

Supplemental information

Widespread mistaken identity in tropical plant collections

Zoë A. Goodwin, David J. Harris, Denis Filer, John R.I. Wood, Robert W. Scotland

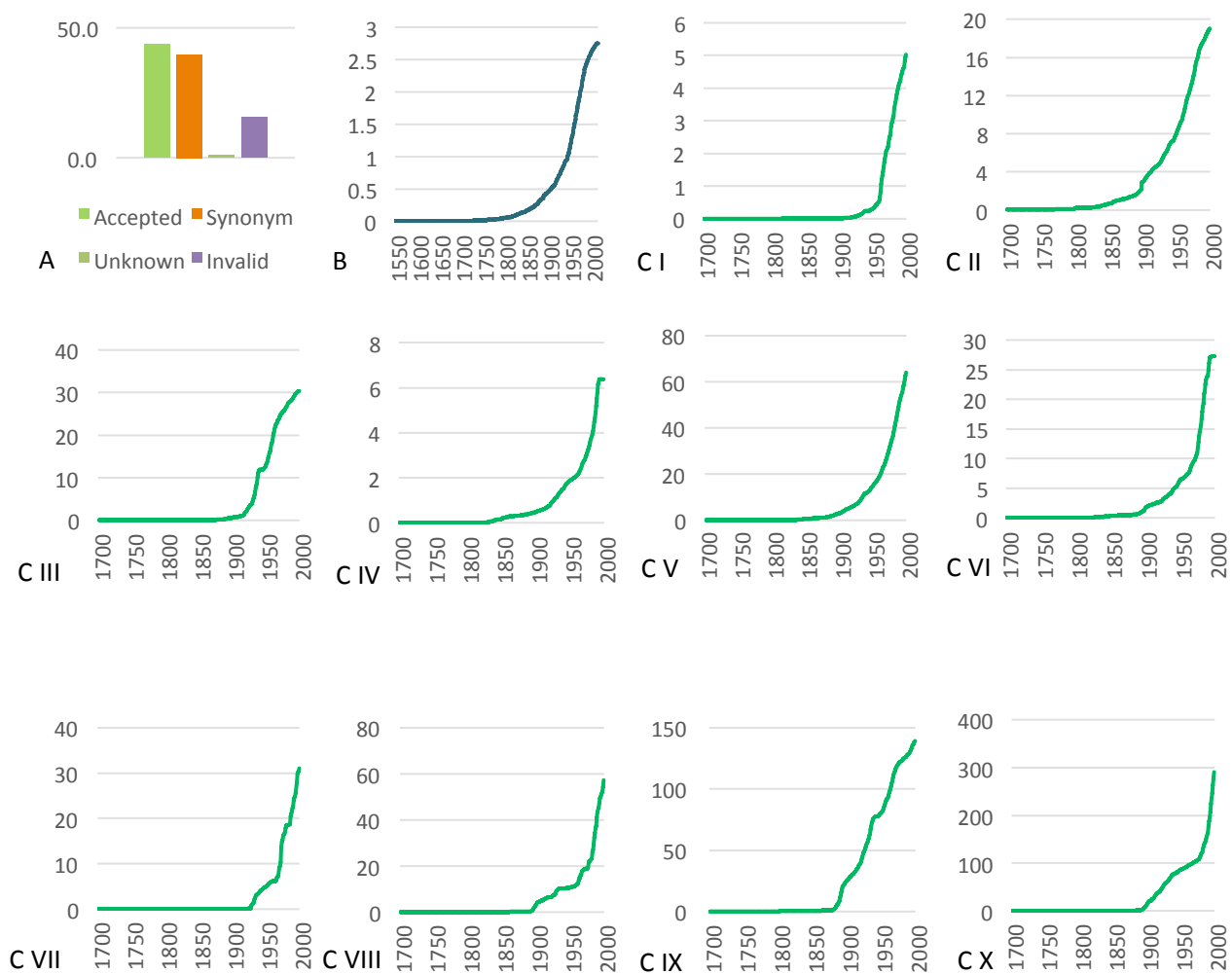
Inventory of Supplemental Information:

Figure S1

Table S1

Supplemental Experimental Procedures

Supplemental References



**Figure S1.** Low quality identifications on specimens are reflected in specimen information in aggregated databases, meanwhile the number of herbaria and specimens in herbaria has doubled between 1970 and 2000.

(A) Percentage of *Ipomoea* (N = 560) names that are currently considered to be accepted, synonyms, or invalid, on specimens in GBIF from North and South America. (B)

Accumulation of herbaria worldwide 1560 to 2014, first herbarium founded in 1569, 50% of herbaria active in 2000 had been founded since 1957, N = 2,750 [S1], number of herbaria in thousands. (C) Accumulation of herbarium specimens between 1700 and 2000 in taxa (CI-VI) and flora (CVII-X); number of specimens in thousands except VI (hundreds). (CI) *Aglaia*

(Meliaceae): 50% of specimens collected since 1976, N = 5,027 [S2]. (CII) Conifers of the World: 50% of specimens collected since 1959, N = 19,011 [S3]. (CIII) Dipterocarpaceae: 50% of specimens collected since 1955, N = 30,367 [S4-12]. (CIV) *Inga* (Fabaceae): 50% of specimens collected since 1977, N = 6,393 in 1994 [S13]. (CV) *Ipomoea* (Convolvulaceae): 50% of specimens collected since 1977, N = 63,996 [S14]. (CVI) *Leucaena* (Fabaceae): 50% of specimens collected since 1978, N = 2,724 in 1996 [S15]. (CVII) Belize: 50% of specimens collected since 1976, N = 31,034 [S16]. (CVIII) Gabon: 50% of specimens collected since 1985, N = 57,334 [S17]. (CIX) Singapore Botanic Garden Herbarium (SING): 50% of specimens collected since 1936, N = 139,398 [S10]. (CX) Rocky Mountain Herbarium (RM, USA only): 50% of specimens collected since 1988, N = 291,167 [S18].

**Table S1.** Widespread misidentifications in herbarium specimens. Mean percentage per decade of determination accuracy of specimens that have dates associated with all new determinations (correct, indeterminate, synonym or wrong) and mean number (N) of specimens extant in that decade (1900-2015) for *Aframomum*.

Decade	Percentage				N
	Correct	Indet	Synonyms	Wrong	
1900-09	12	4	11	73	64
1910-19	13	8	11	68	128
1920-29	10	15	11	65	178
1930-39	11	19	10	60	234
1940-49	12	26	10	52	323
1950-59	12	31	14	44	478
1960-69	13	34	15	38	658
1970-79	25	34	12	29	885
1980-89	37	31	8	23	1132
1990-99	42	30	8	20	1535
2000-09	42	32	8	18	1775
2010-15*	63	21	5	11	1779

\*Percentage of specimens with correct determination reaches 100% in 2014.

### Supplemental Experimental Procedures

Our initial motivation for this study was to document the piecemeal approach and temporal dynamics of the taxonomic discovery process in comparison to a recently completed monograph [S19]. It was only after we had completed the study and we realized the number of wrongly named specimens that we focused on the approach and questions addressed in this paper. To document and describe the complete determination history of a group of tropical plants, the determination data were recorded for specimens from the genus *Aframomum*, including species name, date and name of the author of the determination (Figure 1A) on each specimen. The genus *Aframomum* was revised between 1998 and 2014 by Harris & Wortley [S19], allowing access (2011-2012) to the large volume of specimens (3,176 specimens, 4,550 including duplicates) from 40 herbaria in 21 countries gathered at the Royal Botanic Garden Edinburgh (E) for the monograph. And allowing us a consistent

baseline of taxonomic opinion across all specimens. The scientific names given in all determinations were then classified as either correct, synonym, indeterminate or other relative to the determination in the current monograph [S19]. Basionyms of accepted *Aframomum* names were classified as accepted. The specimen and determination data were filtered to include only specimens for which both a/ the year of collection and b/ the year of every determination are recorded (1,492 specimens, 1,779 duplicates). The accumulation of specimens and the kind of determination were scored and plotted over time.

To assess the percentage of different names on duplicate specimens of the same collections housed in different herbaria, data for 58,860 herbarium specimens of Dipterocarpaceae were acquired from nine herbaria (E, KEP, L, MO, SAN, SAR, SING, U, WAG) [S4-12]. The specimen data were matched for number and collection year to identify specimen records from different herbaria which were duplicates of the same collection. The 9,222 matched collections of 21,075 specimens were then flagged when one or more specimens had a differing determination at genus, or species level. The numbers of specimens matched across different herbaria with the same and different determination were recorded.

To assess the percentage of incorrect (synonym or invalid or non-existing) names from specimen records provided in GBIF, occurrence records for *Ipomoea* (filter=*Ipomoea*, access date 9th December 2014) were downloaded from GBIF (GBIF Data Citations), records were cleaned and filtered by location (restricted to Americas), then the nomenclature was assessed against recent and ongoing taxonomic revision [S20] and scored as accepted, synonym (including basionyms), unknown or invalid. In addition the number of specimens determined to genus only (indeterminate) was counted.

To document plant specimen collection rates over the last three centuries, collection years for eleven specimen data sets were downloaded from online databases [S2-18] for six groups, four regional floras and a global aggregated online dataset. The taxa and floras

cover a range of distributions – northern temperate (Rocky Mountain Herbarium), tropical African (Gabon), tropical Asian (*Aglaia*, and Singapore Botanic Garden Herbarium), tropical American (*Inga*, *Leucaena*, and Belize), pantropical (Dipterocarpaceae, and *Ipomoea*), and global (Conifers of the World) – and habits – trees (*Aglaia*, Conifers, Dipterocarpaceae, *Inga* and *Leucaena*), and climbers (*Ipomoea*). The GBIF data were occurrence data downloaded using two filters: Scientific name = Plantae (Kingdom), Basis of record = Specimen (access date 10th February 2015) (GBIF Data Citations). Cumulative numbers of specimens collected per year between 1700 and 2000 were plotted for each dataset. Data were only plotted to 2000 due to the lag in specimens being accessioned into herbaria and digitized, two datasets (*Inga* and *Leucaena*) only include data up to 1994 and 1996 respectively.

Herbarium foundation dates and status were downloaded from Index Herbariorum [S1] to document temporal variation in the numbers of herbaria worldwide and used to plot the number of herbaria in existence per year between 1560 and 2014.

## **History of *Aframomum***

### **Collecting**

The first extant specimens were collected by Sonnerat from Madagascar in 1772. Low numbers of specimens (one or two a year on average) were consistently collected throughout the 1800s by a small number of collectors across Africa. Rates of specimen collecting start to accelerate (ten to twenty per year) in the first half of the 1900s, however there are two periods of disruption caused by the two World Wars. In the early 1950s rates of specimen collecting accelerated sharply (thirty specimens per year) and continued to increase towards the end of the 20th Century (up to fifty specimens per year).

### **Species publication**

The first species was described by Linnaeus in 1753 [S21] in the genus *Amomum* Roxb., a further 30 African *Amomum* species were described sporadically over the following 140 years. During a twenty year period, 1890-1910, the pace of species description accelerated

dramatically with the publication of 57 new names by several different authors including the Flora of Tropical Africa account [S22], the majority of these names are now considered to be synonyms. During this period, the African species were split from the rest of *Amomum* to form the new genus *Aframomum* K.Schum. in 1904 [S23]. Species publication slowed dramatically between 1910 and 1950, with long periods of no activity, punctuated by the occasional publication of single species, despite featuring in the 1<sup>st</sup> Edition of the Flora of West Tropical Africa [S24]. The rate of species description started to ramp up again in the second part of the 20th Century with the genus featuring in several large Flora projects including the Flora of West Tropical Africa (2nd Edition) [S25] and the Flora of Tropical East Africa [S26]. Two botanists started the process of fully revising the genus, R.M. Lock in the mid-1970s (who also provided the treatment of the genus for the Flora of Tropical East Africa) and M.M. Dhetchuvi in the mid-1990s (a PhD student in Brussels). Both individuals started the process of examining specimens across the range of the species and were able to publish several new species each and synonymise many of the existing names [S26-38]. The genus was completely revised between 1998 and 2014 by Harris & Wortley [S1], allowing us access to the huge volume of specimens gathered at E for the monograph in order to conduct this study into the history of species discovery in this genus.

## **GBIF specimen data citations**

### **Details for Figure 1C**

The full plant specimen dataset for Figure 1C was too large to download directly from GBIF (N = 31,068,510 in the year 2000). To circumvent this, individual queries were made to GBIF using the three filters: Scientific name = Plantae (Kingdom), Basis of record = Specimen and each year of collection between 1700 and 2000 inclusive (access date 10<sup>th</sup> February 2015). The total number of specimen records collected in each year were recorded manually, and the cumulative numbers of specimens collected per year between 1700 and 2000 calculated and plotted.

Thus full citations for the plant specimen year of collection data from GBIF for Figure 1C is that of all data providers who have supplied plant specimen based data to GBIF prior to 10<sup>th</sup> February 2015.

### **Details for Figure S1C (V)**

Full citation information for the *Ipomoea* data is available upon request from the corresponding author.

### **Supplemental References**

- S1. Thiers, B.M. (2015). Index Herbariorum: A global directory of public herbaria and associated staff. (New York Botanical Garden's Virtual Herbarium).  
<http://sweetgum.nybg.org/ih/> Accessed 20 February 2015.
- S2. Pannell, C. (2015). The genus *Aglaia*. <http://herbaria.plants.ox.ac.uk/bol/aglaia>  
Accessed 14 April 2015.
- S3. Farjon, A. (2001). Conifers of the World. <http://herbaria.plants.ox.ac.uk/bol/conifers>  
Accessed 14 April 2015.
- S4. Royal Botanic Garden Edinburgh (2015). Royal Botanic Garden Edinburgh (E) Herbarium Catalogue. <http://elmer.rbge.org.uk/bgbase/vherb/bgbasevherb.php>  
Accessed 20 June 2015.
- S5. Forest Research Institute Malaysia (2015). Forest Research Institute Malaysia (KEP) Herbarium Catalogue. <http://www.tfbc.frim.gov.my> Accessed 20 June 2015.
- S6. Naturalis (2015). Leiden (L), National Herbarium of the Netherlands.  
<http://vstbol.leidenuniv.nl/> Accessed 20 June 2015.
- S7. Missouri Botanical Garden (2015). Missouri Botanical Garden (MO).  
<http://www.tropicos.org/> Accessed 20 June 2015.
- S8. Forest Research Centre (2015). Forest Research Centre (SAN) Herbarium Catalogue.  
Accessed 20 June 2015.



- S9. Department of Forestry Kuching (2015). Department of Forestry Kuching (SAR) Herbarium Catalogue. Accessed 20 June 2015.
- S10. Lee, S. (2015). Singapore Botanic Garden Herbarium (SING). <http://herbaria.plants.ox.ac.uk/bol/sing> Accessed 14 April 2015.
- S11. Naturalis (2015). Utrecht (L), National Herbarium of the Netherlands. <http://vstbol.leidenuniv.nl/> Accessed 20 June 2015.
- S12. Naturalis (2015). Wageningen (WAG), National Herbarium of the Netherlands. <http://vstbol.leidenuniv.nl/> Accessed 20 June 2015.
- S13. Pennington, T.D. (1995). The Genus *Inga*. <http://herbaria.plants.ox.ac.uk/bol/aglaia> Accessed 14 April 2015.
- S14. GBIF (2015). *Ipomoea*. [http://www.gbif.org/occurrence/search?TAXON\\_KEY=2928509](http://www.gbif.org/occurrence/search?TAXON_KEY=2928509) Accessed 09 December 2014.
- S15. Hughes, C.E. (2015). *Leucaena*. <http://herbaria.plants.ox.ac.uk/bol/leucaena> Accessed 21 July 2014.
- S16. Goodwin, Z.A. (2011). Belize. <http://herbaria.plants.ox.ac.uk/bol/belizesavanna> Accessed 14 April 2015.
- S17. Wieringa, J. (2015). The flora of Gabon. <http://herbaria.plants.ox.ac.uk/bol/gabon> Accessed 14 April 2015.
- S18. Hartman, R.L. (2015). Rocky Mountain Herbarium (RM). <http://www.rmh.uwyo.edu> Accessed 07 April 2015.
- S19. Harris, D.J., and Wortley, A.H. (In Press). Monograph of *Aframomum* (Zingiberaceae). Syst. Bot. Monogr.
- S20. Wood, J.R.I., Carine, M.A., Harris, D.J., Wilkin, P., Williams, B., and Scotland, R.W. (2015) *Ipomoea* (Convolvulaceae) in Bolivia. Kew Bull. 70(3): 1 – 124.
- S21. Linnaeus, C. (1753). Species Plantarum. p. 560.
- S22. Baker, J.G. (1898). Order CXXXIV. Scitamineae. In Flora of Tropical Africa, W.T. Thiselton-Dyer, ed. (Lovell Reeve, London), Vol. 7, pp. 293–331.

- S23. Schumann, K.M. (1904). Zingiberaceae. In *Das Pflanzenreich*, A. Engler, ed. (Wilhelm Engelmann, Leipzig), pp. 201–221.
- S24. Hutchinson, J., and Dalziel, J.M. (1936). 1. *Aframomum*. *Flora of West Tropical Africa*, (Crown Agents for the Colonies, London), First Ed., Vol. 2, pp. 329–331.
- S25. Hepper, F.N. (1968). 3. *Aframomum*. *Flora of West Tropical Africa*. (Crown Agents for Oversea Governments and Administrations, London), Second Ed., Vol. 3, pp. 72–77.
- S26. Lock, J.M. (1985). Zingiberaceae. *Flora of Tropical East Africa*, R.M. Polhill, ed. (A.A. Balkema, Rotterdam), pp. 22–37.
- S27. Lock, J.M., and Hall, J.B. (1973). Three new species of *Aframomum* K.Schum. (Zingiberaceae) from Ghana, with notes on spherical geocarpic fruits in the genus. *Kew Bull.* 28, 441–449.
- S28. Lock, J.M., and Hall, J.B. (1975). Taxonomic studies in the genus *Aframomum* (Zingiberaceae). *Boissiera* 24, 225–231.
- S29. Lock, J.M. (1976). Notes on East African species of *Aframomum* (Zingiberaceae). *Kew Bull.* 31, 263–271.
- S30. Lock, J.M. (1978). Notes on the genus *Aframomum* (Zingiberaceae): 1. The species related to *Aframomum polyanthum* (K. Schum.) K. Schum. . *Bull. Jard. Bot. Nat. Belg.* 48, 129–134.
- S31. Lock, J.M. (1978). Notes on the genus *Aframomum* (Zingiberaceae). 2. The Ethiopian species. *Bull. Jard. Bot. Nat. Belg.* 48, 387–391.
- S32. Lock, J.M. (1978). Notes on the genus *Aframomum* (Zingiberaceae). 3. Two new species from West Africa. *Bull. Jard. Bot. Nat. Belg.* 48, 393–398.
- S33. Lock, J.M. (1979). Notes on the genus *Aframomum* (Zingiberaceae). 4. The savanna species. *Bull. Jard. Bot. Nat. Belg.* 49, 179–184.
- S34. Lock, J.M. (1980). Notes on *Aframomum* (Zingiberaceae) in West Africa, with a New Key to the Species. *Kew Bull.* 35(2), 299–313.
- S35. Lock, J.M. (1984). Notes on East African Zingiberaceae. *Kew Bull.* 39, 837–843.

- S36. Poulsen, A.D., and Lock, J.M. (1997). New species and new records of Zingiberaceae and Costaceae from tropical East Africa. *Kew Bull.* 52, 601–616.
- S37. Dhetchuvi, J.-B., and Fischer, E. (2006). A New Species of *Aframomum* (Zingiberaceae) from Nyungwe National Park/Rwanda. *Syst. Geogr. Plants* 76(2), 241–245.
- S38. Dhetchuvi, J.-B., Wortley, A.H., and Harris, D.J. (2011). A new species of *Aframomum* (Zingiberaceae) from Central Africa. *Phytotaxa* 28, 31–34.