

DR. LEHO TEDERSOO (Orcid ID : 0000-0002-1635-1249)

Article type : Report

Co-ordinating Editor : Richard Michalet

**Global database of plants with root-symbiotic nitrogen fixation: NodDB**

Leho Tedersoo, Lauri Laanisto, Saleh Rahimlou, Aurèle Toussaint, Tiit Hallikma & Meelis Pärtel

**Tedersoo, L.** (Corresponding author, leho.tedersoo@ut.ee)<sup>1</sup>

**Laanisto, L.** (lauri.laanisto@emu.ee)<sup>2</sup>

**Rahimlou, S.** (saleh.rahimlou@ut.ee)<sup>3</sup>

**Toussaint, A.** (aurele.toussaint@ut.ee)<sup>3</sup>

**Hallikma, T.** (tiit.hallikma@emu.ee)<sup>2</sup>

**Pärtel, M.** (meelis.partel@ut.ee)<sup>3</sup>

<sup>1</sup>Natural History Museum, University of Tartu, 14a Ravila, 50411 Tartu, Estonia;

<sup>2</sup>Estonian University of Life Sciences

<sup>3</sup>Institute of Ecology and Earth Sciences, University of Tartu, 40 Lai, 51005 Tartu, Estonia.

\*Corresponding author email: leho.tedersoo@ut.ee, tel. +372 56654986

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/jvs.12627

This article is protected by copyright. All rights reserved.

## Abstract

Plants associated with symbiotic nitrogen fixing bacteria play important roles in early successional, riparian and semidry ecosystems. These so-called nitrogen fixing plants are widely used for reclamation of disturbed vegetation and improvement of soil fertility in agroforestry. Yet, available information about plants that are capable of establishing nodulation is fragmented and somewhat outdated. This article introduces the NodDB database of nitrogen fixing plants based on morphological and phylogenetic evidence (available at <http://dx.doi.org/10.15156/BIO/587469>) and discusses plant groups with conflicting reports and interpretation such as certain legume clades and the Zygophyllaceae family. During angiosperm evolution, nitrogen fixing plants became common in the fabid rather than in the 'nitrogen fixing' clade. The global GBIF plant species distribution data indicated that nitrogen fixing plants tend to be relatively more diverse in savanna and semidesert biomes. The compiled and re-interpreted information about nitrogen fixing plants enables accurate analyses of biogeography and community ecology of biological nitrogen fixation.

**Key words:** Rhizobiaceae, Frankiaceae, Nostocaceae, nodulation, Fabaceae, Zygophyllaceae, root symbiosis, nitrogen fixing clade

**Abbreviations:** APG = Angiosperm Phylogeny Group, GBIF = Global Biodiversity Information Facility, GRIN = Germplasm Resources Information Network, LPWG = Legume Phylogeny Working Group

## Introduction

Root symbiotic associations with nitrogen fixing bacteria and mycorrhizal fungi are important evolutionary adaptations of plants to compete for nutrients. Nitrogen fixing plant-bacterial associations are widely distributed across all terrestrial biomes and continents except Antarctica. Nodulated plants form important components of plant communities especially in nitrogen-limited early successional ecosystems, riparian habitats and tropical savanna and shrubland biomes (Cleveland et al. 1999). In early successional habitats, nitrogen fixing plants and their root symbiotic microbes contribute to soil development and facilitate recruitment of other plant species and consumers (Walker et al. 2003). The global symbiotic biological nitrogen fixation amounts roughly 45 Mt annually, which is the main contributor to natural terrestrial nitrogen source (Vitousek et al. 2013).

Nitrogen fixing mutualistic relationships between plant roots and bacteria have evolved multiple times in both partners (Rai et al. 2000; Santi et al. 2013; Werner et al. 2014; Doyle 2016). The differentiated forms of associations occur as root (or additionally stem) nodules, but in multiple instances plants host nitrogen fixing bacteria in undifferentiated leaf, stem or root tissues (Vessey et al. 2005; Santi et al. 2013). Rhizobiaceae ( $\alpha$ -proteobacteria) and Burkholderiaceae ( $\beta$ -proteobacteria) are the most well-known nitrogen fixing bacterial groups that nodulate mostly legumes (Fabaceae; Sprent et al. 2017). A small genus *Parasponia* (Cannabaceae) has evolved independently symbiotic associations with Rhizobiaceae (Trinick 1980). In addition, rhizobial root nodules have been reported in three zygothylaceous genera, *Tribulus*, *Fagonia* and *Zygothylum* (Mostafa and Mahmoud 1951), but this family has received limited attention in recent treatments in spite of monodominance of some species in desert habitats around the world (Sheahan 2007). Frankiaceae (Actinobacteria) form actinorhizal root nodules with genera from multiple eucosid plant families (Chaia et al. 2010). Nostocaceae (Cyanobacteria) form specific root nodules in all examined members of Cycadophyta and inhabit leaves of the angiosperm genus *Gunnera* (Gunneraceae), aquatic fern genus *Azolla* (Salviniaceae; Rai et al. 2000) and shoots of bryophytes (DeLuca et al. 2007). Proteobacterial leaf nodules occur in some species belonging to several genera of Rubiaceae, Myrsinaceae and Dioscoreaceae (Miller 1990).

In spite of the importance of nitrogen fixing symbiotic associations in agroforestry, information about the potential of nitrogen fixation in plants is scattered in multiple sources focusing on different bacterial or plant taxonomic groups. The available information is also largely conflicting (for example, 35.6% nodulation assignments at genus level do not match in Werner et al. 2014 and Li et al. 2015), which could be ascribed to differences in interpretation and level of criticism when assessing original reports or old reviews that contain multiple errors. To be able to assess the role of nitrogen fixing plants in various ecosystems and to reconstruct the evolution of nitrogen fixation in plants, information about the nitrogen fixation capacity must be as accurate as possible. To enable such analyses, we present a freely accessible database about plant genera with nodulated roots (Table S1). This represents a consensus about the nodulation and nitrogen fixing status based on several reviews and accounts for phylogenetic information to interpret the nodulation trait in unstudied groups and to detect potential erroneous reports. We illustrate the distribution of nitrogen fixing plants from the phylogenetic and macroecological perspectives.

## Materials and Methods

We examined five metastudies/reviews (Rai et al. 2000; Sprent 2009; Chaia et al. 2010; Werner et al. 2014; Li et al. 2015) and three databases - GRIN (last updated in February 2009, no longer publicly available; species-level for Fabaceae); Nodulation\_clade (Afkhami et al. 2018; accessed 25.01.2018;

genus-level for Fabaceae); and TRY (Kattge et al. 2011; accessed 25.01.2018; species-level for streptophytes) to obtain the majority of records about nodulation and non-nodulation in terrestrial plant roots. Genera with no reports or conflicting reports in the above data sets were thoroughly searched for nodulation status in the literature. In addition, genera whose nodulation reports did not match expectations based on phylogeny were further searched for additional support. We used a Boolean search combining each genus name AND 'nodulation OR fixation' in Google Scholar (as of 01.06.2017) and studied all hits with relevant matches in the whole text. Records of nodulation in Fabaceae were compared with recent molecular phylogenies (Lavin et al. 2001; Luckow et al. 2003; Bruneau et al. 2008; Bouchenak-Khelladi et al. 2010; Cardoso et al. 2013, 2015; Kyalangalilwa et al. 2013; de la Estrella et al. 2017) to detect potential gains and losses of nitrogen fixation and estimate the likelihood of erroneous records. Following Sprent (2009), legume subfamilies and tribes were considered likely or unlikely to nodulate. Smaller clades (as in Cardoso et al. 2013, 2015; de la Estrella et al. 2017) were considered likely nodulating or not based on information from sources considered reliable. Low reliability was assigned to papers in which >10% or >50% of records out of >20 and >4 records per data set, respectively, were in disagreement with other studies. Due to conflicting reports and missing data, we weighted reports from reliable sources, unreliable sources and phylogeny with a score of 4, 1 and 2, respectively. Genera with scores of >2 were considered sufficiently proven for nodulation status; genera with scores of  $\leq 2$  were considered likely or unlikely to nodulate; score of 0 was not recorded.

Since nodulation in Zygophyllaceae has received conflicting interpretations (Becking 1982; Bond 1983), we searched for direct (morphological, molecular) and indirect ( $\delta^{15}\text{N}$  signature, nitrogenase assays, soil nitrogen content and associated bacterial taxa) evidence for nitrogen fixation in this family. These results are discussed in the ecological and evolutionary perspectives of Zygophyllaceae. The phylogeny of Zygophyllaceae follows Lauterbach et al. (2016) and Sun et al. (2016). For other groups, we follow the recently updated taxonomy of the Plant List (species and genus level; [www.theplantlist.org](http://www.theplantlist.org); accessed 01.06.2017), Angiosperm Phylogeny Group (family and order level; APG IV 2016) and Legume Phylogeny Working Group (LPWG 2017). The latter was preferred in case of inconsistencies in Fabaceae. Species, genus and family names of plants were corrected according to these sources.

In the NodDB database (Table S1), all plant genera were considered associated or unassociated with rhizobia (incl. Rhizobiaceae, Burkholderiaceae), Frankiaceae or Nostocaceae. We also generated extra categories 'likely nodulated' and 'unlikely nodulated' for plant genera that possessed only phylogenetic implications or reports from unreliable sources (see above). In further binary trait-based studies, it is recommended to use these categories as nodulated and non-nodulated, respectively. Notably, all other plant genera absent from this data set are expected to lack the

nodulation capacity, although these may form undifferentiated interactions with free-living nitrogen-fixing bacteria.

## Results and Discussion

### NodDB database

The existing databases cover highly variable amount of information about plants associated with nitrogen fixing bacteria. Sprent (2009), Werner et al. (2014), Li et al. (2015), the Nodulation\_clade, GRIN and TRY databases provided information about nodulation for 391, 360, 469, 505, 490 and 1800 (220 genera in the target group) currently recognized genera, respectively. More specific reviews of Rai et al. (2000) and Chaia et al. (2010) supplemented information about 10 Cyanobacteria-associated genera and 25 actinorhizal plant genera, respectively. Combining these basic sources as well as neglected older and more recent studies increased the available information about plants with or without nitrogen fixing root nodules to 590 genera based on 9446 records. Likelihood for nodulation was assigned to further 234 genera (mainly Fabaceae and Zygophyllaceae) based on their phylogenetic relationships.

Our comparisons across reviews revealed that many of these contain highly conflicting information and/or interpretation of published work (Allen et al. 1981; Halliday and Nakao 1982; Halliday 1984; Li et al. 2015) compared with other studies, which is probably based on personal opinions of these authors. In addition, many case studies provide specific information about nodulation that is in strong disagreement with other studies (Martin and Goodding 1948; Rothschild 1950; Bonnier and Seeger 1958; Corby 1974; Lim 1977; Högberg and Nylund 1981; Faria et al. 1984; Bai et al. 1987; Moreira et al. 1992; Han 1995; Himmat Singh and Pokhriyal 1997; Roggy and Prevost 1999; Diabaté et al. 2005). Many of such incorrect reports could be ascribed to a small amount of material studied, inability to recognize developing nodules, wishful thinking and perhaps misidentification of plant taxa. All these above-cited studies were considered unreliable. The TRY database contained a large proportion of incorrect information, with 29.1% and 2.5% of genera in target and non-target groups, respectively, misassigned. False positive assignments of members of Detarioideae, Caesalpinioideae and Araucariales contribute most to the erroneous data. Revisiting the error-infested sources revealed that potential N fixation assignments were provided overoptimistically based on  $^{15}\text{N}:^{14}\text{N}$  ratio of foliage or no obvious reason whatsoever.

## Fabaceae-rhizobia associations

In general agreement with Werner et al. (2014) and Li et al. (2015), rhizobial associations are missing in the early diverging Fabaceae subfamilies Cercidoideae, Detarioideae (but see below for a potential exception), Dialioideae and Duparquetioideae. Nitrogen fixation has evolved and become widely distributed probably twice in Fabaceae: once in the caesalpinoid-mimosoid interface within Caesalpinioideae and once in the core group of Papilionoideae (Werner et al. 2014; Fig. 1). Unfortunately, the best focused and most up-to-date molecular studies lack resolution in these critical parts of the phylogeny. Within the early diverging Caesalpinioideae, the genus *Tachigali* (including recently synonymised *Sclerolobium*) and phylogenetically distant *Campsiandra* exhibit consistent reports about the presence of nodules, which are in conflict with the overall evolutionary pattern. Furthermore, there are three independent reports of nodulation in *Vouacapoua palliodor*, but only negative reports in *V. americana* and *V. macropetalum*. Similarly in Detarioideae, there are several independent reports of nitrogen fixation in the genus *Humboldtia*. These results suggest that nodulation may have evolved more than twice in Fabaceae, but it has not been related to explosive radiation of taxa in these putative additional cases.

Negative reports about nodulation status suggest that the Papilionoideae clades *Aldina*, *Amphimas*, *Cladrastis*, Exostyleae (Lecointeoid clade), Vataireoid clade, ADA clade (*Angylocalyx* + Dipterygeae + *Amburana*; sensu Cardoso et al. 2012, 2015) probably diverged before the nitrogen fixing associations evolved in this subfamily. It cannot be, however, excluded that some of these groups have secondarily lost nodulation capacity, because the basal divergence patterns in Papilionoideae have limited phylogenetic resolution (Cardoso et al. 2015). It is important to note that the ability to nodulate has been secondarily lost multiple times in the crown group of Papilionoideae: 1) *Chaetocalyx-Nissolia* group in the *Adesmia* clade (Dalbergieae); 2) genus *Bocoa* (Swartzieae tribe) and 3) *Leucomphalos-Baphiopsis* group in the Baphieae tribe. The single negative reports for the papilionoid core group genera *Platysepalum* (Diabate et al. 2005), *Barbieria* (Magalhaes and Silva 1987) and *Grazilodendron* (Faria et al. 1984) warrant independent re-evaluation.

Ancestors of the mimosoid group of Caesalpinioideae were nodulated (Sprent 2009; Werner et al. 2014). Therefore, it is remarkable that the *Adenantha* clade, *Newtonia-Fillaeopsis* group (sensu Kyalangalilwa et al. 2013) and *Cylicodiscus gabunensis* have lost their nitrogen fixing associations. In addition, nodulation capacity may have been secondarily lost in two unrelated mimosoid genera, *Parkia* (*s. stricto*) and *Zapoteca* as based on multiple negative reports. Nodulation is also missing in *Dinizia* that is traditionally considered as a mimosoid genus, but phylogenies place it within the *Dimorphandra* group of caesalpinoid legumes (Bruneau et al. 2008). Some Brazilian species of *Pterocarpus* (Papilionoideae) and American species of *Senegalia* (sect *Monacantha*;

mimosoid Caesalpinioideae) seem to have lost the ability to nodulate, representing the few rare cases of infrageneric differences in nodulation capacity in addition to *Vouacapoua* (Harrier et al. 1997; Sprent 2008).

#### Nodulation in non-legume groups

Phylogenetic studies suggest that rhizobial associations in *Parasponia* (Cannabaceae) evolved in the Pliocene, i.e. much later than in legume groups (Qian and Jin 2016; Sun et al. 2016). The presence of nodulation and nitrogen fixation ability in *Parasponia* was first suggested >100 years ago (Ham 1909), but this was re-visited only several decades later. Nodulation and growth benefits were later confirmed for several species of *Parasponia* (Trinick 1980).

All plant groups previously considered to form actinorrhizal associations with Frankiaceae were confirmed as such in several overviews and more recent case studies. In Rosaceae, the nitrogen fixing trait occurs in the early diverging Dryadeae tribe that evolved in the Late Cretaceous (Chin et al. 2014; Qian and Jin 2016), representing the oldest extant actinorrhizal group. The order Fagales contains three independent actinorrhizal groups, viz. Casuarinaceae, Myricaceae and *Alnus* (Betulaceae). In Myricaceae, only the Neocaledonian endemic *Canacomyrca* has been reported as non-nodulating (Gauthier et al. 2000). These observations, together with its phylogenetic position at the base of Myricaceae suggests that Myricaceae and *Alnus* evolved actinorrhizal associations later (Eocene) than Casuarinaceae (Paleocene; Larson-Johnson 2015; Magallón et al. 2015; Qian and Jin, 2016). Actinorrhizal habit in Rhamnaceae has probably evolved twice in tribes Colletiae and Phyliceae, with both events dating back to the Oligocene-Miocene (Onstein et al. 2015). Nodulation in Elaeagnaceae, *Datisca* (Datiscaceae) and *Coriaria* (Coriariaceae) probably evolved in the Paleocene-Eocene boundary, Miocene and Miocene, respectively (Qian and Jin 2016).

The gymnosperm families Cycadaceae, Stangeriaceae and Zamiaceae represent a single group that is nodulated by Cyanobacteria. This trait is ancestral to all modern members of Cycadophyta, because species of all examined genera host nitrogen-fixing bacteria (Rai et al. 2000). The modern cycads date back to the Late Jurassic (Qian and Jin 2016), rendering this group the oldest extant nitrogen-fixing plant clade. It is not known whether any other members of the large number of extinct gymnosperm lineages were associated with nitrogen fixing microbes.

## Zygophyllaceae, an overlooked nodulating family

Zygophyllaceae is one of the oldest groups of N fixing plants that evolved in Late Cretaceous (Qian and Jin 2016; Sun et al. 2016). Nodulation was first observed for the genus *Tribulus* >100 years ago (Isachenko 1913), but received limited attention for three decades. Nodulation and growth benefits were later confirmed for the Zygophyllaceae genera *Fagonia*, *Larrea*, *Tribulus* and *Zygophyllum* (Sabet 1946; Mostafa and Mahmoud 1951; Athar and Mahmoud 1972; Medan and Tortosa 1983), often accompanied with convincing illustrations about nodule morphology. It has remained unclear, whether Zygophyllaceae associate with rhizobia, Cyanobacteria or a variety of groups. Using electron microscopy, Mahmoud & Athar (1998, 2006) described abundant cells belonging to both Rhizobiaceae and Cyanobacteria in nodules of *Tribulus terrestris*, whereas only Rhizobiaceae spp. were described from multiple zygophyllaceous taxa previously (Sabet 1946; Mostafa & Mahmoud 1951). However, in *Zygophyllum fabago* nodules collected from Iran, we consistently detected abundant *Vibrio* sp. based on molecular analysis of the nitrogenase gene (S. Rahimlou, unpublished).

Inconsistent development of nodules, inability to detect nitrogenase activities in some studies and perhaps limited access to old literature led several authors to exclude Zygophyllaceae from the list of nitrogen fixing plants (e.g. Vessey et al. 2005; Wheeler et al. 2005; Franche et al. 2009; Werner et al. 2014; Li et al. 2015; Doyle 2016) with no commenting. Comparative studies of  $\delta^{15}\text{N}$  in plant tissues do indeed support this opinion (Shearer and Kohl 1983; Omondi et al. 2013). However, the relatively high  $\delta^{15}\text{N}$  values in leaves of Zygophyllaceae may result from non-mycorrhizal habit of this family (Brundrett 2017), which contrasts to legumes, grasses and most other co-occurring plants that form arbuscular mycorrhiza. The presence and type of mycorrhiza has a strong impact on  $\delta^{15}\text{N}$  values in seasonally dry ecosystems (Schmidt and Stewart 2003).

In addition to direct evidence for nodulation in several species of the four Zygophyllaceae genera, multiple times greater soil nitrogen (including nitrate) content occurs under the canopy of *Larrea tridentata* (Parker et al. 1982; Jenkins et al. 1988) and *Porlieria chilensis* (Gutierrez et al. 1993) compared with surrounding vegetation. Although Jenkins et al. (1988) found low nodulation capacity of soil from *L. tridentata* in roots of a legume *Prosopis* sp., the poor nodulation can be ascribed to partner specificity of symbiotic rhizobia (Mostafa and Mahmoud 1951; Bala and Giller 2001). In another study, Demba Diallo et al. (2004) found much higher proportion of Rhizobiaceae under the zygophyllaceous bush *Balanites aegyptiaca* compared with the nodulated legume *Acacia tortilis* in Senegal, suggestive of symbiotic nitrogen fixation. Besides these circumstantial evidence, Zygophyllaceae root nodules do contain rhizobia that are capable of nodulating legume species (Mostafa and Mahmoud, 1951; Athar and Mahmoud 2006). Since all four nodulated genera of Zygophyllaceae are placed in different main clades of the family (Lauterbach et al. 2016; Sun et al. 2016), it is highly probable that the capacity to fix nitrogen is an ancestral trait. Remarkably,



Zygophyllaceae falls outside the ‘nitrogen fixing clade’ of angiosperms (cf. Soltis et al. 1995; Doyle 2016) and may be older than the individual nodulated groups therein (Magallón et al. 2015; Qian and Jin 2016), suggesting that the nitrogen fixing clade should be extended to the entire Fabids (sensu APG IV 2016; Fig. 1).

These results suggest that the evolutionary ecology of plant-rhizobial associations need to be revised. Therefore, it is of particular importance to continue studies on the ecophysiology and specificity of symbiotic root associations in Zygophyllaceae. Our preliminary results indicate that individuals of *Zygophyllum fabago* do indeed associate with Rhizobiaceae (S. Rahimlou, unpublished). Furthermore, species of Zygophyllaceae have been commonly used as reference plants to quantify the proportion of nitrogen symbiotically fixed in roots of legumes (e.g. Shearer et al. 1983; Shearer and Kohl 1986; Omondi et al. 2013). Because of their C4 photosynthesis, non-mycorrhizal, and nitrogen fixing habit, Zygophyllaceae spp. are clearly unsuited for  $\delta^{15}\text{N}$  analysis-based reference for Fabaceae, indicating that the presented values for fixed nitrogen in multiple dryland studies and regional extrapolations need to be re-calculated.

#### Distribution of N fixing plants

We used the above-described nodulation assignments to illustrate the global distribution of nitrogen fixing plants based on GBIF data (<http://gbif.org/>; see figure legend), anticipating the biases in taxonomic and spatial information (Maldonado et al. 2015). Based on GBIF records, the relative species richness of nitrogen fixing plants in relation to other plants is highest in savanna and semidesert biomes across the world (Fig. 2), supporting the suggestion of Cleveland et al. (1999) that nitrogen fixation capacity is most beneficial in seasonally dry habitats and drylands, where N limitation is usually strongest.

#### Conclusive remarks

Here we propose a consensus reference database NodDB about nodulation in plant genera that accounts for critically revised records and plant phylogeny (Table S1). These data can be further used to address evolutionary ecological hypotheses regarding nodulation. Information about the potential nitrogen fixers may greatly improve estimates of nitrogen cycling in plant communities from fine to landscape scales and to plan further research on coevolution and host shifts. Knowledge about nitrogen fixing capacity facilitates development of agroforestry planning and selection of plant species for biofertilisation and reclamation of soil.

Our data raise a hypothesis that rhizobial symbiosis has evolved more than twice (up to six times) in Fabaceae, once in Cannabaceae and once in Zygophyllaceae (Fig. 1). Bacterial associations in Zygophyllaceae and small groups of detarioid and caesalpinoid legumes certainly require further attention. Actinorhizal associations evolved on nine independent occasions in angiosperms, whereas cyanobacterial nodules are known only from all three extant families of cycads and as co-colonisers in Zygophyllaceae. Although the last nitrogen fixing plant groups were discovered in the 1970s, it is possible that some rare nitrogen fixing taxa remain to be discovered from poorly studied habitats.

Understanding the distribution of nitrogen fixation among plants is far from complete. Resolving the multiple potential gains and losses in Fabaceae certainly requires further confirmation of nodulation or the lack of it in critical taxa outlined above. Phylogenomics analyses will provide improved taxonomic resolution and deeper insights into the evolutionary history of nodulation in Fabaceae (Doyle 2016). Comparative genomics of nitrogen fixing bacterial groups are likely to reveal certain functional differences among these symbioses.

#### **Acknowledgements**

We thank M. Schori for communicating nodulation data in GRIN and M. Zobel for intellectual support during project initiation. This project received funding from the Estonian Science Foundation grants PUT1399, PUT1409, MOBJD276, Estonian Ministry of Education and Research (IUT20-29), and by the European Union through the European Regional Development Fund (Centre of Excellence EcolChange). No conflict of interest is declared.

#### **References**

- Afkhami, M.E., Mahler, D.L., Burns, J.H. et al. 2018. Symbioses with nitrogen- fixing bacteria: nodulation and phylogenetic data across legume genera. *Ecology*, in press.
- Allen, O.N. & Allen, E.K. 1981. *The Leguminosae. A source book of characteristics, uses and nodulation*. University of Wisconsin Press, Madison, WI.
- APG IV. 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society* 181: 1-20.
- Athar, M. & Mahmood, A. 1972. Root nodules in some members of Zygophyllaceae growing at Karachi University Campus. *Pakistan Journal of Botany* 4: 209-210.

- Bai, X.L., Tang, D.J. & Xu, X.H. 1987. A survey on the nodulation status of some mimosaceous and caesalpinaceae plants in Guangxi. *Guihaia* 7: 61-65.
- Bala, A. & Giller, K.E. 2001. Symbiotic specificity of tropical tree rhizobia for host legumes. *New Phytologist* 149: 495-507.
- Becking, J.H. 1982. Nitrogen fixation in nodulated plants other than legumes. *Advances in Agricultural Microbiology* 1: 89-110.
- Bond, G. 1983. Taxonomy and distribution of non-legume nitrogen-fixing. In: Gordon, J.C. & Wheeler, C.T. (eds). *Biological nitrogen fixation in forest ecosystems: foundations and applications*, pp. 55-88.
- Bonnier, C. & Seeger, J. 1958. *Symbiose" rhizobium"-légumineuses en région équatoriale: (deuxième communication)*. Institut national pour l'étude agronomique du Congo Belge (INEAC).
- Bouchenak-Khelladi, Y., Maurin, O., Hurter, J. & Van der Bank, M. 2010. The evolutionary history and biogeography of Mimosoideae (Leguminosae): an emphasis on African acacias. *Molecular Phylogenetics and Evolution* 57: 495-508.
- Brundrett, M.C. 2017. Global diversity and importance of mycorrhizal and nonmycorrhizal plants. *Ecological Studies* 230: 533-556.
- Bruneau, A., Mercure, M., Lewis, G.P. & Herendeen, P.S. 2008. Phylogenetic patterns and diversification in the caesalpinoid legumes. *Botany* 86: 697-718.
- Cardoso, D., Pennington, R.T., Queiroz, L.P. et al. 2013. Reconstructing the deep-branching relationships of the papilionoid legumes. *South African Journal of Botany* 89: 58-75.
- Cardoso, D., Queiroz, L.P., Pennington, R.T. et al. 2012. Revisiting the phylogeny of papilionoid legumes: new insights from comprehensively sampled early-branching lineages. *American Journal of Botany* 99: 1991-2013.
- Cardoso, D., São-Mateus, W.M., Cruz, D.T. et al. 2015. Filling in the gaps of the papilionoid legume phylogeny: The enigmatic Amazonian genus *Petaladenium* is a new branch of the early-diverging Amburaneae clade. *Molecular Phylogenetics and Evolution* 84: 112-124.
- Chaia, E.E., Wall, L.G. & Huss-Danell, K. 2010. Life in soil by the actinorhizal nodule endophyte *Frankia*. A review. *Symbiosis* 51: 201-226.
- Cleveland, C.C., Townsend, A.R., Schimel, D.S., Fisher, H., Howarth, R.W., Hedin, L.O., Perakis, S.S., Latty, E.F., Von Fischer, J.C., Elseroad, A. & Wasson, M.F. 1999. Global patterns of terrestrial biological nitrogen (N<sub>2</sub>) fixation in natural ecosystems. *Global Biogeochemical Cycles* 13: 623-645.
- Corby, H.D. 1974. Systematic implications of nodulation among Rhodesian legumes. *Kirkia* 9: 301-329.
- de la Estrella, M., Forest, F., Wieringa, J.J., Fougere-Danezan, M. & Bruneau, A. 2017. Insights on the evolutionary origin of Detarioideae, a clade of ecologically dominant tropical African trees. *New Phytologist* 214: 1722-1735.

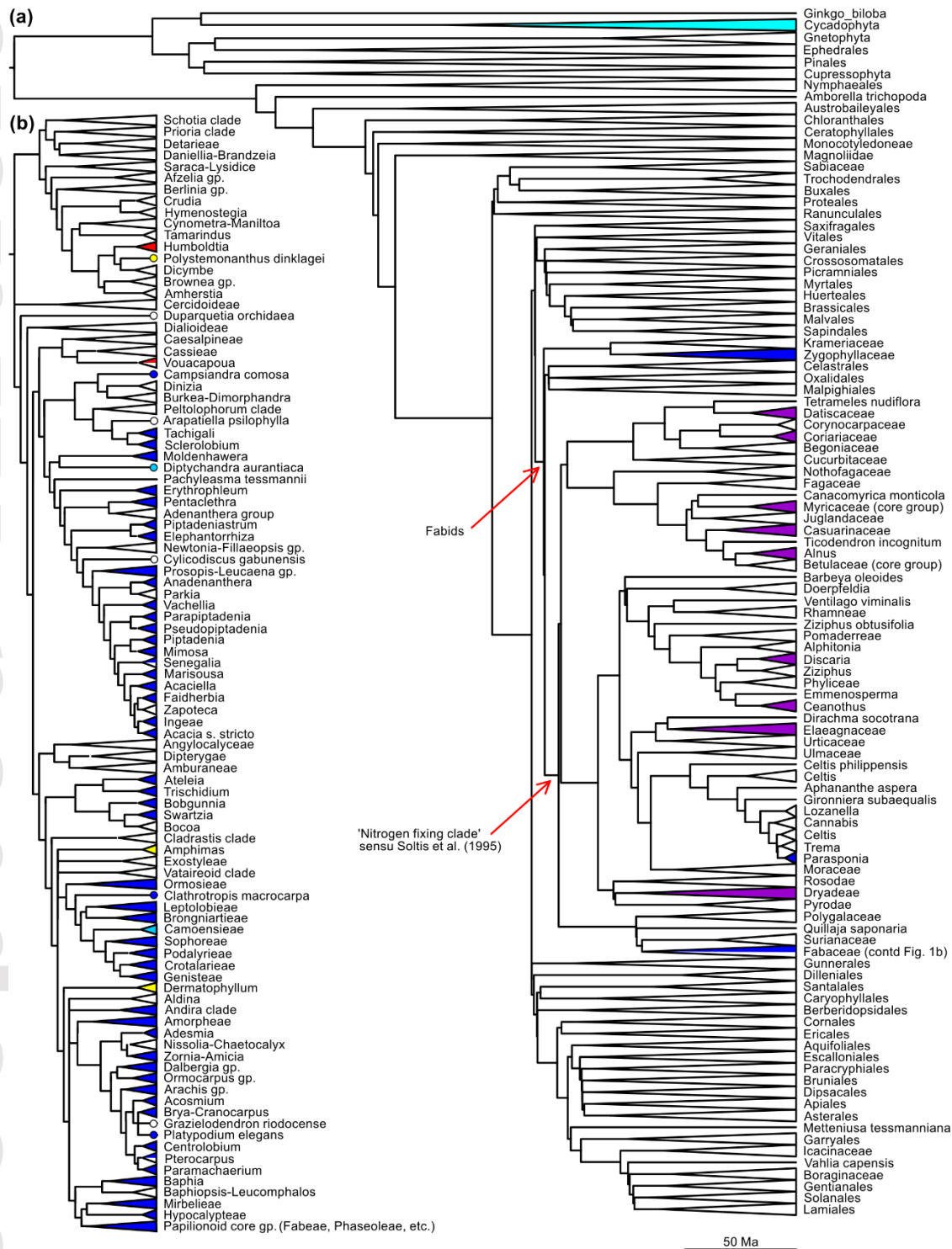
- DeLuca, T.H., Zackrisson, O., Gentili, F., Sellstedt, A. & Nilsson, M.-C. 2007. Ecosystem controls on nitrogen fixation in boreal feather moss communities. *Oecologia* 152: 121–130.
- Demba Diallo, M., Willems, A., Vloemans, N. et al. 2004. Polymerase chain reaction denaturing gradient gel electrophoresis analysis of the N<sub>2</sub>-fixing bacterial diversity in soil under *Acacia tortilis* ssp. *raddiana* and *Balanites aegyptiaca* in the dryland part of Senegal. *Environmental Microbiology* 6: 400-415.
- Diabaté, M., Munive, A., Faria, S.M., Ba, A., Dreyfus, B. & Galiana, A. 2005. Occurrence of nodulation in unexplored leguminous trees native to the West African tropical rainforest and inoculation response of native species useful in reforestation. *New Phytologist* 166: 231-239.
- Doyle, J.J. 2016. Chasing unicorns: Nodulation origins and the paradox of novelty. *American Journal of Botany* 103: 1865-1868.
- Faria, S.M., Franco, A.A., Menandro, M.S. et al. 1984. Levantamento da nodulação de leguminosas florestais nativas na região sudeste do Brasil. *Pesquisa Agropecuaria Brasiliensis* 19: 143-153.
- Franche, C., Lindström, K. & Elmerich, C. 2009. Nitrogen-fixing bacteria associated with leguminous and non-leguminous plants. *Plant and Soil* 321: 35-59.
- Gauthier, D., Jaffré, T. & Prin, Y. 2000. Abundance of *Frankia* from *Gymnostoma* spp. in the rhizosphere of *Alphitonia neocaledonica*, a non-nodulated Rhamnaceae endemic to New Caledonia. *European Journal of Soil Biology* 36: 169-175.
- Gutiérrez, J.R., Meserve, P.L., Contreas, L.C., Vásquez, H. & Jaksic, F.M. 1993. Spatial distribution of soil nutrients and ephemeral plants underneath and outside the canopy of *Porlieria chilensis* shrubs (Zygophyllaceae) in arid coastal Chile. *Oecologia* 95: 347-352.
- Halliday, J. & Nakao, P.L. 1982. Masterlist of woody species under consideration as nitrogen-fixing trees. *Resource Documents on Nitrogen-fixing Trees* 82.3.
- Halliday, J. 1984. Register of nodulation reports for leguminous trees and other arboreal genera with nitrogen fixing members. *Nitrogen Fixing Tree Research Reports* 2: 38-45.
- Ham, S.P. 1909. Disussion in "Groene Bemesting" ("Green manure"). *Indisch Landbouw Syndicaat* 11: 25-27.
- Han S-F. 1995. A nodulation survey of tree legumes in China. *Journal of Nanjing Forestry University* 19: 89-92.
- Harrier, L.A., Whitty, P.W., Sutherland, J.M. & Sprent, J.I. 1997. Phenetic investigation of non-nodulating African species of *Acacia* (Leguminosae) using morphological and molecular markers. *Plant Systematics and Evolution* 205: 27-51.
- Himmat Singh, K.C. & Pokhriyal, T.C. 1997. Survey and identification of some leguminous nitrogen-fixing plants in eastern part of Nepal. *Annals of Forestry* 5: 205-213.
- Högberg, P. & Nylund, J.E. 1981. Ectomycorrhizae in coastal miombo woodland of Tanzania. *Plant and Soil* 63: 283-289.

- Jenkins, M.B., Virginia, R.A. & Jarrell, W.M. 1988. Depth distribution and seasonal populations of mesquite-nodulating rhizobia in warm desert ecosystems. *Soil Science Society of America Journal* 52: 1644-1650.
- Kattge, J., Diaz, S., Lavorel, S. et al. 2011. TRY—a global database of plant traits. *Global Change Biology* 17: 2905-2935.
- Kyalangalilwa, B., Boatwright, J.S., Daru, B.H., Maurin, O. & Bank, M. 2013. Phylogenetic position and revised classification of *Acacia* s.l. (Fabaceae: Mimosoideae) in Africa, including new combinations in *Vachellia* and *Senegalia*. *Botanical Journal of the Linnean Society* 172: 500-523.
- Larson-Johnson, K. 2015. Phylogenetic investigation of the complex evolutionary history of dispersal mode and diversification rates across living and fossil Fagales. *New Phytologist* 209: 418–435.
- Lauterbach, M., van der Merwe, P.D., Keßler, L., Pirie, M.D., Bellstedt, D.U. & Kadereit, G. 2016. Evolution of leaf anatomy in arid environments—a case study in southern African *Tetraena* and *Roepera* (Zygophyllaceae). *Molecular Phylogenetics and Evolution* 97: 129-144.
- Lavin, M., Pennington, R.T., Klitgaard, B.B., Sprent, J.I., de Lima, H.C. & Gasson, P.E. 2001. The dalbergioid legumes (Fabaceae): delimitation of a pantropical monophyletic clade. *American Journal of Botany* 88: 503-533.
- Li, H.L., Wang, W., Mortimer, P.E. et al. 2015. Large-scale phylogenetic analyses reveal multiple gains of actinorhizal nitrogen-fixing symbioses in angiosperms associated with climate change. *Scientific Reports* 5: 14023.
- Lim, G. 1977. Nodulation of tropical legumes in Singapore. *Tropical Agriculture* 54: 135-141.
- LPWG 2017. A new subfamily classification of the Leguminosae based on a taxonomically comprehensive phylogeny. *Taxon* 66, 44-77.
- Isachenko, B.L. 1913. Über die Wurzelknolchen bei *Tribulus terrestris* L. *Bulletin de Jardin Botanique St. Petersburg* 13: 23-31.
- Luckow, M.E., Miller, J.T., Murphy, D.J & Livshultz, T.A. 2003. A phylogenetic analysis of the Mimosoideae (Leguminosae) based on chloroplast DNA sequence data. *Advances in Legume Systematics* 10: 197-220.
- Magalhães, F.M. & Silva, M.F. 1987. Associações *Rhizobium*-leguminosas no estado de Rondônia. *Acta Amazonica* 17: 7-18.
- Magallón, S., Gómez-Acevedo, S., Sánchez-Reyes, L.L. & Hernández-Hernández, T. 2015. A metacalibrated time-tree documents the early rise of flowering plant phylogenetic diversity. *New Phytologist* 207: 437-453.
- Mahmood, A. & Athar, M. 1998. Cyanobacterial root nodules in *Tribulus terrestris* L. (Zygophyllaceae). In: Malik, K.A., Mirza, M.S. & Ladha, J.K. (eds). *Nitrogen Fixation with Non-Legumes*. Pp. 345-350.
- Mahmood, A. & Athar, M. 2006. Scanning electron microscopic observations on micro-organisms in the root nodules of *Tribulus terrestris* L. (Zygophyllaceae). *Scanning* 28: 233-235.

- Maldonado, C., Molina, C.I., Zizka, A. et al. 2015. Estimating species diversity and distribution in the era of Big Data: to what extent can we trust public databases? *Global Ecology and Biogeography* 24: 973-984.
- Martin, W.P. & Goodding, L.N. 1948. Observations on the nodulation of leguminous plants of the southwest. *Plant Study Series* 4: 1-10.
- Medan, P.D. & Tortosa, R.D. 1983. Nodos radicales en *Larrea* (Zygophyllaceae). *Boletín de Sociedad Argentina Botánica* 22: 221-236.
- Miller, I.M. 1990. Bacterial leaf nodule symbiosis. *Advances in Botanical Research* 17: 163-234.
- Moreira, F.A.S, Silva, M.F. & Faria S.M. 1992. Occurrence of nodulation in legume species in the Amazon region of Brazil. *New Phytologist* 121: 563-570.
- Mostafa, M. & Mahmoud, M. 1951. Bacterial isolates from root nodules of Zygophyllaceae. *Nature* 167: 446-447.
- Omondi, S.M., Eunice, G., Jesse, T.N. & Gachene, C.K. 2013. Nitrogen fixation by natural populations of *Acacia senegal* in the drylands of Kenya using  $^{15}\text{N}$  natural abundance. *Arid Land Research and Management* 27: 327-336.
- Onstein, R.E, Carter, R.J., Xing, Y., Richardson, J.E. & Linder, H.P. 2015. Do Mediterranean-type ecosystems have a common history?—Insights from the Buckthorn family (Rhamnaceae). *Evolution* 69: 756-771.
- Parker, L.W., Fowler, H.G., Ettershank, G. & Whitford, W.G. 1982. The effects of subterranean termite removal on desert soil nitrogen and ephemeral flora. *Journal of Arid Environments* 5: 53-59.
- Qian, H. & Jin, Y. 2016. An updated megaphylogeny of plants, a tool for generating plant phylogenies and an analysis of phylogenetic community structure. *Journal of Plant Ecology* 9: 233-239.
- Rai, A.N., Söderbäck, E. & Bergman, B. 2000. Cyanobacterium-plant symbioses. *New Phytologist* 147: 449-481.
- Roggy, J.C. & Prevost, M.F. 1999. Nitrogen-fixing legumes and silvigenesis in a rain forest in French Guiana: a taxonomic and ecological approach. *New Phytologist* 144: 283-294.
- Rothschild, D.E. 1950. Nodulación en leguminosas subtropicales de la flora Argentina. *Revista del Museo Argentino de Ciencias Naturales "Bernardino Rivadavia" e Instituto Nacional de Investigación de las Ciencias Naturales, Botánicas* 3: 267-286.
- Sabet, Y.S. 1946. Bacterial root nodules in the Zygophyllaceae. *Nature* 157: 656.
- Santi, C., Bogusz, D. & Franche, C. 2013. Biological nitrogen fixation in non-legume plants. *Annals of Botany* 111: 743-767.
- Schmidt, S. & Stewart, G.R. 2003.  $\delta^{15}\text{N}$  values of tropical savanna and monsoon forest species reflect root specialisations and soil nitrogen status. *Oecologia* 134: 569-577.
- Sheahan, M.C. 2007. Zygophyllaceae. In: Kubitzky K (ed). *The Families and Genera of Vascular Plants*, pp. 488-500.

- Shearer, G. & Kohl, D.H. 1986. N<sub>2</sub>-fixation in field settings: estimations based on natural <sup>15</sup>N abundance. *Functional Plant Biology* 13, 699-756.
- Shearer, G., Kohl, D.H., Virginia, R. et al. 1983. Estimates of N<sub>2</sub>-fixation from variation in the natural abundance of <sup>15</sup>N in Sonoran desert ecosystems. *Oecologia* 56: 365-373.
- Soltis, D.E., Soltis, P.S., Morgan, D.R., Swensen, S. M., Mullin, B. C., Dowd, J. M., & Martin, P. G. 1995. Chloroplast gene sequence data suggest a single origin of the predisposition for symbiotic nitrogen fixation in angiosperms. *Proceedings of the National Academy of Sciences USA* 92: 2647-2651.
- Sprent, J.I. 2009. *Legume nodulation*. John Wiley and Sons.
- Sprent, J.I., Ardley, J. & James, E.K. 2017. Biogeography of nodulated legumes and their nitrogen-fixing symbionts. *New Phytologist* 215: 40-56.
- Sprent, J.I. 2008. Evolution and diversity of legume symbiosis. In: Dilworth, M.J. *Nitrogen-fixing Leguminous Symbioses*, pp. 1-21.
- Sun, M., Naeem, R., Su, J.X. et al. 2016. Phylogeny of the Rosidae: A dense taxon sampling analysis. *Journal of Systematics and Evolution* 54: 363-391.
- Trinick, M.J. 1980. Growth of *Parasponia* in agar tube culture and symbiotic effectiveness of isolates from *Parasponia* spp. *New Phytologist* 85: 37-45.
- Vessey, J.K., Pawlowski, K. & Bergman, B. 2005. Root-based N<sub>2</sub>-fixing symbioses: Legumes, actinorhizal plants, *Parasponia* sp. and cycads. *Plant and Soil* 274: 51-78.
- Vitousek, P.M., Menge, D.N.L., Reed, S.C. & Cleveland, C.C. 2013. Biological nitrogen fixation: rates, patterns and ecological controls in terrestrial ecosystems. *Philosophical Transactions of the Royal Society B* 368: 20130119.
- Walker, L.R., Clarkson, B.D., Silvester, W.B. & Clarkson, B.R. 2003. Colonization dynamics and facilitative impacts of a nitrogen-fixing shrub in primary succession. *Journal of Vegetation Science* 14: 277-290.
- Werner, G.D.A., Cornwell, W.K., Sprent, J.I., Kattge, J. & Kiers, E.T. 2014. A single evolutionary innovation drives the deep evolution of symbiotic N<sub>2</sub>-fixation in angiosperms. *Nature Communications* 5: 4087.
- Wheeler, C.T., Mark, G.L., Milne, J. & Hooker, J.E. 2005. Para-nodulation of non-nitrogen fixing plants. In: Ramasamy, C. (ed). *Perspectives of Agricultural Research and Development*, pp. 197-203.

Figure legends

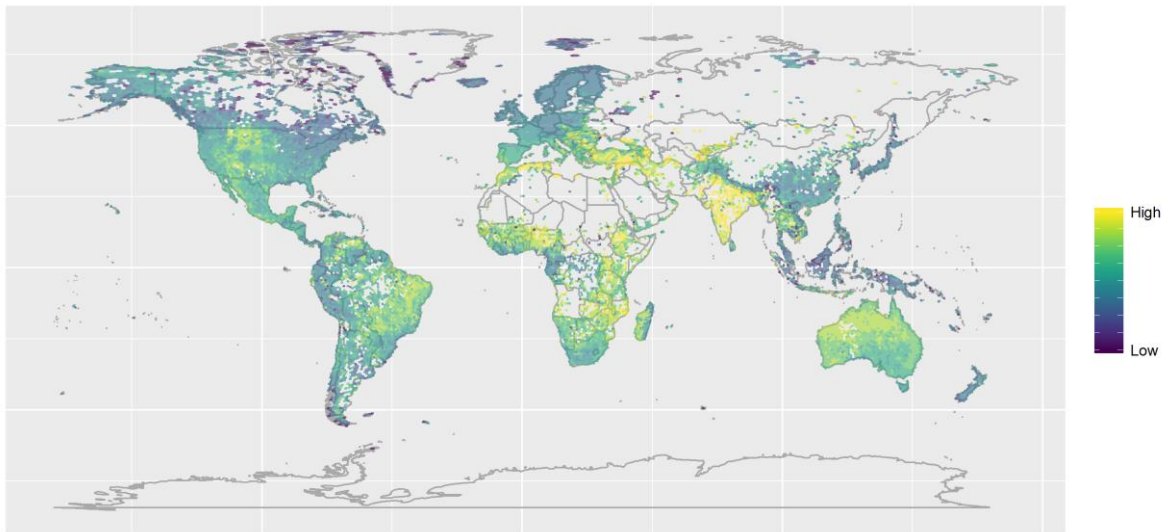


**Fig. 1.** Phylogenetic placement of plants with nitrogen fixing root nodules: (a) overall distribution of nitrogen fixing plant clades in the phylogram of seed plants; and (b) distribution of nitrogen fixing trait in the Fabaceae family. The phylograms are taken from (a) Qian & Jin (2016) and (b) compiled from Bruneau et al. (2008), Cardoso et al. (2015), de la Estrella et al. (2017) and LPWG (2017). Cyan



branches, cyanobacterial associations; violet branches, actinorrhizal associations; blue branches, rhizobial associations. In (b): white branches, non-rhizobial associations; yellow branches, unlikely rhizobial associations; light blue, likely rhizobial associations; red, reported rhizobial associations, which require confirmation.

Percentage of N-fix species (quantiles)



**Fig. 2.** Species richness of nitrogen fixing vascular plants relative to total species richness as based on all GBIF vascular plant records with coordinates (GBIF.org, 10th November 2017, GBIF Occurrence Download <https://doi.org/10.15468/dl.4nqoqv>). Relative richness is calculated per equal-area polygons (ISEA3H) with size ca 7000 km<sup>2</sup>. Colors show quantiles. Only terrestrial polygons hosting more than 50 records are shown.

### Supporting information

Table S1. The NodDB database.