavender-Bares	University of Minnesota	Core	x	Data, science
Steven Dolins	Bradley University	Core	x	IT expert
James Edwards	Encyclopedia of Life	Collaborator		Tools
Stephanie Hampton	NCEAS[2]	Core		Outreach, Education
John Janovecs	Botanical Research Institute of Texas, ATRIUM	Collaborator		Data, tools
Peter Jørgensen	Missouri Botanical Garden	Core	x	Data & taxonomy
Jessie Kennedy	Napier University (UK)	Core		Taxonomy, tools
Nick King	Global Biodiversity Information Facility (Denmark)	Collaborator		Data, tools
James Macklin	Harvard University Herbarium	Collaborator		Taxonomy, tools
Patrick Miles	U.S. Forest Service	Collaborator		Data
Brian McGill	University of Arizona	Core	x	informatics, science
Oliver Phillips	Leeds University (UK)	Core	x	Data, science
Tony Rees	CSIRO[1] (Australia)	Collaborator		Tools
Hans ter Steege	National Herbarium (Netherlands)	Core	x	Data, science
Corine Vriesendorp	Field Museum	Core	x	Outreach, Education
Nathan Swenson	Michigan State University	Core	x	Data, science
David Vieglais	University of Kansas	Core		Tools
Susan Wiser	Landcare Research (New Zealand)	Core	x	Data standards
Kerry Woods	Bennington College	Collaborator	x	Data
Josh Madin	ARC-NZ Research Network for Vegetation Function[3], & Computational Ecology Group, Macquarie University (Australia)	Collaborator		IT, Tools

2 National Center for Ecological Analysis and Synthesis, US
1 Commonwealth Scientific and Research Organization, Australia
3 http://www.vegfunction.net/

Table 2. Plots and other vegetation censuses currently available for the BIEN short-term data integration network. We also list external collaborators (*) who have been contacted and are willing to share data but at this point are satisfied with remaining external to the BIEN group.

Organization	Location	Contact	Number of units	Unit size (ha)
CTFS[8]	S. America	R. Condit	53	1
RAINFOR[9]	S. America	O. Phillips	252	1
ATDN[10]	S. America	H. ter Steege	251	1
SALVIAS[11]	S. America	B. Enquist	233	0.1
TEAM[12]	S. America	S. Andelman	90	1
Missouri Botanical Garden	Bolivia	P. Jørgensen	578	0.1
BRIT/Atrium[13]	Peru	J. Janovecs	81	0.1
US Forest Service, FIA[14]	USA	P. Miles	300,000	0.01
US National Park Service	USA	C. Lea*	5,000	0.1
US Forest Service, NRS[15]	Michigan	K. Woods	445	0.01
VegBank	USA	R. Peet	21,000	0.1
West Virginia Heritage	WV	J. Vanderhorst*	2,900	0.1
Virginia Heritage Program	VA	K. Paterson*	3,900	0.1
US Forest Service, Landfire	USA	D. Long*	365,896	various
Carolina Vegetation Survey	S.E. USA	R. Peet	8200	0.1

8 Center for Tropical Forest Science
9 Amazon Forest Inventory Network
10 Amazon Tree Diversity Network
11 Synthesis and Analysis of Location Vegetation Inventories
12 Tropical Ecology Assessment and Monitoring Network
13 Botanical Research Institute of Texas
14 Forest Inventory and Analysis
15 Northern Research Station

Table 3. Herbarium, museum and other occurrence records presently available for the BIEN data integration program. Most are original sources of data, but GBIF is a secondary collection of the other sources, plus many others. We also list external collaborators (*) who have been contacted and are willing to share data but at this point are satisfied with remaining external to the BIEN group.

Organization	**Contact**	**Number of records**
Missouri Botanical Garden	Peter Jørgensen	3,000,000
New York Botanical Garden	Barbara Thiers	900,000
Smithsonian Institution	Warren Wagner*	1,000,000
University of Arizona	Brad Boyle	173,000
Harbard University Herbaria	James Macklin	200,000
Field Museum	Robin Foster*	90,000
University of Aarhus	Jens Christian Svenning	100,000
Utrecht University	Hans ter Steege	114,000
University of North Carolina	Robert Peet	110,000
GBIF[1]	Nick King	65,626,000 (animals & plants)

[1] Global Biodiversity Information Facility

Appendix B. The following outlines the work-flow and principal applications of the proposed BIEN Data Integration and Discovery Network.

(I). DATA INTEGRATION

1. **Import.** Primary data are mapped and imported to VegX schema-compliant staging tables. Only minimal standardizations needed to match source data to schema are performed at this stage. Plots and other taxon observations will generally require a separate mapping script for each dataset; most specimen dataset can be imported from existing Darwin Core extracts.
2. **Primary standardization.** Standardization is performed independently for each data source. Basic error-checking performed at this stage. No versioning. Some user intervention required via import/error-checking interface. Error reports are generated for correction of source data. Taxonomic and geographic errors must be corrected prior to normalization.
    a. **Taxonomy**. Taxa are checked to ensure they match to nomenclaturally-valid names stored in core database. Only spelling errors and nomenclaturally-invalid names are corrected.
    b. **Geography**. Coordinates checked for numeric errors (out of bounds, etc.). Coordinates checked against locality fields to detect mis-matches. Political divisions are checked against to ensure they match to standard values in core database.
    c. **Missing data**. User is prompted to provide missing data, especially metadata.
3. **Normalization.** Data from staging table are merged field by field to normalized core database.
4. **Secondary standardization & indexing.** These are performed once data is in core database. User intervention required. Changes are versioned from this point forward.
    a. **Taxonomy.** Taxonomic concepts are specified and taxonomic synonymys are adjusted at the discretion of the data owner, according to authority lists linked to core database.
    b. **Georeferencing.** Georeferencing tools may be used to add coordinates to non-georeference specimens at this stage
    c. **Duplicate detection.** Duplicate specimen records are detected and removed.
    d. **Indexing of collectors and determiners.**
5. **Data management.** A rich user interface will allow users to perform ongoing management and update of data directly within core database. Existing data sets may be imported in steps 1-4 above, or entered directly via data management interface.
    a. Correction of existing data.
    b. **Direct entry of new data** within controlled environment of data model, including both addition of new observations to existing data, or entry of entire new data sets.
    c. **Tracking of determinations from voucher specimens.** Updated identifications may be applied to ecological observations by monitoring determination status of voucher specimens deposited as herbarium specimens.
    d. **Adjustment of taxonomy.** Changes in taxonomic status can be tracked and propagated across data sets.
    e. **Upload of additional media.** Additional media linked to observations (images, recordings, environmental measurements) may be uploaded on an ongoing basis.
    f. **Data access.** User sets access level and field embargos (if any) for data Access levels are: 1-hidden; 2-metada visible, data by request only; 3-full data freely available. These permissions may be set as defaults, or assigned by the data owner to particular users for particular datasets. Record and field-specific embargoes may also be applied if necessary (e.g., locality fields hidden to protect threatened and endangered species). See The SALVIAS Project[16] and VegBank[17] for working examples of user-managed data access and field embargoes.

[16] www.salvias.net
[17] www.vegbank.org

(II). DATA DISCOVERY

1. **Data selection**. User may query by any attribute in database, including spatial joins and map browsing.
2. **Data sharing.** User may request access to full data for any restricted datasets or embargoed fields (e.g., endangered species). Interface facilitates direct communication between data owners and requesting parties. Data owners can track data access logs.
3. **Additional taxonomic standardization.** Alternative synonymies and taxonomic concepts may be applied to aggregated data.
4. **Additional data sources.** Data external to BIEN can be linked to primary data and accessed via taxonomy or geography (spatial joins).
    - Phylogeny. Taxa can be mapped to alternative phylogenies, and searched via hierarchically-structured queries capable of retrieving all ancestors (parents) or all descendents (children) of a given taxon.
    - Traits and physiological data
    - Molecular sequence data
    - Climate or other environmental data
    - Satellite imagery
5. **Analysis tools**. Examples include:
    - Mapping
    - Range modeling, under current, past and projected climate scenarios
    - Statistics of diversity and abundance
    - Ordination
    - Phylogenetically structured analyses (e.g., phylogenetic diversity)
6. **Download.** Full data available for download in a variety of formats.
7. **Versioning.** Users may explore alternative versions of data. Content of downloaded datasets are timestamped and archived, and can be retrieved and examine at any time.

Appendix C. Working groups and proposed membership for major components of the BIEN Data Integration and Discovery Network. Each working group consists of both IT experts with a record of relevant application development and domain scientists familiar with the principal cyberinfrastructural challenges. BIEN project leaders and core team members indicated by asterisk.

Role	Name	Institutional affiliation	Relevant applications or activities
1. Specimens, Plots and Observations Working Group			
Domain/IT	Bob Peet*	University of North Carolina	VegBank[18]; SE Floristic Atlas[19]
Domain	Peter Jorgensen*	Missouri Botanical Garden	TROPICOS[20]; The Madidi Project[21]
Domain	Rick Condit	CTFS, Smithsonian	
Domain/IT	Brad Boyle*	University of Arizona	SALVIAS[22], BDII[23]
Domain/IT	Brian Enquist	University of Arizona	SALVIAS
Domain/IT	Bob Magill	Missouri Botanical Garden	TROPICOS
IT	Chris Freeland	Missouri Botanical Garden	TROPICOS
IT	Jesse Kennedy*	Napier University	TDWG[24]
2. Data Model and Exchange Schemas Working Group			
Domain/IT	Bob Peet*	University of North Carolina	VegBank; SE Floristic Atlas
Domain/IT	Brad Boyle*	University of Arizona	SALVIAS, BDII
Domaiin	Susan Wiser*	Landcare Research, New Zealand	NVS[25]
IT	David Hearn	University of Arizona	BDII
IT	Mark Schildhauer*	NCEAS	TDWG-OSR[39]
IT	Josh Madin	Macquarie University	ARC-NZ Research Network for Vegetation Function[39], and Computational Ecology Group
Domain/IT	Nick Spencer	Landcare Research, New Zealand	Vegetation Observations Exchange Schema (VegX)[26], TDWG
Domain/IT	Miguel Cáceres	Universitat de Barcelona, Spain	VegX[26], VegAna[30]
Domain/IT	Martin Kleinkamp	Bundesamt für Naturschutz (BfN),	VegX[26], VegetWeb

18 http://www.vegbank.org
19 http://www.herbarium.unc.edu/seflora/firstviewer.htm
20 http://www.tropicos.org/
21 http://www.mobot.org/MOBOT/Research/madidi/
22 http://www.salvias.net
23 http://loco.biosci.arizona.edu/bdii/
24 http://www.tdwg.org/
25 http://nvs.landcareresearch.co.nz/
39 http://www.vegfunction.net/
26 http://wiki.tdwg.org/twiki/bin/view/Vegetation/WebHome
30 http://biodiver.bio.ub.es/vegana/

		Germany	
IT	Jesse Kennedy*	Napier University	TDWG
IT	Dave Vieglais*	University of Kansas Biodiversity Research Center	Darwin Core[27], DiGIR[28], Specify[29], Plantcollections.org[30]

3. Taxonomy Working Group

Domain	Bob Peet*	University of North Carolina	VegBank
Domain	Jerry Cooper	Landcare Research (New Zealand)	TWDG-Darwin Core[27], TCS[31]
Domain	Peter Jorgensen*	Missouri Botanical Garden	TROPICOS; taxonomic specialist
Domain	Barbara Thiers*	New York Botanical Garden	Taxonomic specialist
IT	Jesse Kennedy*	Napier University	TDWG
IT	Dave Vieglais*	University of Kansas Biodiversity Research Center	Darwin Core, DiGIR, Plantcollections.org
Domain/IT	Bob Magill	Missouri Botanical Garden	TROPICOS, taxonomic specialist
Domain/IT	Brad Boyle*	University of Arizona	SALVIAS, TaxonScrubber[32]
Domain/IT	Tony Rees	CSIRO	TaxaMatch[33]

4. Data System Features and User Interface Working Group

Domain/IT	Bob Peet*	University of North Carolina	VegBank
Domain	Peter Jorgensen*	Missouri Botanical Garden	TROPICOS; The Madidi Project
Domain/IT	John Janovec	BRIT	Atrium[41]
IT	Matthias Tobler	BRIT	Atrium[41]
IT	Dave Vieglais*	University of Kansas Biodiversity Research Center	Specify
IT	Mark Schildhauer*	NCEAS	SEEK[34], EarthGRID

5. Phylogenetic Issues Working Group

Domain/IT	Michael Sanderson	University of Arizona	TOL[35], Phylota Browser[36], BDII
Domain/IT	Cam Webb	The Arnold Arboretum of Harvard University	Angiosperm Phylogeny Website[37], TOL

27 http://wiki.tdwg.org/twiki/bin/view/DarwinCore/WebHome
28 http://digir.sourceforge.net/
29 http://www.specifysoftware.org/Specify
30 http://plantcollections.pathf.com/
31 http://www.tdwg.org/standards/117/
32 http://www.salvias.net/pages/taxonscrubber.html
33 http://www.cmar.csiro.au/datacentre/irmng/
41 http://atrium.andesamazon.org/
41 http://atrium.andesamazon.org/
34 http://seek.ecoinformatics.org/
35 http://www.tolweb.org/tree/
36 http://loco.biosci.arizona.edu/pb/
37 http://www.mobot.org/mobot/research/apweb/welcome.html

Domain	Peter Stevens	Missouri Botanical Garden	Angiosperm Phylogeny Website
Domain	Brad Boyle	University of Arizona	BDII
Domain/IT	Reed Beeman	University of Florida	TOL

6. Trait Integration Working Group

Domain	Brian Enquist*	University of Arizona	SALVIAS, TraitNet
Domain	Jeannine Cavendar-Bares*	University of Minnesota	Traitnet[38]
Domain	Nate Swenson	Michigan State University	SALVIAS, TraitNet
IT	Mark Schildhauer*	NCEAS	TraitNet, TDWG-OSR[39]
IT	Josh Madin	Macquarie University	ARC-NZ Research Network for Vegetation Function, and Computational Ecology Group

7. Education and outreach working group

IT	Matthias Tobler	BRIT	ATRIUM
Domain	Corrine Vriesendorp	Field Musuem	Field Tropical Plant Guides[40]
Domain	Sandy Andelman	TEAM, Conservation International	

8. Science working group

Domain	Brian Enquist	University of Arizona
Domain	Rick Condit	Center for Tropical Forest Science
Domain	Brian McGill	University of Arizona
Domain	Nate Swenson	Harvard University Herbaria
Domain	Jeannine Cavendar-Bares*	University of Minnesota
Domain	Sandy Andelman	TEAM, Conservation International

38 http://www.columbia.edu/cu/traitnet/
39 http://www.tdwg.org/activities/osr/
40 http://fm2.fieldmuseum.org/plantguides/

Appendix D

Proposed initial timeline for the BIEN project.

Relative Timeline											
Preparation Stage											
Create five core working groups	■										
Create synergistic Activities group	■										
Create External Evaluation Team	■										
Build BIEN Network											
Identify data sources	■	■	■								
Compile data	■	■	■								
Develop tools for data standardization				■	■	■	■				
Develop tools for feedback to data providers				■	■	■	■				
Create Data Discovery Environment								■	■	■	■
Education, Outreach and training											
Develop tools for data providers									■	■	
Develop tools for conservation biology									■	■	
Provide links to identification tools											■
Training sessions for users											■
Meetings											
iPlant Challenge/BIEN Team	■		■		■		■		■		■
Working groups (core and synergistic)	■		■		■		■		■		■
External evaluation	■				■				■		■

CVs from BIEN PIs.

Biographical Sketch- Brian J. Enquist

Associate Professor, Dept. of Ecology and Evolutionary Biology, University of Arizona,
benquist@email.arizona.edu, www.salvias.net/~brian

Professional Preparation-

Position, Institution	**Major/area Degree, dates**
Colorado College, Biology,	B.A. (with Distinction) 1991
U. New Mexico Biology,	M.S. (Ecology Program) 1995
U. New Mexico Biology,	Ph.D. (Ecology Program) 1998

Postdoctoral Institution(s) *Area Inclusive Dates (years)*

-NSF Postdoctoral Research Fellowship:

- *The Santa Fe Institute*, Santa Fe, N.M., Aug. 1998 - Aug 1999

- National Center for Ecological Analysis and Synthesis

(NCEAS), U.C. Santa Barbara, CA., Sept. 2000 - Dec. 2000

Appointments-

- **NSF Postdoctoral Fellow,**
  -*National Center for Ecological Analysis and Synthesis,* Sept. 1999 – Sept. 2000
  -*The Santa Fe Institute*. Aug. 1998 - Sept 1999.
- **Research Assistant Professor,** *Department of Biology, University of New Mexico*. 1998 – 2001.
- **Assistant Professor.,** *Department of Ecology and Evolutionary Biology, University of Arizona,* 2001 – 2005.
- **Associate Professor,** *Department of Ecology and Evolutionary Biology, University of Arizona* 2005 – .
- **External Faculty** – *The Santa Fe Institute, 2007 -*

Publications (From 85, Closely Related to the Proposed Project)

Bryant, J., Lamanna, C., Morlon, H., Kerkhoff, A.J., **Enquist, B.J.** and J. L. Green (2008) Microbes on mountainsides: Contrasting elevational patterns of bacterial and plant diversity. PNAS

Enquist, B. J., Kerkhoff, A.J., Stark, S.C., Swenson, N.G., McCarthy, M.C. and C. A. Price (2007) A general model for scaling plant growth, carbon flux, and functional trait spectra. Nature 449:218-222.

Swenson, N.G. and **B.J. Enquist** (2007). Ecological and evolutionary determinants of a key plant functional trait: Wood density and its community-wide variation across latitude and elevation. American Journal of Botany 94:451-459.

Weiser, M.D, **Enquist, B.J.**, Boyle, B., Killeen, T.J., Jorgensen, P.M., Fonseca, G., Jennings, M.D., Kerkhoff, A.J., Larcher, T.E., Monteagudo, A., Nunez Vargus, M.P., Phillips, O.L., Swenson, N.G., and R. Vasquez Martinez (2007). Latitudinal patterns of range size and species richness of New World woody plants. Global Ecology and Biogeography, 16: 679-688.

Enquist, B.J., Haskell, J. P., and Tiffney, B. H. (2002). General patterns of taxonomic diversity and biomass partitioning across tree dominated communities. Nature 419:610-613.

Other significant publications related to project:

Enquist, B.J., Kerkhoff, A.J., Huxman, T.E., and E.P. Economo (2007) Adaptive differences in plant physiology and ecosystem invariants: Insights from Metabolic Scaling Theory. Global Change Biology 13:591-609..

Swenson, N.G., **Enquist, B.J.**, Thompson, J. and J. Zimmerman (2007) The influence of spatial and size scale on phylogenetic relatedness in tropical forest communities. Ecology 88:1770-1780.

McGill, B., Etienne, R., Gray, J., Alonso, D., Anderson, M., Benecha, H.; Dornelas, M., **Enquist, B**., Green, J., He, F., Hurlbert, A., Magurran, A., Marquet, P., Maurer, B., Ostling, A., Soykan, C., Ugland, K., White, E. (2007) Species Abundance Distributions: moving beyond single prediction theories to integration within an ecological framework Ecology Letters 10:995-1015.

Kerkhoff, A.J. Fagan, W.F., Elser, J.J. and **B.J. Enquist** (2006) Phylogenetic and growth form variation

in the scaling of nitrogen and phosphorus in the seed plants. American Naturalist 168:E103-E122.

Enquist, B.J. and K.J. Niklas (2001). Invariant scaling relations across tree-dominated communities. Nature 410:655-660.

Enquist, B.J., Jordan, M.A. & J.H. Brown (1995). Connections between ecology, biogeography and paleobiology: relationship between local abundance and geographic distribution in fossil and recent organisms. Evolutionary Ecology 9:586-604.

Synergistic activities:

- Development of databases: creating global databases on plant community diversity abundance, functional diversity, and distribution bothin within and outside the tropics.
- Maintaining a longterm dataset on the growth, death, and recruitment of all trees within a 17ha long-term study site in Guanacaste, Costa Rica.
- Currently Dr. Enquist has a sizable ecoinformatics outreach. He is currently the developer and curator of a Global EcoInformatics Web Portal (SLAVIAS) (www.salvias.net). This web site was initially funded from Conservation International's Center for Applied Biodiversity Science and EEB. It is sponsored by Arizona Research Labs and the Missouri Botanical Garden. It currently houses an accessible electronic database of global plant communities (approximately three thousand). In addition, the website offers downloadable software to standardize, correct and align botanical taxonomic information.
- Broadening the participation of groups underrepresented in science*:* Collaboration with and training of botanists and ecologists from developing countries. Through my interactions with the development of SALVIAS I currently interact with numerous researchers from throughout Latin America and Europe.

Collaborators (over the last two years)

Dr. Allen, A. P., NCEAS, Prof. Brown, J.H., University of New Mexico, Prof. Elser, J.J. ,Arizona State University, Dr. Ernest, S.K.M., Utah State University, Prof. Fagen, W.F., University of Maryland, Dr. Gillooly J., University of New Mexico, Dr. Jessica Green, University of Oregon, Dr.. Huxman, T., University of Arizona, Prof. Niklas, K.J. ,Cornell University, Dr.. Thompson, J., University of Puerto Rico Prof. Tiffney, B. H. University of California, Santa Barbara, Prof. Peter Reich, University of Minnesota Dr. Savage, V., UCLA, Prof. John Sperry, University of Utah, Dr. West, G.B., The Santa Fe Institute, Dr. White, E.P., Utah State University, Prof. J.K. Zimmerman, National Science Foundation

Graduate and Post Doctoral Advisors.

- PhD Advisor - Brown, J.H. University of New Mexico.
- Postdoctoral- I was supported as a NSF postdoctoral fellow. As a fellow I completed independent study at the Santa Fe Institute and the National Center for Ecological Analysis and Synthesis in Santa Barbara, CA. My fellowship was sponsored by Dr. Karl J. Niklas at Cornell

Biographical Sketch- Richard Condit

Chief Scientist, Center for Tropical Forest Science, Smithsonian Tropical Research Institute
contact: conditr@gmail.com

Undergraduate institution: University of Illinois (B.S., 1978).
Graduate institution: University of California, Santa Cruz (Ph.D., 1984).
Postdoctoral institutions: Melbourne U., U. of Massachusetts, Princeton U.

Appointments:

Staff Scientist, Smithsonian Tropical Research Institute (STRI). 1991-present.
Research Associate, Princeton University. 1988-1991.
Research Associate, University of Massachusetts. 1985-1988.
Lecturer, Antioch/New England graduate school. 1987.
Lecturer, University of California, Santa Cruz. 1985.
Research Associate, University of Melbourne, Australia. 1984.

Five related publications:

Condit, R., Pérez, R., Lao, S., Aguilar, S., and Somoza, A. 2005. Geographic ranges and ß-diversity: discovering how many tree species there are where. Biologiske Skrifter, 55: 57-71.
Condit, R., *et al*. 2002. Beta-diversity in tropical forest trees. Science, 295: 666-669.
Condit, R., *et al*. 1996. Species-area and species-Individual relationships for tropical trees: a comparison of three 50 ha plots. Journal of Ecology, 84: 549-562.
Kenfack, D., Thomas, D.W., Chuyong, G., and Condit, R. 2006. Rarity and abundance in a diverse African forest. Biodiversity and Conservation DOI 10.1007/s10531-006-9065-2.
LaFrankie, J.V., Condit, R., *et al.* 2006. Contrasting structure and composition of the understory in species-rich tropical rain forests. *Ecology*, 87: 2298-2305.

Five other publications:

Condit, R., *et al*. 2006. The importance of demographic niches to tree diversity. Science 313:98-101.
Condit, R., *et al*. 2000. Spatial patterns in the distribution of common and rare tropical tree species: a test from large plots in six different forests. Science, 288: 1414-1418.
Condit, R., R. Sukumar, S.P. Hubbell, and R.B. Foster. 1998. Predicting population trends from size distributions: a direct test in a tropical tree community. American Naturalist, 152: 495-509.
Muller-Landau, H. C., Condit, R., *et al*. 2006. Testing metabolic ecology theory for allometric scaling of tree size, growth, and mortality in tropical forests. Ecology Letters 9: 575-588.
Rees, M., R. Condit, M. Crawley, S. Pacala, D. Tilman. 2001. Long-term studies of vegetation dynamics. Science, 293: 650-655.

Synergistic activities:

Computation methodologies: producing computer programs for wide use, and training in the use of these programs, on demography of trees, seals.
Development of databases: creating databases on tree species of the neotropics, of Central America; taxonomy and demography of 14 permanent tropical forest plots; demography of northern elephant seals.
Broadening the participation of groups underrepresented in science: collaboration with and training of botanists and ecologists from developing countries (especially Ecuador, Congo, Cameroon, Sri Lanka, India, Malaysia, Thailand).
Training: teaching the NSF-funded workshop on the analysis of tree-demographic data, aimed primarily at ecologists, foresters, and botanists from developing countries in the tropics

Collaborators and affiliations: 90 co-authors in the past 48 months; significant and frequent recent collaborators:

Boyle, B.	U. Arizona			Hubbell, S.P.	U. Georgia
Cardenas, D.	COAH, Colomia			Kenfack, D.	U. Missouri, St. Louis
Chave, J.	CNRS, Toulouse, France			Laurance, S.G.	Smithsonian, Panama
Chuyong, G.	U. Buea, Cameroon			Laurance, W.F.	Smithsonian, Panama
Davies, S.	Harvard U.			Le Bouef, B.J.	U.C. Santa Cruz
Duque, A.	U. Nacional, Colombia			Makana, J.-R.	CEFRECOF, Congo
Enquist, B.	U. Arizona			Muller-Landau, H.	Smithsonian, Panama
Foster, R.B.	Field Museum			Oliveira, A.	U. Sao Paolo, Brazil
Green, J.	U. Oregon			Thomas, D.	private consultant
Gunatilleke, N.	U. Peradeniya, Sri Lanka			Thomas, S.	U. Toronto
Gunatilleke, S.	U. Peradeniya, Sri Lanka			Valencia, R.	U. Católica, Quito
Harms, K.E.	Louisiana State U.			Wright, S.J.	Smithsonian, Panama

Graduate and post doctoral advisors: C.L. Ortiz (UC Santa Cruz), B.J. Le Boeuf (UC Santa Cruz), G. Borgia (Maryland), B.R. Levin (UMass-Amherst), S. Hubbell (UCLA).

Students and post-docs advised: L. Comita (Columbia), Y. Chen (Georgia), D. Dent (STRI), K. Harms (Princeton), E. Hooper (McGill), D. Kenfack (Missouri-St.L.), J.R. Makana (Toronto), M. Metz (UC Berkeley), H. Muller-Landau (Princeton), C. Pyke (UC Santa Barbara), T. Robinson (Oregon State), N. Teuschel (UC Santa Cruz), R. Zeno (UC Santa Cruz).

Biographical Sketch - Robert K. Peet

Address: Department of Biology, CB#3280, University of North Carolina
Chapel Hill, NC 27599-3280; Email: peet@unc.edu, Phone: 919-962-6942
Web: http://www.bio.unc.edu/faculty/peet

Education:

Cornell University	Ecology Ph.D.	1975
Univ. Wisconsin Madison	Botany M.S.	1971
Univ. Wisconsin Madison	Botany B.A., honors	1970

Appointments:

2008-09	Fellow, National Evolutionary Synthesis Center
2008	Interim Director, Institute for the Environment, UNC-CH
2003-08	Chairman, Curriculum in Ecology
1989 -	Professor, University of North Carolina, Chapel Hill
2006 -	Research Associate, University of North Carolina Herbarium
2001-02	Fellow, National Center for Ecological Analysis and Synthesis
1980-88	Associate Professor, University of North Carolina
1975-80	Assistant Professor, University of North Carolina

Honors: AAAS, Elected Fellow (1984); Author of a Citation Classic; 1995 Distinguished Service Award - Ecological Society of America; Phi Beta Kappa; Phi Eta Sigma; Phi Kappa Phi; Sigma Xi; President, International Association for Vegetation Science, 2007-2011; Secretary, Ecological Society of America, 2002-2005; Editor-in-Chief, Ecology & Ecological Monographs, 1995-2000.

Five relevant publications:

Franz, N.M. and R.K. Peet. 2009. Towards a language for mapping relationships among taxonomic concepts. *Systematics and Biodiversity (in press).* http://www.bio.unc.edu/faculty/peet/pubs/Syst&Biodiv2009.pdf

Jennings, M.D., D. Faber-Langendoen, O.L. Loucks, R.K. Peet, & D. Roberts. 2009. Characterizing Associations and Alliances of the U.S. National Vegetation Classification. *Ecological Monographs (in press).* http://www.bio.unc.edu/faculty/peet/pubs/EcoMono2008.doc

Franz, N. M., R. K. Peet & A. S. Weakley. 2008. On the use of taxonomic concepts in support of biodiversity research and taxonomy. Symposium Proceedings, In:Wheeler, Q. D.,Ed., *The New Taxonomy*. Systematics Association Special Volume (74). Taylor & Francis, Boca Raton, FL, pp. 63–86. http://www.bio.unc.edu/faculty/peet/pubs/Cardiff.pdf

Fridley, J.D., D.B. Vandermast, D.M. Kuppinger, M. Manthey, and R.K. Peet. 2007. Co-occurrence-based assessment of habitat generalists and specialists: a new approach for the measurement of niche width. Journal of Ecology 95:707-722. http://www.bio.unc.edu/faculty/peet/pubs/jecology95;707.pdf

Peet, R.K. 2006. Ecological classification of longleaf pine woodlands. In: *Longleaf pine ecosystems: ecology, management, and restoration.* S. Jose, E. Jokela and D. Miller. Eds. Springer, New York. Pp 51-93. http://www.bio.unc.edu/faculty/peet/lab/documents/LL_classification.pdf

Five additional publications:

Carr, S.C.,K.M. Robertson,W.J.Platt & R.K.Peet. 2009. A model of geographic, environmental and regional variation in vegetation composition of pyrogenic pinelands of Florida. *Journal of Biogeography (in press).* http://www.bio.unc.edu/faculty/peet/pubs/FL-model.pdf

Xi, W., R.K. Peet, J.K. DeCoster & D.L. Urban.2008. Tree damage risk factors associated with large, infrequent wind disturbances of Carolina forests. *Forestry* 81:317-334. http://www.bio.unc.edu/faculty/peet/pubs/forestry2008.pdf

Fridley, J.D., R.K. Peet, E. van der Maarel, and J.H. Willems. 2006. Integration of local and regional species-area relationships from space-time species accumulation. *American Naturalist* 168:133-143. http://www.bio.unc.edu/faculty/peet/lab/documents/SpaceTime.doc.

Graves, J.H., R.K. Peet, and P.S. White. 2006. The influence of carbon - nutrient balance on herb and woody plant abundance in temperate forest understories. *Journal of Vegetation Science* 17: 217-226. http://www.bio.unc.edu/faculty/peet/lab/documents/graves.pdf.

Fridley, J.D., R.K. Peet, T.R. Wentworth and P.S. White. 2005. Connecting fine- and broad-scale patterns of species diversity: species-area relationships of Southeastern U.S. flora. *Ecology* 86:1172-1177. http://www.bio.unc.edu/faculty/peet/pubs/ecology86;1172.pdf

Synergistic Activities:

1. Established & Chair, EcoInformatics Working Group of the International Association for Vegetation Science. This group is working to establish international standards for exchange of vegetation and species co-occurrence data. 2003 - present. http://www.bio.unc.edu/faculty/peet/vegdata/).
2. Organizer and principal investigator of a collaborative program of the Ecological Society of America, USGS-NBII, NatureServe, FGDC, and the National Center for Ecological Analysis and Synthesis to develop the information system and database architecture to support a US National Vegetation Classification (http://VegBank.org). 1998 - present
3. While Editor-in-Chief of *Ecology* and *Ecological Monographs*, conceived, designed and implemented *Ecological Archives* (http://www.esapubs.org/Archive/), a new digital publication for electronic appendices, supplements and data papers. 1997-2000.
4. While Secretary of the Ecological Society of America, designed a major revision of the governing structures of the Society and wrote the new Constitution and Bylaws. 1994-5.
5. Established (in collaboration with E. van der Maarel) the *Journal of Vegetation Science*, the official organ of the International Association for Vegetation Science. 1990.
6. Established & Chair. Carolina Vegetation Survey, a 22-year multi-institution and agency collaboration to documentation of the natural vegetation of the Carolinas. (http://cvs.bio.unc.edu)

Recent Collaborators & Coauthors: Marc Abrams (Penn. State U), Michael Barbour (UC Davis), James Beach (U Kansas), Forbes Boyle (UNC), Rebecca Brown (Eastern Wash U), Miquel Cáceras (Univ Barcelona), Susan Carr (U FL), Norman Christensen (Duke U), James DeCoster (US Nat Park Serv), D. Faber-Langendoen (NatureServe), Nico Franz (U Puerto Rico), Jason Fridley (Syracuse U), Frank Gillliam (Marshall), Joel Gramling (Citadel), James Graves (Green Mt C.), Cliff Hupp (USGS), Lee Anne Jacobs (Duke U), Michael Jennings (TNC), Matt Jones (NCEAS), Jesse Kennedy (Napier U.), Martin Kleikamp (Bergisch-Gladbach, Germany), Dane Kuppinger (Univ. of the South), Xianhua Liu (NESCent), Orie Loucks (Miami U), Eddy van der Maarel (U. Groenigen), Michael Manthey (Ernst-Moritz-Arndt U), Aaron Moody (UNC), Michael Palmer (Oklahoma St Univ), William Platt (Lousiana St. U), Marcel Rejmank (U Cal Davis), David Roberts (Montana St.), Kevin Robertson (Tall Timbers, FL), Michael Schafale (NC Heritage), Amanda Senft (UNC), Nick Spencer (Landcare New Zealand), Stephen Talbot (US Fish Wildlife Serv), Dave Tart (USFS), Kristin Taverna (Va Heritage), Phil Townsend (U. Wisc), Dean Urban (Duke U), David Vandermast (Elon College), Joan Walker (US Forest Service), Alan Weakley (UNC), Thomas Wentworth (NCSU), Peter White (UNC), Deborah Willard (USGS), Jo Willems (U Utrecht), Bastow Wilson (Otaga U.), Susan Wiser (Lanadcare New Zealand), Weimin Xi (Texas A&M).
Graduate Advisors: Orie L. Loucks (M.S.), Robert H. Whittaker (Ph.D)
Graduate Students (43) and Postdoctoral Scholars (9): Nick Adams, Dorothy Allard, Robert Allen, Peter Avis, William Baker, Forbes Boyle, Kathleen Baker-Brosh, Rebecca Brown, Jeffrey Corbin, Patricia Corry, Phillip Coulling, Orin Pete Council, James DeCoster, James Doyle, Richard Duncan, Nico Franz, Jason Fridley, Cecil Frost, Dan Gafta, Joel Gramling, James Graves, Rachel Hochman, Lee Anne Reilly, Eric Kjellmark, Mark Knott, Robert Knox, Xianhua Liu, Lauro Lopez-Mata, Elizabeth Matthews, Mary Lou May, Jasper McChesney, Claire Newell, Timothy Nifong, K.C. Oh, Jeffrey Ott, Kyle Palmquist, Thomas Philippi, Laura Philipps, Steve Rice, Anne Richards, Janette Schue, Amanda Senft, Stephanie Seymour, Sonja Stiefel, Gary Thorburn, Joan Walker, Brooke Wheeler, Jacqueline White, John White, Susan Wiser, Weimin Xi

Biographical Sketch – Mark P. Schildhauer

National Center for Ecological Analysis and Synthesis E-mail: schild@nceas.ucsb.edu
735 State St., Suite 300 http://www.nceas.ucsb.edu/
Santa Barbara, CA 93101 Tele: (805) 892-2509

PROFESSIONAL PREPARATION

Harvard College	Biology	A.B., 1976
University of California, Santa Barbara	Ecology and Evolutionary Biology	Ph.D.,1991

APPOINTMENTS

- 1995 - present Director of Computing National Center for Ecological Analysis and Synthesis, University of California, Santa Barbara
- 1993 - 1995 Computer Resource Manager and Technical Coordinator Division of Social Sciences, University of California Santa Barbara
- 1990 – 1993 Programmer Analyst Social Sciences Computing Facility, University of California, Santa Barbara
- 1986-1989 Ecological data analyst and laboratory computer coordinator Marine Sciences Institute, University of California Santa Barbara

PUBLICATIONS

- Madin, Joshua S., Shawn Bowers, Mark P. Schildhauer, Matthew B. Jones. 2008. Advancing ecological research with ontologies. Trends in Ecology & Evolution vol. 23, issue 3, pp. 159-168.
- Madin, Joshua S., Shawn Bowers, Mark Schildhauer, Serguei Krivov, Deana Pennington, Ferdinando Villa.. 2007. An ontology for describing and synthesizing ecological observation data. Ecological Informatics, vol. 2, issue 3, pp. 279-296.
- Michener, William K., James H. Beach, Matthew B. Jones, Bertram Ludaescher, Deana D. Pennington, Ricardo S. Pereira, Arcot Rajasekar, and Mark Schildhauer. 2007. A knowledge environment for the biodiversity and ecological sciences. Journal of Intelligent Information Systems, vol. 29, no. 1, pp. 111-126.
- Jones MB, M Schildhauer, OJ Reichman and S Bowers. 2006. The New Bioinformatics: integrating ecological data from the gene to the biosphere. Annual Review of Ecology, Evolution, and Systematics 37, pp. 519-544.
- Gamon J.A., A.F. Rahman, J. Dungan, M. Schildhauer and K.F. Huemmrich. 2006. Spectral Network (SpecNet)-- What is it and why do we need it? Remote Sensing of Environment, vol. 103, pp. 227-235.
- Fegraus, Eric H., S. Andelman, M. B. Jones, and M. Schildhauer. 2005. Maximizing the value of ecological data with structured metadata: an introduction to ecological metadata language (EML) and principles for metadata creation. Bull. Ecol. Soc. Amer. vol. 86, pp.158-168.
- Berkley C., S. Bowers, M.B. Jones, B. Ludaescher, M. Schildhauer, and J Tao. 2005. Incorporating Semantics in Scientific Workflow Authoring. Proceedings of the 17th International Conference on Scientific and Statistical Database Management. IEEE Computer Society.
- Jones, Matthew, C. Berkley, J. Bojilova, M. Schildhauer. 2001. Managing Scientific Metadata. IEEE Internet Computing, vol. 5, no. 5, pp. 59-68.
- Nottrott, R., M.B. Jones, and M. Schildhauer. 1999. Using XML-structured metadata to automate quality assurance processing for ecological data. Proceedings of the Third IEEE Computer Society Metadata Conference. Bethesda, MD. April 6-7, 1999.
- Schildhauer, Mark P. 1998. Virtual Working Groups at NCEAS: Using the Web to Facilitate Scientific Collaboration. In Data and Information Management in the Ecological Sciences: A Resource Guide. Edited by Michener, W.K., J.H. Porter and S.G. Stafford. LTER Network Office, Univ. of New Mexico, Albuquerque, New Mexico.
- Schildhauer, Mark P. 1993. Training people to use the SAS system under UNIX operating systems. In Western Users of SAS Software, Proceedings of the First Regional Conference.

- Hoffman, S. G., M. P. Schildhauer, and R. R. Warner. 1985. The costs of changing sex and the ontogeny of males under contest competition for mates. Evolution 39: 915-927.

SYNERGISTIC ACTIVITIES

- Invited Lecturer. 2007. Ecological Informatics: challenges and approaches, and potential relevance for archaeology. For: Archaeoinformatics Institute, Tucson AZ. Videoteleconference.
- Invited Presentation. 2007. An ontological approach to describing and synthesizing ecological data, using a generalized model for "scientific observations" (with Bowers, Madin, Jones). At: TDWG (Biodiversity Information Standards) Annual Meeting, Bratislava, Slovakia.
- Invited Presentation. 2006. Metadata and Ontologies for the Ecological Sciences. At: International Conference on Hydroscience and Engineering Science (ICHE) Annual Meeting. Philadelphia, PA.
- Invited Panelist. 2005. Empowering ecologists with informatics education and training. At: Ecological Society of America Annual Meeting, Montreal Canada.
- Invited Presentation. 2004. Issues of trust and security for integrative science: experiences at the interface of ecological research and Internet technology. In: Symposium on Trust and Security in Biological Databases, AAAS Annual Meeting. Seattle, WA.
- Invited Presentation. 2003. Trends in Ecoinformatics: SEEK and the KNB. At: Innovations in Ecosystems Research. CEA-CREST Annual Meeting, Los Angeles CA.
- Invited Presentation. 2002. Cyberinfrastructure for the Ecological Sciences. Presentation and panel discussion delivered before the NSF Advisory Committee on Environmental Research and Education. Wash. DC.
- Invited Presentation. 2001. Metadata. Workshop for Long-Term Ecological Data Management. Ecological Society of America Annual Meeting. Madison, Wisc.
- Informatics Instructor. 2006. Scientific Workflows for adaptive management of TPCs at Kruger National Park. Kruger National Park, Republic South Africa.
- Informatics Instructor. 2004. Metadata for Ecological Data at Kruger National Park. Kruger National Park, Republic South Africa.
- Technical Consultant 2005-present. TRENDS project of the LTER Network.
- Technical Advisor 1997. National Index Site Committee.
- NSF Training Workshop (Instructor). 1997. Virtual Working Groups at NCEAS: Using the Web to Facilitate Scientific Collaboration. Ecological Society of America, Annual Meeting. Albuquerque, New Mexico.
- Data Advisory Panel Chairperson. 1999-2005. Multi-Agency Rocky Intertidal Network (consortium), coordinated by the Minerals Management Service of the U.S. Dept. of the Interior.
- NSF Panelist: BDI, ITR, OCI programs.
- Professional Societies: Ecological Society of America, IEEE Computer Society

Biographical Sketch – Barbara Thiers

New York Botanical Garden, Bronx, NY 10458-5126.
Phone: 718/817-8622
Fax: 718/817-8809
e-mail: bthiers@nybg.org.

A. Professional Preparation

San Francisco State University	Biology	B.A. 1977
University of Massachusetts	Botany	Ph.D. 1982

B. Appointments

Postdoctoral Museum Intern New York Botanical Garden: 1981-1982.
Manager of the Cryptogamic Herbarium, New York Botanical Garden, 1982-1985.
Administrative Curator, New York Botanical Garden, 1986-1999.
Associate Director of the Herbarium, New York Botanical Garden 1999-2000.
Acting Vice President for Science, New York Botanical Garden, 2001-2002.
Director of the Herbarium, New York Botanical Garden, 2000-present.

C. Publications

Five publications most closely related to the proposed project:

1983. Index to the genera and species of Hepaticae described by William Mitten. Brittonia 35: 271-300.

1984. Index to species described in Hepaticae Amazonicae et Andinae by Richard Spruce, with nomenclatural updating. Contr. New York Bot. Gard. 15: (1)-(14).

1989. Lejeuneaceae (Hepaticae) of Australia. I. Subfamily Ptychanthoideae of Australia. Mem. New York Bot. Gard. 52: 1-79. (with S. R. Gradstein).

1992. An index to the species of mosses and lichens described by William Mitten. Memoirs of the New York Botanical Garden 67:1-113. (project coordinator).

1993. A monograph of *Pleurozia* (Hepaticae, Pleuroziaceae). Bryologist 96: 517-554.

Five other significant publications:

1983. The fungus herbarium of the Carnegie Museum of Natural History (CM), Pittsburgh, Pennsylvania. Brittonia 35: 367-373. (with Dennis Desjardin and Andrew S. Methven).

1984. Branch characters significant to subfamilial classification of Lejeuneaceae (Hepaticae). Syst. Bot. 9: 33-41.

1984. Fungi from the A. O. Garrett Herbarium, University of Utah (UT). Brittonia 36: 293-296. (with C. T. Rogerson).

1988. Morphological adaptations of the Jungermanniales (Hepaticae) to the tropical rainforest habitat. J. Hattori Bot. Lab. 64: 5-14.

1997. *Cheilolejeunea* in Australia: Description of new taxa and key. J. Hattori Bot. Lab. 82: 321-328.

Synergistic Activities

- New York Botanical Garden Virtual Herbarium. Lead staff member involved in development of electronic data sharing program at the New York Botanical Garden since program inception in 1990. Grants obtained in support of Virtual Herbarium from the National Science Foundation total more than $3 million from 1996—2008. Projects supervised (funded by NSF and other sources) include

The New York Botanical Garden American Bryophyte Catalog. Phases I & II (North America, Central America, Mexico and the West Indies), Macrofungi Type Specimen Catalog; *Flora Borinqueña* (Plants and Fungi of Puerto Rico), Collaborative Digitization of The New York Botanical Garden Specimens from Amazonian Brazil, Vascular Plants of the Intermountain West, and the Global Plants Initiative.

- Editor, Index Herbariorum, 2008-. Index Herbariorum is an on-line directory to the approximately 3900 public herbaria of the world, and 10,000 biodiversity experts and curators associated with those institutions. Editing of the Index involves reviewing new applications for membership, approving updates to current records for member institutions and use of the data for a wide range of scientific and governmental users.
- Global Plants Initiative – NYBG Project Leader. The LAPI project, funded by the Mellon Foundation, is enabling more than 100 herbaria worldwide to digitize type specimens for their own use, and inclusion in a comprehensive digital library (hosted by JSTOR). NYBG not only digitizes its own type specimens as part of this endeavor, but also provides training and support for other participants, and contributes to overall management of the project.
- Encyclopedia of Life – Represents NYBG on the Institutional Member Council since the inception of the project in 2006. The Institutional Council, an international body, advises on the development and growth of the project.
- Panel Member, Natural Sciences and Engineering Research Council of Canada Strategic Network Grant for the Canadian Barcode of Life Netowork. Since 2005, I have been involved in the review and on-going evaluation of this very large grant to the University of Guelph for the development of the Canadian Barcode of Life.
- Visiting Review Committee to the Botany Department, Natural History Museum, London. September, 2007 & for the Royal Botanic Gardens, Edinburgh, Oct 2008. Participated in reviews of the activities of herbarium and collection digitization projects of both institutions.
- Participant, NMNH-NEON Workshop: Curation of Biological Specimens, Physical Samples and Associated Data (20-22 October,2008). The purpose of the workshop was to inform the biodiversity collections community about the current state of the NEON project and to get advice from the collections community on the management of the biodiversity vouchers and samples that will be generated by NEON.
- Collaborator, Botanical Information and Ecology Network (BIEN). Described in this proposal.

E. Collaborators and Other Affiliations

1. Graduate and Postgraduate advisors: Dr. Rudolf M. Schuster

2. Thesis Advisor and Postgraduate-Scholar Sponsors: none

Sample of CVs from some of the key core team members of BIEN.

Biographical Sketch – Steven B. Dolins

Address: Department of Computer Science and Information Systems
Bradley University
Peoria, Illinois 61625
Email: sdolins@bradley.edu
Phone: Voice 309-677-3284; Fax 309-677-4504; Cell 312-925-3956
Web: http://www.bradley.edu/academics/las/cs/dolins.shtml

Education:

Tulane University	Physics, Pol Sci	B.S., B.A.	1978
Tulane University	Comp Sci	M.S.	1982
University of Texas Arlington	Comp Sci	Ph.D.	1989

Appointments:

2007-	Associate Professor, Bradley University
2002-07	Assistant Professor, Bradley University
1989-93	Assistant Professor, University of Wisconsin Parkside

Three relevant publications:

1. Dolins, S.B. & Kero, R.E. Data management challenges for U.S. healthcare providers. 2006 Information Resources Management Association International Conference – Emerging Trends and Challenges in Information Technology Management, Washington D.C., 2006.
2. Dolins, S.B. Using the balanced scorecard process to compute the value of software applications. 28th International Conference on Software Engineering, Shanghai, China, 2006.
3. Silver, M., Sakata, T., Su, H., Herman, C., Dolins, S., and O' Shea, M. Case Study: How to apply data mining techniques in a healthcare data warehouse. Journal of Healthcare Information Management, 15, 2, (2001).

Five additional publications:

1. Dolins, S.B. Analyzing manufacturing processes to determine the placement of diagnostic systems. IEEE Transactions on Components, Hybrids, and Manufacturing Technology, 15, 6, (1992).
2. Dolins, S.B. & Reese, J.D. A curve interpretation and diagnostic technique for industrial processes. IEEE Transactions on Industrial Applications, 28, 1, (1992).
3. Kline, P. & Dolins, S.B. Designing Expert Systems: A Guide to Selecting Implementation Techniques. New York, John Wiley and Sons, July 1989.
4. Dolins, S.B., Srivastava, A., and Flinchbaugh, B. Monitoring and diagnosis of plasma etch processes. IEEE Transactions on Semiconductor Manufacturing, 1, 1, (1988).
5. Kline, P. & Dolins, S.B. Problem features that influence the design of expert systems. Proceedings of the National Conference on Artificial Intelligence (AAAI-86) Philadelphia, PA., 1986.

Synergistic Activities:

1. Filed four patents (two of the patents are related to data warehousing): 1) Improved Process and Apparatus for Detecting Aberrations in Process Operations – U.S. Patent Application, TI-12380A, July 1988 (filed by Texas Instruments, Inc.), 2) Production Process Diagnosis Using Dynamic Time Warping – U.S. Patent Application, TI-12851A, August 1988 (filed by Texas Instruments, Inc.), 3) Rule-based Customization of a Hierarchy in a Data Warehouse (filed by ACNielsen), and 4) Analysis of Massive Data Accumulations Using Patient Rule Induction Method and On-Line Analytical Processing, Patent Number 6,643,646, November 4th, 2003 (filed by Hitachi America, Inc.).
2. Senior Systems Architect for ACNielsen from 1990-1999. Managed one enterprise-wide data warehouse project and designed and built a database for a CRM application.
3. Chief of Technical Development for Hitachi America, Inc. from 1999-2001. Led a team building a data warehouse and business intelligence tools for healthcare providers.
4. Collaborating with the Smithsonian Tropical Research Institute on building a database for storing their plot data starting in 2006. Presented the logical data model and prototype reports at their workshop in December 2006 in Panama.

Biographical Sketch – Brad Boyle

ADDRESS

Dept. of Ecology and Evolutionary Biology
University of Arizona
PO Box 210088
Tucson, AZ 85721-0088
Tel: 520-626-3336
E-mail: bboyle@email.arizona.edu

EDUCATION

December 1996. Ph.D., Biology. Washington University, St. Louis.
May 1986. B.Sc., Zoology. University of British Columbia.

PROFESSIONAL APPOINTMENTS

Research Associate, Dept. of Ecology and Evolutionary Biology, University of Arizona. Sep. 2007 to present. Biodiversity Informatics.

Database Developer and Collections Manager, University of Arizona Herbarium. Jan.-Sep. 2007.

Research Associate, Dept. of Ecology and Evolutionary Biology, University of Arizona. 2001 to 2006. Biodiversity Informatics.

Sessional Course Coordinator, Organization for Tropical Studies, Costa Rica. 1998 to present. Tropical Biology and Tropical Plant Systematics.

Sessional Instructor, Washington University, St. Louis. 1997.

PUBLICATIONS

Boyle, B., Meyer, H.W., Enquist, B. and Salas, S., 2008. Higher taxa as paleoecological and paleoclimatic indicators: A search for the modern analog of the Florissant fossil flora. Special Paper 435: Paleontology of the Upper Eocene Florissant Formation, Colorado: 33-51.

Weiser, M.D, Enquist, B.J., **Boyle, B.**, Killeen, T.J., Jorgensen, P.M., Fonseca, G., Jennings, M.D., Kerkhoff, A.J., Larcher, T.E., Monteagudo, A., Nunez Vargus, M.P., Phillips, O.L., Swenson, N.G., and R. Vasquez Martinez. 2007. Latitudinal patterns of range size and species richness of New World woody plants. Global Ecology and Biogeography, 16: 679-688.

Boyle, B. 2001. Vegetation of two sites in the northern Cordillera Vilcabamba. Pp. 69-79 and 182-216 in: L.E. Alonso, T.S. Schulenberg and F. Dallmeier (eds.), Biological and Social Assessments of the Cordillera Vilcabamba, Peru. RAP Working Papers No.12 / SI/MAB Series 6. Conservation International, Washington, DC.

BIOINFORMATICS SOFTWARE DEVELOPMENT

University of Arizona Biodiversitu Informatics Initiative (BDII). Sept. 2009-present. http://loco.biosci.arizona.edu/bdii/

The University of Arizona Herbarium Website and Collections Database. Jan. 2007 to present. http://loco.biosci.arizona.edu/herbarium/

TaxonScubber: an application for the automated correction and standardization of taxonomic names. 2002 and onward, http://www.salvias.net/pages/taxonscrubber.html.

The SALVIAS Project. March 2001 to present. With B. J. Enquist and M. Weiser. www.salvias.net.

Collections Database & Website, Flora of Oaxaca Project, Sociedad para el Estudio de los Recursos Bióticos de Oaxaca (SERBO, A.C.) , Oaxaca, Mexico.. 2001. http://serboax.org/

SYNERGISTIC ACTIVITIES

Botanical surveys in North, Central and South America, 1989 to present. Examples: Santa Catalina elevational gradient (Arizona), Los Fresnos Desert Grassland Preserve (Sonora, Mexico), El Cielo Biosphere Reserve (Tamaulipas, Mexico), Sierra Juárez (Oaxaca, Mexico), Braulio Carrillo & Guanacaste National Parks (Costa Rica), Cerro Golondrinas (Carchi Province, Ecuador), Cordillera Vilcabamba (Peru). Research supported by Missouri Botanical Garden, NSF, Conservation International, The Nature Conservancy, National Geographic Society.

Coordinator and instructor, Organization for Tropical Studies course "Tropical Plant Systematics", Costa Rica. 1998 to present. In Spanish and English. Over 200 Ph.D and Masters students from 17 countries.

GRANTS AND FELLOWSHIPS

University of Arizona Biodiversity Travel Grant. January 2008.
National Parks Service Research Award, 2007
National Parks Service Research Award, 2003-2004
National Geographic Society Research Grant, 1994.
Graduate Fellowship, Washington University, 1989-1996

GRADUATE AVISORS AND POSTDOCTORAL SPONSORS

Michael Sanderson (Postdoc)
Brian Enquist (Postdoc)
Alwyn H. Gentry (Ph.D)

Biographical Sketch – Peter Møller Jørgensen

Associate Curator, Missouri Botanical Garden, St. Louis, USA,

Aarhus Universitet, Denmark	Botany	Cand. scient. (=MS) in botany, 1985.
Aarhus Universitet, Denmark	Botany	Ph.D. in sciences, 1993.
Missouri Botanical Graden	Botany	Post Doctoral 1992–1998

APPOINTMENTS

2004–p.p. Missouri Botanical Garden, St. Louis, USA, Associate Curator.

RELEVANT PUBLICATIONS

- Pitman, N.C.A & **P.M. Jørgensen**. 2002. Estimating the Size of the World's Threatened Flora. — Science 298: 989.
- Pitman, N.C.A., **P.M. Jørgensen**, R.S.R. Williams, S. León-Yánez & R. Valencia. 2002. Extinction-Rate Estimates for a modern Neotropical Flora. Conservation Biology 16(5): 1427–1431.
- **Jørgensen, P. M.** & S. León-Yánez (eds.) 1999. Catalogue of the Vascular plants of Ecuador. — Monograph. Syst. Bot. Missouri Bot. Gard. 75: i–vii, 1–1181. http://www.mobot.org/MOBOT/research/ecuador/welcome.htm
- **Jørgensen, P. M.** & C. Ulloa Ulloa, R. Valencia R., & J.E. Madsen. 1995. A Floristic Analysis of the High Andes of Ecuador. pp. 221–237. In: S.P. Churchill, H. Balslev, E. Forero & J.L. Luteyn, (eds.) Biodiversity and Conservation of the Neotropical Montane Forests. — The New York Botanical Garden, New York.
- Ulloa Ulloa, C. & **P.M. Jørgensen**, 1993. Arboles y arbustos de los Andes del Ecuador. — AAU Rep. 30: 1–264.

OTHER SIGNIFICANT PUBLICATIONS

- **Jørgensen, P.M.** & C. Ulloa Ulloa. 2002. 30B. Olacaceae. In: G. Harling & L. Andersson, (eds.) Flora of Ecuador. 69: 59–104.
- Picker, T. & **P.M. Jørgensen** 2001. Epiphyte diversity in Ecuador – a Geographical Information System (GIS) project. In: J. Nieder & W. Barthlott (eds.) Epiphytes and Canopy Fauna of the Otonga Rain Forest (Ecuador). J. Nieder & W. Barthlott. Bonn.
- **Jørgensen, P. M.** & C. Ulloa Ulloa. 1994. Seed plants of the high Andes of Ecuador – a checklist. — AAU Rep. 34: 1–460.
- Valencia R., R. & **P. M. Jørgensen**, 1992. Composition and structure of a humid montane forest on Pasochoa, Ecuador. — Nord. J. Bot. 12: 239–247.
- Holm-Nielsen, L. B., **P. M. Jørgensen**, & J. E. Lawesson, 1988. 126. Passifloraceae. In: Flora of Ecuador (G. Harling & L. Andersson, eds.) 31: 1–130.

SYNERGISTIC ACTIVITIES

- Senior editor of the catalogue of the vascular plants of ecuador (jørgensen & león-yanéz, 1999; http://www.mobot.org/mobot/research/ecuador/welcome.htm) and of the upcoming catalogo de las plantas vasculares de bolivia. both projects brought together about 250 contributing taxonomic specialists.
- Contribution editor of a red list of the Ecuador endemic species using iucn standards (valencia et al. 2000; http://www.puce.edu.ec/herbario/publirojo.htm) and member of the iucn recognized expert group evaluation Ecuadorian plants.
- Founding partner in the andes biodiversity consortium (http://www.andesbiodiversity.org/abc); promoting multidisciplinary research in the tropical andes with the goal of conservation of biodiversity. active participant in the spatial analysis of local vegetation inventories across scales (salvias; http://www.salvias.net) a portal for managing plant plot data.
- Teacher and mentor for 22 students from belize, ecuador, mexico, usa, and bolivia. were involved, for three, years in building the largest herbarium in ecuador. most of my students are still involved with biology from high school teaching to controlling international trade in plants and animals to managing the distribution of research fund in biology in ecuador or as junior officers in international conservation organizations.

COLLABORATORS & OTHER AFFILIATIONS

Collaborators and Co-Editors: A. Acebey (LPB), P. Acevedo (US), C. Aedo (MA), I. Al-Shehbaz (MO), W. Alverson (F), L. Andersson (GB), C. Antezana (BOLV), A. Arbelaez (USZ), M. Arbo (CTES), S. Arrázola (BOLV), L. Arroyo (USZ), M. Atahuachi (BOLV), D. Atha (NY), G. Aymard (PORT), N. Bacigalupo (SI), V. Badillo (MY), H. Ballard (BHO), E. Balslev (AAU), H. Balslev (AAU), V.E. Barney, G. Barriera (G), S. Beck (LPB), C. Berg (BG), P. Berry (WIS), R. Bianchini (SP), C. Bonifaz (GUAY), F. Borchsenius (AAU), T. Borsch (BONN), B. Boyle, R.J. Burnham, J. Caballero (USZ), E. Cabral (CTES), R. Callejas (HUA), K. Camelbeke (GENT), J. Cardiel (MA), A. Carretero (HSB),

L. Chatrou (U), E. Chávez (USZ), T. Cochrane (WIS), R. Condit, T. Consiglio (MO), G. Coppens d'Eeckenbrugge, T. Croat (MO), L. Cusato (BAA), D. Daly (NY), A. Davis (K), B. Deghan (FLAS), M. Dillon (F), C. Dodson (QCNE), S. Dressler (FR), R. Duno (VEN), J. Dutilh, E. Emschwiller (F), B.J. Enquist, R. Eriksson (GB), E. Fernandez (BOLV), J. Fernández Casas (MA), M. Ferrucci (CTES), G. Fonseca (CI), E. Forero (NY), A. Freire (MO), I. Friis (C), P. Fritsch (CAS), P. Fryxell (TEX), A. Fuentes (MA), I. Galarza (LPB), E. Garcia (LPB), M. Garvizu (USZ), D. Geltman (LE), R. Gereau (MO), P. Goetghebeur (GENT), F. Gonzalez (COL), J. Gonzalez (LPB), S. Graham (MO), E. Guaglianone (SI), J.E. Guevara, R. Guillén (USZ), R. Harley (K), G. Harling (GB), J. Hedin (Ariz), N. Hensold (F), P. Hiepko (B), B. Holst (SEL), H. Huaylla (HSB), P. Ibisch (FAN), H. Iltis (WIS), A. Jardim (USZ), M. Jennings, C. Jordan (USZ), J. Justiniano (USZ), H. Kennedy (UBC), A.J. Kerkhoff, M. Kessler (GOET), R. Kiesling (SI), T. Killeen (CI), R. Kral (BRIT), J. Kress (US), T. Kroemer (GOET), J. Kuijt (UVIC), S. Lægaard (AAU), T. Lacher, T. Lammers (OSH), S. Landivar R. (USZ), E. Landolt (ZT), B. León (USM), S. León-Yánez (QCA), M. Liden (GB), R. Liesner (MO), L. Lohmann (MO), P. Loizeau (G), J. Lombardi (BHCB), R. Lopez (LPB), R. Lundin (S), J. Luteyn (NY), H. Luther (SEL), P. Maas (U), M. Macía (MA), F. Mamani (USZ), J.M. MacDougal, A. Meerow (FLAS), M. Menacho (USZ), M. Mendoza (USZ), R. Meneses (LPB), M. Mercado (BOLV), J. Mickel (NY), J. Miller (MO), J. Mitchell (NY), H. Mogollón, A. Monro (BM), J. Montero (USZ), L. Mora (COL), M. Moraes (LPB), F. Morales (INBIO), R. Moran (NY), S. Mori (NY), M. Múlgura (SI), M. Nee (NY), M. Negrito (CORD), D. Neill (QCNE), B. Nordenstam (S), E. Norman (DLF), P. Nuñez-Vargas, B. Øllgaard (AAU), C. Orozco (COL), R. Ortiz Gentry (MO), T. Ortuño (LPB), N. Paniagua (LPB), C. Parra-Osorio (COL), C. Paz (LPB), O.L. Phillips, T. Picker (BONN), J. Pirani (SPF), M. Pirie (U), J. Pither, N.C.A. Pitman (ACA), D. Porter (VPI), A. Portugal (LPB), G. Prance (RNG), L. Prather (MSC), J. Pringle (HAM), A. Quevedo (USZ), L. Rea (LPB), S. Renner (MO), S. Renvoize (K), C. Reynel (MOL), A. Reznicek (MICH), N. Ritter (NHA), H. Robinson (US), Y. Roca (USZ), Z. Rodgers (MO), P. Rojas (BOLV), K. Romoleroux (QCA), Z. Rúgolo (SI), T. Ruiz (USZ), V. Salazar (USZ), M. Saldias (USZ), L. Sanchez (MO), C. Sastre (P), J. Schneider (FR), R. Seidel (LPB), M. Serrano (HSB), S. Simmons (TEX), B. Simpson (TEX), D. Simpson, D. (K), P. Sklenar (PRG), L. Skog (US), N. Snow (GREE), D.D.A. Smith (WIS), C. Stace (LTR), B. Stahl (GB), W. Stevens (MO), D. Stevenson (NY), L. Struwe (NY), C. Taylor (MO), W. Thomas (NY), M. Toledo (USZ), G. Torrico (BOLV), R. Tortosa (BAA), A. Tye, C. Ulloa Ulloa (MO), E. Valenzuela (LPB), R. Valencia (QCA), H. van der Werff (MO), E. Vargas (LPB), I. Vargas (USZ, FAN), J.H. Vargas (QCNE), R. Vasquez (MO), R. Vasquez (USZ), G. Villa, X. Villavicencio (LPB), B. Wallnöfer (W), D. Wasshausen (US), F. Weberling (ULM), G. Webster (DAV), M. Weigend (NY), M.D. Weiser, A. Weitzman (US), G. Wheeler (MIN), J. Wiersema (BARC), R.S.R. Williams, J. Wood (LPB), C. Xifreda (SI), M. Zarate B. (BOLV), E. Zardini (MO), M. Zeballos (LPB).

Graduate and Postgraduate Advisors: Lauritz B. Holm-Nielsen, Aarhus University (now at the World Bank); Henrik Balslev, Aarhus University; Henk van der Werff, Missouri Botanical Garden.

Thesis Advisor and Postgraduate-Scholar Sponsor:

Biographical Sketch - Hans ter Steege

Address: Institute of Environmental Ecology, Plant Ecology and Biodiversity

Sorbonnelaan 14, 3584 CA Utrecht, the Netherlands

Email: h.tersteege@uu.nl

Web: http://www3.bio.uu.nl/peb/staff/personal/htersteege/htersteege.html

Education: 1993 Utrecht University; PhD Tropical Ecology

1987 Utrecht University; MSc Tropical Ecology

Employment

2004 – pr. Utrecht University, Associate Professor 'Plant Ecology and Biodiversity'

2002-04 Utrecht University, Assistant Professor 'Biodiversity Assessment'

2001-02 International Institute for Geo-information Science and Earth Observation (ITC). Associate Professor 'Spatial Analysis of Forest Conservation and Rehabilitation'.

2001 Consultant for NRI, UK. Preparation of vegetation maps of Guyana. Database support for Guyana Forestry Commission Herbarium.

1998-2000 Utrecht University, Post-doc.

1994-98 Tropenbos-Guyana Programme, Programme Team-leader.

1992-93 PhD student, Utrecht University

1989-1992 Guyana, Tropenbos-Guyana Programme, Project Leader

1987-89 Guyana, Utrecht University/University of Guyana, 'Forest Project Mabura Hill', Coordinating Scientist.

Five relevant publications:

Stropp, J., **ter Steege, H.**, Malhi, Y., ATDN, RAINFOR. (accepted). Disentangling regional and local tree diversity in the Amazon. *Ecography*.

Hubbell, S.P., He, F., Condit, R., Borda-de-Água, L., Kellner, J. & **ter Steege, H.** (2008) How Many Tree Species Are There in the Amazon and How Many of Them Will Go Extinct? *PNAS* 105: 11498-11504.

ter Steege, H., Pitman, N.C.A., Phillips, O.L., Chave, J., Sabatier, D., Duque, A., Molino, J.-F., Prévost, M.-F., Spichiger, R., Castellanos, C., von Hildebrand, P., & Vasquez, R. (2006). Continental-scale patterns of canopy tree composition and function across Amazonia. *Nature* 443: 444-447.

ter Steege, H., Pitman , N.C.A., Sabatier, S., Castellanos, H., van der Hout, P., Daly, D.C., Silveira, M., Phillips, O., Vasquez, R. van Andel, T., Duivenvoorden, J., de Oliveira, A.A., Ek, R.C., Lilwah, R., Thomas, R.A., van Essen, J., Baider, C., Maas, P.J.M., Mori, S.A., Terborgh J., Nuñez-Vargas, P Mogollón, H. & Morawetz, W. (2003). A spatial model of tree α-diversity and -density for the Amazon Region. *Biodiversity and Conservation* 12: 2255-2276.

ter Steege, H. & Hammond, D.S. (2001). Character convergence, diversity, and disturbance in tropical rain forest in Guyana. *Ecology* 82: 3197-3212.

Five additional publications:

Phillips, O.L., Aragão, L., Fisher, J.B., Lewis, S.L., Lloyd, J., López-González, G., Malhi, Y,. Monteagudo, A., Peacock, J., Quesada, C., van der Heijden, G., Almeida, S., Amaral, I., Arroyo, L., Aymard, G., Baker, T.R., Bánki, O.S., Blanc, L., Bonal, D., Brando, P., Chave, J., de Oliveira, A.C., Dávila Cardozo, N., Espejo, J., Feldpausch, T., Aparecida Freitas, M., Higuchi, N., Jiménez, E., Meir, P., Mendoza B. C., Morel, A., Neill, D., Nepstad, D., Patiño, S., Peñuela, M.C., Prieto, A., Ramírez, F., Schwarz, M., Silveira, M., Sota, A., **ter Steege, H.**, Stropp, J., Vásquez, R., Zelazowski, P., Alvarez Dávila, E., Andelman, S., Erwin, T., di Fiore, A., Chao, K.-J., Honorio, E., Keeling, H., Killeen, T., Laurance, W., Nascimento, H., Peña Cruz, A., Pitman, N., Núñez Vargas, P., Ramírez, H., Rudas, A., Salamão, R., Silva, N., Terborgh, J. & Torres, A. (in press). Drought sensitivity of the Amazon rainforest. *Science*.

Raes, N. & **ter Steege, H.** (2007). A null-model for significance testing of presence-only species distribution models. *Ecography* 30: 727-736.

Chave, J. Muller-Landau, H.C., Baker, T.R., Easdale, T., **ter Steege, H.** & Webb, C.O. (2006). Regional and phylogenetic variation of wood density across 2456 neotropical tree species. *Ecological Applications* 16: 2356-2367.

ter Steege, H., Jansen-Jacobs, M., & Datadin, V. (2000). Can botanical collections assist in a National Protected Area Strategy in Guyana? *Biodiversity and Conservation* 9: 215-240.

ter Steege, H. (1998). The use of forest inventory data for a National Protected Area Strategy in Guyana. *Biodiversity and Conservation* 7: 1457-1483.

Synergistic Activities:

1. Established & Coordinated, Amazon Tree Diversity Network. 2000 - present. A loose e-network of c. 55 scientist working across the Amazon. http://www.bio.uu.nl/~herba/Guyana/ATDN/

2. Participant and (co-)organizer of workshop in biodiversity assessment and monitoring in Venezuela, Guyana, Suriname, Brazil, Bolivia, in the Framework of Tropenbos and Panamazonia.

3. While programme team-leader for Tropenbos-Guyana, responsible for programme development (d. 55 staff at max), monitoring and management, supervision of research, integration and dissemination of results to policy and management.

Recent Collaborators & Coauthors: see papers

PhD Students (11) and Postdoctoral Scholars (1)

Biographical Sketch - Sandy J. Andelman

Vice President and Director
Tropical Ecology, Assessment and Monitoring Network
Conservation International
2011 Crystal Drive, Suite 500
Arlington, Virginia 22202
Tel 703-341-2690 Email: sandelman@conservation.org

PROFESSIONAL PREPARATION
Lewis and Clark College, Biology, B.A.,1977.
University of Washington, Behavioral Ecology, Ph.D., 1985.

APPOINTMENTS
- Vice President, TEAM Network, Conservation International (2006-present)
- Senior Director, TEAM Network, Conservation International (2005-2006)
- Editor, Biological Conservation (2004-2006)
- Deputy Director, National Center for Ecological Analysis & Synthesis, University of California, Santa Barbara (1999 – 2005).
- Adjunct Professor, Department of Ecology, Evolution & Marine Biology, University of California, Santa Barbara (2003 – present).
- Adjunct Professor, Bren School of Environmental Science & Management, University of California, Santa Barbara (2000 – 2002).
- Research Associate, National Center for Ecological Analysis & Synthesis, University of California, Santa Barbara (1998).
- Visiting Professor, Bren School of Environmental Science & Management, University of California, Santa Barbara (1997 – 1998).
- Director of Conservation Science, The Nature Conservancy of Washington (1994 – 1997).
- Research Assistant Professor, Institute for Environmental Studies, University of Washington (1990 – 1994).

FIVE RECENT PUBLICATIONS (from more than 50)
- DeFries, R., F. Rovero, P. Wright, J. Ahumada, S. Andelman, K. Brandon, J. Dempewolf, A. Hansen, J. Hewson, and J. Liu. Linking plot-level biodiversity measurements with human influences over multiple spatial scales in the tropics: A conceptual framework. Frontiers in Ecology and the Environment, in press.
- Pyke, C.R. and S.J. Andelman. 2007. Land use and land cover tools for climate adaptation. Climatic Change 80:239-251.
- Andelman, S.J., C.M. Bowles, R. Waide and M.R. Willig. 2004. Disentangling biocomplexity through knowledge networking. BioScience 54:240-246.
- Andelman, S.J. and M.R. Willig. 2004. Networks by design: A revolution in ecology. Science 305: 1565-1567.
- Rodrigues, A.L., S.J. Andelman, M.I. Bakarr, L. Boitani, T.M. Brooks, R.M. Cowling, L.D.C. Fishpool, da Fonseca, K.J. Gaston, M. Hoffmann, J. Long, P.A. Marquet, J.D. Pilgrim, R.L. Pressey, J. Schipper, W. Sechrest, S. N. Stuart, L.G. Underhill, R.W. Waller, M.E.J. Watts and X. Yan. 2004. Effectiveness of the global protected area network in representing species diversity. Nature 428: 640-643.

FIVE OTHER RELATED PUBLICATIONS
- Wilson, K.A., E.C. Underwood, S.A. Morrison, K.R. Klausmeyer, W.W. Murdoch, B. Reyers, G. Wardell-Johnson, P.A. Marquet, P.W. Rundel, M.F. McBride, R.L. Pressey, M. Bode, J. Hoekstra, S. J. Andelman, M. Looker, C. Rondini, P. Kareiva, M.R. Shaw and H.P. Possingham. 2007. Conserving biodiversity efficiently: What to do where and when. PLOS Biology, 5(9):e223.
- Grace, J.B., T.M. Anderson, M.D. Smith, E. Seabloom, S.J. Andelman, G. Meche, E. Weiher, L.K. Allain, H. Jutila, M. Sankaran, J. Knops, M. Ritchie and M.R. Willig. 2007. Implications of a multiprocess view of Pyke, C.R., S.J. Andelman and G. Midgley. 2005. Identifying priority areas for bioclimatic representation under climate change: A case study for Proteaceae in the Cape

Floristic Region, South Africa. Biological Conservation, 125:1-9.
- Meir, E., S.J. Andelman and H.P. Possingham. 2004. Does conservation planning matter in a dynamic and uncertain world. Ecology Letters 7:615-622.
- Rodrigues, A.L., H.R. Akcakaya, S.J. Andelman, M.I. Bakarr, L. Boitani, T.M. Brooks, J.S. Chanson, L.D.C. Fishpool, G. da Fonseca, K.J. Gaston, M. Hoffmann, P.A. Marquet, J.D. Pilgrim, R.L. Pressey, J. Schipper, W. Sechrest, S. N. Stuart, L.G. Underhill, R.W. Waller, M.E.J. Watts and X. Yan. 2004. Global gap analysis: Priority regions for expanding the global protected-area network. BioScience 54:1092-1100.
- Andelman, S.J. and M.R. Willig. 2003. Present patterns and future prospects for biodiversity in the Western Hemisphere. Ecology Letters 6:818-824.

SYNERGISTIC ACTIVITIES
- Design and implementation of a global network of research sites to monitor biodiversity in tropical forests (2005-present).
- PI, Biodiversity conservation in dynamic landscapes network. National Science Foundation Research Coordination Network (NSF) (2005-2010).
- Panel Member, National Science Foundation Panel for NEON Preliminary Design Review 2007.
- Science Advisory Board Member, Department of Defense Strategic Environmental Research and
- Development Program (2008-present).
- Deputy Director, National Center for Ecological Analysis & Synthesis, University of
- California, Santa Barbara (1999 – 2005).
- CO-PI, KDI: A knowledge network for biocomplexity: Building and evaluating a metadata-based framework for integrating heterogeneous ecological data (NSF) (1999 – 2004).
- Developed (w/P. Kareiva) a pedagogical model for new national graduate seminar used for three NCEAS projects, involving 40 universities and over 400 graduate students (1997 – 2005).

COLLABORATORS
Harry Biggs, South African National Parks; Michael Bode, University of Melbourne; Richard Cowling, Nelson Mandela University; Matthew Jones, NCEAS; Judith Kruger, South African National Parks; Steve Polasky, University of Minnesota; Hugh Possingham, University of Queensland; Jim Reichman, NCEAS; Mark Schildhauer, NCEAS; Robert Waide, University of New Mexico; Michael Willig, University of Connecticut; Kerrie Wilson, The Nature Conservancy.

GRADUATE AND POSTDOCTORAL STUDENTS
David Chalcraft, East Carolina University; Stephen Cox, Texas Tech University; Cory Craig; Eric Fegraus, Conservation International; Paola Gomez-Priego; Michael McCarthy, University of Melbourne; Christopher Pyke; Helen Regan, University of California, Riverside; Mellyn Reuling; Elizabeth Sandlin; Melinda Smith, Yale University; John Williams, University of Wisconsin.

GRADUATE ADVISORS
Ph.D. Advisors: Michael Beecher, University of Washington; Gordon Orians, University of Washington

Biographical Sketch - Oliver Phillips

Chair in Tropical Ecology, University of Leeds

Contact: o.phillips@leeds.ac.uk

Undergraduate institution: University of Cambridge (1987).
Graduate institution: Washington University (Ph.D., 1993).
Postdoctoral institutions: Missouri Botanical Garden, University of Leeds.

Appointments:

2006-	University of Leeds, School of Geography, Chair in Tropical Ecology
2003-2006	University of Leeds, School of Geography, Reader in Tropical Ecology
1999-2003	University of Leeds, School of Geography, Lecturer
1996-9	Natural Environment Research Council Fellow, School of Geography, University of Leeds. Remote sensing, GIS, tropical forest floristic inventory and dynamics.
1995-6	Research Fellow in Biodiversity, School of Geography, University of Leeds.
1994-5	Project Coordinator for the Missouri Botanical Garden, U.S.A.: Gentry tropical forest diversity project.

Five related publications:

Butt, N., Malhi, Y., Phillips, O.L, New, M. 2008. Floristic & functional affiliations of woody plants with climate in Western Amazonia. *Journal of Biogeography* 35:939–50.

Phillips, O.L., et al. 2008. The changing Amazon forest. *Philosophical Transactions of the Royal Society, Ser.B.* 363: 1819-1828. DOI: 10.1098/rstb.2007.0026.

Pitman, N.C.A., Mogollón, H., Dávila, N., Ríos, M., García-Villacorta, R., Guevara, J., Baker, T.R., Monteagudo, A., Phillips, O.L., Ahuite, M., Aulestia, M., Cardenas, D., Cerón, C., Neill, D.A., Núñez, P., Palacios, W., Spichiger, R., Valderrama, E., & R. Vásquez. Tree community change across 700 km of lowland Amazonian forest from the Andean foothills to Brazil. 2008. *Biotropica* 40, 525-535.

Parmentier, I., Y. Malhi, B. Senterre, R.J. Whittaker, A.T.D.N., A. Alonso, M.P.B. Balinga, A. Bakayoko, F. Bongers, C. Chatelain, J. Comiskey, R. Corta, M.-N. Djuikouo Kamdem, J.-L. Doucet, L. Gautier, W.D. Hawthorne, Y.A. Issembe, F.N. Kouamé, L.A. Kouka, M.E. Leal, J. Lejoly, S. Lewis, D. Newbery, L. Nusbaumer, M. Parren, K. S.-H. Peh, O.L. Phillips, L. Poorter, D. Sheil, B. Sonké, M. Sosef, T. Sunderland, J. Stropp, H. ter Steege, M. D. Swaine, M. Tchouto, B. van Gemerden, J. L. van Valkenburg, H. Wöll. 2007. The odd man out? Might climate explain the lower tree α-diversity of African rain forests relative to Amazonian rain forests? *Journal of Ecology* 95: 1058–71.

Phillips, O.L., S. Rose, A. Monteagudo, P. Núñez Vargas. 2007. Resilience of south-western Amazon forests to anthropogenic edge effects. *Conservation Biology* 20:1698–1710.

Five other publications:

Phillips, O.L., et al. Drought sensitivity of the Amazon rainforest. 2009 *Science* (in press).

ter Steege H., Pitman N., Phillips O., Chave J., Sabatier D., Duque A., Molino J., Prévost M, Spichiger R, Castellanos H, van Hildebrand P & Vásquez R. 2006. Continental- scale patterns of canopy tree composition and function across Amazonia. *Nature* 443:444-7.

Bunker, D., De Clerck, F., Bradford, J., Colwell, R., Garden, P., Perfecto, I., Phillips, O.L., Sankaran, M., & Naeem S. 2005. Carbon sequestration and biodiversity loss in a tropical forest. *Science* 310: 1029-1103.

Phillips, O.L., et al. 2005. Large lianas as hyperdynamic elements of the tropical forest canopy. *Ecology* 86:1250-1258.

Phillips, O.L., et al.. 2004. Pattern and process in Amazon forest dynamics, 1976-2001. *Philosophical Transactions of the Royal Society, Series B* 359: 381-407.

Synergistic activities:

Science leadership: RAINFOR (Amazon Forest Inventory Network), with Y.Malhi and J.Lloyd
Development of databases: creating a database of forest plots for the tropics

Broadening the participation of groups under-represented in science: collaboration with and training of botanists and ecologists from developing countries (especially Peru, Bolivia, Colombia, Brasil, Ecuador, Ghana).

Collaborators and affiliations**: >200 co-authors in the past 48 months, most frequently:**

Almeida, S.	Museu Emilio Goeldi, Belem, Brasil		Honorio, E.	Instituto de Investigaciones de la Amazonia Peruana, Iquitos
Arroyo, L.	Museo Noel Kempff, Bolivia		Davila, N.	Universidad Nacional Amazonia Peruana, Iquitos
Baker, T	University of Leeds, UK		Quesada, C.	INPA, Brasil
Chao, K.	University of Leeds, UK		Pena Cruz, A	Jardin Botanico Missouri, Peru
Chave, J.	CNRS, Toulouse, France		Ramirez, H.	Universidad de los Andes, Merida, Venezuela
Higuchi, N.	INPA, Brasil		Rudas, A.	Universidad Nacional Colombia, Bogotá
Killeen, T.	Conservation International, Bolivia		Swaine, M	University of Aberdeen, UK
Lewis, S.	University of Leeds, UK		Stropp, J.	Utrecht University, Netherlands
Lopez-Gonzalez, G.	University of Leeds, UK		Anderson, LO	University of Oxford, UK
Malhi, Y.	University of Oxford, UK		Laurance, W.F.	Smithsonian, Panama
Monteagudo, A.	Jardin Botanico Missouri, Peru		Prieto, A.	Instituto von Humboldt, Boyaca, Colombia
Neill, D.	Missouri Botanical Garden, Ecuador		Alvarez, E.	Universidad Nacional Colombia, Medellin
Peacock, J.	University of Leeds, UK		Patino, S.	Universidad Nacional Colombia, Leticia
Pitman, N.	Los Amigos, Peru		Gloor, E	University of Leeds, UK
Silva, N.	CIFOR, Brasil		Aragao, LEOC.	University of Oxford, UK
ter Steege, H.	Utrecht University, Netherlands		Torres-Lezama, A.	Universidad de los Andes, Merida, Venezuela
Nunez Vargas, P.	Universidad San Antonio Abad del Cusco, Peru		R. Vásquez	Jardin Botanico Missouri, Peru
van der Heijden, G.	University of Leeds, UK		Mendoza, A	Jardin Botanico Missouri, Peru

Graduate advisors: A.H. Gentry (Missouri Botanical Garden), W.L. Lewis (Washington University), P.H. Raven (Missouri Botanical Garden)

Students and post-docs advised since 2000: 13 PhD students, and 9 PDRAs.

Appendix D.

Letters of support from a sample of BIEN collaborators who represent various networks of data providers, institutions, and computational initiatives.

Dr. Richard Condit
Global Forest Observatory Network
Smithsonian Tropical Research Institute

GBIF

Global Biodiversity Information Facility Secretariat

Universitetsparken 15
DK-2100 Copenhagen Ø
Denmark

Tel.: +45 35 32 14 71
Fax: +45 35 32 14 80
Email: nking@gbif.org
Web: www.gbif.org

04 February 2009

Dear Dr Condit,

GBIF SUPPORT FOR BIEN PROJECT

This letter serves to confirm our support for the Botanical Information and Ecology Network (BIEN) submission to iPLANT.

The Global Biodiversity Information Facility (GBIF) is a multi-lateral initiative set-up to provide free and open access to biodiversity data worldwide via the Internet, to facilitate scientific research. Priorities, with an emphasis on promoting participation and working through partners, include promoting in-country activities to mobilize biodiversity data, developing the protocols and standards to ensure scientific integrity and interoperability, building an informatics architecture to allow the interlinking of diverse data types from disparate sources, promoting capacity building and catalyzing development of analytical tools for improved decision-making based on these data.

The BIEN project will develop and contribute an invaluable resource for a wide array of ecological change analyses. This will significantly enhance and complement the many other activities underway in the US and worldwide in biodiversity data publishing, discovery, access and analysis and contribute to our combined ability to respond to urgent global, regional, national and local issues such as climate change, food security, invasive species, and conservation of biodiversity.

The US is a founding member of GBIF and numerous US research institutions, government agencies etc. already make extensive use of the infrastructure, tools and services developed by GBIF as global standards – which is the reason for GBIFs establishment. As such, GBIF promotes, supports and encourages initiatives such as BIEN which can adopt rapidly and build upon the global standards etc developed by GBIF for these very purposes, and thus leverage the investments in GBIF by many countries, including the US.

GBIF requests iPLANT to give the BIEN project the full support proposed and we look forward to productive engagement and collaboration with the project team.

Yours Sincerely

Dr Nicholas King
Executive Secretary

UF UNIVERSITY of FLORIDA

Florida Museum of Natural History
Informatics, Department of Natural History

358 Dickinson Hall
PO Box 117800
Gainesville, FL 32611-7800
352-273-1980
rbeaman@ufl.edu

Dr. Brian Enquist
Dept of Ecology and Evolutionary Biology
University of Arizona
Tempe, AZ

Dear Brian,

I am writing in interest of your Grand Challenge proposal to iPlant, the Botanical Information and Ecology Network (BIEN). This is a highly desirable project that could address cyberinfrastucture support across domains of ecology, taxonomy, phylogenetic systematics, and biogeography.

As you may know, I have shared this common interest throughout my career, and am currently a PI and informatics lead on several NSF supported projects, including the Angiosperm AToL, the Liverwort AToL, the Gymnosperm AToL, the pPOD AToL and the Euphorbia PBI. All are higly relevant to CI development through iPlant. I was also founding developer for the BioGeomancer and HERBIS projects, both of which address high throughput and automated data capture for botanical information. I am currently serve on the TDWG (Biological Standards) Executive as one of the North American representatives and am convener of the Geospatial Interest Group for TDWG. I also have the dubious honor of serving as the lead PI on the final competing proposal (based at UF) for iPlant, and though unsuccessful, gave me a strong sense of CI and community requirements across a broad swath of the plant sciences.

It is through the network of my collaborations that I might best benefit the BIEN project and iPlant. Beyond my own participation, I can also assist in engaging other members of biodiversity, phylogenetics, and informatics research communities.

With best wishes,

Reed S. Beaman

Reed S. Beaman
Associate Curator, Informatics

Dave Vieglais
Biodversity Research Center
1345 Jayhawk Blvd
Lawrence, Kansas 66045
T 785 979 0530 F 785 864 5335
vieglais@ku.edu

February 2, 2009

Mark Schildhauer
NCEAS
735 State Street
Suite 300
Santa Barbara, California 93101

Dear Mark,

I am delighted to hear about the BIEN (Botanical Information and Ecology Network) proposal to iPlant, which is aimed at constructing a globally confederated information resource of plant specimen and occurrence data. The successful construction of such a cyberinfrastructure resource for the plant biodiversity and ecological community would not only provide an outstanding new service for a large number of ecological researchers, but it would enable plant biology researchers for the first time to address a number of critical large-scale questions, such as the impact of climate change on plant distributions, and the potential implications of such changes on global plant productivity.

I would enjoy being involved with several of the technology efforts of this project, and feel that with my professional background I could provide a positive contribution to the project. I have long been involved with the creation of protocols and standards for the exchange of natural history specimen and occurrence information including botanical and living plant collections (e.g. www.plantcollections.org). I am one of the principal authors of the Darwin Core standard with provides a set of well defined semantic terms for working with biodiversity in much the same way as the Dublin Core. I have also been PI, Co-PI or a lead in numerous projects whose charge was federation of specimen and observation data from a variety of taxonomic disciplines (e.g. Fishnet, ORNIS, MaNIS, HerpNet) and which required development of new data discovery and distribution protocols (e.g. DiGIR and TAPIR). I have also lead the architecture design for various species distribution projects such as LifeMapper (version 1 and 2) which required implementation of various spatial analyses and data handling tools operating in a globally distributed processing environment (LM 1) or on compute clusters (LM 2).

In summary I feel that my familiarity with biodiversity issues and computer science and technology would be beneficial to your project, and the proposed project is well aligned with my current and ongoing professional interests.

Sincerely yours,

Dave Vieglais

RAINFOR (Red Amazónica de Inventarios Forestales)
www.rainfor.org
Earth and Biosphere Institute
School of Geography
University of Leeds
Leeds LS2 9JT
U.K.

4 February 2009

Dear Brian,

On behalf of researchers involved in the RAINFOR project, I would like to express strong support of your iPlant proposal that will enable the creation of a Botanical Information and Ecology Network (BIEN). A network of core databases and data networks to create a novel resource for quantitative plant biodiversity science. The proposed Grand Challenge project will create a data resource of unprecedented size and scope together with the tools for its use, thereby empowering botanists to address fundamental issues in plant ecology and global change biology. Indeed, linking the treasure trove of geo-referenced biological record data with the plot samples of vegetation is one of the big challenges in biology, and the work proposed combining the most important groups working in both areas will help shed light on key patterns and processes across the planet. For example, by putting in place the necessary cyber-infrastructure this proposal will contribute substantially to our understanding of large-scale patterns of biodiversity across the world's most species-rich and threatened ecosystems.

The main focus of the projects coordinated by RAINFOR is to understand the variation in the structure, dynamics, composition and diversity of tropical forests across space and time and in the face of atmospheric changes. Our plot network links numerous ecological observations throughout Amazonia (www.rainfor.org), and involves more than 50 investigators world-wide. RAINFOR-associated studies incorporate extensive fieldwork to reveal the structure and processes of biodiversity and forest dynamics, and use this to describe spatial and environmental patterns, monitor change, and assess causality. As coordinator of the plot research program working with our partners in Europe, South America, and the United States, I am glad to help bring the RAINFOR plot network into your grand, collaborative effort.

Yours sincerely

Prof. Oliver Phillips
Chair in Tropical Ecology
University of Leeds

Universiteit Utrecht

Faculty of Science

Sorbonnelaan 16, 3584 CA Utrecht, The Netherlands

Institute of Environmental Biology
Plant Diversity and Biodiversity

Faxnr
+31 (0)30 251 8061
Telephone
+31 (0)30 253 6877
E-mail
h.tersteege@uu.nl

Dr. Brian Enquist
University of Arizona
Department of Ecology and Evolutionary Biology
BioSciences West,
Tucson, AZ 85721

Date
04/02/2009
Subject

Dear Brian,

I would like to express my support for the IPlant proposal "IBiodiversity Ecological Network and Information'. Within the Amazon Tree Diversity Network we have worked mainly on tree-alpha diversity and the proposed cyber infrastructure will no doubt allows us and many others to make better use of the specimen and plot composition data. While alpha-diversity patterns are relatively well described (but still less understood) patterns in composition and turn-over are far less well known, and therefore even less understood.

Linking museum data with plot data, generating range maps for the thousands of species of e.g. the Amazon will be a fantastic achievement, for which at this moment the digital support is simply not available.

I am very happy with this proposal and I am looking forward to collaborating in it with our European and American (north and south) partners.

Yours sincerely/,Met vriendelijke groeten,

Hans ter Steege

TROPICAL ECOLOGY
ASSESSMENT AND MONITORING

3 February, 2009

Dr. Brian Enquist
University of Arizona
Tucson, Arizona 85721

Dear Brian,

On behalf of the Tropical Ecology, Assessment and Monitoring Network, I am very pleased to participate in BIEN and efforts to develop an integrated botanical information network to investigate the impacts of global climate change on plant biodiversity. TEAM's mission is to generate real time data for monitoring long-term trends in tropical biodiversity through a global network of field stations, providing an early warning system on the status of biodiversity.

The question addressed by the proposed activities is compelling. We are particularly interested in the cyberinfrastructure challenges associated with understanding how dynamic processes such as climate change and land cover change are influencing patterns of plant distribution, abundance and community composition. Addressing these questions requires integration of plant data with environmental data (at a minimum, current climate; land cover; and projected future climate and projected future land cover), and currently there are no simple solutions to these problems because of the technical challenges of dealing with large volumes of data, multiple spatial and temporal scales, and differences in data resolution.

TEAM has invested substantial resources in cyberinfrastructure to capture, manage, synthesize and distribute global biodiversity, climate and land cover data, however we have barely scratched the surface. Given the scarcity of resources for plant ecology research and for plant conservation science, it makes sense to tackle these challenges as a community in a coordinated fashion.

Resolving the challenges outlined in the proposal will represent a major contribution, not only to fundamental plant science, but also to some of the earth's pressing environmental challenges.

We look forward to working with you on this.

Sincerely,

Sandy J. Andelman, Ph.D.
Vice President

The Center for Applied Biodiversity Science (CABS) at Conservation International
2011 Crystal Drive Suite 500 t 703.341.2400 www.teamnetwork.org
Arlington, VA 22202 f 703.979.2980 team@conservation.org

Support letter from J. Macklin

Hi Rick,

Sure, I am still quite interested in the potential of uniting these data sources and making them more available. There are some serious challenges here, from an informatics perspective, and will do what I can to lend advice and participate in whatever capacity makes sense. I would be interested in seeing (and contributing to) the pre-proposal as it evolves.

I am not sure if I mentioned to you my current NSF project called the Filtered-Push (other PIs are Paul Morris at Harvard and Bob Morris at UMASS-Boston). This is an annotation service which takes advantage of Google's map-reduce technology (using Hadoop) in a peer-to-peer network. Thus, the data sets on the network are connected and benefit through annotations by users. Although conceived using natural history databases as the target this technology is applicable to any related data sets. I think this could be a core network infrastructure for iPlant. Perhaps this is something I can bring to the table here. For the techies in the crowd you can see some of what we are up to on our wiki at http://mantis.cs.umb.edu/wiki/index.php/Main_Page

Look forward to ongoing discussion,

James Macklin, Ph.D.
Director of Collections and Informatics
Harvard University Herbaria
22 Divinity Avenue
Cambridge, MA 02138
Ph. (617) 496-1566
Fax. (617) 495-9484

MACQUARIE
UNIVERSITY ~ SYDNEY

Mark Schildhauer
National Center for Ecological Analysis and Synthesis
Santa Barbara, CA 93101
Phone: 805-892-2509

3^{th} February, 2009

Dear Mark,

I am writing to express my support for your motion to bring together a cross-disciplinary team of experts to help define, direct and advise the future of informatics in the biodiversity sciences under the iPlant cyberinfrastructure initiative.

I believe that it is important for new informatics efforts to enlist the help of existing ones in order to reduce redundancy and move forward rapidly—especially at a time when improved data software is fundamental to understanding, predicting, and solving large-scale questions including those of climate change. Your team has the experience and scope to navigate the heterogeneous informatics landscape.

I would be delighted to contribute to the program. My group at Macquarie University has expertise in quantitative ecology and evolution and ecoinformations. More specifically, I have a strong background in knowledge representation and an understanding of the utility and function of ontologies in the ecological and environmental sciences.

Best wishes.

Sincerely,

Josh Madin

AUSTRALIA'S INNOVATIVE UNIVERSITY

Joshua Madin, Lecturer, Department of Biological Sciences
MACQUARIE UNIVERSITY ~ NEW SOUTH WALES 2109 AUSTRALIA
TELEPHONE: 61 2 9850 4258 ~ FACSIMILE: 61 2 9850 9748
E-MAIL: jmadin@bio.mq.edu.au WEB: http://www.bio.mq.edu.au/computational_ecology

Literature Cited

1. Brown, J.H., *On the Relationship Between Abundance and Distribution of Species.* American Naturalist, 1984. **124**(#2): p. 255-279.
2. Lomolino, M.V., B.R. Riddle, and J.H. Brown, *Biogeography*. 2006: Sinauer.
3. McGill, B.J., Etienne, R.S., Gray, J.S., Alonso, D. , Anderson, M.J., Benecha, H.K., Dornela, M., Enquist, B.J. , Green, J.L., He, F., Hurlbert, A.H., Magurran, A.E., Marquet, P.A., Maurer, B.A., Ostling , A., Soykan, C.U., Ugland, K.I. and E. P. White, *Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework.* Ecology Letters, 2007: p. 995-1015.
4. Woodward, F.I., *Climate and plant distribution*. 1987, Cambridge: Cambridge University Press.
5. Nobel, P.S. and E.G. Bobich, *Plant frequency, stem and root characteristics, and CO_2 uptake for Opuntia acanthocarpa: elevational correlates in the northwestern Sonoran Desert.* Oecologia, 2002. **130**: p. 165-172.
6. Osmond, C.B., et al., *Stress physiology and the distribution of plants.* Bioscience, 1987. **37**: p. 38-48.
7. Feild, T.S., T. Brodribb, and M. Holbrook, *Hardly a relict: Freezing and the evolution of vesselless wood in winteraceae.* Evolution, 2002. **56**(3): p. 464-478.
8. Swenson, N.G. and B.J. Enquist, *Ecological and evolutionary determinants of a key plant functional trait: Wood density and its community-wide variation across latitude and elevation.* American Journal of Botany, 2007. **94**(3): p. 451-459.
9. Sebastiaan, V., et al., *Life-history traits are correlated with geographical distribution patterns of western European forest herb species.* Journal of Biogeography. **34**: p. 1723-1735.
10. Condit, R., S.P. Hubbell, and R.B. Foster, *Changes in tree species abundance in a Neotropical forest: Impact of climate change.* Journal of Tropical Ecology ;, 1996. **12**(pt.2)): p. 231-256.
11. Eckhart, V.M., M.A. Geber, and C.M. McGuire, *Experimental studies of adaptation in Clarkia zantiana I. Sources of variation across a subspecies border. .* Evolution, 2003. **58**: p. 59-70.
12. Wilczek, A.M., et al., *Effects of Genetic Perturbation on Seasonal Life History Plasticity.* Science, 2009. DOI: 10.1126/science.1165826.
13. Pauwelsa, N.M., et al., *When population genetics serves genomics: putting adaptation back in a spatial and historical context.* Current Opinion in Plant Biology, 2008(2): p. 129-134.
14. Leimu, R. and M. Fischer, *A meta-analysis of local adaptation in plants.* PLoS ONE, 2008. **3**: p. e4010.
15. Hooper, D.U., et al., *Effects of biodiversity on ecosystem functioning: A consensus of current knowledge.* Ecological Monographs, 2005. **75**(1): p. 3-35.
16. Bergengren, J.C., et al., *Modeling global climate-vegetation interactions in a doubled CO2 world.* Climatic Change, 2001. **50**(1-2): p. 31-75.
17. Williams, J.W., et al., *Rapid vegetation responses to past climate change. .* Geology 2002(30): p. 971-974. .
18. Williams, J.W. and S.T. Jackson, *Novel Climates, No-Analog Plant Communities, and Ecological Surprises: Past and Future.* Frontiers in Ecology and Evolution 2007(5): p. 475-482. .
19. Shafer, S.L., P.J. Bartlein, and R.S. Thompson, *Potential changes in the distributions of western North America tree and shrub taxa under future climate scenarios.* Ecosystems ;, 2001. **4**(3): p. 200-215.
20. Stephen P. Hubbell, S.P., et al., *How many tree species and how many of them are there in the Amazon will go extinct?* PNAS, 2008. **105**: p. 11498-11504.
21. Colwell, R.K., et al., *Lowland Biotic Attrition in the Wet Tropics Global Warming, Elevational Range Shifts, and Lowland Biotic Attrition in the Wet Tropics.* Science, 2008. **322**: p. 258-261.

22. Loiselle, B.A., et al., *Predicting species distributions from herbarium collections: does climate bias in collection sampling influence model outcomes?* 2008. **35**: p. 105-116.
23. Myers, N., et al., *Biodiversity hotspots for conservation priorities.* . Nature, 2000(403): p. 853-858.
24. Pitman, N.C.A., et al., *Dominance and distribution of tree species in upper Amazonian terra firme forests.* Ecology, 2001. **82**: p. 2101-2117.
25. Weiser, M.D., et al., *Latitudinal patterns of range size and species richness of New World woody plants.* Global Ecology and Biogeography, 2007. **16**(5): p. 679-688.
26. Williams, P.H., K.J. Gaston, and C.J. Humphries, *Mapping biodiversity value worldwide: Combining higher-taxon richness from different groups.* Proceedings of the Royal Society of London Series B-Biological Sciences, 1997. **264**(1378): p. 141-148.
27. NRC, *NEON: Addressing the nation's environmental challenges.* . 2003, Washington, DC: National Research Council.
28. Bortolus, A., *Error Cascades in the Biological Sciences: The Unwanted Consequences of Using Bad Taxonomy in Ecology.* Ambio, 2008. **37**: p. 114-118.
29. Reich, P.B., et al., *The evolution of plant functional variation: Traits, spectra, and strategies.* International Journal of Plant Sciences, 2003. **164**(3): p. S143-S164.
30. Gentry, A.H., *Changes in plant community diveristy and floristic composition on environmental and geographic gradients.* Annals of the Missouri Botanical Garden., 1988. **75**: p. 1-34.
31. Gentry, A.H., *Tropical forest biodiversity: distributional patterns and their conservational significance.* Oikos, 1992. **63**: p. 19-28.
32. ter Steege, H., et al., *Continental-scale patterns of canopy tree composition and function across Amazonia.* Nature, 2006. **443**: p. 444-447.
33. Berendsohn, W.G., *A taxonomic information model for botanical databases the IOPI Model.* . Taxon, 1997. **46**: p. 283-309.
34. Kennedy, J., R. Kukla, and T. Paterson, *Scientific names are ambiguous as identifiers for biological taxa: their context and definition are required for accurate data integration*, in *Data Integration in the Life Sciences: Proceedings of the Second International Workshop,*, B.L.R. Ludäscher, Editor. 2005, DILS 2005, LNBI: San Diego, CA.
35. Franz, N.M., R.K. Peet, and A.S. Weakley, *On the use of taxonomic concepts in support of biodiversity research and taxonomy.* , in *The New Taxonomy. Systematics Association Special Volume (74) Symposium Proceedings,*, Q.D. Wheeler, Editor. 2008, Taylor & Francis: Boca Raton, FL, . p. 63-86.
36. Enquist, B.J. and K.J. Niklas, *Invariant scaling relations across tree-dominated communities.* Nature, 2001. **410**: p. 655-660.
37. Swenson, N.G., et al., *The influence of spatial and size scale on phylogenetic relatedness in tropical forest communities.* Ecology, 2007. **88**(7): p. 1770-1780.
38. Rabeler, R. and R.K. Macklin, *Herbarium networks in the United States: Towards creating a toolkit to advance specimen data capture.* Collection Forum, 2007. **21**: p. 223-231.
39. Stephan, M.H. and J.H.J. Schaminée, *TURBOVEG, a comprehensive data base management system for vegetation data* Journal of the Vegetation Science, 2001. **12**: p. 589-591.
40. Violle, C., et al., *Let the concept of trait be functional!* Oikos, 2007. **116**: p. 882-892.
41. McGill, B., et al., *Rebuilding community ecology from functional traits.* Trends in Ecology & Evolution, 2006. **21**: p. 178-185.
42. Westoby, M., et al., *Plant ecological strategies: Some leading dimensions of variation between species.* 2002.
43. Wright, I.J., et al., *The worldwide leaf economics spectrum.* Nature, 2004. **428**(6985): p. 821-827.
44. Poorter, L., et al., *Are functional traits good predictors of demographic rates? Evidence from five neotropical forests.* Ecology, 2008. **89**(7): p. 1908-1920.
45. Franz, N., R.K. Peet, and A.S. Weakley, *On the use of taxonomic concepts in support of biodiversity research and taxonomy.* 2008.

46. Franz, N.M. and R.K. Peer, *Toward a language for mapping relaionships among taxonomic concepts.* Systematics and Biodiversity, 2008. doi:10.1017/S147720000800282X
47. Berendsohn, W.G., *A taxonomic information model for botanical databases: the IOPI Model. .* Taxon, 1997. **46**: p. 283-309.
48. Graham, C.H., et al., *New developments in museum-based informatics and application in biodiversity analysis.* Trends in Ecology and Evolution 2004. **19**: p. 497-503.
49. Elith, J., et al., *Novel methods improve prediction of species' distributions from occurrence data.* Ecography, 2006. **29**: p. 129-151.
50. Randin, C.F.e.a., *Are niche-based species distribution models transferable in space? .* Journal of Biogeography, 2006(33): p. 1689-1703.
51. Philips, S.J., *Transferability, sample selection bias and background data in presence-only modelling: a response to Peterson et al. .* Ecography, 2008. **31**: p. 272-278.