

# SCIENTIFIC DATA

**OPEN**

**SUBJECT CATEGORIES**

- » Ecology
- » Phylogenetics
- » Fungal ecology
- » Arbuscular mycorrhiza
- » Fungal evolution

Received: 09 February 2016

Accepted: 23 March 2016

Published: 10 May 2016

## Data Descriptor: MycoDB, a global database of plant response to mycorrhizal fungi

V. Bala Chaudhary *et al.*<sup>#</sup>

Plants form belowground associations with mycorrhizal fungi in one of the most common symbioses on Earth. However, few large-scale generalizations exist for the structure and function of mycorrhizal symbioses, as the nature of this relationship varies from mutualistic to parasitic and is largely context-dependent. We announce the public release of MycoDB, a database of 4,010 studies (from 438 unique publications) to aid in multi-factor meta-analyses elucidating the ecological and evolutionary context in which mycorrhizal fungi alter plant productivity. Over 10 years with nearly 80 collaborators, we compiled data on the response of plant biomass to mycorrhizal fungal inoculation, including meta-analysis metrics and 24 additional explanatory variables that describe the biotic and abiotic context of each study. We also include phylogenetic trees for all plants and fungi in the database. To our knowledge, MycoDB is the largest ecological meta-analysis database. We aim to share these data to highlight significant gaps in mycorrhizal research and encourage synthesis to explore the ecological and evolutionary generalities that govern mycorrhizal functioning in ecosystems.

<b>Design Type(s)</b>	database creation objective • observation design • data integration objective
<b>Measurement Type(s)</b>	symbiosis, encompassing mutualism through parasitism • phylogenetic analysis objective
<b>Technology Type(s)</b>	digital curation • computational phylogenetic analysis
<b>Factor Type(s)</b>	
<b>Sample Characteristic(s)</b>	mycorrhizal samples

Correspondence and requests for materials should be addressed to V.B.C. (email: veerbala@hotmail.com).  
<sup>#</sup>A full list of authors and their affiliations appears at the end of the paper.

## Background & Summary

Plant performance is largely a function of the plant-symbiotic microbiome<sup>1</sup>. As a result, ecosystem functions and the vital services humans derive from them (e.g., food and fiber production, carbon sequestration) are fundamentally dependent on the interactions plants have with symbionts. Although symbioses are common, our knowledge of their impact on ecosystem functions and services is relatively incomplete. Broad generalizations about the relationships between plants and their symbionts are limited due to the context-dependent nature of such symbioses, which exist along a continuum of possible outcomes, from mutualistic to parasitic. Examining the results of many experiments through meta-analysis allows for larger-scale generalizations than individual experiments can provide independently and can lead to synthesis-generated evidence<sup>2</sup>. Furthermore, including phylogenetic information in meta-analyses can account for correlated evolutionary relationships among taxa used in multiple studies. Understanding ecological outcomes of symbioses, and the environmental and evolutionary context contributing to such outcomes, is crucial to maintaining and restoring the ecosystem functions and services upon which humans depend.

Mycorrhizal fungi form an ancient symbiosis with most plants on Earth<sup>3,4</sup>. The host plant and associated fungi form a trading partnership where the fungi increase the effective absorptive capabilities of the plant, delivering nutrients and water in exchange for plant-derived photosynthates<sup>5</sup>. Since most plants associate with mycorrhizal fungi, the outcome of this symbiosis can influence ecosystem structure, function, and services mediated by plant productivity. Several empirical studies have individually demonstrated how host plant traits and identities, fungal partners, soil biotic and abiotic conditions, and experimental conditions can alter the structure and function of mycorrhizal symbioses<sup>6–9</sup>. The important role of mycorrhizal fungi in global change dynamics such as N deposition, climate change, and invasive species has also been documented<sup>10–12</sup> as well as the key role that mycorrhizal fungi play in the restoration and conservation management of ecosystems<sup>13–15</sup>. Understanding the generality of how environmental context impacts the relationship of plants with their mycorrhizal symbionts should also affect how we manage terrestrial ecosystems.

Statistical methods to simultaneously examine multiple ecological and evolutionary factors in a meta-analysis framework have been recently developed<sup>16–18</sup>, facilitating computational analytic approaches to studying global patterns of mycorrhizal symbioses. Large datasets with multiple moderators are vital to such approaches because ecological meta-analyses based on small datasets can be vulnerable to spurious interpretations resulting from unbalanced data distributions, correlated moderators, and unavailable information regarding potentially important predictors. For example, a meta-analysis of 51 studies<sup>19</sup> suggested the absence of synergy between mycorrhizal fungi and nitrogen-fixing bacteria likely resulted from an overrepresentation of annual and agricultural species in the meta-analysis database; by contrast, separate tests of late successional legumes demonstrated strong synergism<sup>20</sup>. Another meta-analysis<sup>21</sup> suggested that the paradoxical result of declining plant growth response to mycorrhizal inoculation with nitrogen but not phosphorus fertilization was due to the overrepresentation of studies in the dataset with high soil phosphorus and the potential correlation among moderators. The problem of spurious interpretations within ecological meta-analyses can be reduced by simultaneous testing of multiple predictors, which is made possible by larger databases.

Here we present MycoDB, a large database of mycorrhizal inoculation experiments, linked with plant and fungal phylogenies (Data Citation 1), to facilitate tests on the ecological and evolutionary contexts in which the addition of mycorrhizal fungi is beneficial or parasitic to plant hosts. MycoDB focuses on studies of two dominant types of mycorrhizal fungi, ectomycorrhizal (EM) fungi and arbuscular mycorrhizal (AM) fungi, because they predominate among published studies on mycorrhizal symbioses. MycoDB contains data on plant productivity response to mycorrhizal fungi from 4,010 studies (from 438 unique publications) and is organized in a hierarchical fashion such that a single publication can contain multiple discrete experiments and a single experiment can contain multiple studies. The ecological and evolutionary context of studies can be explored with 24 additional explanatory variables (e.g., plant functional group, inoculum complexity, plant or fungal origin; Table 1 (available online only) and Table 2) and mycorrhizal fungal and plant host phylogenetic trees (Figs 1 and 2). MycoDB can be used to model phylogenetic heritability of plant response to mycorrhizal fungi in host plant lineages, fungal lineages, and their interaction, as well as explore the relationship among explanatory variables and plant response to mycorrhizal fungi, while controlling for the influence of plant and fungal phylogenies.

## Methods

### Overview and literature searches

MycoDB (Data Citation 1) contains data from three main phases of data collection and validation. Phase I occurred in 2005, when we identified 1852 publications by conducting an initial literature search of the ISI Web of Science database using the key words *mycorrhiz\** and *inocul\** (on January 22, 2005). From this initial list, 134 publications were selected, in random order, as having met our inclusion criteria for meta-analysis such as reporting plant biomass response, use of a mycorrhizal addition treatment, and inclusion of a non-inoculated control (see ‘Criteria for inclusion’ below). More publications from the initial list of 1852 publications likely met our criteria, but were excluded from Phase I of database construction because of time constraints. Data from 49 additional publications on EM fungi were added from a previous meta-analysis<sup>22</sup> to reduce dominance of the data by studies on AM symbioses. This process

resulted in a total of 183 publications summarized in MycoDB during Phase I. In 2010, as part of an NCEAS (National Center for Ecological Analysis and Synthesis) Distributed Graduate Seminar conducted across nine institutions, we began Phase II of the data collection to dramatically increase the size of MycoDB with several targeted literature searches. On September 21, 2010, we conducted searches of the ISI Web of Science database using the following search terms: (1) (*mycorrhiz\** or *ectomyc\** or *endomyc\** or *arbuscul\** or *vesicular\**) and *inocul\** resulting in 4,013 papers; (2) search terms from (1) AND *restoration* or *rehabilitation* or *reclamation* or *revegetation* or *reforestation* resulting in 305 papers; (3) search terms from (1) AND *local adaptation* or *strain* or *isolate* or *genotype* or *ecotype* or *geograph\** resulting in 627 papers; (4) search terms from (1) AND *tissue P* or *tissue N* or *shoot P* or *shoot N* or *leaf P* or *leaf N* resulting in 793 papers; (5) search terms from (1) AND *Gigaspor\** or *Acaulospor\** or *Scutellospora* or *Archaeospora* resulting in 387 papers. Searches 2–5 were designed to enrich the database in studies relevant to restoration ecology, local adaptation, influences of nutrients, and AM fungi besides *Glomus*, because these topics were identified either as being of interest for planned focused meta-analyses (restoration, local adaptation) or as under-represented in the Phase I search results (influences of nutrients, AM fungi besides *Glomus*). The results from all five searches were collated, duplicates (as well as publications already included in MycoDB) were removed, and papers were selected that appeared in at least one of the four focused searches (2–5), resulting in a list of 1,768 publications. Again, from this larger list, 255 publications were selected, in random order, as having met our inclusion criteria for meta-analysis such as reporting plant biomass response, use of a mycorrhizal addition treatment, and inclusion of a non-inoculated control (see below ‘Criteria for inclusion’), and their data were added to MycoDB. After Phase II, MycoDB contained data from a total of 4,010 studies from 438 publications. Phase III of the creation of MycoDB consisted of extensive data validation and the creation of phylogenetic trees for all plant species and fungal genera in the database.

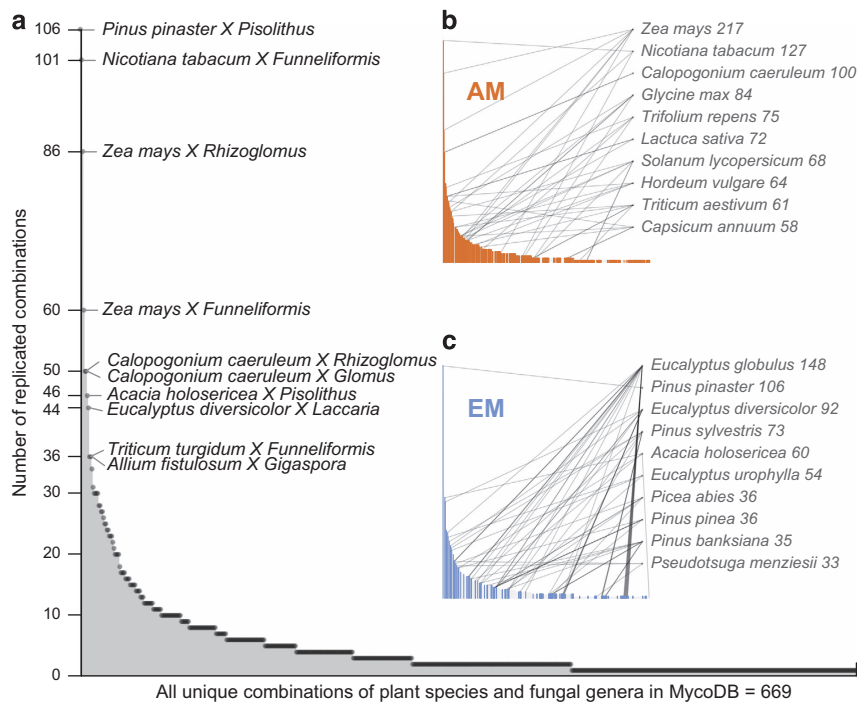
A subset of MycoDB was used to study how edaphic properties, plant functional groups, and microbial community complexity determine the outcome of mycorrhizal symbioses<sup>21</sup>. Different subsets of these data have been used to explore local adaptation among plants, mycorrhizal fungi, and soils (Rúa *et al.* in review) as well as how partner identity, colonization levels, and P fertilization impact plant host response to EM associations<sup>22</sup>.

### Criteria for inclusion

Prior to inclusion in MycoDB, publications were screened for meta-analysis appropriateness and for an experimental design that was amenable to our research questions. It was required that all studies compare results of a mycorrhizal inoculation treatment (or several treatments) to a non-inoculated control. In other words, studies must compare plant response for some addition of mycorrhizal fungi to no addition. The method of inoculation varies among studies in MycoDB and can include the addition of spores, roots, mycelia, pot culture, field soil, or any combination thereof. Studies that apply mycorrhizal fungi to all treatments, eliminate fungal presence (e.g., using fungicide application), or otherwise manipulate ecological factors to promote or suppress fungi were not included. We included studies with unsterilized background soil, many of which likely contained propagules of mycorrhizal fungi in all treatments, though this could not be confirmed because fungal colonization data was not consistently reported. If an experiment contained the manipulation of a factor in addition to mycorrhizal fungi (e.g., fertilizer treatment, soil amendment), the results are included in MycoDB as separate studies within the same experiment. All studies report mean plant biomass data as the response variable. Studies could report shoot biomass, total biomass, or root biomass and shoot biomass. Studies must report means, but were still included if measures of dispersion (e.g., standard error, standard deviation, or error bars on figures) were not given as was the case in 91% of studies. If sample size was not given, the associated parameter ‘n’ was coded as 1 for both inoculated treatments and non-inoculated controls, reducing the weight of the study relative to what it would be if the sample size was known. Data presented in tables were extracted directly, but data presented in figures were extracted using Engauge Digitizer software version 4.1<sup>23</sup>. Data on additional explanatory variables (e.g., plant family, plant functional group) were also extracted from the publication text when available or looked up using supplementary peer-reviewed resources. Both lab studies and field studies are included in MycoDB and coded separately using the variable LOCATION (Table 1 (available online only)). No limitation was placed on the duration of study for inclusion in MycoDB; in the case of studies that examined plant biomass over a time series, only data from the last sampling event was included. Data were then entered into MycoDB using a custom web-based data entry interface and database that matched inoculated treatments with non-inoculated controls<sup>24</sup>.

### Plant and fungal phylogenetic tree construction

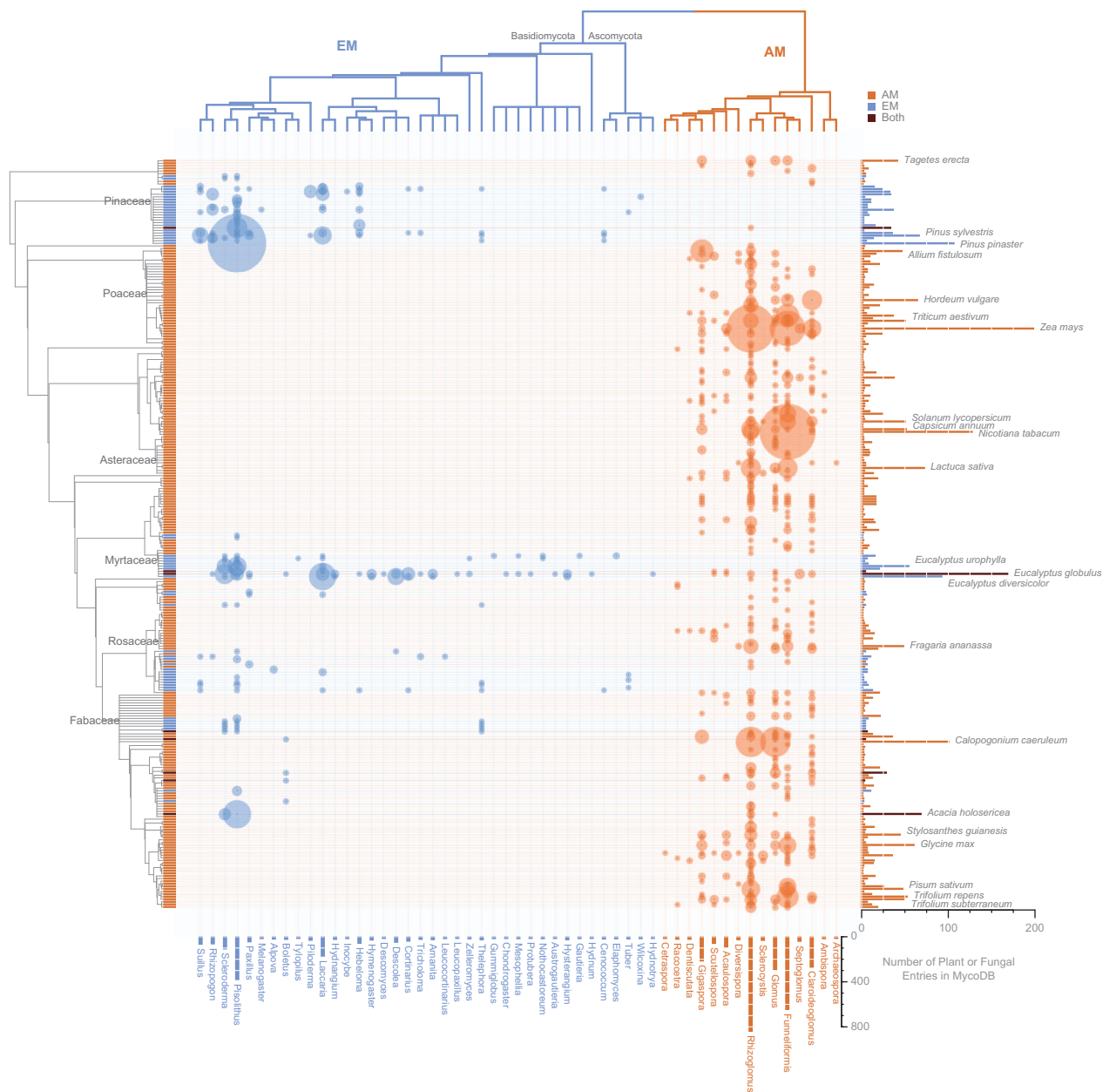
We constructed plant and fungal phylogenetic trees for all the species in MycoDB using a composite phylogeny approach, which combines taxonomic and phylogenetic information into a single tree<sup>16,25,26</sup>. For plant phylogeny, we derived phylogenetic topology from existing ‘supertrees’ and assigned well-supported divergence times to all possible internal bifurcations (evolutionary divergence event) using TimeTree<sup>27</sup> as a source of published divergence times. The remaining unknown branch lengths were rooted with known divergence dates<sup>28</sup> and arbitrarily set and scaled to yield an ultrametric tree wherein all extant species were lined up at the present date. In cases where taxonomic nomenclature has changed



**Figure 1.** The distribution of unique plant species and fungal genus combinations contained in MycoDB. Larger figure (a) shows the overall distribution of unique plant fungal combinations (669 total) in the database with the most common combination being *Pinus pinaster* inoculated with *Pisolithus* (106 studies). Inset graphs separate AM (b) and EM (c) fungal studies and highlight the 10 most common plant species in each subset. Numbers next to plant names on insets indicate the quantity of studies in MycoDB that utilize that plant host. Lines next to plant names indicate the number of fungal genera used to inoculate a particular plant species across studies.

LA_Code	Plant	Fungi	Soil	#studies
A	1	1	1	96
B	1	1	2	8
C	1	2	1	57
D	2	1	1	14
E	1	2	3	104
F	1	1	X	64
G	1	2	X	77
H	1	1	Z	30
I	1	2	Z	140
J	1	X	1	11
K	1	X	2	63
L	X	1	1	96
M	X	1	2	299
N	Z	1	1	77
O	Z	1	2	265

**Table 2.** Description of LA\_Codes such that components of studies—plants, fungi, and soil—share the same number if they originate from the same known location. Unknown locations are indicated by ‘X’. Artificial soils (e.g., peat moss) or non-wild plant varieties (e.g., cultivar or hybrid variety) are indicated by ‘Z’. For example, the database contains 299 studies, coded as ‘M’, where the source of the plant is unknown and the soil and AM fungi came from different locations.



**Figure 2.** Heat map showing the frequency of studies in MycoDB according to unique plant host and fungal genus combinations and their phylogenetic relationships. Larger circles indicate a larger number of studies that exist in MycoDB for that unique plant species-fungal genus pair. The left side shows a composite phylogenetic tree for plant species, and a composite phylogenetic tree for mycorrhizal fungal genera is on the top of the graph. Different colors indicate plants grown in association with different mycorrhiza types (blue = EM fungi, red = AM fungi, black = either AM or EM).

since the creation of the original ‘supertrees’ or publication of papers used in the database, names were changed manually to reflect current consensus taxonomy. Pairwise shared branch-lengths were then used to calculate a variance-covariance matrix, which can be used in mixed multifactor meta-analyses.

For the fungal composite phylogeny, we manually reconstructed the evolutionary relationships among different genera based on known or commonly accepted taxonomy using information from previously published reports. Fungal taxonomy, particularly of AM fungi, has undergone major revisions during the duration of the compilation of MycoDB. We traced the evolution of fungal taxa into current consensus systematics<sup>29</sup>; however, because of ambiguity in species identification of AM fungi, we only included fungal taxonomic identification to the genus level. Even so, some taxa formerly named *Glomus* could not

be placed definitively within current genera and were therefore left as *Glomus*. In the case of the AM fungal phylogenetic branch, the composite tree topologies between different genera of this clade were informed by taxonomic position of the type species of each genus (when possible) in relationship with the taxonomic position of the other type species of another genus<sup>29–31</sup>. For EM fungi, the position of each species in the phylogenetic tree follows the online version of Index Fungorum ([www.indexfungorum.org](http://www.indexfungorum.org)) and recent taxonomic literature. The phylogenetic framework and tree topology were based on the 2007 AFTOL classification of Fungi<sup>32</sup> and other recent efforts in fungal systematics<sup>33,34</sup>.

Construction and manipulation of composite phylogenies was conducted using R Statistical Software<sup>35</sup> (version 3.0.2), the *ape* package in R<sup>36</sup>, Phylocom<sup>37</sup>, and Phylomatic<sup>37</sup>. The files contain the fungal and plant composite phylogenies in the Newick tree format. The FungalTree\_version1.txt file (Data Citation 1) represents the evolutionary relationships among different fungal genera that exist in the MycoDB database. Similarly, the PlantTree\_version1.txt file (Data Citation 1) represents the evolutionary relationships between different plant species that exist in this database, with each node of the plant's composite tree labeled with corresponding higher taxonomic classification.

### Data attributes

Data in MycoDB are organized in a hierarchical manner as a single publication often contained data from multiple discrete experiments and studies (i.e., trials) on multiple plant hosts. As such, the 438 publications in MycoDB contain data for 4,010 studies (Data Citation 1). A study is defined as a comparison of average plant performance between plants that were inoculated with mycorrhizal fungi (AM or EM, never both) and plants that were not inoculated. Table 1 (available online only) contains detailed meta-data for all variables in MycoDB including descriptions of variables and levels. Figure 1 demonstrates the frequency and distribution of unique plant species and fungal genus combinations (669 total) contained in MycoDB. For example, the most frequently reported mycorrhizal combination was *Pinus pinaster* (maritime pine) inoculated with species from the EM fungal genus *Pisolithus* (106 studies). The two inset graphs represent the most common plant species in MycoDB, separated by mycorrhizae type (AM vs EM). Lines to plant species indicate the number of fungal genera in association with each plant species. For example, the most common plant species in MycoDB are *Zea mays* (corn, 217 studies) and *Eucalyptus globulus* (blue gum, 148 studies), in association with AM fungi and EM fungi, respectively. Figure 2 is a heat map representing the frequency of studies according to unique plant host and fungal genus combinations and their phylogenetic relationships. For EM fungi, the most commonly represented plant-fungal combinations in MycoDB occur between plants within the Pinaceae growing in association with *Pisolithus* fungi. For AM fungi, hotspots occur within the Poaceae, Solanaceae, and Fabaceae grown in association with *Rhizoglosum* and *Funneliformis* fungi. As these plant families are important to forestry and agriculture, their prevalence in the literature makes sense, but the tropics and thus a large portion of plant and fungal biodiversity are underrepresented. Figure 2 suggests that empirical work thus far regarding the mycorrhizal symbiosis is not only limited with respect to the plant and fungal species examined, but also relatively poorly represented among phylogenetically diverse clades of plants and fungi.

Although this database represents the efforts of over 80 people distributed over 10 years, these data still only represent a fraction of the total available data on plant response to mycorrhizal fungi. The number of papers published each year that fit our search criteria has grown exponentially in the time since our initial search. The 351 plant species in our database represent a small proportion of the 450,000 total plant species on Earth<sup>38</sup>, the majority of which likely associate with mycorrhizal fungi. Moreover, as might be expected, the plant taxa represented are heavily biased toward species important for agriculture and forestry (e.g., corn, tobacco, pine, eucalyptus). Similarly, the fungal taxa that are best represented in the database are taxa that have been commercially marketed, such as the ectomycorrhizal fungus *Pisolithus tinctorius*<sup>39</sup>. Given this uneven representation of plant and fungal species and potential correlations among closely related plant and fungal species, it is important to analyze these data using phylogenetic mixed models even when testing environmental moderators of plant responsiveness to mycorrhizal fungal inoculation.

### Statistical considerations

MycoDB was prepared in anticipation of common technical problems (and statistical solutions thereof) encountered in meta-analyses. In particular, a difficulty in many ecological meta-analyses is the lack of independence of the estimates. Multiple estimates ('studies') extracted from the same publication may be more similar to each other than those arising from different publications due to similarities in experimental methods or context within the same publication. Multilevel meta-analytic models (with estimates nested within publications) can be used to account for such correlated data structures<sup>18</sup>. Similarly, multiple estimates may represent contrasts of different treatments that are compared against a common control condition, leading to statistical dependencies in the estimates due to reuse of information from the control condition<sup>40</sup>. Hence, identification of estimates that share a common control condition is of crucial importance (i.e., variable CTLTRTSETID in Table 1 (available online only)). Moreover, multiple studies may use the same species or different species that are phylogenetically related, and such taxonomic overrepresentation may limit the scale of inference of the meta-analysis. Inclusion of information on phylogenetic relations within a mixed-effects model can account for these correlations

and allow tests of generality of results. This can be done through inclusion of categorical taxonomic (e.g., family, genus, species) variables as random effects<sup>21</sup> or through an analysis that includes the full phylogeny<sup>41,42</sup>. Including phylogenetic information in meta-analyses can account for correlated evolutionary relationships as well as allow inference on the rates and constraints of evolution on the phenotypic character being considered<sup>18</sup>. Finally, the availability of phylogenies for both plant species and fungi offers the possibility of modeling potential coevolution and ecological interactions using appropriate mixed-effects models<sup>43</sup>.

As the largest database concerning this important symbiosis, MycoDB may prove particularly useful for the development of meta-analysis educational curriculum or statistics tutorials using ecological data. In a classroom setting, for example, the database could be used in demonstrations of single predictor meta-analyses, which could then be followed up with comparison of multiple moderator meta-analyses to demonstrate the consequences of correlated predictors. Subsets of the data could support investigation of the advantages of larger datasets in overcoming problems of correlated predictors, thereby promoting exploration of context-dependency in future meta-analyses.

## Data Records

### Data record 1

The database file in csv format, titled 'MycoDB\_version1.csv' (February 5, 2016 version), was uploaded to the Dryad Digital Repository (Data Citation 1). Detailed meta-data for each column is located in Table 1 (available online only) of this Data Descriptor article. Taxonomic information on plant species and fungal genera contained in the database are included in Supplementary File 1 of this article.

### Data record 2

The phylogenetic tree of mycorrhizal fungal genera present in Data record 1, in txt format and titled 'FungalTree\_version1.txt' (February 5, 2016 version), was uploaded to the Dryad Digital Repository (Data Citation 1). The file contains the fungal genera composite phylogeny in the Newick tree format.

### Data record 3

The phylogenetic tree of plant hosts present in Data record 1, in txt format and titled 'PlantTree\_version1.txt' (February 5, 2016 version), was uploaded to the Dryad Digital Repository (Data Citation 1). The file contains the plant composite phylogeny in the Newick tree format.

## Technical Validation

We devised several layers of methods to ensure the quality of the data in MycoDB. First, random sampling of publications that resulted from our initial searches was conducted to reduce bias in which data were included in the database. Second, on the front end, data collection was conducted using a web-based custom data entry system with organized fields and drop down menus to reduce data entry error<sup>24</sup>. This approach also allowed data collection to be conducted simultaneously and remotely by multiple users. After front-end data entry by users, database administrators conducted back-end database content management to validate data integrity by examining distributions and outliers as well as iteratively hand-checking random subsets of papers for accuracy. We used the United States Department of Agriculture PLANTS Database<sup>44</sup> and The Plant List version 1.1<sup>45</sup> to verify and update scientific names and life histories for each species included in MycoDB. Database administrators hand checked outliers and returned to original papers when data were missing, as well as added moderator variables and edited moderator levels to facilitate specific meta-analyses. Finally, for the local adaptation study subset of MycoDB, all entries were compared with the original paper and corrected when necessary. The data validation methods used to create MycoDB satisfy all data-related compliance criteria designed to promote methodological quality in ecological meta-analyses<sup>46</sup>.

## Usage Notes

MycoDB is deposited in the Dryad Digital Repository at <http://dx.doi.org/10.5061/dryad.723m1> and publicly available under the CC0 public domain dedication, given proper scholarly citation of the version used and this data descriptor. We recommend that, prior to publication, users validate data subsets against original publications.

## References

1. Turner, T. R., James, E. K. & Poole, P. S. The plant microbiome. *Genome Biol.* **14**, 209 (2013).
2. Cooper, H. M. in *Synthesizing research: A guide for literature reviews* (Sage, 1998).
3. Brundrett, M. C. Coevolution of roots and mycorrhizas of land plants. *New Phytol.* **154**, 275–304 (2002).
4. Wang, B. & Qiu, Y.-L. Phylogenetic distribution and evolution of mycorrhizas in land plants. *Mycorrhiza* **16**, 299–363 (2006).
5. Smith, S. E. & Read, D. J. in *Mycorrhizal symbiosis* (Academic press, 1996).
6. Hetrick, B., Wilson, G. & Todd, T. Differential responses of C3 and C4 grasses to mycorrhizal symbiosis, phosphorus fertilization, and soil microorganisms. *Canadian Journal of Botany* **68**, 461–467 (1990).
7. Johnson, N. C., Tilman, D. & Wedin, D. Plant and soil controls on mycorrhizal fungal communities. *Ecology* **73**, 2034–2042 (1992).
8. Klironomos, J. N. Variation in plant response to native and exotic arbuscular mycorrhizal fungi. *Ecology* **84**, 2292–2301 (2003).
9. Lekberg, Y. & Koide, R. T. Integrating physiological, community, and evolutionary perspectives on the arbuscular mycorrhizal symbiosis 1. *Botany* **92**, 241–251 (2013).

10. Pringle, A. *et al.* Mycorrhizal symbioses and plant invasions. *Annual Review of Ecology, Evolution, and Systematics* **40**, 699–715 (2009).
11. Treseder, K. K. *et al.* Integrating microbial ecology into ecosystem models: challenges and priorities. *Biogeochemistry* **109**, 7–18 (2012).
12. Nuñez, M. A. & Dickie, I. A. Invasive belowground mutualists of woody plants. *Biol. Invasions* **16**, 645–661 (2014).
13. Middleton, E. L. & Bever, J. D. Inoculation with a native soil community advances succession in a grassland restoration. *Restor. Ecol.* **20**, 218–226 (2012).
14. Turrini, A. & Giovannetti, M. Arbuscular mycorrhizal fungi in national parks, nature reserves and protected areas worldwide: a strategic perspective for their in situ conservation. *Mycorrhiza* **22**, 81–97 (2012).
15. Maltz, M. R. & Treseder, K. K. Sources of inocula influence mycorrhizal colonization of plants in restoration projects: a meta-analysis. *Restor. Ecol.* **23**, 625–634 (2015).
16. Lajeunesse, M. J. Meta-Analysis and the Comparative Phylogenetic Method. *Am. Nat.* **174**, 369–381 (2009).
17. Viechtbauer, W. Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software* **36**, 1–48 (2010).
18. Nakagawa, S. & Santos, E. S. Methodological issues and advances in biological meta-analysis. *Evol. Ecol.* **26**, 1253–1274 (2012).
19. Larimer, A. L., Bever, J. D. & Clay, K. The interactive effects of plant microbial symbionts: a review and meta-analysis. *Symbiosis* **51**, 139–148 (2010).
20. Larimer, A. L., Clay, K. & Bever, J. D. Synergism and context dependency of interactions between arbuscular mycorrhizal fungi and rhizobia with a prairie legume. *Ecology* **95**, 1045–1054 (2014).
21. Hoeksema, J. D. *et al.* A meta-analysis of context-dependency in plant response to inoculation with mycorrhizal fungi. *Ecol. Lett.* **13**, 394–407 (2010).
22. Karst, J., Marczak, L., Jones, M. D. & Turkington, R. The mutualism-parasitism continuum in ectomycorrhizas: a quantitative assessment using meta-analysis. *Ecology* **89**, 1032–1042 (2008).
23. Mitchell, M. Engauge Digitizer. A free open-source software to extract data points from a graph image. <http://markumitchell.github.io/engauge-digitizer/> (2002).
24. Chaudhary, V. B., Walters, L. L., Bever, J. D., Hoeksema, J. D. & Wilson, G. W. Advancing synthetic ecology: a database system to facilitate complex ecological meta-analyses. *Bull. Ecol. Soc. Am.* **91**, 235–243 (2010).
25. Hadfield, J. & Nakagawa, S. General quantitative genetic methods for comparative biology: phylogenies, taxonomies and multi-trait models for continuous and categorical characters. *J. Evol. Biol.* **23**, 494–508 (2010).
26. Lajeunesse, M. J., Rosenberg, M. S. & Jennions, M. D. Phylogenetic nonindependence and meta-analysis *Handbook of meta-analysis in ecology and evolution* 284–299 (Princeton Univ. Press, 2013).
27. Hedges, S. B., Dudley, J. & Kumar, S. TimeTree: a public knowledge-base of divergence times among organisms. *Bioinformatics* **22**, 2971–2972 (2006).
28. Webb, C. O., Ackerly, D. D. & Kembel, S. W. Phylocom: software for the analysis of phylogenetic community structure and trait evolution. *Bioinformatics* **24**, 2098–2100 (2008).
29. Redecker, D. *et al.* An evidence-based consensus for the classification of arbuscular mycorrhizal fungi (Glomeromycota). *Mycorrhiza* **23**, 515–531 (2013).
30. Schüßler, A. & Walker, C. The Glomeromycota: a species list with new families and new genera. The Royal Botanic Garden Kew, Botanische Staatssammlung Munich, and Oregon State University (2010).
31. Oehl, F., Sieverding, E., Palenzuela, J., Ineichen, K. & da Silva, G. A. Advances in Glomeromycota taxonomy and classification. *IMA Fungus. The Global Mycological Journal* **2**, 191 (2011).
32. Hibbett, D. S. *et al.* A higher-level phylogenetic classification of the Fungi. *Mycol. Res.* **111**, 509–547 (2007).
33. Hibbett, D. S. *et al.* in *Systematics and Evolution* 373–429 (Springer, 2014).
34. Tedersoo, L., May, T. W. & Smith, M. E. Ectomycorrhizal lifestyle in fungi: global diversity, distribution, and evolution of phylogenetic lineages. *Mycorrhiza* **20**, 217–263 (2010).
35. R Core Team. *R: A language and environment for statistical computing* (R Foundation for Statistical Computing: Vienna, Austria, 2014).
36. Paradis, E., Claude, J. & Strimmer, K. APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics* **20**, 289–290 (2004).
37. Webb, C. O. & Donoghue, M. J. Phylomatic: tree assembly for applied phylogenetics. *Molecular Ecology Notes* **5**, 181–183 (2005).
38. Pimm, S. L. & Joppa, L. N. How Many Plant Species are There, Where are They, and at What Rate are They Going Extinct? *Ann. Mo. Bot. Gard.* **100**, 170–176 (2015).
39. Schwartz, M. W. *et al.* The promise and the potential consequences of the global transport of mycorrhizal fungal inoculum. *Ecol. Lett.* **9**, 501–515 (2006).
40. Lajeunesse, M. J. On the meta-analysis of response ratios for studies with correlated and multi-group designs. *Ecology* **92**, 2049–2055 (2011).
41. Maherali, H. & Klironomos, J. N. Influence of phylogeny on fungal community assembly and ecosystem functioning. *Science* **316**, 1746–1748 (2007).
42. Reinhart, K. O., Wilson, G. W. & Rinella, M. J. Predicting plant responses to mycorrhizae: integrating evolutionary history and plant traits. *Ecol. Lett.* **15**, 689–695 (2012).
43. Hadfield, J. D., Krasnov, B. R., Poulin, R. & Nakagawa, S. A tale of two phylogenies: comparative analyses of ecological interactions. *Am. Nat.* **183**, 174–187 (2014).
44. USDA, NRCS. *The PLANTS Database* <http://plants.usda.gov> (National Plant Data Team, Greensboro, North Carolina 27401, 2013).
45. The Plant List. Version 1.1. <http://www.theplantlist.org/> (2013).
46. Koricheva, J. & Gurevitch, J. Uses and misuses of meta-analysis in plant ecology. *J. Ecol.* **102**, 828–844 (2014).

## Data Citation

1. Chaudhary, V. B. *et al.* *Dryad Digital Repository* <http://dx.doi.org/10.5061/dryad.723m1> (2016).

## Acknowledgements

We are thankful to all participants who contributed to the creation of MycoDB: Lynette Abbott, Chris Allen, Brady Allred, Markandu Anputhas, Kristine Averill, Allison Baker, Wes Beaulieu, Reed Couch, Ahn-Heum Eom, Hui Fang, Christopher Fernandez, Aaron Godin, Mitchell Greer, Jill Grenon, Kristin Haider, Brynn Heckel, Kris Hennig, Nicole Hergott, Taylor Holland, David Johnson, John Klironomos, Alexander Koch, Liz Koziol, Andy Krohn, Anna Larimer, Becky Mau, Ann McCauley, Luke McCormack, Elizabeth Middleton, Meghan Midgley, R. Michael Miller, Peter Moutoglis, Bailey Nicholson, Brian Ohsowski, Gregory Penn, Michael Perkins, Krittika Petprakob, Anne Pringle, Stephen



Russell, Mark Schwartz, Ben Sullivan, Bill Swenson, Miranda Welsh, Raymond West, Steven Wilkening, and Catherine Zabinski. This project was supported financially by a working group and Distributed Graduate Seminar through the National Center for Ecological Analysis and Synthesis, which in turn is supported by the National Science Foundation (NSF; DEB-0072909), the University of California at Santa Barbara and the state of California; The Radcliffe Institute for Advanced Study at Harvard University; and the National Evolutionary Synthesis Center, which receives support from an NSF grant (EF-0423641), Duke University, the University of North Carolina, Chapel Hill and North Carolina State University. V.B.C. was supported by the NSF (grants DGE-0549505, DGE-0742483) and Loyola University Chicago. MAR was supported by an NSF Postdoctoral Research Fellowship in Biology under Grant No. DBI-12-02676 and a postdoctoral fellowship at NIMBioS, sponsored by NSF grant no. DBI-1300426 and the Univ. of Tennessee, Knoxville. J.D.H. was supported by NSF DEB Award #1119865.

### Author Contributions

V.B.C. participated in establishing the criteria used in and the vetting of papers in Phase I and II, helped design, create, and maintain MycoDB during Phase I and Phase II, helped with training in Phase II, entered and validated data in all three phases, and wrote the paper. M.A.R. entered data during Phase II, participated in extensive data validation and constructed new columns during Phase III, and commented on the manuscript. A.A. entered data during Phase II and commented on the manuscript. J.D.B. participated in establishing the criteria used in and the vetting of papers in Phase I and II, was involved in all phases of database development, including initiation of database, data entry during Phase I, training in Phase II, and data quality control, and commented on the manuscript. J.C. entered data during Phase II, participated in data validation during Phase III, and commented on the manuscript. A.C. entered data during Phase II and commented on the manuscript. J.D. entered data during Phase II and commented on the manuscript. A.F. entered data during Phase II and commented on the manuscript. M.G. helped construct plant and fungal phylogenetic trees and validate data during Phase III and commented on the manuscript. C.G. participated in establishing the criteria used in and the vetting of papers in Phase I, entered data during phase I and II, helped with training in Phase II, and commented on the manuscript. Michelle Ha entered data and conducted data validation during Phases II and III and commented on the manuscript. Miranda Hart entered data and helped with training during Phase II and commented on the manuscript. J.H. participated in data validation and quality control during Phase III and commented on the manuscript. B.J. entered data during Phase II and commented on the manuscript. N.C.J. helped acquire funding for Phase I, participated in establishing the criteria used in and the vetting of papers in Phase I, entered data into MycoDB during phase I and II, helped with training in Phase II, and commented on the manuscript. W.K. helped construct plant and fungal phylogenetic trees, validate data during Phase II and III, and commented on the manuscript. J.K. added data from EM studies into MycoDB in Phase I and commented on the manuscript. R.T.K. participated in establishing the criteria used in and the vetting of papers in Phase I, entered data during phase I and II, helped with training in Phase II, and commented on the manuscript. L.J.L. entered data during Phase II and commented on the manuscript. J.M. entered data during Phase II, conducted data validation during Phases II and III, helped construct plant and fungal phylogenies, created figures, contributed R code, and commented on the manuscript. B.G.M. participated in training during Phase II, helped curate and validate data during Phase II and III, and commented on the manuscript. J.C.M. participated in establishing the criteria used in and the vetting of papers in Phase I, entered data during Phase I, and commented on the manuscript. T.H.P. IV entered data during Phase II and commented on the manuscript. B.P. participated in the search and selection of papers in Phase II, entered data in Phase II, managed students and helped with training in Phase II, entered and validated data in Phases II and III, and commented on the manuscript. B.R. entered data during Phase II and commented on the manuscript. S.Si. participated in establishing the criteria used in and the vetting of papers in Phase I, entered data during Phase I, and commented on the manuscript. S.Sh. entered data during Phase II and commented on the manuscript. J.U. participated in establishing the criteria used in and the vetting of papers in Phase I, entered data during phase I and II, helped with training in Phase II, participated in data validation during Phase III, and commented on the manuscript. W.V. participated in data validation during Phase III and commented on the manuscript. L. W. helped design, create, and maintain MycoDB and the web-based data collection system used during Phase I and Phase II and commented on the manuscript. G.W.T.W. participated in establishing the criteria used in and the vetting of papers in Phase I, entered data during Phase I and II, helped with training in Phase II, and commented on the manuscript. P.C.Z. entered data during Phase II, participated in data validation during Phase III, and commented on the manuscript. J.D.H. helped to acquire funding for all phases of the work, helped design MycoDB, participated in data entry during Phase I, helped oversee and train during Phase II, participated in data validation and quality control during all three phases, and commented on the manuscript.

### Additional Information

Table 1 is only available in the online version of this paper.

Supplementary Information accompanies this paper at <http://www.nature.com/sdata>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Chaudhary, V. B. *et al.* MycoDB, a global database of plant response to mycorrhizal fungi. *Sci. Data* 3:160028 doi: 10.1038/sdata.2016.28 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0>

Metadata associated with this Data Descriptor is available at <http://www.nature.com/sdata/> and is released under the CC0 waiver to maximize reuse.

V. Bala Chaudhary<sup>1</sup>, Megan A. Rúa<sup>2,3</sup>, Anita Antoninka<sup>4</sup>, James D. Bever<sup>5</sup>, Jeffery Cannon<sup>6</sup>, Ashley Craig<sup>7</sup>, Jessica Duchicela<sup>8,9</sup>, Alicia Frame<sup>10</sup>, Monique Gardes<sup>11</sup>, Catherine Gehring<sup>7</sup>, Michelle Ha<sup>3</sup>, Miranda Hart<sup>12</sup>, Jacob Hopkins<sup>8</sup>, Baoming Ji<sup>13</sup>, Nancy Collins Johnson<sup>7,14</sup>, Wittaya Kaonongbua<sup>15</sup>, Justine Karst<sup>16</sup>, Roger T. Koide<sup>17</sup>, Louis J. Lamit<sup>18</sup>, James Meadow<sup>19,20</sup>, Brook G. Milligan<sup>21</sup>, John C. Moore<sup>22</sup>, Thomas H. Pendergast IV<sup>23</sup>, Bridget Piculell<sup>3</sup>, Blake Ramsby<sup>3</sup>, Suzanne Simard<sup>24</sup>, Shubha Shrestha<sup>25</sup>, James Umbanhowar<sup>26</sup>, Wolfgang Viechtbauer<sup>27</sup>, Lawrence Walters<sup>28</sup>, Gail W. T. Wilson<sup>29</sup>, Peter C. Zee<sup>30</sup> & Jason D. Hoeksema<sup>3</sup>

<sup>1</sup>Department of Environmental Science and Studies, DePaul University, Chicago, Illinois 60614, USA. <sup>2</sup>National Institute for Mathematical and Biological Synthesis, University of Tennessee, Knoxville, Tennessee 37996-3410, USA. <sup>3</sup>Department of Biology, University of Mississippi, University, Mississippi 38677, USA. <sup>4</sup>School of Forestry, Northern Arizona University, Flagstaff, Arizona 86011, USA. <sup>5</sup>Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, Kansas 66045, USA. <sup>6</sup>Colorado Forest Restoration Institute, Colorado State University, Fort Collins, Colorado 80523-1472, USA. <sup>7</sup>Department of Biological Sciences, Northern Arizona University, Flagstaff, Arizona 86011, USA. <sup>8</sup>Department of Biology, Indiana University, Bloomington, Indiana 47405, USA. <sup>9</sup>Departamento de Ciencias de la Vida, Universidad de las Fuerzas Armadas—ESPE, Sangolquí 1715231B, Ecuador. <sup>10</sup>US Environmental Protection Agency, Office of Solid Waste and Emergency Response, Washington DC 20004, USA. <sup>11</sup>Université Toulouse 3 Paul Sabatier, CNRS, ENFA; UMR5174 EDB (Évolution & Diversité Biologique); F-31062 Toulouse, France. <sup>12</sup>Department of Biology, University of British Columbia Okanagan, Kelowna BC, Canada V1V1V7. <sup>13</sup>College of Forestry, Beijing Forestry University, Beijing 100083, China. <sup>14</sup>School of Earth Sciences and Environmental Sustainability, Northern Arizona University, Flagstaff Arizona 86011, USA. <sup>15</sup>Department of Microbiology, Faculty of Science, King Mongkut's University of Technology Thonburi, Bangkok 10140, Thailand. <sup>16</sup>Department of Renewable Resources, University of Alberta, Edmonton, Canada T6G 2E3. <sup>17</sup>Department of Biology, Brigham Young University, Provo, Utah 84602, USA. <sup>18</sup>School of Forest Resources and Environmental Science, Michigan Technological University, Houghton, Michigan 49931-1295, USA. <sup>19</sup>Institute of Ecology and Evolution, University of Oregon, Eugene, Oregon 97403, USA. <sup>20</sup>Department of Land Resources and Environmental Sciences, Montana State University, Bozeman, Montana 59717, USA. <sup>21</sup>Department of Biology, New Mexico State University, Las Cruces, New Mexico 88003, USA. <sup>22</sup>Department of Ecosystem Science and Sustainability, and the Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, Colorado 80523, USA. <sup>23</sup>Department of Plant Biology, University of Georgia, Athens, Georgia 30602, USA. <sup>24</sup>Department of Forest and Conservation Sciences, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z4. <sup>25</sup>Department of Biological Sciences, Winston Salem State University, Winston-Salem, North Carolina 27110, USA. <sup>26</sup>Department of Biology, University of North Carolina, Chapel Hill, North Carolina 27599, USA. <sup>27</sup>Department of Psychiatry and Neuropsychology, Maastricht University, 6200 MD Maastricht, The Netherlands. <sup>28</sup>Software Engineering, Enova International Inc., Chicago, Illinois 60604, USA. <sup>29</sup>Department of Natural Resource Ecology and Management, Oklahoma State University, Stillwater, Oklahoma 74078, USA. <sup>30</sup>Department of Biology, California State University—Northridge, Northridge, California 91330, USA.