



<https://doi.org/10.11646/megataxa.10.1.6>

<https://zoobank.org/urn:lsid:zoobank.org:pub:B90211AE-07D7-4D02-A1F2-844070995C63>

Towards digital descriptions of all extant reptile species

PETER UETZ^{1,*}, YAA ADARKWA DARKO^{1,2} & DUSTIN ZELIFF^{1,3}

¹Center for Biological Data Science, Virginia Commonwealth University, Richmond, Virginia 23284, USA

²[✉ adarkwadarkya@vcu.edu](mailto:adarkwadarkya@vcu.edu); [ORCID: https://orcid.org/0000-0001-8972-0140](https://orcid.org/0000-0001-8972-0140)

³[✉ zeliffdj@vcu.edu](mailto:zeliffdj@vcu.edu); [ORCID: https://orcid.org/0009-0000-4790-7396](https://orcid.org/0009-0000-4790-7396)

*Corresponding author: [✉ uetz@vcu.edu](mailto:uetz@vcu.edu); [ORCID: https://orcid.org/0000-0001-6194-4927](https://orcid.org/0000-0001-6194-4927)

Abstract

Vertebrate databases have been slow to digitize species descriptions. One of them, the Reptile Database (<http://www.reptile-database.org>), has accumulated ~8,000 species descriptions for ~3,000 species of snakes, ~5,000 species of lizards, and ~150 species of turtles and crocodiles. Here we discuss how this data contributes to character analysis, species identification, but also to integration with other data sources such as citizen science observations (which depend on correct identifications). Importantly, the data described here may serve as training data for machine learning projects and we present examples of species comparisons using ChatGPT. While these AI-driven comparisons are still erroneous, we expect substantial improvements in the near future. We request the herpetological community to help complete our public collection of species descriptions and suggest that other species databases follow suit and provide similar data for their taxa.

Key words: diagnosis, Squamata, Sauria, Serpentes, Testudines, Crocodylia, ChatGPT

Introduction

About 20,000 new species are described every year, including about 200 new reptiles. Species and their definitions are the basis of all biological research, ranging from medical experiments in mice to the ecology of insects in tropical forests. With the advent of digital publishing, most species descriptions are now available online. However, having a paper with a new species on a website does not mean that the information in that paper is easily accessible. It may be behind a paywall or simply difficult to find. Equally important, species information needs to be accessible in a format that can be reused. In other words, species data needs to be FAIR (findable, accessible, interoperable, and reusable).

There are surprisingly few databases that systematically collect species descriptions. Even among vertebrates, we are not aware of any database or website that has a comprehensive collection of species descriptions. For instance, **Fishbase** has morphological information for “only” 19,130 out of ~35,000 species (55%) but

this statistic has not been updated on their website since 2015 (<https://fishbase.se/home.htm>). **AmphibiaWeb** currently has 3,753 species accounts which contain descriptions, out of 8,654 species total (43%) (Uetz *et al.* 2021, Michelle Koo, pers. comm.). While there are huge amounts of data for birds, relatively little of that seems to be available in a structured database such as **AVONET** and even the latter focuses on a few key traits such as wing and beak measurements (Tobias *et al.* 2022). The situation for mammals is not much better, with no species descriptions in the **Mammal Diversity Database** (Mammal Diversity Database 2023). In defense of birds and mammals, they always had excellent comprehensive books and monographs with very detailed descriptions, *e.g.*, the Lynx editions of bird and mammal handbooks (del Hoyo *et al.* 2011, Wilson *et al.* 2019).

The Reptile database started out as a species list, like many other databases (Uetz & Etzold 1996). Before it started to add actual descriptions, it first completed reference lists for all original species descriptions (Uetz 2010; Uetz & Stylianou 2018). Having references to original descriptions is not sufficient though. As indicated above, the data from these descriptions need to be accessible and they need to be in a useful format. Although it is a particular challenge to structure free text descriptions as in most species descriptions, we have started to copy species descriptions from the original and secondary literature into the Reptile Database where they are made publicly available. Here we summarize the status of this project and how it contributes to reptile taxonomy. Importantly, we will discuss how that information can be made more useful by employing artificial intelligence, such as implemented by ChatGPT.

This report should be of interest to all taxonomists, as all of them work with species descriptions at some level, especially when specimens need to be identified in the field or in collections. We suggest that similar projects are carried out on all other taxa, and our preliminary estimates suggest that descriptions of all ~2 million known species could be completed for US \$10 to \$50 million, or as low as US \$5 per species (although lower costs may require substantial automation and division of labor).

Materials and Methods

We have collected **descriptions** from the **original literature** (Uetz 2010, Uetz & Stylianou 2018) as well as the **secondary literature**, such as revisions, books, or new descriptions that redescribed (or redefined) previously described species. The sources for these descriptions are provided in the Reptile Database. Species descriptions are usually copied verbatim into the database, but frequently slightly edited, often for OCR errors, but also in other ways (such as merging hyphenated words at the end of lines). There are certainly still many OCR errors and other problems left, but they will most likely not affect our results in a significant way.

Data format: At this point, descriptions and diagnoses are captured as free text, only delimited by headers, as in the following example for the pipe snake *Cylindrophis ruffus*, the type species of the genus *Cylindrophis* (which is also the type genus of the family Cylindrophiidae):

Diagnosis (genus): (1) relatively blunt head with minute eyes, head not distinct from neck, bearing a mental groove; (2) absence of true gastrosteges, with ventral scales only slightly larger than or equal in size to dorsal scales; (3) presence of a pair of pelvic spurs (= cloacal spurs) in both sexes; (4) a very short tail, often with conspicuous ventral coloration; and (5) contrasting light and dark ventral blotching (Kieckbusch *et al.* 2016).

Original description: “Corpore aequali, ruffo, lineis trans-versalibus albis interruptis; abdomine vario” (Laurenti 1768) translated to “Body uniform, red, broken white transverse bands; abdomen various” by (Adler *et al.* 1992).

Diagnosis: *Cylindrophis ruffus* is distinguished from all congeners by having the following characters: 19 midbody scale rows (vs. 17 in *C. engkariensis*; 21 in *C. isolepis*, *C. lineatus*, *C. maculatus*, *C. yamdena*; 23 in *C. aruensis*, *C. opisthorhodus*), 186–197 ventrals (vs. 233–275 in *C. melanotus*; 201–225 in *C. burmanus*), wide and constant bands encircling dark body (vs. dorsum uniform black with no cross bands in *C. boulengeri*; narrow and alternating bands on paler body in *C. burmanus*), an interrupted and wide band on the nape (vs. no ring on the nape in *C. boulengeri*; a complete and narrow ring encircling the nape in *C. burmanus*) (Amarasinghe *et al.* 2015).

Diagnosis, description, treatment. We distinguish between *diagnosis* and *description*, with “*diagnosis*” being explicitly used as such in the source publication. Ideally, a *diagnosis* lists the traits that distinguish two or more species, even that is often not the case when additional species or specimens are described or when the definition or range of a species changes. A “*description*” could be any description, without further specification in a source publication, and in fact, is often called differently, such as “identification”, “characteristics”, “definition” etc. (if there is such an explicit statement, we try to use that). However, in many older publications, there is no such differentiation, and we just call these texts “description”.

The content of the description field in the database may vary, and may contain a diagnosis or description of the species or diagnostic characters of genera (included in the entry of the type species of a genus, as in the example

above). Ideally, **original descriptions** are given too, but this is only true in a minority of cases (although true for practically all recently described species). A **treatment** in this context is a block of text such as a diagnosis, the description of a type, a section on comparisons or variation etc. We subsume all these treatments under one *description* per species.

Character counts and traits. A custom Python script was used to extract the various *treatments* for each species and their character counts, using the April 2023 version of the Reptile Database as input. We measure information content as *character* count, that is, the number of letters (or “bytes”) in a description. This must not be confused with a morphological **character**, which we call “**trait**” in this text. Given that species descriptions are usually lists of traits, the number of letters is a rough approximation for the information content of a description (although not always true, e.g., when a description has an extended discussion of geographic variation of one or a few traits).

Language statistic. The language of each entry was determined using a separate python script. The language scan script used a predefined list of words for each of the four most common languages in the document (English, German, French, and Spanish).

Taxonomy. The taxonomic backbone for this study was derived from the June 2023 version of the Reptile Database (<http://www.reptile-database.org>).

Images. We used the number of images available in the Reptile Database in June 2023. For this study we ignored other image sources, such as those listed in Marshall *et al.* (2020).

ChatGPT. Queries were submitted to <https://chat.openai.com> in July and August 2023 as described in Results and Supplementary Data.

Results

Species descriptions. The Reptile Database has collected descriptions of about 8,000 species over the past 20 years (Fig. 1A). While we do distinguish between *diagnoses* and *descriptions*, the distinction is simply based on what they are called in their source publication, often without clear differences. Hence we have not attempted to analyze the difference between the two categories in detail. In order to provide an overview of their text content, we provide a summary of their sizes in Fig. 1B. The “size” of descriptions (including diagnoses) varies greatly, with very short descriptions, as commonly used in many older publications, whether these diagnoses are useful or not. For instance, the diagnosis of *Chalcides striatus* has only 86 characters and reads “*characterized by nine, sometimes 11 brownish dorsal, somewhat fuzzy lines*” (Orsini & Cheylan 1981). On the other end of the spectrum, 72 species currently have more than 10,000 characters of text. The largest one at the time of writing was that of *Sceloporus poinsetti*, with 33,482 characters, corresponding to 5,439 words or more than 7 pages of text (at a font size of 11). However, this description also contains morphological details about various subspecies.

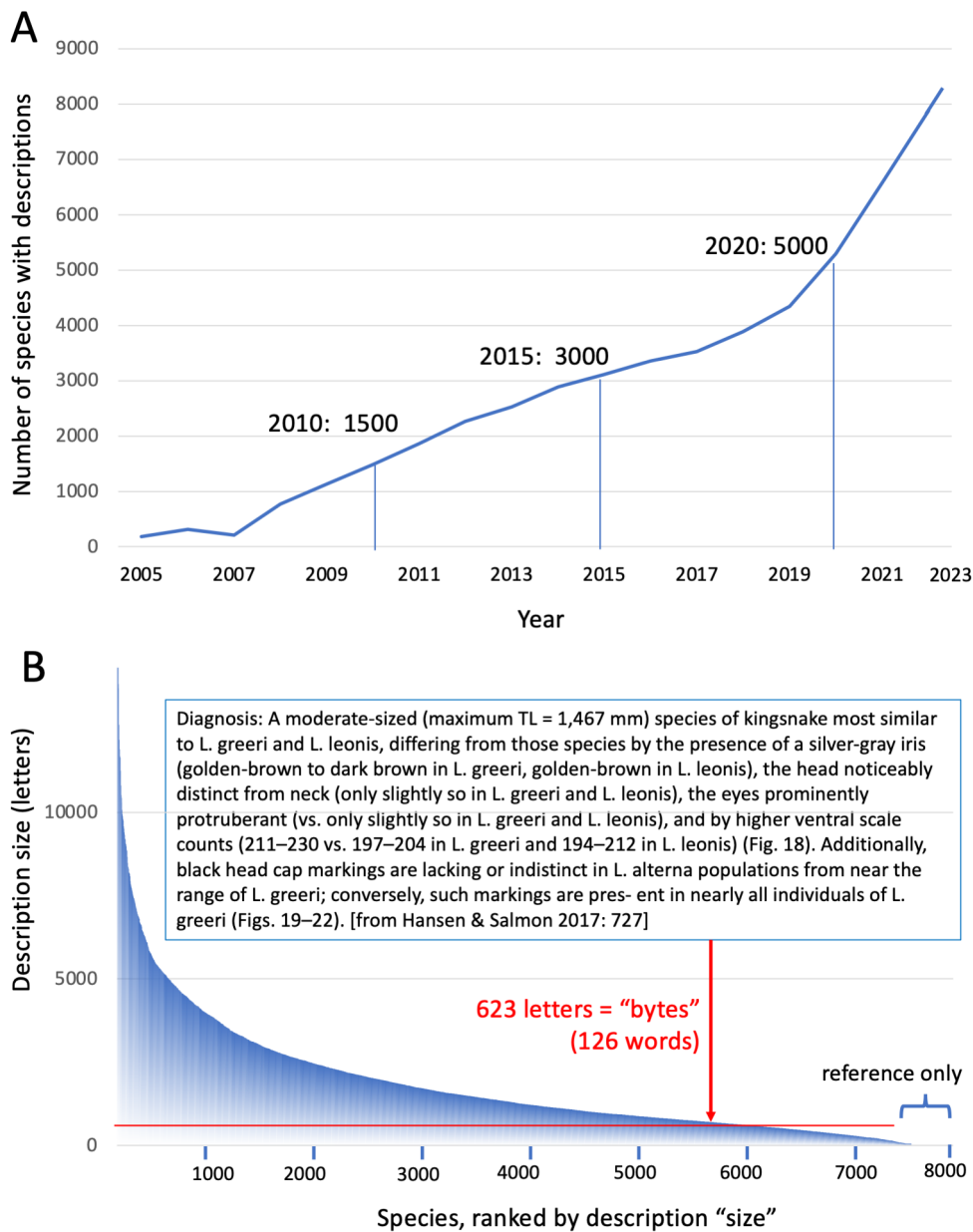


FIGURE 1 Species descriptions in the Reptile Database over time and ranked by information content. **(A)** Timeline over which descriptions have been added to the Reptile Database. **(B)** Descriptions ranked by “size”. Each species with a description is shown as a vertical bar, with the height indicating its size in characters (i.e., “bytes” or letters without spaces, Y axis). The X axis shows the ranked series of species. One example is shown in the insert, namely the description of *Lampropeltis alterna*. This description has 623 characters and is on the shorter end of the spectrum. The tail end of this graph shows ~700 species in which the description only consists of a reference, e.g., “Description: McDowell 1979: 51” (here: for *Candoia aspera* Günther 1877), hence they are very short.

Our descriptions also include ~700 species that have only references in their “description” field. We have not attempted to provide precise counts for very short descriptions vs. references only, because both data types are currently collected in the same data field. As a rule of thumb, descriptions that have less than about 120 characters almost always consist of one or two references *only*. For instance, 623 species have less than 100 characters and are thus mostly references. The median text size of descriptions currently is 1,272 characters and the

average is 1,914 characters. For an example see **Figure 1** and for a summary statistic see **Table 1**.

Overall, the Reptile Database has diagnoses and descriptions for 964 and 5,736 species, respectively (**Table 1**). This distinction is important because diagnoses only contain diagnostic traits and are thus much shorter than descriptions, which can be very detailed and may include traits that are not informative for species identification. However, they may become useful for specific studies. Notably, variation is not included in many species

descriptions, given the small number of specimens usually collected from a limited number of localities, hence trait data may substantially change with additional sampling.

TABLE 1. Summary statistic for description information in the Reptile Database (April 2023 release). Note that body size is also collected in a separate data field in the database. “Definition” and “diagnosis” are often used synonymously. A block or section of text, such as “comparisons” is called “treatment” (see text for definitions).

Treatment	Number of Species	Average Body size (bytes)
Color, coloration	847	248
Description (species)	964	2,175
Diagnosis (species)	5,736	244
Description (genus)	104	315
Diagnosis (genus)	931	374
Comparisons	1,170	1,778
Variation	246	92
Measurement	77	173
Size	131	128
Definition	203	209

Languages. Of the 8,223 entries with descriptions, 6,451 were in English, 54 in French, 107 in German, and 38 in Spanish. 43 were determined to be mixed language articles, and 810 were determined to not have enough words to reliably indicate the language. Of those that did not have enough words, 716 were short entries that only contained 150 or less characters, and thus likely had no description at all (but rather consisted of a short literature reference only, as mentioned in Fig. 1). Some descriptions contain text in English and a second (usually original)

language. We found 43 such mixed-language cases. This is relevant for downstream text analysis, which is almost always tailored to a specific language (see discussion).

Multi-species descriptions. We wondered how many descriptions contain information on other species. That is, if species A does not have a diagnosis, maybe another species B may have information or diagnostic characters (traits) that distinguish the two species. In fact, 4,206 species descriptions mention at least one other species and almost 2,000 of those mentioned species do not have a description on their own. This leaves less than 2,000 species without any descriptive information (Fig. 2 A).

Images. Images are often more informative than text descriptions, given that images can contain much more information than text. Hence, the Reptile Database has also collected images of reptiles, including both live images (Marshall *et al.* 2020) as well as standardized images of museum specimens. The latter project is described in more detail in a separate paper (Uetz *et al.* 2023).

Within the last 3 years, the number of species with images in the Reptile Database grew from 5,144 (Marshall *et al.* 2020) to 6,326 in July 2023 (52.7% of all species) (Fig. 2B). Hence, we wondered how many species may have a description *and/or* photos. Interestingly, of the ~8,000 species with descriptions, 36% also have at least one photo, but about 4,000 have only a description but no photo. By contrast, almost 2,000 species have photos but no description (Fig. 2B). Overall, only ~1,800 species have neither a description, nor a photo. These species are our main targets to collect either descriptions or photos.

Description of type specimens. Ideally, each species description should be based on a type series, with detailed descriptions of the primary type and the paratypes. However, this is often not the case, and especially in older descriptions the type specimen is often not explicitly identified, e.g., by a catalogue number (although the specimen may be identified based on implicit information).

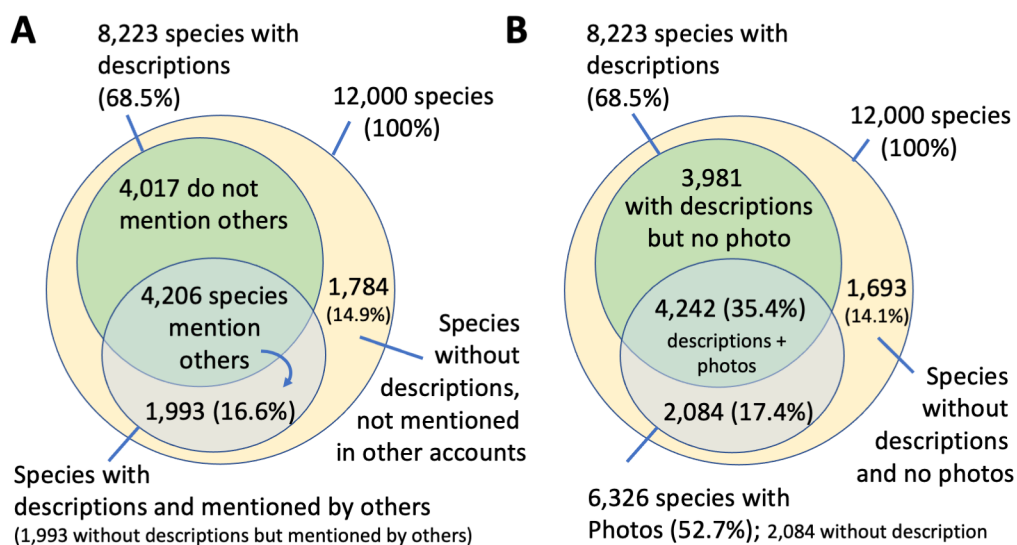


FIGURE 2. Number of species with descriptions and images in the Reptile Database (A) 4,206 species descriptions contain information about other species, so even though the latter species do not have a description, information about them may be derived from those species that do. (B) Coverage of species by descriptions and/or photos.

In any case, we have included the explicit descriptions of types, when available.

Taxonomic breakdown. Overall, the Reptile Database has descriptions for about 70% of all reptile species. For 27 out of the 92 families currently recognized, the database has descriptions of all species, even though most of these families have a relatively small number of species. The largest family with a complete set of descriptions is the Homalopsidae (57 species). However, for some groups, especially for turtles, representation is more limited. For example, the Emydidae (58 species), Geoemydidae (71), and Testudinidae (47) have only 15–20% of their species described. Nevertheless, the largest families are relatively well represented: the Colubrids (*sensu lato*, 2,100 species) have 72%, the skinks (1,740 species) 65%, and the geckos (*s.l.*, Gekkonidae, 1,553 species) 81% of their species described.

Similarly, out of 1,252 genera, 526 have descriptions for all species, but 148 genera have no descriptions at all (all with fewer than 10 species). The remainder is somewhere in between. The largest genus with all species described is *Hemiphyllodactylus* with 56 species (not counting *Stenocercus* with 80 species, many of which only have references to papers with descriptions).

We have mapped all descriptions to a phylogenetic tree (Fig. 3) which shows the clades and their absolute and relative number of descriptions. A complete list of species is provided in Supplementary Table S1.

Species comparisons using ChatGPT: One possible application of species descriptions is to identify a species or to compare multiple species side by side. This is difficult with text, given that human readers have to jump back and forth and find the correct characters (traits). However, with tools like ChatGPT we can compare traits directly, *e.g.* by asking ChatGPT to place those traits side by side in a table. For instance, ChatGPT can compare the diagnosis of *Cylindrophis ruffus* to the diagnosis of *C. subocularis*, two pipe snakes that occur on the island of Java. The query may look like this:

“Extract morphological characters from the following 2 species descriptions and compare them in a table: The first species is *C. ruffus* with the following diagnosis [insert diagnosis from above]. The second species, *Cylindrophis subocularis*, “can be readily distinguished from all congeners by the following combination of characters: (1) presence of a single subocular scale, positioned between 3rd and 4th or 4th and 5th supralabial, contacting postocular and separating 4th or 5th supralabial from orbit (Fig. 4B); (2) prefrontal in very narrow contact with or separated from orbit; (3) 19 smooth dorsal scale rows at midbody; (4) 6–7 supralabials; (5) 6–7 infralabials; (6) 190–196 ventrals; (7) 6–7 subcaudals; (8) 40–48 transverse light ventral blotches, and (9) light blotches on lateral surfaces of prefrontals.” (Kieckbusch *et al.* 2018).

The output in this example is shown in Table 2.

TABLE 2. Table created by ChatGPT from the two descriptions of *Cylindrophis ruffus* and *C. subocularis*. Dashes indicate that no data was provided in the input diagnoses, which is a deficiency of the diagnoses, not necessarily of ChatGPT. Compare to Table 3.

Trait	<i>Cylindrophis ruffus</i>	<i>Cylindrophis subocularis</i>
Midbody Scale Rows	19	19
Ventrals	186–197	190–196
Body Bands	Wide and constant	-
Nape Band	Interrupted and wide	-
Subocular Scale	-	Single subocular
Prefrontal-Orbit Contact	-	Very narrow contact or separated from orbit
Supralabials	-	6–7
Infralabials	-	6–7
Subcaudals	-	6–7
Ventral Blotches	-	40–48
Prefrontal Lateral Blotches	-	Light blotches present

The *Cylindrophis* example shows that many descriptions are often not using the same traits, so that a direct comparison is difficult, if not impossible (in this example no trait is clearly unique or simply not given). However, if descriptions use a similar format and more or less the same traits, ChatGPT can produce useful comparisons, as shown in Table 3 (comparing species of the snake genus *Ialtris*). However, even in the latter case, many coloration features were left out from the table, although ChatGPT is able to return them when asked explicitly for colors.

We have carried out comparisons of many other species using ChatGPT and provide input and output for 4 genera (*Anomalepis*, *Dendragama*, *Eurydactylodes*, *Ialtris*) in Supplementary Data. The only criteria for their selection were that they have a small number of species (four each), that they represent two snakes and two lizards and that they had diagnoses in the Reptile Database (from the same or different sources).

None of the comparisons were error-free, but two of them (*Eurydactylodes*, *Ialtris*) were good or even very good, with few errors (see Supplementary Data for details). The main determinant for good comparisons was that the diagnosis or description comes from the same source and/or uses the same traits, the same terms, and ideally the same order and amount of information. A detailed analysis of the performance of ChatGPT and other LLMs is beyond the scope of this study though and we will revisit this topic in future papers.

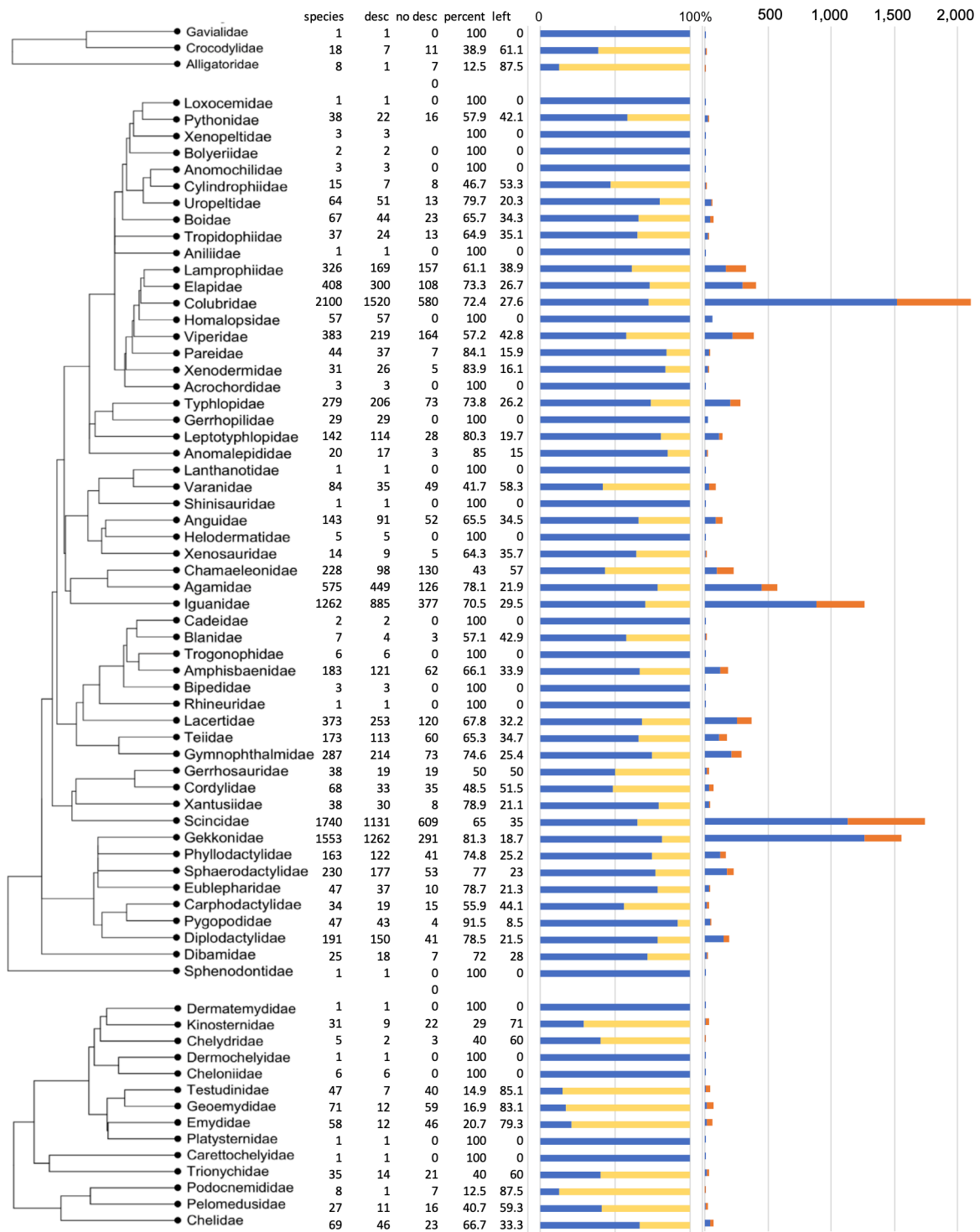


FIGURE 3. Currently available descriptions across different families. The 5 number columns show the number of species in each family (species), the number of species with descriptions or without (desc, no desc), the percentage of species with descriptions (percent), and the number of species without (left). The left bar chart shows the percentage of species with/out descriptions (blue/yellow). The right bar chart shows the absolute number of species with/out descriptions (blue/orange). Note that Iguanidae contains all 8 families now considered Iguanidae (*s.l.*), *e.g.* Opluridae *etc.* The Lamprophiidae (*s.l.*) also contain Atractaspididae, Psammophiidae, and Pseudoxyrhophiidae. Gymnophthalmidae (*s.l.*) contain Alopoglossidae and Anguidae (*s.l.*) contain Diploglossidae.

TABLE 3. A comparison of the four species of the snake genus *Ialtris*, generated from text descriptions by ChatGPT. Color traits are not shown. For complete table and descriptions see **Supplementary data.**

Trait	Species			
	<i>Ialtris agyrtes</i>	<i>Ialtris dorsalis</i>	<i>Ialtris haetianus</i>	<i>Ialtris parishi</i>
SVL (mm)	Up to 576	Up to 990	Up to 305	Up to 745
Scale rows at midbody	19	19	19	19
Dorsal scales	Smooth	Smooth	Without apical pits	Smooth
Ventrals (females)	160–175	180–192	-	163–172
Ventrals (males)	-	179–191	132–144	163–167
Subcaudals (females)	57–63	98–109	40–51	-
Subcaudals (males)	-	99–115	45–54	77 (1 male)
Anal	Divided	-	-	-
Loreal	1	-	Absent	-
Preocular	1	1	1/1	1
Postocular	Usually 2	2	2/2	Strongly modally 2
Supralabials	Usually 7	Usually 7	7	7
Infralabials	Usually 9	Usually 9	8	Usually 9

Discussion

Species descriptions are a critical component of a species' definition. However, with changing species concepts and increasing amounts of data, species boundaries constantly change. This is critical for all downstream usages or species names, which critically depend on constantly updated species descriptions.

The species descriptions collected by us are not only incomplete—they are far from perfect, as the comparisons with ChatGPT show. However, the database is constantly updated, and descriptions are added, extended, or replaced. In the early years of the database, we have focused on collecting diagnoses, as they appeared to be sufficient to identify a specimen. However, it turned out that these are too limited for many uses. For instance, the diagnosis published in a new species description may be based only on a few specimens and thus may not cover the variation within a species (*i.e.*, additional specimens). Hence, we have moved on to provide full descriptions, including the original description, as well as more recent, updated descriptions, and sections on variation and comparisons, whenever available.

It will require a detailed analysis of the descriptions in the database to determine which **diagnoses or descriptions are sufficient** to diagnose the species within a genus or to diagnose the genera within a family. This is beyond the scope of this report.

In the future, descriptions will need to be imported into databases like the Reptile Database in an automated fashion. **Plazi.org** has developed tools to do that (Agosti *et al.* 2019). While they extract descriptions and other information (“treatments”) into their TreatmentBank (Guidoti *et al.* 2021), their coverage is currently too small for most reptile genera with data for only about 2,500 squamates. However, their process is a model for

the future and we hope that they can provide descriptions of all new species described in the future and possibly retroactively for older descriptions. Equally important, it will be critical that Plazi provides their data to specialized species databases such as the Reptile Database for further critical analysis and improvement.

Will species databases be required in the future, given the rapid development of **artificial intelligence** and Large Language Models (LLMs)? ChatGPT can already read a practically unlimited number of papers and extract species descriptions, process and compare them, and produce analytical output in many formats. At this point (October 2023), LLMs are not quite ready for routine use and their database is still too small for truly comprehensive analyses but this will quickly change. In any case, our database will provide an excellent training data set for ChatGPT and other LLMs, so we are confident that our data will be useful for some time.

This leaves the question how to **integrate species descriptions with other online data**, such as observations (*e.g.*, those from iNaturalist), museum specimens (VertNet), or DNA sequence data. This is yet another issue beyond the scope of this study, but an important aspect to keep in mind. We are confident that the scientific community will manage to organize these entities with the help of stable identifiers and thus to organize them in a large metaverse of FAIR biological information.

Outlook

Species descriptions and especially diagnoses are never “final”, given that new species and specimens are being described and taxonomy is changing, *e.g.*, by phylogenetic re-organization. Hence, we ask the herpetological community to submit corrections or updated species

descriptions to the Reptile Database. In the future, automated analysis of species descriptions may often require as many traits as possible, as most descriptions will be read by automated systems but not by humans, hence few people need to read all those details. Although human-curated databases will be required for some time, the long-term goal will be a largely automated, “living monograph” on all reptiles, hopefully expanded to all living animal and plant species in the not-too-distant future. That said, we also expect new species to be described by semi-automated LLMs which will generate template descriptions which can be modified for specific characters (Güntsch *et al.* 2019). This will speed up species descriptions but hopefully also improve efforts to integrate and constantly evaluate threat status and thus nature protection in a world of rampant human overpopulation, habitat destruction, and biodiversity loss.

Author contributions

PU collected most of the data, carried out the ChatGPT analysis, and wrote most of the manuscript. YAD analyzed and broke down the data into “treatments”. DZ did the language analysis and helped with the species name identification across species descriptions.

Acknowledgments

We thank Trang Ly and Sahil Sewani for an earlier version of a script to identify species names in descriptions. Michelle Koo provided data on AmphibiaWeb.

References

- Adler, K., Zhao, E. & Darevsky, I.S. (1992) First records of the pipe snake (*Cylindrophis*) in China. *Asiatic Herpetological Research*, 4, 37–41.
- Agosti, D., Catapano, T., Sautter, G. & Egloff, W. (2019) The Plazi Workflow: The PDF prison break for biodiversity data. *Biodiversity Information Science and Standards*, 3, e37046. <https://doi.org/10.3897/biss.3.37046>
- Amarasinghe, A.A.T., Campbell, P.D., Hallermann, J., Sidik, I., Supriatna, J. & Ineich, I. (2015) Two new species of the genus *Cylindrophis* Wagler, 1828 (Squamata: Cylindrophidae) from Southeast Asia. *Amphibian & Reptile Conservation*, 9, 34–51.
- Guidoti, M., Sokolowicz, C., Simoes, F., Gonçalves, V., Ruschel, T., Alvares, D. & Agosti, D. (2021) TreatmentBank: Plazi’s strategies and its implementation to most efficiently liberate data from scholarly publications. *Biodiversity Information Science and Standards*, 5, e75690. <https://doi.org/10.3897/biss.5.75690>
- Güntsch, A., Luther, K., Müller, A., Kohlbecker, A., Piltzner, P., Kilian, N., Henning, T., Von Mering, S., Von Raab-Straube, E., Kusber, W.-H. & Berendsohn, W.G. (2019) The Platform for Cybertaxonomy: Standards, services and tools. *Biodiversity Information Science and Standards*, 3, e37177. <https://doi.org/10.3897/biss.3.37177>
- Hansen, R.W. & Salmon, G.T. (2017) Distribution analysis, taxonomic updates, and conservation status for the *Lampropeltis mexicana* group (Serpentes: Colubridae). *Mesoamerican Herpetology*, 4 (4), 700–758. [<http://mesoamericanherpetology.com/uploads/3/4/7/9/34798824/mh4-4hansen-and-salmon.pdf>]
- del Hoyo, J., Elliott, A. & Christie, D.A. (Eds) (2011) *Handbook of the Birds of the World, volume 16. Tanagers to New World Blackbirds*. Lynx Edicions, Barcelona, 894 pp.
- Kieckbusch, M., Mader, F., Kaiser, H. & Mecke, S. (2018) A new species of *Cylindrophis* Wagler, 1828 (Reptilia: Squamata: Cylindrophidae) from Boano Island, northern Maluku Province, Indonesia. *Zootaxa*, 4486, 236–250. <https://doi.org/10.11646/zootaxa.4486.3.2>
- Mammal Diversity Database (2023) Mammal Diversity Database (Version 1.11) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.7830771>
- Marshall, B.M., Freed, P., Vitt, L.J., Bernardo, P., Vogel, G., Lotzkat, S., Franzen, M., Hallermann, J., Sage, R.D., Bush, B., Duarte, M.R., Avila, L.J., Jandzik, D., Klusmeyer, B., Maryan, B., Hošek, J. & Uetz, P. (2020) An inventory of online reptile images. *Zootaxa*, 4896, 251–264. <https://doi.org/10.11646/zootaxa.4896.2.6>
- Orsini, J.P.G. & Cheylan, M. (1981) *Chalcides chalcides* - Erzschiele. In: *Böhme, W. (ed.) Handbuch der Reptilien und Amphibien Europas, Band 1, Echsen I*. Akademische Verlagsgesellschaft, Wiesbaden, pp. 318–337.
- Tobias, J.A., Sheard, C., Pigot, A.L., Devenish, A.J.M., Yang, J., Sayol, F., Neate-Clegg, M.H.C., Alioravainen, N., Weeks, T.L., Barber, R.A., Walkden, P.A., MacGregor, H.E.A., Jones, S.E.I., Vincent, C., Phillips, A.G., Marples, N.M., Montaña-Centellas, F.A., Leandro-Silva, V., Claramunt, S., Darski, B., Freeman, B.G., Bregman, T.P., Cooney, C.R., Hughes, E.C., Capp, E.J.R., Varley, Z.K., Friedman, N.R., Korntheuer, H., Corrales-Vargas, A., Trisos, C.H., Weeks, B.C., Hanz, D.M., Töpfer, T., Bravo, G.A., Remeš, V., Nowak, L., Carneiro, L.S., Moncada R., A.J., Matysioková, B., Baldassarre, D.T., Martínez-Salinas, A., Wolfe, J.D., Chapman, P.M., Daly, B.G., Sorensen, M.C., Neu, A., Ford, M.A., Mayhew, R.J., Fabio Silveira, L., Kelly, D.J., Annorbah, N.N.D., Pollock, H.S., Grabowska-Zhang, A.M., McEntee, J.P., Carlos T. Gonzalez, J., Meneses, C.G., Muñoz, M.C., Powell, L.L., Jamie, G.A., Matthews, T.J., Johnson, O., Brito, G.R.R., Zyskowski, K., Crates, R., Harvey, M.G., Jurado Zevallos, M., Hosner, P.A., Bradfer-Lawrence, T., Maley, J.M., Stiles, F.G., Lima, H.S., Provost, K.L., Chibesa, M., Mashao, M., Howard, J.T., Mlamba, E., Chua, M.A.H., Li, B., Gómez, M.I., García, N.C., Päckert, M., Fuchs, J., Ali, J.R., Derryberry, E.P., Carlson, M.L., Urriza, R.C., Brzeski, K.E., Prawiradilaga, D.M., Rayner, M.J., Miller, E.T., Bowie, R.C.K., Lafontaine, R., Scofield, R.P., Lou, Y., Somarathna, L., Lepage, D., Illif, M., Neuschulz, E.L., Templin, M., Dehling, D.M., Cooper, J.C., Pauwels, O.S.G., Analuddin, K., Fjeldsá, J., Seddon, N., Sweet, P.R., DeClerck, F.A.J., Naka, L.N., Brawn, J.D., Aleixo, A., Böhning-Gaese, K., Rahbek, C., Fritz, S.A., Thomas, G.H. & Schleuning, M. (2022) AVONET: morphological, ecological and geographical data for all birds T. Coulson

- (Ed). *Ecology Letters*, 25, 581–597.
<https://doi.org/10.1111/ele.13898>
- Uetz, P. (2010) The original descriptions of reptiles. *Zootaxa*, 2334, 59–68.
<https://doi.org/10.11646/zootaxa.2334.1.3>
- Uetz, P., Uetz, P., Koo, M.S., Aguilar, R., Brings, E., Catenazzi, A., Chang, A.T., Chaitanya, R., Freed, P., Gross, J., Hammermann, M., Hošek, J., Lambert, M., Sergi, Z., Spencer, C.L., Summers, K., Tarvin, R., Vredenburg, V.T. & Wake, D.B. (2021) A Quarter Century of Reptile and Amphibian Databases. *Herpetological Review*, 52, 246–255.
- Uetz, P. & Etzold, T. (1996) The EMBL/EBI Reptile Database. *Herpetological Review*, 27, 174–175.
- Uetz, P., Stoope, S. & Gbdamosi, Z. (2023) Reference image database of reptiles. *Manuscript in preparation*.
- Uetz, P. & Stylianou, A. (2018) The original descriptions of reptiles and their subspecies. *Zootaxa*, 4375 (2), 257–264.
<https://doi.org/10.11646/zootaxa.4375.2.5>
- Wilson, D.E., Mittermeier, R.A. & Lacher, T.E. (editors) (2019) *Handbook of the Mammals of the World*. Lynx Edicions. [<https://www.lynxeds.com/product/handbook-of-the-mammals-of-the-world-volumes-1-9/>]

Supplementary information

Examples for 2 species diagnoses, as cited in the main text:

Diagnosis: *Cylindrophis ruffus* is distinguished from all congeners by having the following characters: 19 midbody scale rows (vs. 17 in *C. engkariensis*; 21 in *C. isolepis*, *C. lineatus*, *C. maculatus*, *C. yamdena*; 23 in *C. aruensis*, *C. opisthorhodus*), 186–197 ventrals (vs. 233–275 in *C. melanotus*; 201–225 in *C. burmanus*), wide and constant bands encircling dark body (vs. dorsum uniform black with no cross bands in *C. boulengeri*; narrow and alternating bands on paler body in *C. burmanus*), an interrupted and wide band on the nape (vs. no ring on the nape in *C. boulengeri*; a complete and narrow ring encircling the nape in *C. burmanus*) (Amarasinghe *et al.* 2015).

Diagnosis: *Cylindrophis subocularis*, “can be readily distinguished from all congeners by the following combination of characters: (1) presence of a single subocular scale, positioned between 3rd and 4th or 4th and 5th supralabial, contacting postocular and separating 4th or 5th supralabial from orbit (Fig. 4B); (2) prefrontal in very narrow contact with or separated from orbit; (3) 19 smooth dorsal scale rows at midbody; (4) 6–7 supralabials; (5) 6–7 infralabials; (6) 190–196 ventrals; (7) 6–7 subcaudals; (8) 40–48 transverse light ventral blotches, and (9) light blotches on lateral surfaces of prefrontals.” (Kieckbusch *et al.* 2018).

See Table 2 in the main text for a comparison generated by ChatGPT.

Additional ChatGPT queries

Note that ChatGPT queries are not always reproducible as OpenAI seems to change their algorithms or their LLMs.

Text in red is data that is misrepresented or organized erroneously in the output file of ChatGPT.

Query 1

Compare the characters of the following 4 species of the genus *Anomalepis*, with their characters in rows and the species in columns, and save as spreadsheet: [all inputs are diagnoses, without sources added]

Species 1 (*Anomalepis aspinosus*): Teeth in both upper and lower jaws. Nostril between two nasals; dorsal scales 320–343 from rostral to terminal scute of tail. Scale rows about middle of body 24–26, frontal with area about that of a prefrontal; four upper and three lower labials. No trace of a spine on terminal plate. Scales with minute pits, usually single. (Taylor 1939: 92)

Species 2 (*Anomalepis colombia*): *Anomalepis colombia* is distinguished from its congeners by having 31/29/27 scales in rows around anterior / midbody / posterior body, and 387 scales on dorsal middle row (vs. 26/24/22 and 283 scales in *A. aspinosus*; 26/26/24 and 304–308 in *A. flavapices*; and 25–26/22–24/21–23 and 277–312 *A. mexicanus*). (MARRA-SANTOS & REIS 2019)

Species 3 (*Anomalepis flavapices*): This species can be distinguished from all other members of the genus by the absence of brown pigment on the yellowish white head and end of the tail, and by the number of scale rows from the rostral to the tail tip, which are 304–308. It is distinguished from all species except *aspinosus* by the number of scale rows around the body which is 26. (Peters 1957: 3)

Species 4 (*Anomalepis mexicana*): A tiny uniformly glossy brown snake, without enlarged ventral scales and with the minute eyes hidden under the head scales. (Savage 2002: 554).

Output: (red values missing from output table, hence they have been added manually; output had gaps instead, blue values are in wrong rows or duplicated in extra rows)

Characters	<i>aspinosus</i>	<i>colombia</i>	<i>flavapices</i>	<i>mexicana</i>
Teeth	Present	-	-	-
Nostril position	Between two nasals	-	-	-
Dorsal scales	320-343 from rostral to terminal scute of tail	-	-	-
Scale rows (around middle of body)	24-26	-	26	-
Frontal scale area	About size of prefrontal	-	-	-
Upper labial scales	4	-	-	-

...Continued on the next page

TABLE. (Continued)

Characters	<i>aspinosus</i>	<i>colombia</i>	<i>flavapices</i>	<i>mexicana</i>
Lower labial scales	3	-	-	-
Terminal plate spine	Absent	-	-	-
Pitted scales	Usually single	with minute pits	-	-
Scale counts (Anterior/Midbody/ Posterior)	26/24/22	31/29/27	26/26/24	25–26/22– 24/21–23
Dorsal middle row scale count	283	387	304–308	277–312
Head and tail coloration	-	-	Absence of brown pigment on yellowish white	-
Scale count (Rostral to tail tip)	-	-	304–308	-
Scale rows (around body)	-	-	26	-
Enlarged ventral scales	-	-	-	Absent
Eye position	-	-	-	Minute eyes hidden under head scales

Query 2: Compare the following 4 species and put their characters in a table with species in columns and characters in rows:

Diagnosis: The following combination of characters distinguishes *Dendragama boulengeri* from its congeners: (1) enlarged scales below eye separated from supralabials by continuous row of small lorilabials; (2) enlarged tympanic scale intermediate in size (7–14% of head length) and set far back on head (orbit–meatus 24–32% of head length); (3) sublabial tubercular scales 0–3, usually 1, in front of postriental modified scale; (4) gulars small, 30–38; (5) dorsal crest of 11–18 projecting scales, serrate with obvious gaps between enlarged projecting scales; (6) dorsolateral crest of more or less continuous posttympanic and dorsolateral series; (7) scales around midbody 62–84; (8) scales of lower flanks heterogenous in size, smooth to feebly keeled; (9) ventrals heavily keeled, 50–63; (10) no white sublabial stripe extending from below eye to level of ear, however 1–3 white spots usually present in this region; (11) most specimens with prominent black prescapular blotch edged posteriorly in white; (12) proximal half of tail with 7–9 dark brown or green bands; (13) buccal epithelium and tongue yellow or orange [Harvey *et al.* 2017: 83]

Diagnosis: The following combination of characters distinguishes *Dendragama australis* from its congeners: (1) enlarged scales below eye broadly contacting supralabials; (2) enlarged tympanic scale relatively small (6–10% of head length) and set relatively close to eye (orbit–meatus 21–36% of head length); (3) sublabial tubercular scales 3–5, usually 4, in front of postriental modified scale; (4) gulars large, 15–30; (5) dorsal crest of 19–37 projecting scales, serrate and continuous along back without obvious gaps; (6) dorsolateral crest broken: posttympanic series of enlarged, modified scales straight or curving upward, dorsolateral series curving downward; (7) scales around midbody 61–94; (8) scales

of lower flanks heterogenous; (9) ventrals smooth or very feebly keeled, 44–60 from preaxial border of arm to vent; (10) sharp white stripe or line of broken white spots on sublabials from level of eye to below or behind ear; (11) **brown band on neck**, but large black prescapular blotch absent; (12) proximal half of tail with 8–13 dark brown or green bands; (13) buccal epithelium and tongue yellow or orange. (Harvey *et al.* 2017).

Diagnosis: The following combination of characters distinguishes *Dendragama dioidema* from its congeners: (1) enlarged scales below eye broadly contacting supralabials; (2) tympanum relatively large (11–16% of head length) and close to eye (orbit–meatus 23–29% of head length); (3) sublabial tubercular scales 0–2, usually 1, in front of postriental modified scale; (4) gulars large, 16–22; (5) dorsal crest of 9–19 projecting scales, serrate with obvious gaps between enlarged projecting scales; (6) dorsolateral crest of more or less continuous posttympanic and dorsolateral series; (7) scales around midbody 57–77; (8) scales of lower flanks heterogenous in size, smooth to feebly keeled; (9) ventrals smooth or very feebly keeled, 49–63; (10) no white sublabial stripe; however, one or two white or pale yellow spots (**corresponding to sublabial tubercular scale and postriental modified scales**) usually present in area below rictus and ear; (11) no distinctive prescapular blotch edged posteriorly in white; band crossing neck interrupted by narrow white or yellow line; (12) proximal half of tail with 6–10 dark brown or green bands; (13) buccal epithelium cream, tongue pink to red. (Harvey *et al.* 2017)

Diagnosis: *Dendragama schneideri* is A species reaching **at least 201 mm in total length (SVL)** and distinguished from congeners by the following characters: (1) midbody scales 58–67; (2) **dorsal scales** heterogeneous across flanks (Fig 6); (3) strongly keeled white/yellow scales randomly distributed **along flanks** (more numerous and distinct in females); (4) ventral scales 48–59; (5) **banding pattern along flanks** often muddled, but typically

vertical when present; (6) mouth and tongue pink to red in life; (7) narrow, vertical black **stripes across dorsal crest, limbs, digits and most of tail**; (8) female color in life dark brown, yellow and black with amber coloration on underside, while males green and lacking amber coloration along ventral surface; (9) dorsal and nuchal crest clearly separated by 5–9 dorsal scales; (10) dorsal crest serrate, extending to base of the tail, comprised of 23–31 projecting, triangular scales; (11) a series of 3–4

enlarged tubercles present along the chin of males and females; (12) A series of 12–18 strongly keeled, white/yellow femoral spines present (combined count on both sides) (Shaney *et al.* 2020).

Output with **character numbers from Harvey *et al.* 2017 added manually in red**. The section mostly derived from Shaney *et al.* 2020 is indicated by a table separator and **green character numbers**. The cells with **red text** represent erroneously copied text from *D. schneideri*.

Characters	<i>Dendragama bouleengeri</i>	<i>Dendragama australis</i>	<i>Dendragama dioidema</i>	<i>Dendragama schneideri</i>
(1) Enlarged scales below eye	Separated from supralabials by continuous row of small lorilabials	Broadly contacting supralabials	Broadly contacting supralabials	-
(2) Enlarged tympanic scale	Intermediate in size (7–14% of head length), set far back on head (orbit–meatus 24–32% of head length)	Relatively small (6–10% of head length), set relatively close to eye (orbit–meatus 21–36% of head length)	Relatively large (11–16% of head length), close to eye (orbit–meatus 23–29% of head length)	-
(3) Sublabial tubercular scales	0–3, usually 1, in front of postrictal modified scale	3–5, usually 4, in front of postrictal modified scale	0–2, usually 1, in front of postrictal modified scale	-
(4) Gulars	Small, 30–38	Large, 15–30	Large, 16–22	-
(5) Dorsal crest	11–18 projecting scales, serrate with obvious gaps between enlarged projecting scales	19–37 projecting scales, serrate and continuous along back without obvious gaps	9–19 projecting scales, serrate with obvious gaps between enlarged projecting scales	-
(6) Dorsolateral crest	More or less continuous posttympanic and dorsolateral series	Broken: posttympanic series of enlarged, modified scales straight or curving upward, dorsolateral series curving downward	More or less continuous posttympanic and dorsolateral series	-
(7) Scales around midbody	62–84	61–94	57–77	58–67
(8) Scales of lower flanks	Heterogeneous in size, smooth to feebly keeled	Heterogeneous	Heterogeneous in size, smooth to feebly keeled	Dorsal scales heterogeneous across flanks
(9) Ventrals	Heavily keeled, 50–63	Smooth or very feebly keeled, 44–60 from preaxial border of arm to vent	Smooth or very feebly keeled, 49–63	Ventral scales 48–59
(10) White sublabial stripe	No white sublabial stripe extending from below eye to level of ear, however 1–3 white spots usually present in this region	Sharp white stripe or line of broken white spots on sublabials from level of eye to below or behind ear	No white sublabial stripe; however, one or two white or pale yellow spots usually present in area below rictus and ear	-

...Continued on the next page

TABLE. (Continued)

Characters	<i>Dendragama boulengeri</i>	<i>Dendragama australis</i>	<i>Dendragama dioidema</i>	<i>Dendragama schneideri</i>
(11) Prominent black prescapular blotch	Most specimens with prominent black prescapular blotch edged posteriorly in white	Large black prescapular blotch absent	No distinctive prescapular blotch edged posteriorly in white; band crossing neck interrupted by narrow white or yellow line	-
(12) Proximal half of tail bands	7–9 dark brown or green bands	8–13 dark brown or green bands	6–10 dark brown or green bands	-
(13) Buccal epithelium and tongue color	Yellow or orange	Yellow or orange	Buccal epithelium cream, tongue pink to red	Mouth and tongue pink to red
(8) Additional Female Color (Life)	-	Dark brown, yellow and black with amber coloration on underside	-	Female color in life dark brown, yellow and black with amber coloration on underside
(8) Additional Male Color (Life)	-	Green and lacking amber coloration along ventral surface	-	Males green and lacking amber coloration along ventral surface
(5) Dorsal and nuchal crest separation	-	-	-	Dorsal and nuchal crest clearly separated by 5–9 dorsal scales
(6, 5, 10) Dorsal crest extension and composition	Extends to base of the tail, comprised of 23–31 projecting, triangular scales	Extends to base of the tail, comprised of 23–31 projecting, triangular scales	Extends to base of the tail, comprised of 23–31 projecting, triangular scales	Dorsal crest serrate, extending to base of the tail, comprised of 23–31 projecting, triangular scales
(11) Chin tubercles	Series of 3–4 enlarged tubercles in males and females	-	Series of 3–4 enlarged tubercles in males and females	A series of 3–4 enlarged tubercles present along the chin of males and females
(12) Femoral spines	-	-	-	A series of 12–18 strongly keeled, white/yellow femoral spines present (combined count on both sides)

Query 3: Compare the following 4 species and put their characters in a table with species in columns and characters in rows:

Diagnosis of *Eurydactyloides symmetricus*: To 60.3 mm SVL (CAS 232007). nape with cruciform patch of raised tubercles; head scales very large, symmetrically arranged, **not separated by small granules**; body scalation relatively homogeneous, **without much-enlarged flattened scales**; postlabial and subauricular slits continuous with one another. (Bauer *et al.* 2009)

Diagnosis of *Eurydactyloides vieillardii*: To 60.3 mm SVL (CAS 231983). Scales of nape not raised and

tubercular; enlarged head scales irregular, separated from one another by tiny granules; dorsal body scalation highly heterogeneous, postlabial and subauricular slit separated from one another by a scaled patch of skin. (Bauer *et al.* 2009)

Diagnosis of *Eurydactyloides agricolae*: To 60.0 mm SVL (CAS 231995). enlarged head scales irregular, separated from one another by tiny granules; nape without patch of raised tubercles; body scalation relatively homogeneous, without enlarged flattened scales; postorbital and subauricular slits continuous with one another. (Bauer *et al.* 2009)

Diagnosis of *Eurydactyloides occidentalis*:

Eurydactyloides occidentalis can be distinguished from all congeners on the basis of the following combination of characters: size small (to 53.4 mm SVL), no tubercular scales on nape; postlabial slit and subauricular groove not continuous, separated by a narrow column of scaled skin; head scales, particularly in frontonasal and loreal regions, large, angular and relatively symmetrical (Figs 5, 7); male precloacal pores 60 or more in multiple rows; dorsal pattern of three irregular dark bands on a pale background. *Eurydactyloides occidentalis* may be distinguished from both *E. symmetricus* and *E. agricolae* by its discontinuous postlabial-subauricular slit-groove. it is very similar to *E. vieillardii* in most respects but it possesses **larger, more**

symmetrical and angular dorsal head scales, especially in the frontonasal and loreal regions. in addition, based on the limited sample of specimens known (n = 4) it is also considerably smaller than *E. vieillardii* (maximum SVL 53.4 mm versus 60.3 mm SVL) (radiographs reveal that all four *E. occidentalis* are probably sexually mature and that the male specimens have fused long bone and digital epiphyses, suggesting they are near full size). it also differs substantially from all of its three congeners with respect to genetic sequence. (Bauer *et al.* 2009)

Output

Characters	<i>Eurydactyloides symmetricus</i>	<i>Eurydactyloides vieillardii</i>	<i>Eurydactyloides agricolae</i>	<i>Eurydactyloides occidentalis</i>
SVL (mm)	Up to 60.3 (CAS 232007)	Up to 60.3 (CAS 231983)	Up to 60.0 (CAS 231995)	Up to 53.4 (SVL)
Nape scales	Patch of raised tubercles	Absence of raised tubercles	Absence of raised tubercles	No tubercular scales
Head scales arrangement	Symmetric, very large scales, not separated	Irregular, separated by tiny granules	Irregular, separated by tiny granules	Large, angular, relatively symmetrical scales
Body scalation	Relatively homogeneous	Highly heterogeneous	Relatively homogeneous	Relatively homogeneous, no enlarged flattened scales
Postlabial and subauricular slits	Continuous	Separated	Continuous	Not continuous, separated by a narrow column of scaled skin
Male precloacal pores	-	-	-	60 or more in multiple rows
Dorsal pattern	-	-	-	Three irregular dark bands on a pale background
Additional Distinctions	-	-	-	Smaller size, larger and more symmetrical angular dorsal head scales, genetic sequence differences

Query 4: Compare the following 4 species and put their characters in a tab-delimited table with species in columns and characters in rows:

DESCRIPTION of *Ialtris agyrtes*: Maximum SVL 576 mm; 19 scale rows at midbody; dorsal scales smooth; ventrals 160–175 in females (males unknown); subcaudals 57–63 in females; anal divided; 1 loreal; 1 preocular, usually 2 postoculars; supralabials usually 7, infralabials usually 9. Dorsum brown to tan without pattern in adults; a cream to white line from angle of jaws completely around snout, bordered above by dark brown (less well bordered below); venter patternless or with moderately heavy stippling, white anteriorly grading to pale buffy posteriorly and on underside of tail; **chin and throat longitudinally streaked with brown; in juveniles,**

upper surface of head rich brown; dorsal ground dull tan with 40 (on each side) vertically elongate bars each about 3 scales high and 1 scale long, more or less symmetrically arranged and continuing onto upper surface of tail; lower 3 scale rows grayish and paler than rest of dorsum; white postocular line; subcaudals immaculate pale gray (Schwartz & Henderson 1991: 614).

DESCRIPTION of *Ialtris dorsalis*: Maximum SVL 990 mm; 19 dorsal scale rows at midbody; dorsal scales smooth; ventrals 179–191 in males, 180–192 in females; subcaudals 99–115 in males, 98–109 in females; 1 preocular, **usually 2** postoculars; supralabials usually 7; infralabials usually 9; dorsal pattern extremely variable, from very dark, due to **extensive black pigmentation**, to olive or gray with prominent single or paired black blotches or zigzag middorsal band anteriorly, to “**salt-and-**

pepper” black and yellow, or more or less uniformly gray or brown posteriorly; some specimens give impression of being longitudinally striped (Île de la Tortue); head usually (barely discernible in some specimens) with a prominent, black W, the arms of W beginning behind eyes and extending onto neck, the angulate central portion of W extending anteriorly between parietal scales and with accessory transverse bar on frontal and supraoculars; venter bluish gray or brown, with or without dark markings anteriorly on chin and throat. (Schwartz & Henderson 1991: 614).

DESCRIPTION of *Ialtris haetianus*: Maximum SVL 305 mm; ventrals 132–144; subcaudals 45–54 in males, 40–51 in females; **ventrals + subcaudals 178–193 in males, 174–191 in females**; dorsal scale rows at midbody 19; dorsal scales without apical pits; supralabials usually 7; infralabials usually 8; loreal absent; preoculars usually 1/1; postoculars usually 2/2; **temporals usually 1+2 (rarely 1+1)**; head short, blunt; eye relatively large. Dorsal ground color tan to dark brown; dorsum patternless, strongly lineate, or with scattered, pale, paramedian dorsal dots **(each encompassing 3–4 dorsal scales) outlined with dark**

brown or black; venter ontogenetically variable: orange in juveniles to black in full adults, intermediate stages with intermediate conditions (Schwartz & Henderson 1991: 598).

DESCRIPTION of *Ialtris parishi*: Maximum SVL 745 mm; 19 dorsal scale rows at midbody; dorsal scales smooth; ventrals 163–167 in males, 163–172 in females; subcaudals 77 in 1 male, unknown for females; 1 preocular; postoculars strongly modally 2; supralabials 7; infralabials usually 9; dorsal pattern of 39–48 light, vertical bars or blotches on brown to reddish brown ground color; head unpatterned but with a fine cream to white vitta, bordered above by dark brown, **extending from ventrolateral portion of eye across supralabials 4–7 to end at angle of jaws**; venter yellow to yellow-orange with gray, brown, or black squares or rectangles on each scale giving venter a bilineate pattern and **underside of tail a unilineate pattern; ventral pattern may be faint or distinct**; chin speckled with black (Schwartz & Henderson 1991: 615).

Output:

Characters	<i>Ialtris agyrtes</i>	<i>Ialtris dorsalis</i>	<i>Ialtris haetianus</i>	<i>Ialtris parishi</i>
SVL (mm)	Up to 576	Up to 990	Up to 305	Up to 745
Scale rows at midbody	19	19	19	19
Dorsal scales	Smooth	Smooth	Without apical pits	Smooth
Ventrals (females)	160–175	180–192	-	163–172
Ventrals (males)	-	179–191	132–144	163–167
Subcaudals (females)	57–63	98–109	40–51	-
Subcaudals (males)	-	99–115	45–54	77 (1 male)
Anal	Divided	-	-	-
Loreal	1	-	Absent	-
Preocular	1	1	1/1	1
Postocular	Usually 2	2	2/2	Strongly modally 2
Supralabials	Usually 7	Usually 7	7	7
Infralabials	Usually 9	Usually 9	8	Usually 9
Dorsal pattern	Brown to tan without pattern in adults	Extremely variable, ranging from very dark to gray or brown with prominent single or paired black blotches or zigzag middorsal band	Brown Tan to dark brown, patternless, lineate, or with scattered pale paramedian dorsal dots	39-48 light vertical bars or blotches on brown to reddish brown ground color
Head pattern	A cream to white line from angle of jaws completely around snout, bordered above by dark brown (less well bordered below)	Prominent black W shape; W-pattern beginning behind eyes and extending onto neck, central portion of W extends anteriorly between parietal scales	W-pattern beginning behind eyes and extending onto neck, central portion of W extends anteriorly between parietal scales	Unpatterned, fine cream to white vitta bordered above by dark brown

...Continued on the next page

TABLE. (Continued)

Characters	<i>Ialtris agyrtes</i>	<i>Ialtris dorsalis</i>	<i>Ialtris haetianus</i>	<i>Ialtris parishi</i>
Venter	Patternless or with moderately heavy stippling, white anteriorly grading to pale buffy posteriorly and on underside of tail	Bluish gray or brown, with or without dark markings anteriorly on chin and throat	Orange in juveniles to black in full adults	Venter yellow to yellow-orange with gray, brown, or black squares or rectangles on each scale, giving a bilineate pattern
Additional Distinctions	-	Some specimens give impression of being longitudinally striped	Head short, blunt; eye relatively large	Chin speckled with black

Supplementary Materials. The following supporting information can be downloaded at the DOI landing page of this paper.

Python scripts to extract the language for each description and to generate treatment statistics:

Supplementary Information

Python script 1 TreatmentExtraction.py
Python script 2 LanguageScan.py

Supplementary Table S1. List of species with descriptions.