White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of York.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/247372/

Version: Accepted Version

Article:
Thomas, Chris (Accepted: 2026) Curation of digital biological records substantially increases knowledge of the distributions of rare and endemic species. Insect conservation and diversity. ISSN: 1752-458X (In Press)

Reuse
This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:
https://creativecommons.org/licenses/

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

Title	Curation of digital biological records substantially increases knowledge of the distributions of rare and endemic species
Título (Español)	*La manejo y validación de registros biológicos digitales aumenta sustancialmente el conocimiento de la distribución de especies raras y endémicas*

Author	Chris D Thomas
Address	Leverhulme Centre for Anthropocene Biodiversity, Department of Biology, University of York, York, UK.
Corresp. Author	chris.thomas@york.ac.uk
ORCID	0000-0003-2822-1334

Running title	Curating endemic species records
Key words	biodiversity, butterfly, Costa Rica, distributions, endemism, Hesperiidae, iNaturalist, Panama
Palabras clave	*biodiversidad, Costa Rica, distribuciones, endemismo, Hesperiidae, iNaturalist, mariposas, Panamá*

Abstract

1. Research quality distributional records remain sparse for most insect species, despite high and increasing levels of citizen science recording. A challenge is to upgrade unidentified or incorrectly identified online observations to research quality records.

2. Taking seven species of skipper butterflies endemic to montane regions of Costa Rica and Panama, this study considers the extent to which additional potential research-quality records might exist: as photographs uploaded to iNaturalist (not identified to species, identification not confirmed, or incorrectly identified), from online museum records, and other sources that could complement existing data in the Global Biodiversity Information Facility (GBIF).

3. The curation process took approximately 1.5 to 2 person days of effort per species, over a 44-day period, despite limited initial knowledge by the author.

4. Additional records were found for all seven species, and errors in existing records for five of them. In total, the available research quality records increased five-fold; from 24 originally correct records to 123 at the end of the curation process. Although the number of records per species remains modest (2 to 50, at an average of 17.6 records per species), the additions revealed major new parts of the species distributions for all the species.

5. Correct identifications ensure that recorders have larger numbers of accurate online images (and future computer-assisted identifications), hence increasing the likelihood of future correct identifications. This form of curation can make major contributions to knowledge of the distributions of rare, endemic and potentially threatened species, informing research and conservation programmes.

Extracto

1. *Los registros de distribución con calidad de investigación siguen siendo escasos para la mayoría de las especies de insectos, a pesar del progresivo incremento de la ciencia ciudadana. Uno de los retos consiste en mejorar la calidad de las observaciones en línea no identificadas o identificadas incorrectamente convirtiéndolas en registros con calidad de investigación.*

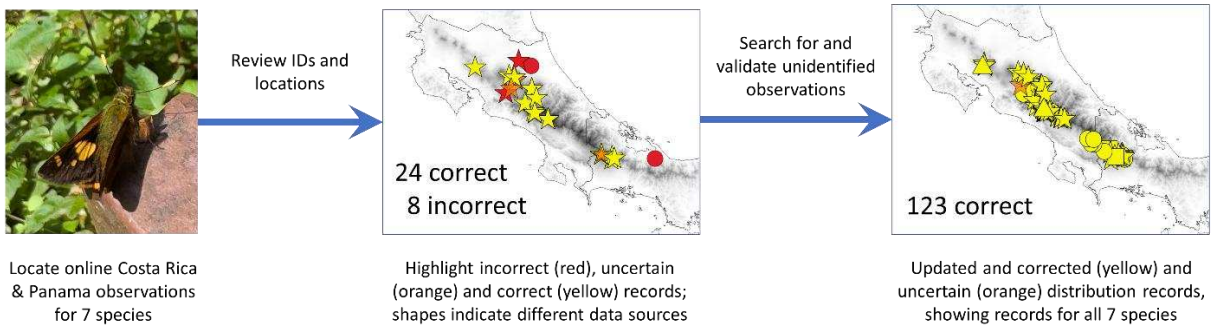
2. *Tomando siete especies de mariposas hespéridos endémicas de Costa Rica y Panamá, este estudio considera hasta qué punto podrían existir registros potenciales adicionales de calidad de investigación: como fotografías cargadas en iNaturalist, de registros de museos en línea y de otras fuentes que podrían complementar los datos existentes en el GBIF.*

3. *El proceso de revisión y validación de los registros tomó aproximadamente entre 1.5 y 2 días-persona de esfuerzo por especie, durante un período de 44 días.*

4. *Se encontraron registros adicionales para las siete especies y se identificaron errores para cinco de ellas. La calidad de los registros disponibles para la investigación se quintuplicó: de 24 registros originalmente correctos a 123 al final del proceso de curación. Las adiciones revelaron nuevas e importantes zonas de las distribuciones de todas las especies.*

5. *El esfuerzo por especie representa un nivel viable que podrían alcanzar naturalistas e investigadores interesados con poca experiencia previa. Esta forma de manejo y validación de registros puede contribuir significativamente al conocimiento de la distribución de especies raras, endémicas y potencialmente amenazadas, aportando información a los programas de investigación y conservación.*

Graphical Abstract



Introduction

Understanding the distributions of rare, endemic and potentially threatened species is facilitated by good distributional data, from which it is also possible to extract preliminary information on species' environmental limits and the potential consequences of climate change and land use alterations. Citizen science recording provides opportunities to increase this knowledge through apps and online platforms such as iNaturalist, iRecord (in the UK) and Merlin/eBird, and by extracting information from museum and other online records (e.g., through volunteer reading of museum labels) - generating ever-increasing numbers of biological records available for research and conservation (e.g., Sullivan et al., 2014; Pocock et al., 2015; Ballard et al., 2017; Nugent, 2018; Callagham et al., 2022; Dell Rocca et al., 2024; Mason et al., 2025). These are then made available via the Global Biodiversity Information Facility (GBIF), as well as through the individual platforms. The number of 'observations' made through the iNaturalist platform, as of 30 March 2025, was 232 million, of which 61 million belonged to 209 thousand species of Insecta. These numbers keep growing, reaching 321 million (up 38%), with 85 million Insecta records (up 39%) covering 232 thousand species (up 11%), by 2 June 2026. While recording is skewed towards well-known taxonomic groups and regions (see below) and most insect species still lack *any* such records, there is nonetheless considerable and increasing potential to derive valuable distributional and environmental information from these databases. This article considers how to increase the number of research quality records of the species for which there are *some* observations, but few, here considering species with up to 12 research quality records prior to the curation of digital records (hereafter 'curation').

The iNaturalist platform defines records as "Research Grade" at the species level if there is a two-thirds or greater consensus on the identity of a posted, normally photographic, observation by iNaturalist users (iNaturalist, 2025; e.g., Araya, 2025) (Figure 1). In practice, the level of expertise of independent identifiers is typically quite high for relatively poorly known taxa that lack field guides, such as those considered here. Research Grade is the level of iNaturalist record then included in GBIF datasets, where they are combined with records from other authenticated national and international platforms, including museum inventories and records from the scientific literature. I refer to "Research Quality" for all

records (including iNaturalist Research Grade records) that are regarded as sufficiently authenticated to be made available through GBIF (see Glossary).

Several analyses suggest that the accuracy of “Research Grade” identifications in iNaturalist are typically quite high, for example 98.3% accuracy for the large milkweed bug (*Oncopeltus fasciatus*) in the USA (Garretson et al., 2023), 84.2% for species in 10 plant families in the southeastern USA (White et al., 2023), and 97% for vascular plants in Western Australia (Mesaglio et al., 2025). Not surprisingly, studies of less widely known groups and ‘problem taxa’ find lower accuracy (e.g., non-marine molluscs by Barbato et al., 2021; lichens by Munzi, Isocrono & Ravera, 2023; potentially hybridising *Castilleja* plant species by Bürger & Chory, 2024). Hence, interpretation of records needs care, particularly for certain taxa and regions of the world with few specialist identifiers.

Citizen science (and professional) records tend to be skewed towards well known taxa and accessible locations in well-recorded parts of the world (e.g., Barbato et al., 2021; Di Cecco et al., 2021; Mesaglio et al., 2023; Dell Rocca et al., 2024; Díaz-Calafat et al., 2024; Alfeus, Irish & Birkhofer, 2025). The above authors have, however, noted that there are additional records available on the iNaturalist platform that are still classed as “Needs ID” for several reasons. The observations may be: (i) currently unconfirmed (i.e., the correct species is listed, but an additional identifier is required to achieve Research Grade), (ii) two or more identifiers disagree on the identification, (iii) only identified to the order, family or genus level, and (iv) incorrectly identified at Needs ID (or even at Research Grade) level. Hence, there is a need for additional ‘identifier’ contributions to online records (Callagham et al., 2022), with considerable potential for ‘expert’ identification ‘blitzes’ to increase the quality and quantity of data available (Mesaglio et al., 2025).

Here, I evaluate the opportunity to increase the numbers of research quality records, using seven species of skipper butterfly (Figure 2) endemic to montane regions of Costa Rica and Panama - to illustrate the challenge. The need for additional records and identifications is evident for this exemplar group. As of 30 March 2025, there were 18,377 iNaturalist skipper (Family Hesperiidae) observations for this region, encompassing 547 species at an average of 34 records per species. However, most observations are of widespread and distinctive species, such that 337 (62%) of these species had been identified (Research Grade +

Needs ID) five or fewer times in this region. While iNaturalist does provide digital identification suggestions, it requires 100 photos from 60 or more observations before including a species within the computer-learning suggestions (at the species level). No skipper species that is endemic (i.e., confined) to Costa Rica and western Panama achieved this level of observation (as of 09 April 2025) - the platform only offered potential species-level identifications for those species that are more widespread in the Americas and pass the 100 photos / 60 observations threshold.

Nonetheless, I found that it was possible by curating existing online unidentified iNaturalist images (e.g., that the recorder had simply listed as 'skippers'), along with online museum and other records, to generate an approximately five-fold increase in research quality records, hence greatly enhancing knowledge of the distributions and elevation-ranges of these species. Since the author was by no means an 'expert' at tropical skipper identification (and only became an 'expert' on the genera considered by the end of this study; see Supplementary Materials), this suggests that other citizen scientists and researchers that lack prior knowledge of specific taxa could make major contributions to biological knowledge by curating existing observations.

Materials and Methods

The study region and species. I consider here seven skipper species endemic to the mountains of Costa Rica and Panama (Cordillera de Talamanca in both countries; the Cordillera Central and Cordilleras de Tileran and Guanacaste towards the north-west of Costa Rica; and Panama's Cordillera Central) (Figure 4). This montane region (highest point Cerro Chirripó in Costa Rica, at 3,821 m asl) is isolated by lower elevations to the East and to the North-West, and thus it contains large numbers of endemic butterfly species (DeVries, 1987, 1997).

I consider the potential availability of georeferenced records for two species of *Dalla* (Skipperlings, Subfamily Heteropterinae) endemic to this region, a third species of *Dalla* that is 'likely endemic' (see Supplementary Materials), two endemic species of *Buzyges* (Grass Skippers, Subfamily Hesperiinae), and two endemic species of *Lon* (also Subfamily Hesperiinae) (Figure 2). For *Lon*, I additionally consider the implications of taxonomic

changes recognised by recent genetic/phylogenetic analyses (Cong et al., 2019). These seven species represent a specific type of challenge for citizen science identification. While their taxonomy is increasingly well understood, they are localised and there is no full field guide to the skippers of Costa Rica or Panama (see Identification options for recorders; Supplementary Materials). This leaves iNaturalist as the most readily available source of identification information for recorders in the field.

Period under consideration. I reviewed all records for the study species in GBIF and iNaturalist, and additional museum and online sources (see below). These were divided into a 'before' intervention period, 'during' intervention (when my iNaturalist handle @chrisdt was engaging iteratively with observers and online experts at iNaturalist and GBIF, and searching sources online), and an 'after' intervention period. The 'during' period was 25 February 2025 to 08 April 2025 inclusive. 'After' was taken as 09 April 2025. There were three exceptions, which I included within the 'during' period. Prior to travel, I added one historic record of my own to iNaturalist for *Dalla octomaculata* (on 14 January 2025), and two community records of *Lon poa* were updated after 08 April 2025 to clarify the recording locations. These additions were also treated as @chrisdt interventions and included in the 'after' totals.

GBIF. I checked all georeferenced records from the GBIF portal for species listed to the end of February 2025. The majority were sourced from iNaturalist, but others were from museum-provided online georeferenced records. I evaluated how many such records there were (and from where) and evaluated how many were correct (and started engagements with GBIF where errors were identified).

Museum, online and literature records. In principle, georeferenced records from museums are already included within GBIF, but I also separately searched the Museo Nacional de Costa Rica online Biodiversity catalogue for georeferenced specimens (https://biodiversidad.museocostarica.go.cr/), and the Smithsonian National Museum of Natural History records (USA, https://collections.nmnh.si.edu/search/iz/), where the Gordon B Small collection (Panama) and papers are held. I also examined the metadata associated with the photographs (specimens and wild) illustrated as part of the online Butterflies of America resource (Warren et al. 2024); and contacted the author to determine the location of a photograph of one species in Glassberg (2017). I also searched for relevant literature on

Google Scholar, using the current and past genera and species names in turn, and the six combinations of Costa Rica/Panama, skipper/'Hesperi'/'butterfl').

iNaturalist. Most iNaturalist records are assigned to one of two categories: Needs ID (Latin binomial not yet agreed) or Research Grade (Figure 1, see Glossary). A series of processes and algorithms are used within iNaturalist for a record to achieve Research Grade (see Achieving iNaturalist Research Grade, Supplementary Materials, Figure 1).

Observers uploading an observation that they cannot immediately identify at the species level (which applies to most skipper observations in the region considered) commonly make use of the iNaturalist platform, which "*uses computer vision systems trained on users' photos and identifications in order to provide automated taxon identification suggestions*" (iNaturalist 2024). When a naturalist uploads a photograph, the platform offers identification 'suggestions', indicating a level of confidence at different taxonomic ranks (for example, species, genus, subfamily, or family). Across multiple taxonomic groups, these suggestions are estimated to be accurate 88.3% of the time (using 16 February 2025 data, and model version 2.21), although the identification suggestions are not necessarily to species level (Loarie, 2025).

Since none of the endemic skipper species met the 100 photos / 60 observations threshold for automated taxon identification assistance, many observers who have photographed an endemic species end up listing it at a higher taxonomic level (from family down to genus, depending on the iNaturalist suggestions and the observer's confidence) or, on occasion, selecting an incorrect identification (e.g., by accepting one of the iNaturalist 'top suggestions'; see Results). A few of the incorrect identifications had then wrongly been upgraded to Research Grade by another iNaturalist user.

These misidentified Needs ID and Research Grade images can mislead other recorders and can also be incorporated into the set of images used to train the identification algorithms (e.g., expanding the apparent phenotypic variation of a geographically widespread species to include characteristics of localised endemics that belong to different species). This can result in a 'positive feedback' whereby errors propagate subsequent errors. None of this is a criticism of iNaturalist. The purpose here is to address the challenge of how to increase the

fraction of online images that are ‘languishing’ as broad higher-taxon records (and errors) into accurate species-level records, for taxa that can in principle be identified from photographs.

Having learned (for each genus in turn) how to identify the relevant species (from Glassberg, 2017; Warren et al., 2024; iNaturalist species accounts/photographs; and feedback from other iNaturalist online identifiers, as @chrisdt), I scanned every iNaturalist photograph of a skipper from the region of interest (excluding some lowland areas where no correct record of any of the species has been made). Despite there being a potential 547 skipper species recorded on iNaturalist for this region, it is relatively easy to pick out a single species (or a genus) at a time, once an identifier has learned the critical diagnostic features.

I started with each species of interest (i.e., records that already bore the relevant binomial) and reviewed every record. I could then either agree or disagree with identifications and offer an alternative identification (at an appropriate taxonomic level) if I was able to. I then repeated the search for all observations listed for the three genera, and then (once I was confident in the identification of all species in all three genera - endemic or not) scanned every skipper observation within the montane regions (above ~750m) and adjacent lowlands to search for additional observations not yet listed at the genus or species level. In terms of time, the species and genus searches were accomplished in anything from an hour to half a day each (depending on the number of records and ease of identification) while the all-skipper-records scans took a few hours.

Overall, the online search, checking, and iNaturalist discussions with other identifiers likely summed to about two person-weeks of effort - about 1.5 to 2 person days per species - spread over the period (including some preparatory online searching prior to 25 February 2025 and iNaturalist communications after 09 April 2025; subsequent and ongoing @chrisdt contributions to the identification of new observations of these genera posted after 09 April 2025 are not included in the totals reported here).

Locations and elevations. iNaturalist records come with Latitude and Longitude estimates (and an evaluation of location accuracy). Elevations (m asl) were from iNaturalist topographic maps supplemented by https://en-za.topographic-map.com/. Some museum

and other records that specified locations but were not previously georeferenced were checked for Lat/Long estimates; the latter were only included if the specimen had a named location *and* elevation, such that any environmental data used in subsequent studies would be for the correct elevation.

'Before' and 'after' elevation estimates are provided, as a single parameter indicator of the extent to which upgrading and checking records and locations might influence environmental parameters associated with species records, and potential priority conservation zones. I highlight where the elevation difference is ≥350 m (for minimum, median and maximum known elevations). All elevations rounded to the nearest 50 m.

Results

Identification trials. As an iNaturalist identification trial, I reuploaded my own already-identified photographs of species in the three genera considered, to evaluate the 'identification advice' provided by iNaturalist, as of 30 March 2025 (Figure 3 shows screenshots for the three uploads). The *Dalla octomaculata* specimen was correctly identified as belonging to the genus *Dalla* (Figure 3, left), but not to species. *Buzyges idothea* was also (incorrectly in this instance) identified as belonging to the genus *Dalla*" (Figure 3, middle). A photograph of *Lon co* (not included in this study because it is not confined to this region, but it is visually similar to *Lon ma*) was correctly identified to Subtribe Hesperiina, but the 'Top suggestions' species were not in the correct genus (Figure 3, right).

I then evaluated the species identification suggestions for all 91 Research Grade iNaturalist records (after curation) on 08 July 2025 (Table S1). I first considered the primary iNaturalist suggestion "*WE'RE PRETTY SURE THIS IS*…." (see Figure 3). Of these, 92% were correct to the level specified: 48 (53%) were correct genus-level suggestions and 35 (38%) were correct higher taxonomic level suggestions (Subtribe to Family). However, 7 (8%) were incorrect genus-level suggestions, and no suggestion was made for one *Lon poa* record (Table S1a). While 92% correct at higher taxonomic levels is admirable, zero records (0%) were identified at the species level. This is not surprising (or a specific criticism), given that none of the seven species had sufficient records for individual computer-assisted

identification. Nonetheless, it helps explain the substantial numbers of unidentified observations listed at higher taxonomic levels (see below).

The first-named of the "*HERE ARE OUR TOP SUGGESTIONS*" list (see Figure 3) were less successful. Only 10 (11%) of these suggestions were correct (to genus, not species, level), while 81 (89%) were incorrect species- and genus-level suggestions (Table S1b). Considering the top *four* suggestions for every observation record, none (0%) of the suggestions was for the correct species. Other butterfly families, moths and even brightly-coloured homopterans were suggested on occasion.

Dalla lethaea was identified most reliably to genus *Dalla*, and the other two *Dalla* species were the only ones where the first-listed top suggestion was ever in the correct genus. For the other four species (*Buzyges* and *Lon*), the top suggestions were *all* wrong at the taxonomic level suggested (Table S1b).

Overall curation results. A total of 24 correct results across the seven species was increased to 123 overall during the curation process, a 5.1-fold increase in the number of georeferenced research quality records, and there is now sufficient information available (Tables S2-S15) that could enable GBIF mapping for all seven (previously four) of the species (Table 1). Of these, the 24 Research Grade iNaturalist records increased to 84 as a result of curation (by 09 April 2025). Equally important, eight errors were removed, several of which were previous georeferencing errors, as well as misidentifications - leaving the recognised distributions of the species more tightly associated with mid-high elevations (Figure 4). For whatever reason, many records from the Museo Nacional de Costa Rica had not been captured within GBIF records, generating additional location data.

Following curation, three of the species 'added a country' to their georeferenced ranges (*Dalla octomaculata* in Panama as well as Costa Rica; *Buzyges rolla* and *Lon ma* in Costa Rica as well as Panama). These country-level range expansions represent increases in georeferenced data available to the global biodiversity and conservation communities, rather than new discoveries *per se*.

Considering the individual species briefly, *Dalla* records continued to be extremely localised: *Dalla octomaculata* added a country, *Dalla lethaea* had an identification error corrected (reducing the known range), and *Dalla pulchra* extended its range, based on museum records (Figure 5). *Buzyges idothea* went from three incorrect records (two incorrectly georeferenced, one an identification error) to 18 research quality records by 09 April 2025, while *Buzyges rolla* went from 4 to 14 research quality records, and added Costa Rica to its distribution (Table 1, Figure 6). Following taxonomic revisions, the endemic *Lon poa* showed a considerable increase in numbers of records in central and northern Costa Rica, and Museo Nacional de Costa Rica records suggest that *Lon ma* is established in easternmost Costa Rica as well as in Panama (Figure 7). Individual species summaries are provided in *Supplementary Materials*, and lists of the individual records are available from Supplementary Tables S2-S15.

These records have improved knowledge of the altitudinal ranges (and hence environmental associations) of all the species. Considering before-after curation comparisons for the lowest elevation records, median elevations and highest elevation records across the seven species (i.e., 21 elevation comparisons in total), the curated records differ by 350 m (~1,150 ft) or more in 11 instances, 52% of the records (with the pre-curation *Buzyges idothea* georeferencing errors generating the largest differences) (Table 2).

Discussion

Both GBIF and iNaturalist are ever-improving enterprises. However, gathering information for rarely recorded and poorly known taxa is genuinely difficult, and may come with a higher error rate than analyses of error rates across ‘all species’, which are biased towards a subset of species, taxonomic groups and regions that are more familiar to recorders and identifiers. For example, the top 25 most recorded species in iNaturalist (06 June 2026) are 15 birds, two mammals, six insects (two ladybirds, two bees, two large migratory butterflies), and two plants; and they all occur in the North Temperate zone (some inhabit additional regions too); ranging from 243,000 to 854,000 records each. Of the seven species considered here, six had fewer than five correct research quality records (Table 1), and only four already had separate GBIF pages/accounts, constraining the ability of biologists and

environmental planners to utilise all available information (museum, online, literature, and iNaturalist Research Grade). Biodiversity analysts should be wary because missing species with localised distributions may result in major underestimates of levels of endemism (potentially resulting in lower conservation priorities than would be appropriate), although the opposite is also possible when species thought to be restricted to a particular region have not yet been observed elsewhere. For the four species that did have GBIF accounts/maps: *D. octomaculata* only had one georeferenced record; *D. pulchra* had six of which two were incorrect; *B. idothea* had three out of three incorrect; and *B. rolla* had four records, all correct. For this subsample of four species, five of the 14 records were incorrect (36%), and the remaining information was insufficient for any kind of robust ecological or conservation analysis.

The curation process, summed across the seven species, resulted in 24 research quality records that appear to be accurate being upgraded to 123, a five-fold increase. During this process, eight 'mistakes' were removed. Of the mistakes, six were georeferencing errors, two mistaken identifications whereby other species of skipper had been assigned Research Grade for these species (one for *B. idothea*, one for *D. lethaea*), and there were three cases where *L. ma* observations were still listed as *L. zabulon* Research Grade records (I have ignored these from the 'mistakes' statistics since it represents a taxonomic split rather than misidentification *per se*). This gives an overall error rate estimate of ~24% amongst the original 'research quality' records.

The additional iNaturalist records were fairly evenly split between three categories: 15 records were already listed under the 'correct' species (as Needs ID); 21 were listed as broader taxonomic categories such as 'Skippers', 'Grass skippers' or the correct genus without a species-level identification; and 19 records were listed as Needs ID under an incorrect species or genus name. For some reason, the Museo Nacional de Costa Rica online georeferenced records had not been picked up by GBIF either. The combination of corrections and additional records substantially increased knowledge of the species' elevation range and, thus, the environmental conditions where they occur. Half the elevation range estimates (Table 2) changed by at least 350 m, which corresponds to a ~2.1^{o}C differences in expected mean temperature; given a lapse rate of ~0.6^{o}C per 100 m elevation

increase (cf. projected regional warming of ~2.6ºC represents an ‘expected’ uplift of species elevation ranges by >400 m; Birkel et al., 2022).

The broader message is that a considerable amount of additional information is available to draw conclusions about the distributions, elevational ranges and (with further data) environmental bounds of different species that are currently localised and (fairly) poorly recorded. This turned out to be surprisingly quick to achieve per species - although writing the paper has taken a bit longer! Whilst I had visited the region previously (in 1979, 1981, 1984-1987), and I am an experienced butterfly biologist (which did help with the interpretation of phylogenetic revisions), I had no particular prior expertise in tropical skipper identification - in the 1980s they were widely regarded as ‘too difficult to identify’. So, I approached the identification of online images in much the same way that any enthusiastic naturalist with experience of butterfly identification elsewhere would be able to achieve. It was feasible because it is only necessary to know the identification characteristics of one species at a time (and of that species’ closest relatives). The task is to identify the target species, not to attempt to identify every photograph of every skipper (in this instance) species. Perhaps surprisingly, nearly all the photographs could be identified as yes / no for each target species. There were hardly any ‘maybes’ left at the end.

Identification was done progressively, starting with ‘Research Grade’ records of each species (of which there were very few), followed by the ‘Needs ID’ records for the same species, then the ‘Needs ID’ records of the target genus, followed by scanning records at higher taxonomic levels. This was highly efficient. The fieldwork that I undertook (targeting all butterfly species, not just skippers) generated approximately one record of these skipper species per week in both 1979 and 2025: two georeferenced location records in ~2 weeks in relevant high-elevation environments in 1979, and one record in 2025. Curation was approximately 50 times faster, generating 99 additional records that reached research quality levels and removing 8 errors in two person-weeks of effort.

The median durations from the date when naturalists posted their observation online to achieve Research Grade status ranged from ~8 months to nearly 3 years (depending on the species), with maximum durations from ~15 months up to ~8 years (Table 3; iNaturalist tracks ‘interventions’ so that dates and durations can be extracted). Most of these would still

be waiting to achieve Research Grade status had I not intervened. In contrast, a median of only 1 day was required to reach Research Grade level after curation interventions by @chrisdt (or by an alternative expert). Once an observation had become an 'old observation' (say a few months old), the rate of these subsequently being identified correctly appeared to be very low. For those listed at a higher taxonomic level or assigned the wrong name, it seems likely that they would languish indefinitely as unidentified records.

While the iNaturalist system already has numerous designated curators who check records, oversee taxonomic revisions and species boundary issues, etc., they are generally swamped by the number of records, rates of taxonomic revisions and other identification challenges. As @chrisdt, I am not formally registered as an iNaturalist curator, but it was nonetheless easy to intervene to upgrade the records for particular species/region combinations. Thus, it would be quite feasible for an individual research project (such as PhD research) to upgrade 100 or more species, enabling such a project to, for example, analyse endemism patterns, project climate change impacts, and undertake systematic conservation planning.

It might be assumed that museum record identifications are more robust, and they certainly provide invaluable historical information. However, the exact locations where specimens were obtained may be less accurate than modern amateur records that are typically accompanied by high resolution gps or equivalent data (two iNaturalist location errors were associated with duplicate place names-now corrected following curation). The worst example was for *Buzyges idothea*, where one museum record had become associated with the georeferenced centroid of the shape file for the Rio Sucio in Costa Rica, the location named on the specimen label. This resulted in it being placed in a location at low elevation, while the montane distribution of the species suggests that the specimen must have been obtained closer to the river's headwaters. The other *B. idothea* had mysteriously become associated with the centroid of a different Panamanian Province to the one from which the specimen was obtained (Supplementary Materials). Therefore, one should not automatically presume that the georeferencing of location data is correct for museum records - and some online museum records reflected the names of species prior to taxonomic revisions (again, these were not counted as 'errors' here).

Curation is an iterative process because it results in larger numbers of identified specimen images becoming available online, which should make it easier for subsequent recorders to identify their observations, and enable additional computer-assisted identification guidance once sufficient records are available. As increasing numbers of species models become available (Loarie, 2025), it should in principle be possible to revisit all records of each species to highlight possible geographical and environmental outliers that may require reconsideration. In due course, it will become possible for iNaturalist automatically to run all existing identifications back through the image identification algorithms, to highlight possible identifications (among higher taxa) and misidentifications in past records, which can then be reassessed by the original observers and expert identifiers.

A key question is how to increase the number of people carrying out these low-level curation activities. Possibilities include to: (i) generate an iNaturalist system whereby recorders posting *and* identifying observations for less well-known species and regions, receive messages encouraging them to use the explore function within iNaturalist to provide additional identifications for particular uncommonly-recorded taxa/region combinations (asking them to consider going back through the historical records), (ii) target members of the iNaturalist community who already provide large numbers of identifications (which are still valuable) to encourage them to target taxa and regions for which further identifications will bring the greatest increases in knowledge (relative to providing 'third or fourth identifications' of heavily recorded species), and (iii) focus on capacity building for the regions and taxa of interest (involving, e.g., wildlife guides and NGOs as well as museums and academic institutions). Quite modest shifts in the recording and identification behaviour of naturalists so that they place more emphasis on the curation of records could provide a mini-revolution in the number and accuracy of records of poorly known taxa that are globally localised, rare and/or threatened.

Acknowledgements

I thank all online naturalist recorders, identifiers and those responsible for online recording and museum data platforms, who are collectively transforming knowledge of species distributions. Thanks to Jack Hatfield for comments on the manuscript. Thanks to Jack Hatfield, Andrew Hill and Inês Martins for QGIS advice, and to Kian Hayles-Cotton for

iNaturalist input. This work was supported by a Leverhulme Trust Research Centre grant—the Leverhulme Centre for Anthropocene Biodiversity— Grant number: RC-2018-021.

Data Availability Statement

All iNaturalist data are available under the species names (along with the observer and identifier curation suggestions and changes) from the iNaturalist web sites. All Museum-sourced and other web-based (GBIF, Butterflies of America) materials were obtained from open-access web portals, the links to which are provided in the manuscript. Each individual source to individual observation record sources is provided in Tables S02-S015, and all of the data are provided in the supplementary Tables (Table S02 onwards), available from Thomas (2026), at https://zenodo.org/records/21218626.

Conflict of Interest

The author acknowledges that they have no conflict of interest

References.

Alfeus, M., Irish, J. & Birkhofer, K. (2025) Recognition and completeness metrics from iNaturalist and GBIF can inform future citizen science and research projects: a case study on arthropods in Namibia. *Biodiversity and Conservation*, **34**(2), 467-480. Available from: https://doi.org/10.1007/s10531-024-02981-z

Araya, F.A.S. (2025) Yellow-clubbed Skipper *Buzyges idothea. iNaturalist, Observed* by @fabs1009, *1st June 2025.* Last accessed 10 October 2025. Available from: https://www.inaturalist.org/observations/286099131

Ballard, H.L., Robinson, L.D., Young, A.N., Pauly, G.B., Higgins, L.M., Johnson, R.F. & Tweddle, J.C. (2017) Contributions to conservation outcomes by natural history museum-led citizen science: Examining evidence and next steps. *Biological Conservation*, **208**, 87-97. Available from: https://doi.org/10.1016/j.biocon.2016.08.040

Barbato, D., Benocci, A., Guasconi, M. & Manganelli, G. (2021) Light and shade of citizen science for less charismatic invertebrate groups: quality assessment of iNaturalist nonmarine mollusc observations in central Italy. *Journal of Molluscan Studies*, **87**(4), p.eyab033. Available from: https://doi.org/10.1093/mollus/eyab033

Beccaloni, G.W., Viloria, Á.L., Hall, S.K. & Robinson, G.S. (2008) *Catalogue of the hostplants of the Neotropical butterflies*. Monografías Tercer Milenio, Vol. 8, S.E.A., Zaragoza. Available from: https://www.researchgate.net/publication/262099000

Birkel, C., Dehaspe, J., Chavarría-Palma, A., Venegas-Cordero, N., Capell, R. & Durán-Quesada, A.M. (2022) Projected climate change impacts on tropical life zones in Costa Rica. *Progress in Physical Geography: Earth and Environment*, **46**(2), 180-200. Available from: https://doi.org/10.1177/03091333211047046

Bürger, M. & Chory, J. (2024) Identification challenges of *Castilleja* (Orobanchaceae) on iNaturalist. *PLoS ONE*, **19**(10), e0311088. Available from: https://doi.org/10.1371/journal.pone.0311088

Callaghan, C.T., Mesaglio, T., Ascher, J.S., Brooks, T.M., Cabras, A.A., Chandler, M. *et al.* (2022) The benefits of contributing to the citizen science platform iNaturalist as an identifier. *PLoS Biology*, **20**(11), e3001843. Available from: https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.3001843

Cong, Q., Zhang, J., Shen, J. & Grishin, N.V. (2019) Fifty new genera of Hesperiidae (Lepidoptera). *Insecta mundi*, **0731**, 1-56. Available from: 10.5281/zenodo.3677235 and https://pmc.ncbi.nlm.nih.gov/articles/PMC8791444/

Della Rocca, F., Musiani, M., Galaverni, M., & Milanesi, P. (2024) Improving online citizen science platforms for biodiversity monitoring. *Journal of Biogeography*, **51**(12), 2412–2423. Available from: https://doi.org/10.1111/jbi.15000

DeVries, P. J. (1987) *The butterflies of Costa Rica and their natural history. Papilionidae, Pieridae, Nymphalidae*. Princeton University Press, Princeton.

DeVries, P. J. (1997) *Butterflies of Costa Rica and their natural history. Vol 2. Riodinidae*. Princeton University Press, Princeton.

Díaz-Calafat, J., Jaume-Ramis, S., Soacha, K., Álvarez, A. & Piera, J. (2024) Revealing biases in insect observations: A comparative analysis between academic and citizen science data. *Plos ONE*, **19**(7), e0305757. Available from: https://doi.org/10.1371/journal.pone.0305757

Di Cecco, G.J., Barve, V., Belitz, M.W., Stucky, B.J., Guralnick, R.P. & Hurlbert, A.H. (2021) Observing the observers: How participants contribute data to iNaturalist and implications for biodiversity science. *BioScience*, **71**(11), 1179-1188. Available from: https://doi.org/10.1093/biosci/biab093

Garretson, A., Cuddy, T., Duffy, A.G. & Forkner, R.E. (2023) Citizen science data reveal regional heterogeneity in phenological response to climate in the large milkweed bug,

Oncopeltus fasciatus. *Ecology and Evolution*, **13**(7), e10213. Available from: https://doi.org/10.1002/ece3.10213

Glassberg, J. (2017) *A Swift Guide to Butterflies of Mexico and Central America*. Second Edition. Princeton University Press, Princeton.

iNaturalist (2024) What is computer vision? *iNatHelp*, accessed 13 October 2025. Available from: https://help.inaturalist.org/en/support/solutions/articles/151000170364-what-is-computer-vision-

iNaturalist (2025) What is the Data Quality Assessment and how do observations qualify to become "Research Grade"? *iNatHelp*, accessed 10 October 2025. Available from: https://help.inaturalist.org/en/support/solutions/articles/151000169936

Loarie, S. (2025) New computer vision model (2.21) with over 1,000 new taxa. *iNaturalist*, posted April 24, 2025 by @loarie. Accessed 10 October 2025. Available from: https://www.inaturalist.org/blog/110120-new-computer-vision-model-2-21-with-over-1-000-new-taxa

Mason, B.M., Mesaglio, T., Heitmann, J.B., Chandler, M., Chowdhury, S., Gorta, S.B.Z., Grattarola, F., Groom, Q., Hitchcock, C., Hoskins, L., Lowe, S.K., Marquis, M., Pernat, N., Shirey, V., Baasanmunkh, S., & Callaghan, C.T. (2025) iNaturalist accelerates biodiversity research. *BioScience*, biaf104. Available from: https://doi.org/10.1093/biosci/biaf104

Mesaglio, T., Callaghan, C.T., Samonte, F., Gorta, S.B. & Cornwell, W.K. (2023) Recognition and completeness: Two key metrics for judging the utility of citizen science data. *Frontiers in Ecology and the Environment*, **21**(4), 167–174. Available from: https://doi.org/10.1002/fee.2604

Mesaglio, T., Shepherd, K.A., Wege, J.A., Barrett, R.L., Sauquet, H. & Cornwell, W.K. (2025) Expert identification blitz: A rapid high value approach for assessing and improving iNaturalist identification accuracy and data precision and confidence. *Plants, People, Planet*, **7**(5), 1469–1484. Available from: https://doi.org/10.1002/ppp3.70005

Munzi, S., Isocrono, D. & Ravera, S. (2023) Can we trust iNaturalist in lichenology? Evaluating the effectiveness and reliability of artificial intelligence in lichen identification. *The Lichenologist*, **55**(5), 193–201. Available from: https://doi.org/10.1017/S0024282923000403

Murillo-Hiller, L.R., Segura-Bermúdez, O.A., Barquero, J.D. & Bolaños, F. (2019) The skipper butterflies (Lepidoptera: Hesperiidae) of the Reserva Ecológica Leonelo Oviedo, San José, Costa Rica. *Revista de Biología Tropical*, **67**(2), 228-248. Available from: http://dx.doi.org/10.15517/rbt.v67i2supl.37236 and https://www.scielo.sa.cr/scielo.php?pid=S0034-77442019000200228&script=sci_arttext

Murillo-Hiller, L.R. (2025) Butterflies and Moths of Costa Rica. Comstock Publishing Associates, Cornell University Press, Ithaca and London. https://www.cornellpress.cornell.edu/book/9781501781605/butterflies-and-moths-of-costa-rica/#bookTabs=1

Pocock, M.J., Roy, H.E., Preston, C.D. and Roy, D.B. (2015) The Biological Records Centre: a pioneer of citizen science. *Biological Journal of the Linnean Society*, **151**(3), 475-493. Available from: https://doi.org/10.1111/bij.12548

Sullivan, B.L., Aycrigg, J.L., Barry, J.H., Bonney, R.E., Bruns, N., Cooper, C.B. *et al.* (2014) The eBird enterprise: An integrated approach to development and application of citizen science. *Biological Conservation*, **169**, 31-40. Available from: https://doi.org/10.1016/j.biocon.2013.11.003

Thomas, C. & Cheverton, M. (1979) *Cambridge Butterfly Expedition to Central America: Panama 1979, Final Report*. Corpus Christi College, Cambridge, UK.

Thomas, C. D. (1983) *Dalla octomaculata*, the Eight-spotted skipper. In: Wells, S.M., Pyle, R.M. & Collins, N.M. (compilers). The IUCN Invertebrate Red Data Book. IUCN, Gland, Switzerland and Cambridge, UK.

Thomas, C. (2026). CSV data files for "Curation of digital biological records substantially increases knowledge of the distributions of rare and endemic species" manuscript [Dataset]. In Insect Conservation and Diversity (pp. 1–37). Zenodo. https://doi.org/10.5281/zenodo.21218626

Viloria, A.L., Austin, G.T. and Warren, A.D. (2008) A spectacular new *Dalla* Mabille, 1904, from Venezuela-Colombia (Hesperiidae: Heteropterniae). **Bulletin of the Allyn Museum**, **156**, 1-12. Available from: https://www.floridamuseum.ufl.edu/wp-content/uploads/sites/56/2017/05/McGuire-AME156.pdf

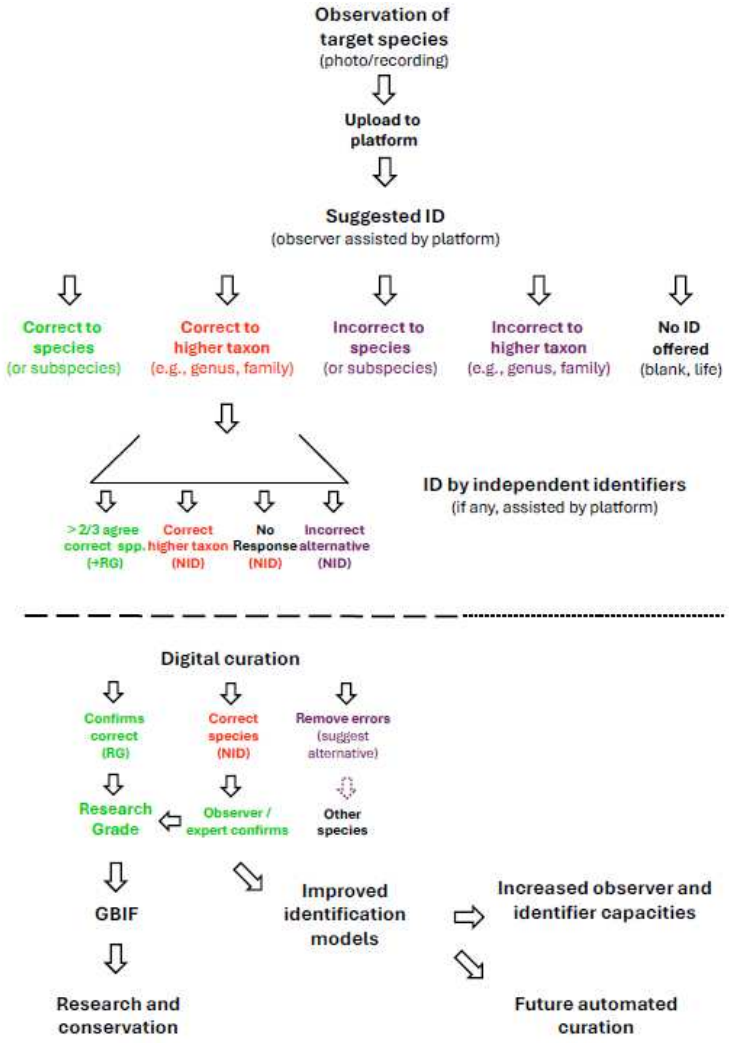
Warren, A.D., Davis, K.J., Stangeland, E.M., Pelham, J.P., Willmott K.R. & Grishin N.V. (2024) *Illustrated Lists of American Butterflies (North and South America)*. Version 9-III-2024. Available from: https://www.butterfliesofamerica.com/L/Neotropical.htm

Warren, A.D., Gott, R.J., Dorado, O. & Legal, L. (2020) A new species of *Dalla* Mabille, 1904, from Morelos, Mexico (Lepidoptera: Hesperiidae: Heteropterinae). *Tropical Lepidoptera Research*, **30**, 46-51. Available from: https://lepscience.com/wp-content/uploads/2020/06/122847-article-text-189787-1-10-20200619.pdf

White, E., Soltis, P. S., Soltis, D. E., & Guralnick, R. (2023). Quantifying error in occurrence data: Comparing the data quality of iNaturalist and digitized herbarium specimen data in flowering plant families of the southeastern United States. PLoS ONE, 18(12), e0295298. Available from: https://doi.org/10.1371/journal.pone.0295298

World Conservation Monitoring Centre (1996) *Dalla octomaculata*. The IUCN Red List of Threatened Species 1996: e.T6230A12586267. Accessed 28 June 2025. Available from: https://dx.doi.org/10.2305/IUCN.UK.1996.RLTS.T6230A12586267.en.

Zhang J., Cong Q., Shen J., Song L., & Grishin N.V. (2023) Genomic analysis reveals new species and subspecies of butterflies. *The Taxonomic Report of the International Lepidoptera Survey*, **11**(6), 1-89. Available from: 10.5281/zenodo.10372378



691
692

693 **Figure 1.** Process of online citizen science recording, as illustrated by the iNaturalist
694 platform. Observers initially suggest identifications (if they can), and then other members of
695 the naturalist community (identifiers) assess whether these are correct (illustrated for records
696 initially identified at a higher taxon level). Most unrecognised records of rare species remain
697 in this category (correct to higher taxon, but not at species level). Curation of these digital
698 records and communication via the platform with observers and other identifiers (see main
699 text) can convert most additional observations from Needs Identification (NID) into Research
700 Grade (RG) records. These RG records can then be uploaded into the Global Biodiversity
701 Information Facility (GBIF) and can improve future iterations of species identification models.
702

703



704

705 **Figure 2.** Seven species of skipper butterfly (Hesperiidae) endemic to the Costa Rica-
706 Panama highlands. *Dalla octomaculata*, Eight-spotted skipperling (© @sandralamberts /
707 Sandra Lamberts); *Dalla lethaea*, Schaus' skipperling (© @heinerziegler
708 / Heiner Ziegler); *Dalla pulchra*, Costa Rica skipperling (© @Birdernaturalist / Richard C.
709 Hoyer); *Buzyges idothea*, Yellow-clubbed skipper (© @chrisdt / Chris D Thomas); *Buzyges*
710 *rolla*, Rolla skipper (© @squiresk / Kai Squires); *Lon poa*, Poas skipper (© @markb249 /
711 Mark Brown); *Lon ma*, Chiriqui skipper (© @hippobosca / John Vallender).

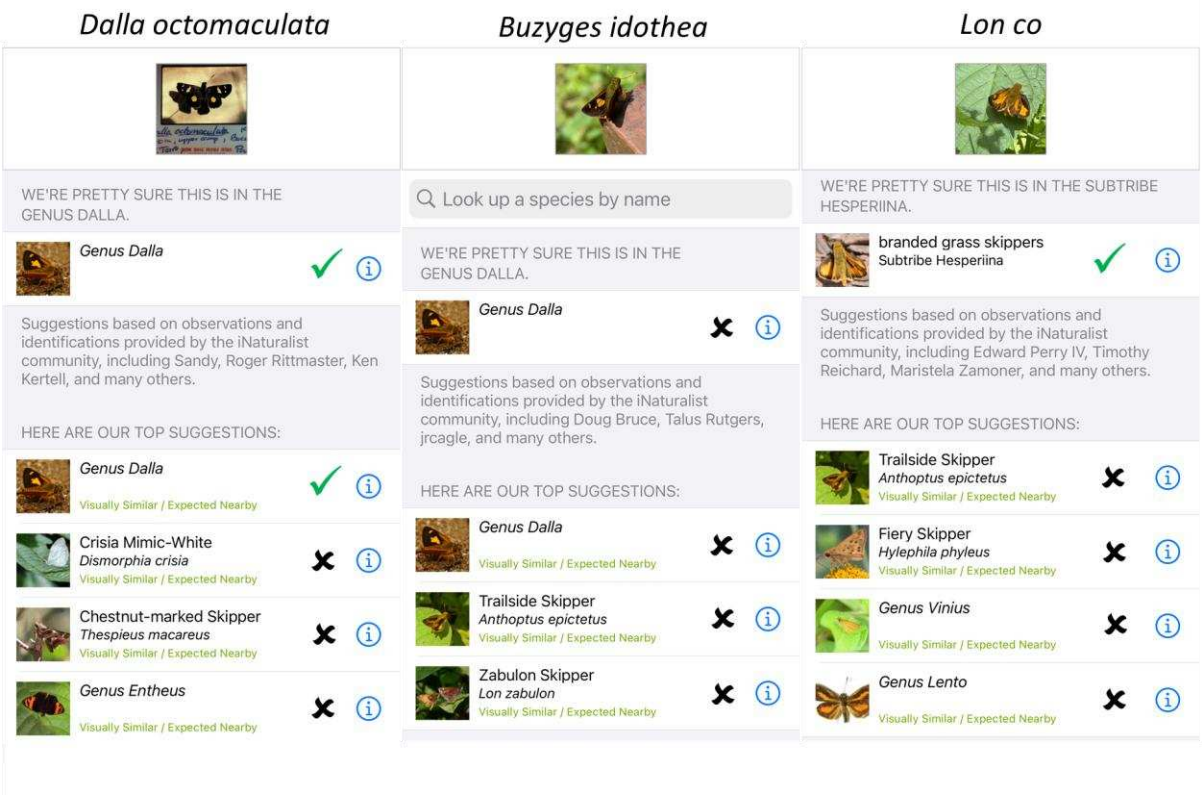
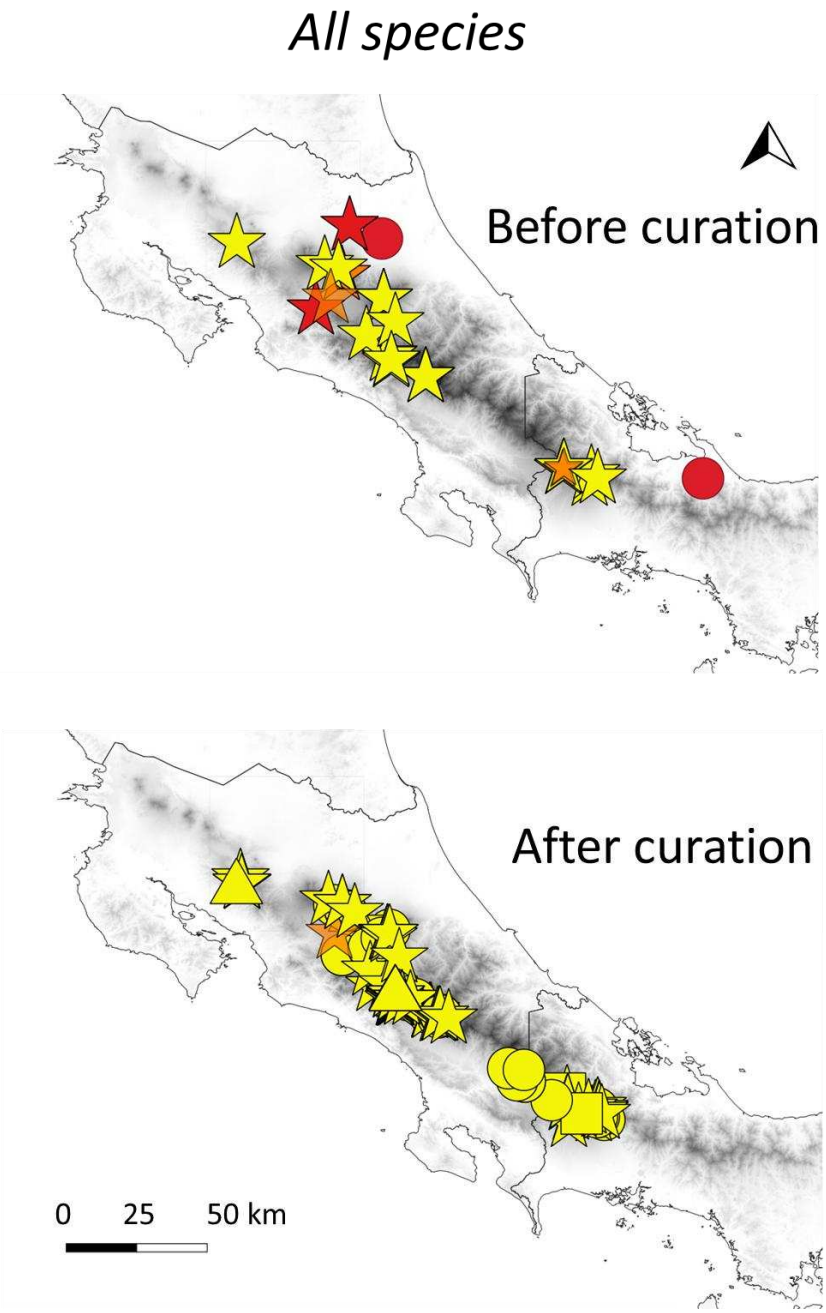


Figure 3. Screenshots from iNaturalist phone app, following uploads of images of species in the genera *Dalla*, *Buzyges* and *Lon* on 30 March 2025. Green ticks indicate correct identification suggestions; black crosses show incorrect suggestions. *Lon co* Grishin 2023 occurs in the Central Highlands of Costa Rica but is not endemic, occurring throughout Central America.



723
724
725 **Figure 4.** Combined distribution records of seven endemic *Dalla*, *Buzyges* and *Lon* species, before
726 and after curation. Yellow indicates validated records, Red errors of species identification or location,
727 and Orange correct species but location requiring further verification. Shapes indicate different types
728 of data source - Circles: museum, literature, online information sources available 25 Feb 2025 to 08
729 Apr 2025; Stars: Research Grade iNaturalist records - before curation at 24 Feb 2025, after curation
730 at 09 Apr 2025; Triangles: Research Grade iNaturalist observations uploaded 25 Feb 2025 to 08 Apr
731 2025; Squares: iNaturalist records still classed as 'Needs ID' on 20 Jun 2025, but where two or more

732 independent specialists have validated the record. Darker shading indicates higher elevations (0-3821
733 m asl).

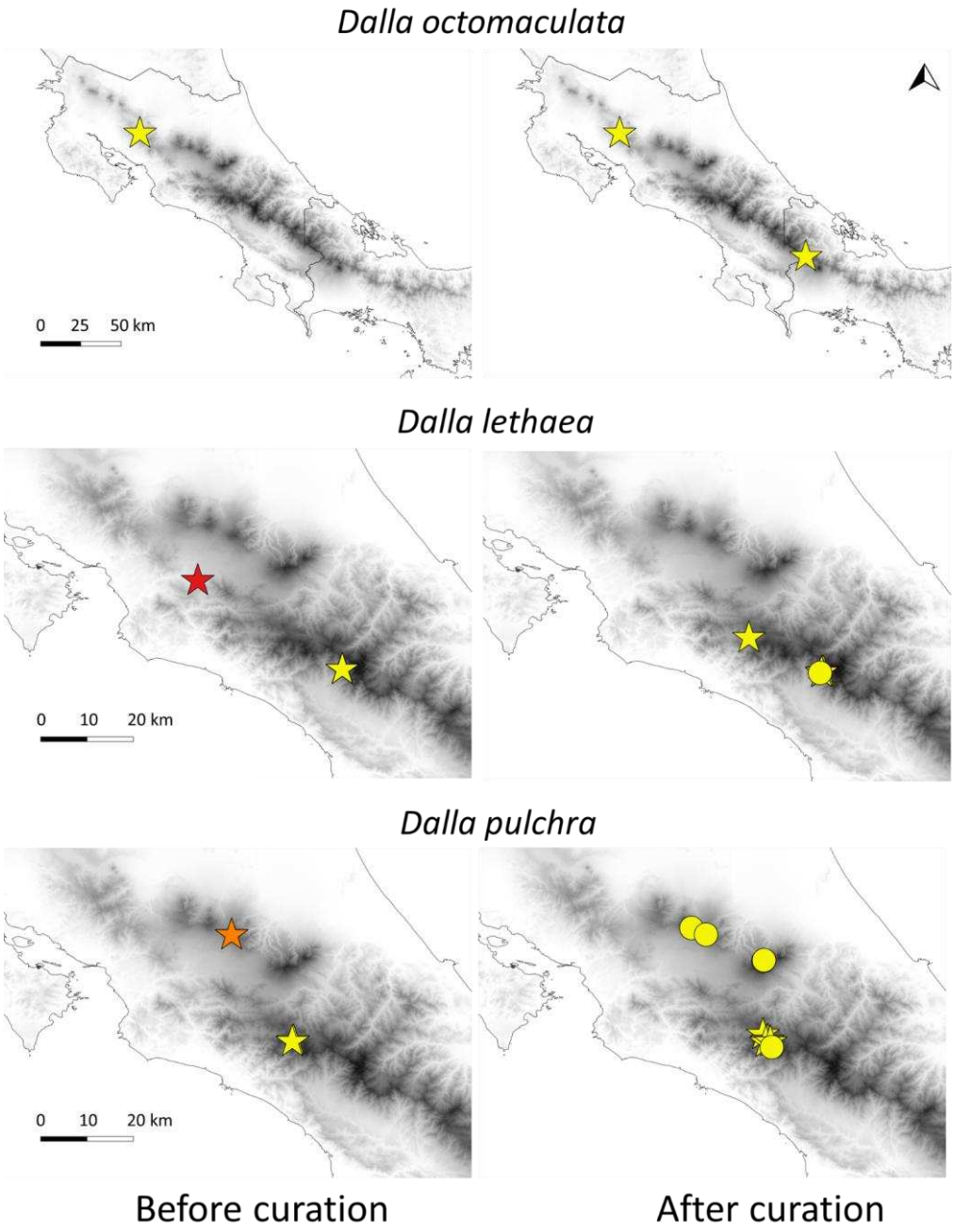


Figure 5. Records of three *Dalla* species before and after curation. Yellow indicates validated records, Red errors of species identification or location, and Orange correct species but location requires further verification. Sources - Circles: museum, literature, online information sources available 25 Feb 2025 to 08 Apr 2025; Stars: Research Grade iNaturalist records - before curation at 24 Feb 2025, after curation at 09 Apr 2025. Darker shading indicates higher elevation (0-3821 m asl). Dalla pulchra: @Birdernaturalist/Richard C. Hoyer

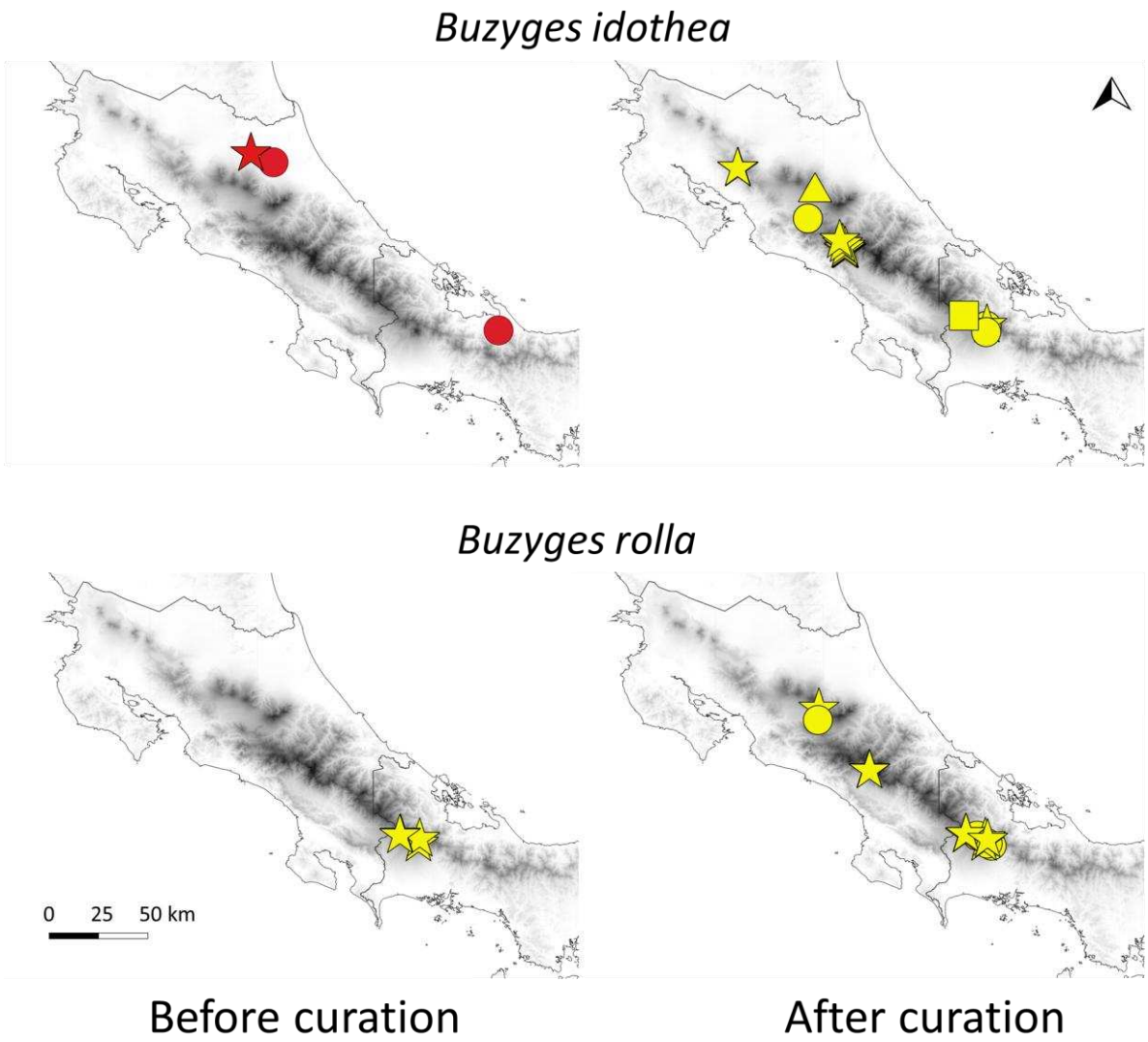


Figure 6. Records of two *Buzyges* species before and after curation. Yellow indicates validated records, Red errors of species identification or location. Sources - Circles: museum, literature, online information sources available 25 Feb 2025 to 08 Apr 2025; Stars: Research Grade iNaturalist records - before curation at 24 Feb 2025, after curation at 09 Apr 2025; Triangles: Research Grade iNaturalist observations uploaded 25 Feb 2025 to 08 Apr 2025; Squares: iNaturalist records still classed as 'Needs ID' on 20 Jun 2025, but where two or more independent specialists have validated the record. Darker shading indicates higher elevation (0-3821 m asl).

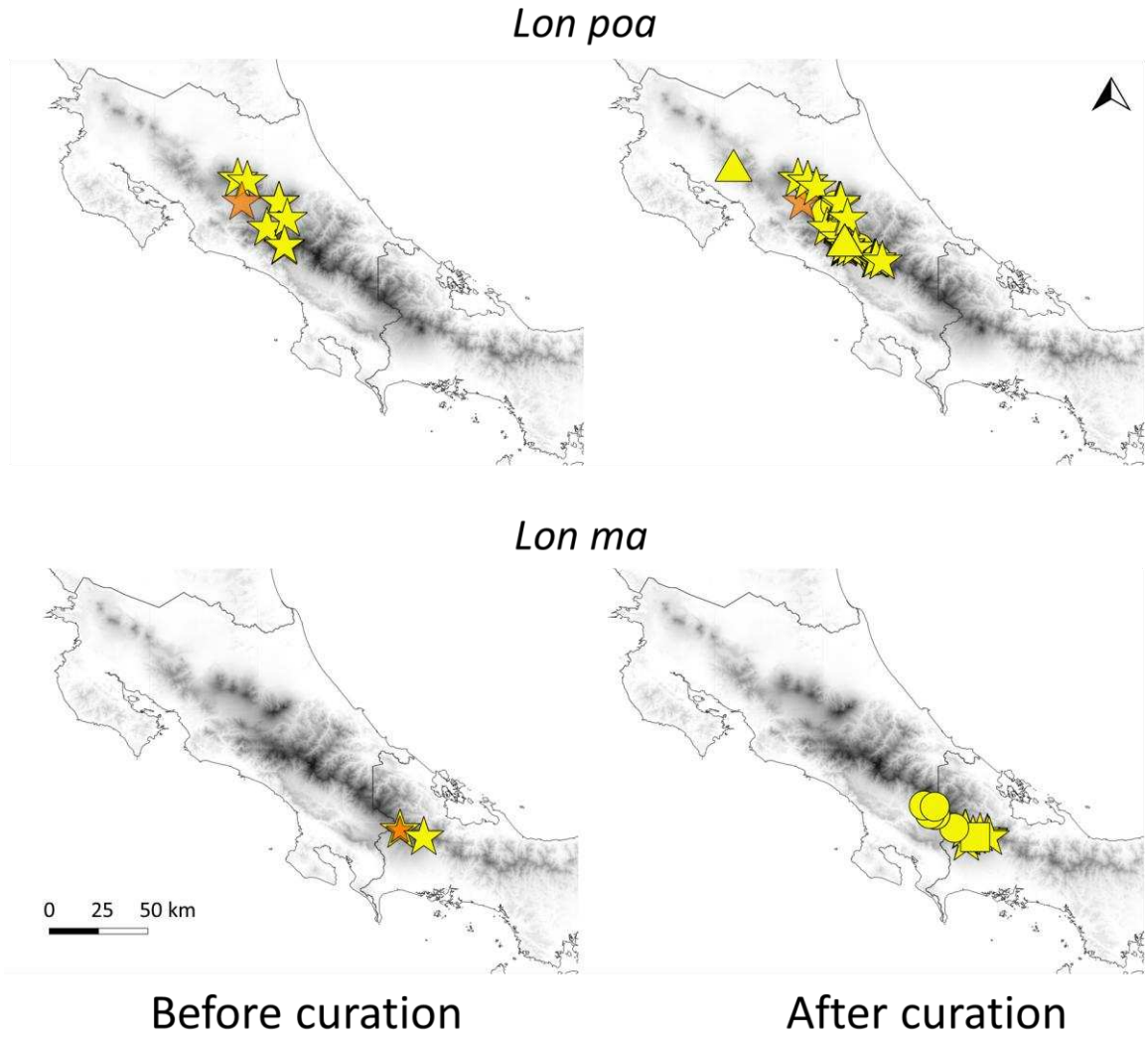


Figure 7. Records of two *Lon* species before and after curation. Yellow indicates validated records, Red errors of species identification or location, and Orange correct species but location requires further verification (remaining 'orange' record for *Lon poa* verified as correct location, 12 Jul 2025). Sources - Circles: museum, literature, online information sources available 25 Feb 2025 to 08 Apr 2025; Stars: Research Grade iNaturalist records - before curation at 24 Feb 2025, after curation at 09 Apr 2025; Triangles: Research Grade iNaturalist observations uploaded 25 Feb 2025 to 08 Apr 2025; Squares: iNaturalist records still classed as 'Needs ID' on 20 Jun 2025, but where two or more independent specialists have validated the record. Darker shading indicates higher elevation (0-3821 m asl).

768 **Table 1.** Effects of online curation on the numbers of research grade georeferenced records
769 recognised for endemic skipper butterflies in the Costa Rica-western Panama mountains.
770 iNaturalist English names adopted, but with new English names for *Buzyges rolla*, *Lon poa*
771 and *Lon ma* suggested here.
772

		Numbers of correct georeferenced records							
		BEFORE CURATION				**AFTER CURATION(**			
Scientific name	**English name**	**GBIF[a]**	**iNat[b] Res Grade**	**iNat[b] Needs ID**	**TOTAL[c]**	**Museo[d]**	**iNat Res Grade**	**iNat Needs ID**	**TOTAL**
Dalla octomaculata Godman 1900	Eight-spotted skipperling	1	1	0	**1**	0	2	0	**2**
Dalla lethaea[e] Schaus 1913	Schaus' skipperling	No entry	4	1	**4**	1	7	0	**8**
Dalla pulchra Godman 1900	Costa Rica skipperling	2	2	1	**2**	5	5	0	**10**
Buzyges idothea Godman 1900	Yellow-clubbed skipper	0	0	2	**0**	4	13	1[f]	**18**
Buzyges rolla Mabille 1883	Rolla skipper	4	4	0	**4**	4	10	0	**14**
Lon poa Evans 1955	Poas skipper	No entry	12	8	**12**	17	33	0	**50**
Lon ma Grishin 2023	Chiriqui skipper	No entry	1[g]	3	**1**	6	15	0	**21**
Total		7	24	15	**24**	**37**	**85**	**1**	**123**

773
774 Notes:

[a] GBIF records include most iNaturalist Research Grade records as well as authenticated museum records. GBIF numbers do not include records that were discovered to be errors during curation.

[b] Numbers shown are correct 'Research Grade' and 'Needs ID' observations that were confirmed as Research Grade during curation (excludes iNaturalist identification errors).

[c] Sum of unique correct research quality records across sources (none of the Museum records included by GBIF were found to be correct, so the total is the same as the numbers of iNaturalist Research Grade records).

[d] mostly Museo Nacional de Costa Rica (MNCR) online catalogue records, but also from additional museums, online identification and taxonomic works, and publications. Museum records often contained multiple specimens from a single location, research visit (from one day up to ~two months) and collector. Each site-time-recorder combination was reduced into a single georeferenced record to avoid duplicates, as follows: 15 MNCR *D. pulchra* specimens reduced to 2 separate georeferenced records; 5 *B. idothea* specimens reduced to 4; 28 *B. rolla* specimens/observations reduced to 4 (25 observations over a two-year period at one locality treated as a single record); 55 *L. poa* specimens reduced to 17; and 30 *L. ma* specimens reduced to 6 separate records.

[e] Costa Rica distribution only - excluding Mexican records that likely belong to a different species.

[f] Treated here as an additional research quality identification (ID by Mr G B Small in 1979 and by BMNH - Natural History Museum London - staff in 1980-81).

[g] Three additional Research Grade records ascribed to *Lon zabulon* were reallocated to *L. ma* during curation (taxonomic change).

797

798 **Table 2.** Elevation ranges estimated for species before and after curation. Eleven (52%) of
799 the original min-median-max estimates were ≥350 m different (underlined) from the after-
800 curation estimates .

801

	ELEVATION RANGE BEFORE CURATION				ELEVATION RANGE AFTER CURATION		
Species	**min**	**median**	**max**		**min**	**median**	**max**
Dalla octomaculata	1600	1600	*1600*		1600	1800	2000
Dalla lethaea	1550	1550	*1550*		1500	1550	1900
Dalla pulchra	*2500*	2800	3000		1950	2750	3000
Buzyges idothea	*50*	*100*	*150*		1550	2250	2900
Buzyges rolla	*1750*	1975	2500		1200	1750	2500
Lon poa	*900*	*3200*	*3200*		1300	2750	3500
Lon ma	*1900*	1900	*1900*		1300	1650	2250

802

803

804

Table 3. Durations from observation submission to achieving iNaturalist Research Grade status (for the correct species). All durations in days.

Species	MAXIMUM DURATION	MEDIAN DURATION	MEDIAN DURATION AFTER CURATION[a]
Dalla octomaculata	1258	629	1
Dalla lethaea[a]	1435	294.5	0
Dalla pulchra	452	249	2.5
Buzyges idothea	2183	1047	1
Buzyges rolla	1721	507.5	0.5
Lon poa	1836	393	1
Lon ma	2881	707	1

[a] after @chrisdt interventions (AND/OR after @jesusgarciad interventions for *D. lethaea*)

Glossary

Curation of digital records: using any available online data and images (including through apps) to study, catalogue, update (e.g., following new genetic phylogenetic studies and taxonomic works), correct any errors, and identify records to the taxonomic resolution possible.

iNaturalist: iNaturalist (https://www.inaturalist.org/) is a citizen science database, app and online social network where users record species to document global biodiversity (See Figure 1).

GBIF: Global Biodiversity Information Facility (https://www.gbif.org/) is an international open access database that includes species records, available open access to researchers, conservationists and any other interested parties.

Research Grade: Classification of a biological photographic / audio record within iNaturalist with at least two-thirds agreement by iNaturalist users of the identity of the species (normally).

Research Quality: Records from museums, the scientific literature and equivalent resources, validated national recording schemes, and iNaturalist Research Grade records, which are regarded as validated records that are uploaded into the GBIF database.

Needs ID: Classification of a biological photographic / audio record within iNaturalist that lacks two-thirds agreement of the identity of the species (normally).