Botanical Journal of the Linnean Society, 2022, **XX**, 1–11. With 6 figures.

The largest digital database of fern and lycopod records from Honduras: spatial, temporal and collector biases

SVEN P. BATKE[1,2,*], THOM DALLIMORE[1,2,4], JOHAN REYES-CHÁVEZ[1,2], RINA FABIOLA DIAZ MARADIAGA[2], EDWARD SOMERS[1], INDIANA JONES[1], WENDY ATKINSON[4], LILIAN FERRUFINO ACOSTA[3] and GERALDINE REID[4]

[1]*Biology Department, Edge Hill University, Ormskirk, UK*
[2]*Centro Zamorano de Biodiversidad, Departamento de Ambiente y Desarrollo, Escuela Agrícola Panamericana, Francisco Morazán, Honduras*
[3]*Universidad Nacional Autónoma de Honduras, Escuela de Biología, Honduras*
[4]*World Museum, National Museums Liverpool, UK*

Received 22 October 2021; revised 11 December 2021; accepted for publication 16 February 2022

Honduras is one of the least botanically studied countries in Central America. Most of the scientific botanical information for Honduras is housed in globally distributed herbaria, an often-under-used resource. A recently published checklist of ferns and lycopods from Honduras indicated that for the 713 taxa, often few distribution data are available and that we still do not know where fern and lycopod collections have taken place in the past. Therefore, the aims of this work were to (1) bring together for the first time a comprehensive inventory of fern and lycopod records from international herbaria and to (2) identify spatial, temporal and collector biases of these collections. Published and unpublished herbarium inventories of ferns and lycopod records were accessed from 2212 global herbaria. Of these, 39 hosted Honduran fern and lycopod collections. The final database included 22 194 herbarium records. Spatial and idiosyncratic collection biases are shown, with collections hotspots in areas such the Department of Francisco Morazán near the TEFH and Escuela Agrícola Panamericana (EAP) herbaria, in Celaque National Park near the city of San Pedro Sula and Tela and Lancetilla Botanical Garden. This unique database deposited at EAP, TEFH, LIV and Edge Hill University will enable Hondurans to share information to support the protection, restoration and sustainable use of their ecosystems.

ADDITIONAL KEYWORDS: distribution – herbarium – protected areas – rain forest – sampling.

INTRODUCTION

Herbaria are the cornerstone of plant science, documenting plants since the 1500s when they were developed by Luca Ghini (1490–1550) in Italy as a teaching aid (Edmondson, 2014). The 3426 active internationally recognized botanical herbaria from around the world contain > 396 million plant specimens (Thiers, 2021). This is an invaluable scientific resource that has been used for > 400 years, progressing our understanding of taxonomy (Heberling, Prather & Tonsor, 2019), ecology and global change (Lavoie, 2013; Meineke, Davis & Davies, 2018; Lang *et al.*, 2019; Miller *et al.*, 2020), biogeography (Lavoie, 2013), medicine (Souza & Hawkins, 2017), biotechnology, conservation (Nualart *et al.*, 2017), pathology (Malmstrom *et al.*, 2007), phenology (Pearson *et al.*, 2020) and, more recently, genomics (Beck & Semple, 2015; Bakker, 2017; Alsos *et al.*, 2020; see also https://academic.oup.com/botlinnean/pages/herbarium-samples-evolutionary-botany). Herbaria have also been an inspirational source for the arts and social and computer sciences, and are used in education (Soltis, 2017; Bakker *et al.*, 2020). Although the importance of herbaria in enabling novel research has largely been acknowledged, they are still regarded as an under-used resource (Lang *et al.*, 2019), potentially more so as a result of the international decline in the number of botanical taxonomists (Hopkins & Freckleton, 2002; Orr *et al.*, 2020). In addition, the treatment of these services can often differ from country to country, depending on the

*Corresponding author. E-mail: sven.batke@edgehill.ac.uk

economic status of the region in which the herbaria are located (Meineke *et al.*, 2018).

Over the past decade, advancements in scanning technology have allowed plant collections to be digitized (i.e. images and online indexes) at an exponential rate. For example, the New York Botanical Garden has doubled the number of digitized herbarium specimens in the last few years, to 4.7 million as a result of new rapid imaging technologies (Thiers, Tulig & Watson, 2016; Payel *et al.*, 2020). However, due to financial constraints, accessibility of resources such as imaging equipment and trained staff has limited this process for some developing countries such as Honduras. Despite this, even where digitized collections are not available, many herbaria hold at least inventory databases of their collections, some of which have recently become available on request.

Several recent studies using digitized herbarium data suggest large-scale sampling biases (Daru *et al.*, 2018). This raises the question, what are the implications of these biases on our understanding of regional and global floristic diversity? This is important, as many national and international conservation strategies are reliant on a solid understanding of local, regional and national species inventories, to which herbaria contribute significantly (Schatz, 2002; Chong *et al.*, 2012). Such biases can appear in spatial sampling, dominant idiosyncratic collecting by individuals and temporal sampling efforts (Daru *et al.*, 2018). When considering the use of herbarium collection data at different scales, these biases first need to be quantified and accounted for to validate the quality of the available information, the results of which can be used to direct future work (Williams & Lutterschmidt, 2006).

Honduras is a Central American country that has shown a recent increase in botanical interest (Martin *et al.*, 2021). Much work is currently underway to digitize a significant proportion of local herbaria collections, complimenting the recently digitized USA collections from the country (e.g. the Pteridopyte Collection Consortium: www.pteridoportal.org). Additionally, national botanical inventories are now being taxonomically updated, including the most recent revision of the ferns and lycopods (Reyes-Chávez, Tarvin & Batke, 2021a). Relative to this, efforts by the authors of this article are being undertaken to DNA barcode all known Honduran species of ferns and lycopods using new collections and material from herbaria, as part of the Honduran Fern Flora project. The recent boom in using herbarium material from Honduras justifies a need to have a broader understanding of the distribution of available material and data stored in national and international collections.

There are currently 713 taxa of Honduran ferns and lycopods (hereafter collectively referred to as ferns) known from Honduras (Reyes-Chávez *et al.*, 2021a). In their review, Reyes-Chávez *et al.* (2021a) indicated that a significant proportion of Honduran vouchers are in herbaria in the USA, and there are some indications that the fern collections have been biased by dominant collectors. Spatial and idiosyncratic targeted sampling could therefore be heavily skewing the collection-based data from the region. Historically, key expeditions led by plant collectors, including Antonio Molina Rositto (1926–2012), Truman George Yuncker (1891–1964) and Cyril (Cirilo) Hardy Nelson Sutherland (1938–2020), are believed to have targeted specific areas in Honduras, and from these expeditions collections were deposited in national and international herbaria (Reyes-Chávez *et al.*, 2021a). However, the exact number and location of where these collections have been deposited has not been collectively assessed. With regards to sample bias, we do not know whether there is a quantifiable bias with regards to where collections have taken place (e.g. in protected areas, proximity to infrastructures, certain elevations and specific habitats). An assessment of these biases is required to help prioritize future resources and gain a better understanding of national inventories. Additionally, it will help to identify gaps in local knowledge on biodiversity that can inform local conservation efforts. Here we propose that the development of large, complete and coherent herbarium databases could be of vital importance as a baseline to inform conservation efforts in the future.

The focus of this research therefore sets out to bring together, for the first time, the largest comprehensive database of Honduran herbarium specimens of ferns and to investigate spatial, temporal and collector biases in this database. This assemblage was selected as they currently represent the most systematically organized group of Honduran plants (Reyes-Chávez *et al.*, 2021a).

MATERIAL AND METHODS

Sources and description of data

To develop the database, 2212 global herbaria and publicly available index records were investigated for digitized herbarium specimens, of which 39 herbaria produced 22 194 data points (Table 1). Herbarium collection inventories were compiled from the Pteridophyte Collections Consortium in June–July 2021 (PCC, 2021), Tropicos (Tropicos, 2021) and individual digitized herbarium inventories that were sent to us by curators on request (Fig. 1). It should be noted that other databases such as GBIF only held *c.* 8000 fern records, of which all were accounted for in the database. For each record, the international

Table 1. Summary of the total number of Honduran fern records and the number of records used in the spatial analysis for each of the 39 herbaria

Code	Herbarium name	Total number of records	Number of records used in spatial analysis	Used in spatial analysis (%)
TEFH	Universidad Nacional Autónoma de Honduras	4820	3080	64
MO	Missouri Botanical Garden	4542	4385	97
EAP	Escuela Agrícola Panamericana	3621	3498	97
F	Field Museum of Natural History	3084	2700	88
US	Smithsonian Institution	2931	1991	68
NY	The New York Botanical Garden	881	671	76
CR	Museo National de Costa Rica	655	610	93
UC	University of California	512	200	39
TEX	University of Texas at Austin	329	29	9
Naturalis	Combined collection of WAG, L, U & AMD	213	28	13
VT	University of Vermont	141	24	17
MICH	University of Michigan	93	89	96
WIS	University of Wisconsin	71	47	66
BRIT	Botanical Research Institute of Texas	59	58	98
DUKE	Duke University	51	0	0
PH	Academy of Natural Sciences	40	1	3
MSC	Michigan State University	30	27	90
CHRB	Rutgers University	24	18	75
Y	Yale University Herbarium	17	17	100
NLU	University of Louisiana at Monroe	15	15	100
WTU	University of Washington	12	5	42
AAU	University of Aarhus	10	10	100
BM	The Natural History Museum	8	8	100
MEXU	Universidad Nacional Autónoma de México	8	7	88
B	Botanischer Garten und Botanisches Museum Berlin-Dahlem, Zentraleinrichtung der Freien Universität Berlin	5	5	100
IND	Indiana University	4	4	100
VDB	Vanderbilt University in Nashville, Tenessee	3	3	100
FLAS	Florida Museum of Natural History	2	1	50
K	Royal Botanic Gardens	2	1	50
MU	Miami University	2	2	100
ASU	Arizona State University	1	1	100
BISH	Herbarium Pacificum, Bishop Museum	1	1	100
ENCB	Instituto Politecnico Nacional	1	1	100
G	Conservatoire et Jardin botaniques de la Ville de Geneve	1	0	0
P	Muséum National d'Histoire Naturelle	1	1	100
S	Swedish Museum of Natural History	1	1	100
UAMIZ	Universidad Autónoma Metropolitana, Iztapalapa	1	0	0
USZ	Universidad Autónoma Metropolitana Iztapalapa	1	0	0
Z	Universität Zurich	1	0	0
Total		22 194	17 539	79

herbarium code, taxonomic assignment (family, genus, species and author), herbarium reference number (if applicable), individual collection number, year of collection, collection recorder and location information (latitude/longitude, major and minor area and city/village) was collated. Approximately 10% of the records did not have GPS coordinates. To spatially place each individual record, latitude/longitude coordinates were manually added using the herbarium label descriptions and Google Earth where possible. However, in cases

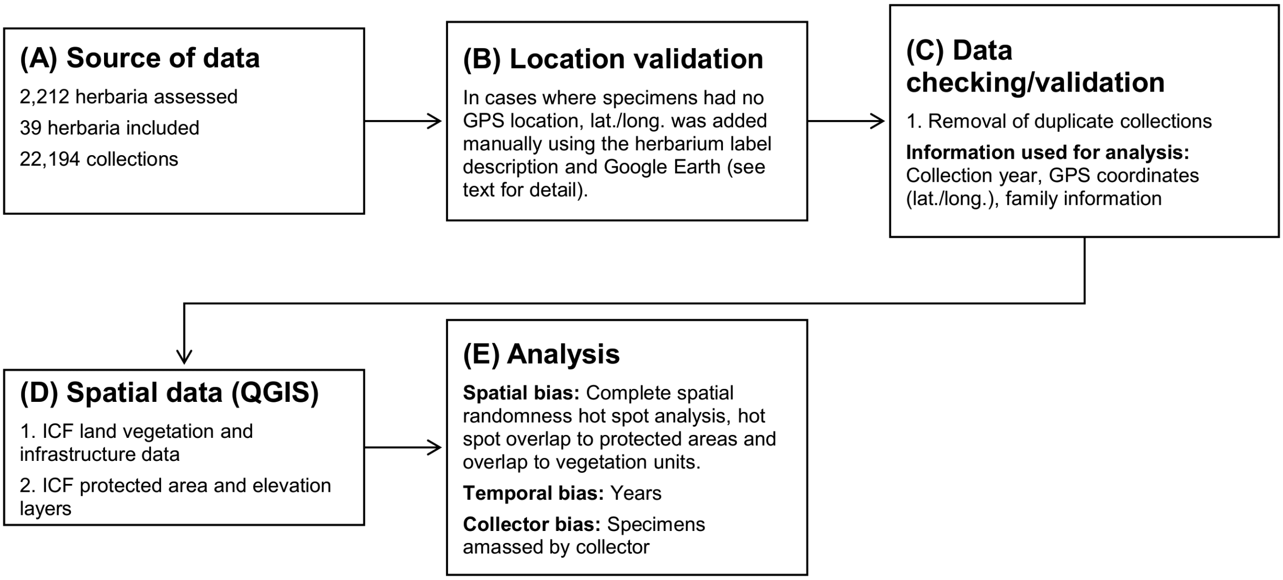

Figure 1. Analytical workflow representing different steps in the development of the study data set and analysis.

(N = 2351) where the herbarium label descriptions were not available or were describing the collection location poorly (e.g. where specimens were labelled 'Honduras'), these records were excluded from any subsequent spatial analysis. The database was checked for duplicate collections (i.e. that had the same collector and collection number), and individual duplicate entries were subsequently removed (N = 4662). Following the validation of our database, 17 539 (79%) of the original records were retained for further spatial analysis.

Herbarium database analysis

All data were projected using WGS85 in QGIS (QGIS, 2021). To identify spatial bias, herbarium records were spatially binned in 5 × 5 km rasters using the 'create a grid' tool in QGIS. In each polygon grid, the number of herbarium records was counted using the 'count points in polygon' tool and the data spatially analysed using a complete spatial randomness hot spot analysis using Gi* local statistic in Python (Spatial Analysis Library, PySAL) (Rey & Anselin, 2010). The statistical hotspot analysis identified spatial relationships between polygon features, which were weighted by the number of herbarium records in each polygon. A hotspot is defined as a location of statistical significance in which high numbers of collections have been made. To identify any spatial sample biases, the most recent vegetation maps, elevation layers and protected area layers were downloaded from the Forest Conservation Institute (ICF) (ICF, 2018). The frequency and cover of statistically significant hotspot polygons was calculated from our hotspot analysis that were overlapping different vegetation units and protected area boundaries. The number of records for different elevation ranges (i.e. 50 m bins) was compared by plotting the density distribution of the individual herbarium and the hotspot analysis records. This was done by extracting the data for each point layer in QGIS across a 50 m contour elevation layer of Honduras. The data was plotted using kernel density estimation in R. The density curves represent a non-parametric way to estimate the probability density functions of the fern collections. Finally, infrastructure bias was calculated using the minimum distance of each collection locality to the nearest major road and village using the QGIS 'distance to nearest hub (points)' tool. To test the relationship between the number of fern records per grid cell and total length of each major road (i.e. roads that are passable by a vehicle) and the number of major villages (i.e. permanent settlements) per grid cell, a Spearman's rank correlation coefficient was calculated in R (v.1.2.5033). The temporal bias of collections was assessed by plotting the historical accumulation of records across herbaria and by calculating a ten-year moving average.

RESULTS

Of the 2212 herbaria assessed, only 39 herbaria had Honduran fern collections. Following the removal of duplicates for the spatial analysis only 35 herbaria remained, from which a total of 17 539 collections were included (Table 1). Over half (58%) of Honduran fern collections are in herbaria located in the USA, with Honduran herbaria holding the second largest number (38%). Other international herbaria only hold *c.* 4% of Honduran fern collections (Fig. 2). Of the 4662

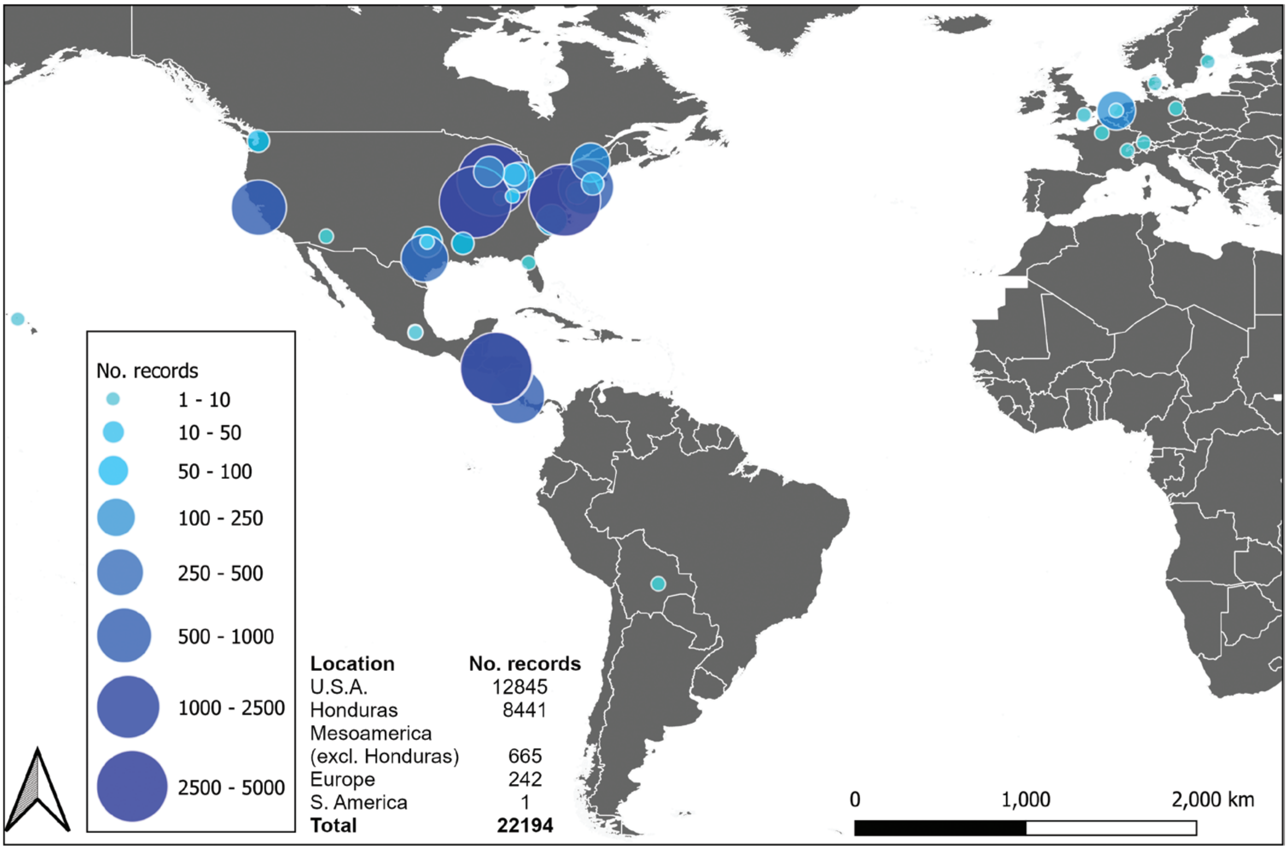

Figure 2. Map showing the distribution of herbaria that hold Honduran fern collections. Larger and darker circles indicate higher numbers of records (including duplicates).

duplicates that were removed from the database, *c.* 55% were from herbaria in the USA. It was assumed that the specimen was first deposited in either EAP or TEFH, and thus any subsequent collection was classed as a duplicate.

A hotspot analysis revealed 19 statistically significant fern collection hotspots in Honduras (Fig. 3). Most collection hotspots were in the western departments of Honduras, with departments Francisco Morazán, El Paraíso, Intibucá, Lempira, Santa Bárbara, Cortés and Atlántida having the highest number of hotspots (Fig. 3). The largest fern collection cluster was in the Department of Francisco Morazán near the TEFH and EAO herbaria, in Celaque National Park, near the city of San Pedro Sula and Tela and Lancetilla Botanical Garden (Supporting Information, Fig. S1). However, when assessing the frequency of hotspot polygons that overlapped with protected areas in Honduras, only 33.6% of hotspots fell in protected areas (Supporting Information, Fig. S1). In addition, the hotspots were located within 32 out of 68 different ICF recognized vegetation types (Supporting Information, Table S1). The highest number of grid cells of the hotspot analysis (30%) overlapped with vegetation types broadly defined as lower montane forest (Table 2); 18.5% of grid cells fell in lowland forest and only 5% of hotspot grids fell in upper montane forest category (Table 2). See Martin *et al.* (2021) for a definition of these vegetation types.

A Spearman's rank correlation showed that there was a moderate positive correlation between the number of records found per grid cell and the total length of road per grid cell ($P < 0.05$, $\rho = 0.41$). Similarly, there was a moderate positive correlation between the number of records found per grid cell and the total number of villages per grid cell ($P < 0.05$, $\rho = 0.39$). Herbarium record density decreases with distance from major roads and villages (Fig. 4). Herbarium records were greatest/most frequently found within *c.* 100 m of a major village and 250 m of a major road (Fig. 4). The highest density of specimens was collected between 800 and 1000 m a.s.l., and our analysis showed that in the identified hotspots, 22% were between 900 and 1250 m a.s.l. (Fig. 5).

The largest contribution of fern specimens to herbaria collections occurred in 1940–1950 and 1980–1990 (Fig. 6). These specimens were collected by 599 different collectors. Of these, the five collectors with

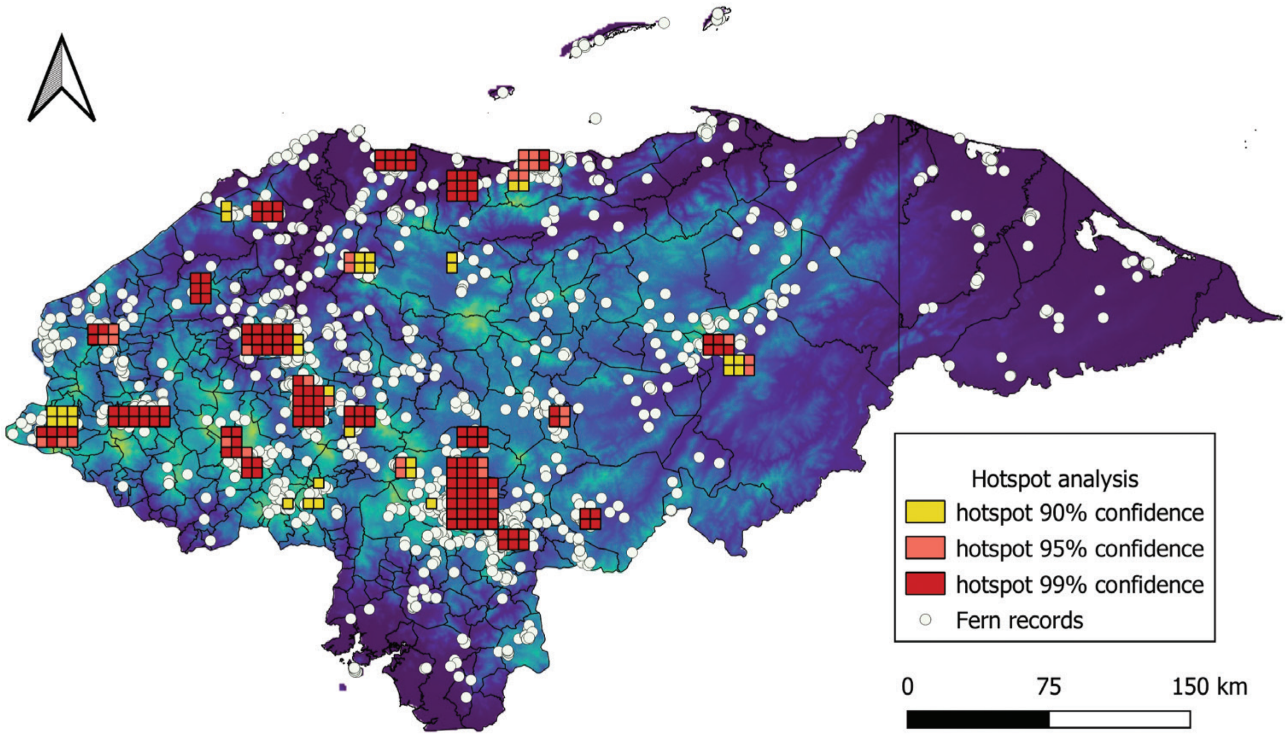

Figure 3. Hotspot analysis of Honduran fern herbarium records (*N* = 17 539). The individual fern records (white dots) and hotspot analysis rectangles (red and yellow rectangles) were overlaid on an elevation map of Honduras (dark blue to yellow = 0–2725 m a.s.l.). The rectangles show areas that were identified from the hotspot analysis as being statistically significant hotspots; the redder the rectangle, the higher the confidence level.

Table 2. Summary of the vegetation overlap with the hotspot analysis, showing percentage overlap for each broader vegetation type. See Supporting Information, Table S1 for a more detailed breakdown of vegetation categories according to ICF

Vegetation categories	Total grid cells (%)
Lower montane forest	30.0
Lowland forest	18.5
Shrub	15.5
Dunes and Savannah	12.4
Submontane forest	9.3
Aquatic/Coastal	5.9
Upper montane forest	5.0
Others	2.3
Swamp	1.0
Urban	0.1

the highest contribution of specimens were Molina (*N* = 1920), Nelson (*N* = 1533), Standley (*N* = 1496), Williams (*N* = 1114) and Yuncker (*N* = 712). The most active collection years observed in Figure 6 correspond with the highest collection years of these five collectors. For example, Yuncker collected almost 99% of his samples between 1934 and 1938, Louis Otho (Otto) Williams (1908–1991) and Paul Carpenter Standley (1884–1963) collected 68 and 72%, respectively, of their samples between 1945 and 1955, Molina collected 73% of his samples between 1955 and 1975, and Nelson collected 89% of his samples between 1975 and 1995.

DISCUSSION

The findings from this study indicate that fern collections have clearly been influenced by historical botanical events in Honduras. The earliest account of botanical dates back to 1836–1842 and was published by Bentham (1844), who described botanical collections as part of his natural history explorations of Central America. These first collections were made by Andrew Sinclair (1794–1861) the ship's surgeon on *HMS Sulphur*, and George Barclay from Tiger Island in the Gulf of Fonseca, in 1838. However, in these early collections no ferns were collected. Later work by Hemsley (1879–1888) documented biodiversity of Central American regions, which included some collections of ferns from Honduras by Gaumer from 1885 to 1886. These specimens are housed at the Royal Botanic Gardens, Kew.

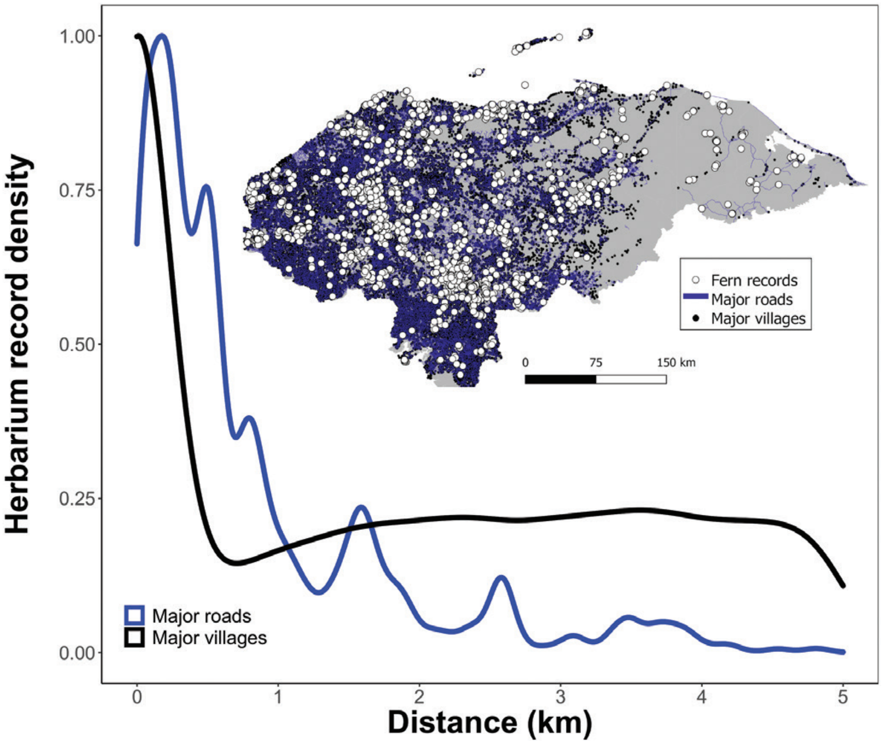

Figure 4. Kernel density distribution of herbarium records with distance to major roads (blue) and major villages (black) in Honduras. The inserted map shows the major road network (blue lines; N = 145 754), major villages (black dots; N = 27 969) and herbarium records (white dots; N = 17 539).

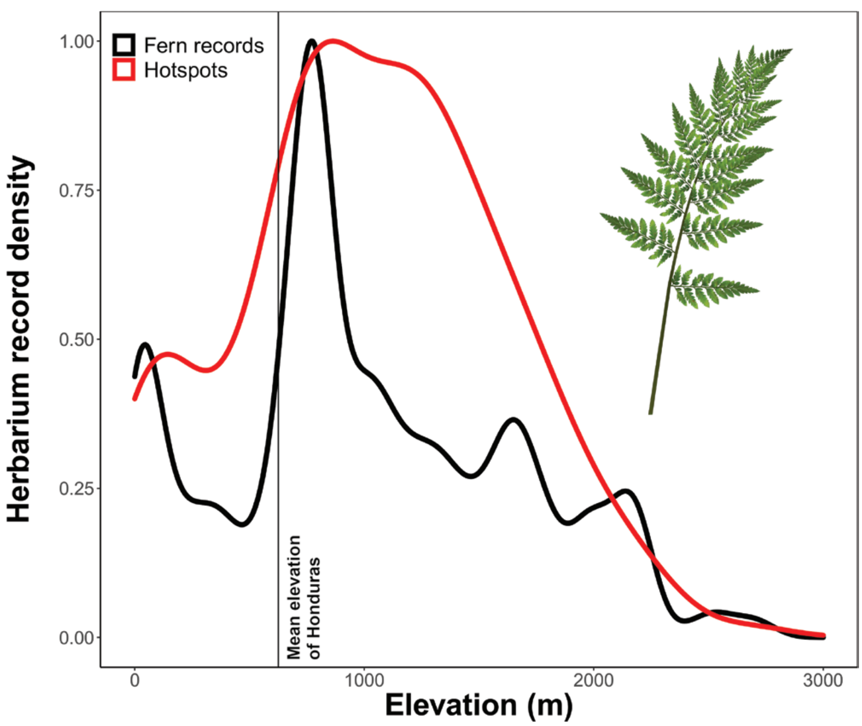

Figure 5. Kernel density distribution of herbarium records (red line; N = 17 539) and grid areas that were identified from the hotspot analysis (black line; N = 203) with elevation. The black vertical line indicates the mean elevation of Honduras.

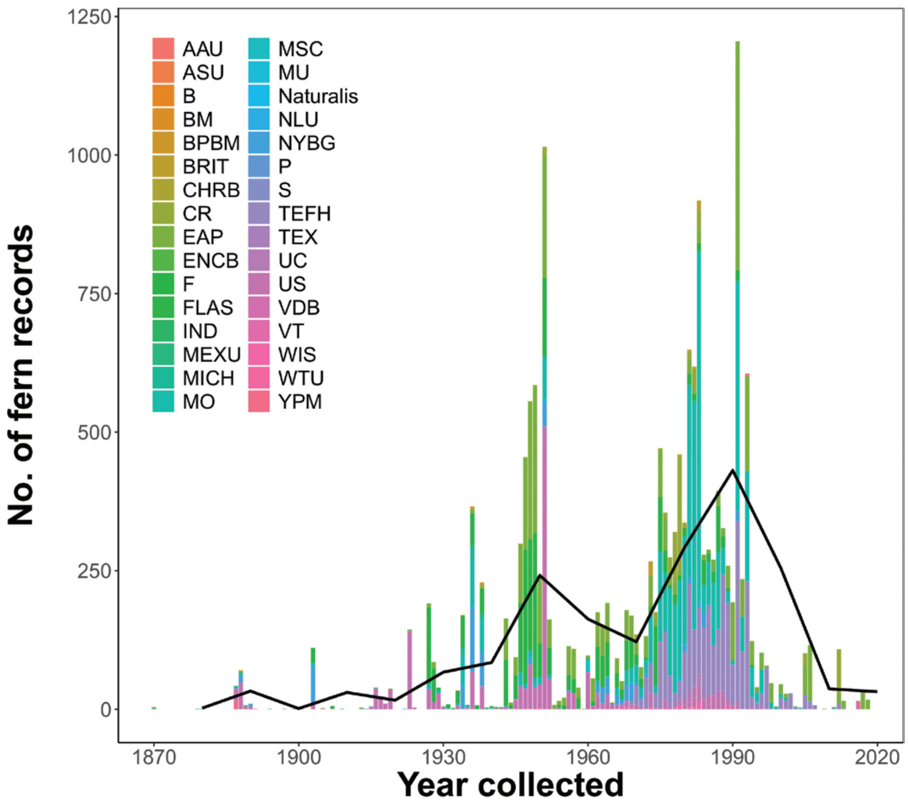

Figure 6. Total number of Honduran fern records (excluding duplicates; N = 17 206) lodged in herbaria between 1870 and 2018. The different colours represent collections in different herbaria (see Table 1 for herbarium code definition). The black line is the calculated ten-year moving average.

From the beginning of the 20th century companies based in the USA showed a strong interest in the cultivation and breeding of bananas in the northern departments of Honduras. As a result, botanical specialists were brought to Honduras from the USA to oversee the breeding and cultivation of banana crops. However, their strong interest in the study of the natural history, also resulted in several large botanically focused expeditions, including the work by Standley on the flora of the Lancetilla valley in 1931 (Standley, 1931) and by Yuncker on the flora of the Aguan valley in 1940 (Yuncker, 1940) (Fig. 6). Shortly after this, the first herbarium in Honduras was established at the EAP in 1943 following the establishment of the Zamorano University by Samuel Zemurray, the former president of the United Fruit Company (Andrews & Monroy, 1993; Pilz, 2011). In the 1950s Molina, Williams, Standley and Paul Hamilton Allen (1911–1963) dominated Honduran botanical works (Popenoe, 1964; Pitty, 1995; Pilz, 2011), and these collectors were primarily based at EAP (Fig. 6). It is therefore not surprising to see that three of these collectors disproportionally contributed to Honduran fern records, with Molina being the only Honduran botanist. Many of their duplicate collections, including type specimens (> 200 including ferns), were later donated to US herbaria in 1956, with isotypes retained in Honduras. However, the donation of duplicates to the USA did not contribute to the description of new species. At this point EAP was the largest herbarium in Central America (Pilz, 2011). Several years later, the National Herbarium was established at the Universidad Nacional Autónoma de Honduras (UNAH-TEFH) in 1969 (Nelson & Sandoval, 2005), which hosted prominent Honduran botanists including Nelson and Paul House (1961–2016). In 2008, the herbarium was named after Cyril Hardy Nelson (TEFH). In the early 1990s, we see a sudden increase in fern collections (Fig. 6), as a result of collections that were undertaken for *Flora Mesoamericana* (Moran, 1997). In addition, since the 1980s students in UNAH had to collect *c.* 20 specimens (including ferns) for their Plant Taxonomy I and Plant Taxonomy II courses, many of which have also been included into the herbarium at TEFH. However, due to difficulty of pressing larger specimens (e.g. many tree ferns), student collections are believed to be biased to smaller fern species (Ferrufino AL, pers. comm.).

Our data showed that 55% of Honduran fern collections have duplicate specimens in US herbaria that hold *c.* 58% of all Honduran fern specimens. Given the history of collectors from the USA in Honduras, this is not surprising. Nelson, in particular, donated large numbers of collections to USA herbaria, including MO (the herbarium of the Missouri Botanical Garden). It was not until 2007 and 2011 that the Honduran government regulated plant collections in Honduras and introduced a law that requires all international collectors to deposit duplicate collections in local herbaria (acuerdo número 045-2011, Diario-oficial, 2012). Currently, there are 39 global herbaria identified in our study that hold Honduran fern collections. However, we are aware of additional herbaria that hold small Honduran fern collections, but these were not included in our database. This was because they either had no inventories of their collections (lost or are in the process of digitizing their collections) or on contacting the relevant listed herbarium curator, we received no response. For example, the herbarium at Trinity College Dublin holds a small number of duplicate fern records that are also deposited at TEFH, but they have not yet digitized their collections. We also found several smaller collections in other European herbaria, which are probably the result of botanical non-fern specific expeditions to Honduras.

Of the > 22 000 fern records, 21% could not be included in any spatial analysis because limited information was available to satisfactorily allocate a specific spatial reference. Many of the early collections made use of simple location descriptions, which in many cases were not accurate enough to retrospectively allocate a GPS coordinate. Similarly, many of the recent TEFH collections were made by university students who did not report GPS positions or provided only limited site descriptions on their herbarium labels. As a result, we could only use 64% of herbarium records from TEFH, in comparison to 97% from EAP. Good collection practices are therefore crucial to allow representative analysis of herbarium records. This was also recently discussed by Kozlov *et al.* (2021), who showed that there has been a significant decrease in the collection practices of herbaria records since the 16th century.

When assessing the spatial bias of our samples we found 19 collection hotspots, of which only 33.6% fell in protected areas (Fig. 3, Supporting Information, Fig. S1). This is not surprising, as many protected areas in Honduras are at > 1800 m a.s.l. and were only declared as protected areas after 1987 (House, Cerrato & Vreugdenhil, 2002a, b). We found that many collections in hotspot areas were in lowland montane and lowland forest (combined total of 48.5%). This was also indicated by our elevation analysis, which showed that many hotspots and fern records that did not fall inside sample hotspots were collected around 800–1250 m a.s.l. Furthermore, most records were collected within *c.* 100 m of a major village and 250 m of a major road (Fig. 4). This indicates that many of the collectors carried out surveys close to infrastructure. For example, there is anecdotal evidence that Standley collected most of his collections, when he worked at EAP, along road sites near the university campus. Similarly, students at UNAH were only allowed to collect near roads and villages, as they often fell outside protected areas. It is therefore likely that some of this bias is a result of convenience of sampling. However, in some cases, these results might also reflect the fact that some more recent collections at high elevation sites, e.g. Reyes-Chávez *et al.* (2021b), and further away from infrastructures might not be included in current databases, as many collections are yet to be integrated in local herbaria. For example, although the ferns that were recorded in the most recent work of Reyes-Chávez *et al.* (2021b) have been submitted to EAP, they are still awaiting full integration into the database of EAP. The work of Reyes-Chávez *et al.* (2021b) was focused on fern surveys at Celaque National Park, the highest mountain in Honduras. If integrated, their collections would add an additional 140 records to the EAP index record. This is not an uncommon example in Honduras, as both herbaria are under resourced.

There are areas of Honduras, including large parts of the east, that have received less attention during botanical expeditions, which might possibly be the result of the lower density of roads and villages. The Mosquitia region, Pico Bonito National Park, Sierra de Agalta and Cordillera de Nombre de Dios are some of the most biodiverse hotspots in Honduras, but they have proportionally received little fern collection consideration. It is therefore likely that additional botanical surveys would reveal new additions to the Honduran flora, as suggested in previous work (Reyes-Chávez *et al.*, 2021a).

CONCLUSIONS

The Honduran fern specimens distributed throughout the herbaria of the World represent a unique and irreplaceable resource that are now brought together as a research infrastructure. This will enable informed conservation planning and the strategical targeting of under-collected areas to generate a baseline data mapping of Honduran fern biodiversity. The combined dataset has demonstrated spatial biases in sampling and idiosyncratic collecting that can now be addressed with targeted collecting efforts. This database will enable Hondurans to share information to support the protection, restoration and sustainable use of their ecosystems.

ACKNOWLEDGEMENTS

We would like to thank all the 599 plant collectors that have contributed over the last 150 years to Honduran botany. Without your hard work and love for ferns, our work would not have been possible. Many thanks also go to people that have supported us during this project including Eric Van Den Berghe, Paul Ashton and all the herbarium curators that have sent us species inventories.

FUNDING

This project was financially supported by Edge Hill University Research Investment Funding and a Chevening Scholarships, the UK government global scholarship programme, funded by the Foreign, Commonwealth and Development Office (FCDO) and partner organizations.

DATA AVAILABILITY

A copy of the database has been deposited at four institutes and can be obtained on request from Edge Hill University (sven.batke@edgehill.ac.uk), Escuela Agrícola Panamericana (EAP) (rfdiaz@zamorano.edu), TEFH (lilian.ferrufino@unah.edu.hn) and LIV (geraldine.reid@liverpoolmuseums.org.uk).

REFERENCES

Alsos IG, Lavergne S, Merkel MKF, Boleda M, Lammers Y, Alberti A, Pouchon C, Denoeud F, Pitelkova I, Puscas M, Roquet C, Hurdu BI, Thuiller W, Zimmermann NE, Hollingsworth PM, Coissac E. 2020. The treasure vault can be opened: large-scale genome skimming works well using herbarium and silica gel dried material. *Plants* **9:** 1–18.

Andrews KL, Monroy JA. 1993. Cincuenta años de la investigacióny desarrollo en Zamorano. *Ceiba* **34:** 104–184.

Bakker FT. 2017. Herbarium genomics: skimming and plastomics from archival specimens. *Webbia* **72:** 35–45.

Bakker FT, Antonelli A, Clarke JA, Cook JA, Edwards SV, Ericson PGP, Faurby S, Ferrand N, Gelang M, Gillespie RG, Irestedt M, Lundin K, Larsson E, Matos-Maravi P, Muller J, von Proschwitz T, Roderick GK, Schliep A, Wahlberg N, Wiedenhoeft J, Kallersjo M. 2020. The Global Museum: natural history collections and the future of evolutionary science and public education. *PeerJ* **8:** e8225.

Beck JB, Semple JC. 2015. Next-generation sampling: pairing genomics with herbarium specimens provides species-level signal in *Solidago* (Asteraceae). *Applied Plant Science* **3:** apps.1500014.

Bentham G. 1844. *The botany of the voyage of H.M.S. Sulphur, under the command of Captain Sir Edward Belcher, R.N., C.B., F.R.G.S., etc. during the year 1836-1842*. London: Smith, Elder & Co.

Chong KY, Lee SM, Gwee AT, Leong PK, Ahmad S, Ang WF, Lok AF, Yeo CK, Corlett RT, Tan HT. 2012. Herbarium records do not predict rediscovery of presumed nationally extinct species. *Biodiversity and Conservation* **21:** 2589–2599.

Daru BH, Park DS, Primack RB, Willis CG, Barrington DS, Whitfeld TJ, Seidler TG, Sweeney PW, Foster DR, Ellison AM. 2018. Widespread sampling biases in herbaria revealed from large-scale digitization. *New Phytologist* **217:** 939–955.

Diario-oficial. 2012. *Instituto nacional de conservacion y desarrollo forestal, areas protegidas y vida silvestre (045-2011)*. La Gaceta, Tegucigalpa: Republica de Honduras; 1–2: **8**.

Edmondson J. 2014. William H. Higdon's *Hortus Siccus*. *Webbia* **69:** 107–116.

Heberling JM, Prather LA, Tonsor SJ. 2019. The changing uses of herbarium data in an era of global change: an overview using automated content analysis. *Bioscience* **69:** 812–822.

Hemsley WB. 1879–1888. *Biologia Centrali-Americana*. London: R.H. Porter & Dulau & Co.

Hopkins GW, Freckleton RP. 2002. Declines in the numbers of amateur and professional taxonomists: implications for conservation. *Animal Conservation* **5:** 245–249.

House P, Cerrato C, Vreugdenhil D. 2002a. Rationalization of the Protected Areas System of Honduras. *Biodiversity of Honduras, Vol. II*. Tegucigalpa: World Institute for Conservation and Environment.

House R, Cerrato C, Vreugdenhil D. 2002b. Rationalisation of the protected areas system of Honduras. *Vol. II: Biodiversity of Honduras, English version*. http://www.birdlist.org/cam/honduras/Rationalisation_Vol_2_Biodiversity.pdf.

ICF. 2018. Geoportal del Sector Forestal de Honduras. http://geoportal.icf.gob.hn/geoportal/main:ForestConservationInstitute.

Kozlov MV, Sokolova IV, Zverev V, Zvereva EL. 2021. Changes in plant collection practices from the 16th to 21st centuries: implications for the use of herbarium specimens in global change research. *Annals of Botany* **127:** 865–873.

Lang PLM, Willems FM, Scheepens JF, Burbano HA, Bossdorf O. 2019. Using herbaria to study global environmental change. *New Phytologist* **221:** 110–122.

Lavoie C. 2013. Biological collections in an ever changing world: herbaria as tools for biogeographical and environmental studies. *Perspectives in Plant Ecology, Evolution and Systematics* **15:** 68–76.

Malmstrom CM, Shu R, Linton EW, Newton LA, Cook MA. 2007. Barley yellow dwarf viruses (BYDVs) preserved in herbarium specimens illuminate historical disease ecology of invasive and native grasses. *Journal of Ecology* **95:** 1153–1166.

Martin T, Jones S, Hoskins H, McCann N, Batke S, Kelly D, Kolby J, Downing R, Zelaya S, Green S. 2021. A review of the ecological value of Cusuco National Park: an urgent call for conservation action in a highly threatened Mesoamerican cloud forest. *Journal of Mesoamerican Biology* **1:** 6–50.

Meineke EK, Davis CC, Davies TJ. 2018. The unrealized potential of herbaria for global change biology. *Ecological Monographs* **88:** 505–525.

Miller EA, **Lisin SE**, **Smith CM**, **Van Houtan KS. 2020.** Herbaria macroalgae as a proxy for historical upwelling trends in central California. *Proceedings of the Royal Society B: Biological Sciences* **287:** 20200732.

Moran R. 1997. *Flora Mesoamericana - Pteridophyta*. The Bronx: New York Botanical Garden.

Nelson CH, **Sandoval G. 2005.** Las colecciones tipo del Herbario de la Universidad Nacional Autónoma de Honduras (TEFH). *Ceiba* **46:** 1–11.

Nualart N, **Ibáñez N**, **Soriano I**, **López-Pujol J. 2017.** Assessing the relevance of herbarium collections as tools for conservation biology. *The Botanical Review* **83:** 303–325.

Orr M, **Ascher J**, **Bai M**, **Chesters D**, **Zhu C-D. 2020.** Three questions: how can taxonomists survive and thrive worldwide? *Megataxa* **1:** 19–27.

Payel P, **Sayantan D**, **Chowdhury M**, **Das D. 2020.** *Herbarium technique: evolution from conventional to digitization*. Smriti Nagar: OrangeBooks Publication.

PCC. **2021.** Pteridophyte Collections Consortium. https://pteridoportal.org/portal/index.php.

Pearson KD, **Nelson G**, **Aronson MFJ**, **Bonnet P**, **Brenskelle L**, **Davis CC**, **Denny EG**, **Ellwood ER**, **Goeau H**, **Heberling JM**, **Joly A**, **Lorieul T**, **Mazer SJ**, **Meineke EK**, **Stucky BJ**, **Sweeney P**, **White AE**, **Soltis PS. 2020.** Machine learning using digitized herbarium specimens to advance phenological research. *Bioscience* **70:** 610–620.

Pilz GE. 2011. Paul C. Standley Herbarium, a Pan-American Institution. *Ceiba* **52:** 107–110.

Pitty A. 1995. Antonio Molina R., botánico centroamericano. *Ceiba* **36:** 169–174.

Popenoe W. 1964. Paul C. Standley, an appreciation. *Ceiba* **10:** 1–4.

QGIS. **2021.** QGIS Development Team. 3.16.3 ed. http://qgis.osgeo.org: Open Source Geospatial Foundation Project.

Rey SJ, **Anselin L. 2010.** PySAL: a Python library of spatial analytical methods. In: Fischer M, Getis A, eds. *Handbook of applied spatial analysis*. Berlin, Heidelberg: Springer, 175–193. https://doi.org/10.1007/978-3-642-03647-7_11

Reyes-Chávez J, **Quail M**, **Tarvin S**, **Kessler M**, **Batke S. 2021b.** Nowhere to escape–diversity and community composition of ferns and lycophytes on the highest mountain in Honduras. *Journal of Tropical Ecology* **37:** 72–81.

Reyes-Chávez J, **Tarvin S**, **Batke S. 2021a.** Ferns and lycophytes of Honduras: a new annotated checklist. *Phytotaxa* **506:** 1–113.

Schatz GE. 2002. Taxonomy and herbaria in service of plant conservation: lessons from Madagascar's endemic families. *Annals of the Missouri Botanical Garden* **89:** 145–152.

Soltis PS. 2017. Digitization of herbaria enables novel research. *American Journal of Botany* **104:** 1281–1284.

Souza ENF, **Hawkins JA. 2017.** Comparison of herbarium label data and published medicinal use: Herbaria as an underutilized source of ethnobotanical information. *Economic Botany* **71:** 1–12.

Standley PC. 1931. *Flora of the Lancetilla Valley, Honduras*. Chicago: Field Museum of Natural History.

Thiers BM. 2021. *The World's herbaria 2020: a summary report based on data from index herbariorum*. 5th edn. The Bronx: New York Botanical Gardens. sweetgum.nybg.org/science/ih.

Thiers BM, **Tulig MC**, **Watson KA. 2016.** Digitization of the New York Botanical Garden Herbarium. *Brittonia* **68:** 324–333.

Tropicos. **2021.** *The Tropicos database*. St Louis: Missouri Botanical Garden. https://www.tropicos.org/.

Williams JK, **Lutterschmidt WL. 2006.** Species-area relationships indicate large-scale data gaps in herbarium collections. *Lundellia* **2006:** 41–50.

Yuncker TG. 1940. *Flora of the Aguan Valley and the coastal regions near la Ceiba Honduras*. Chicago: Field Museum of Natural History.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Supplementary Fig. S1. Hotspot analysis of Honduran fern herbarium records (N = 17 539). The hotspot analysis rectangles (red and yellow rectangles) were overlaid on a protected area shape file layer of Honduras (green polygons). The rectangles show areas that were identified from the hotspot analysis as being statistically significant hotspots. The redder the rectangle, the higher the confidence level. The blue dots show specific location mentioned in the main text.

Table S1. Summary of the vegetation overlap with the hotspot analysis, showing maximum and mean percentage overlaps for each vegetation type. The last column shows the percentage total number of grids in each vegetation type that overlap with the hotspots