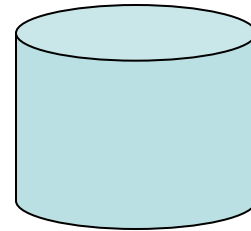
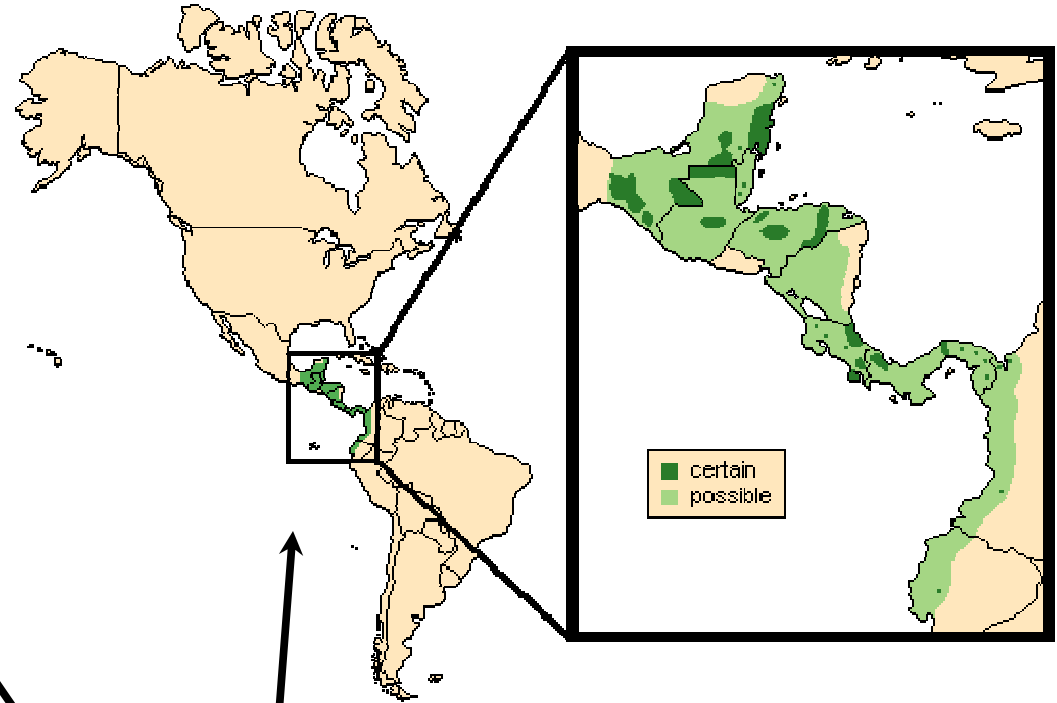
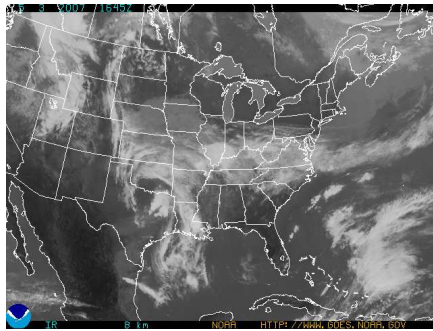
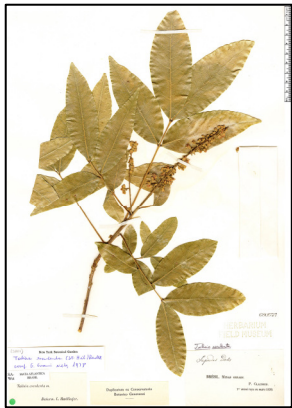


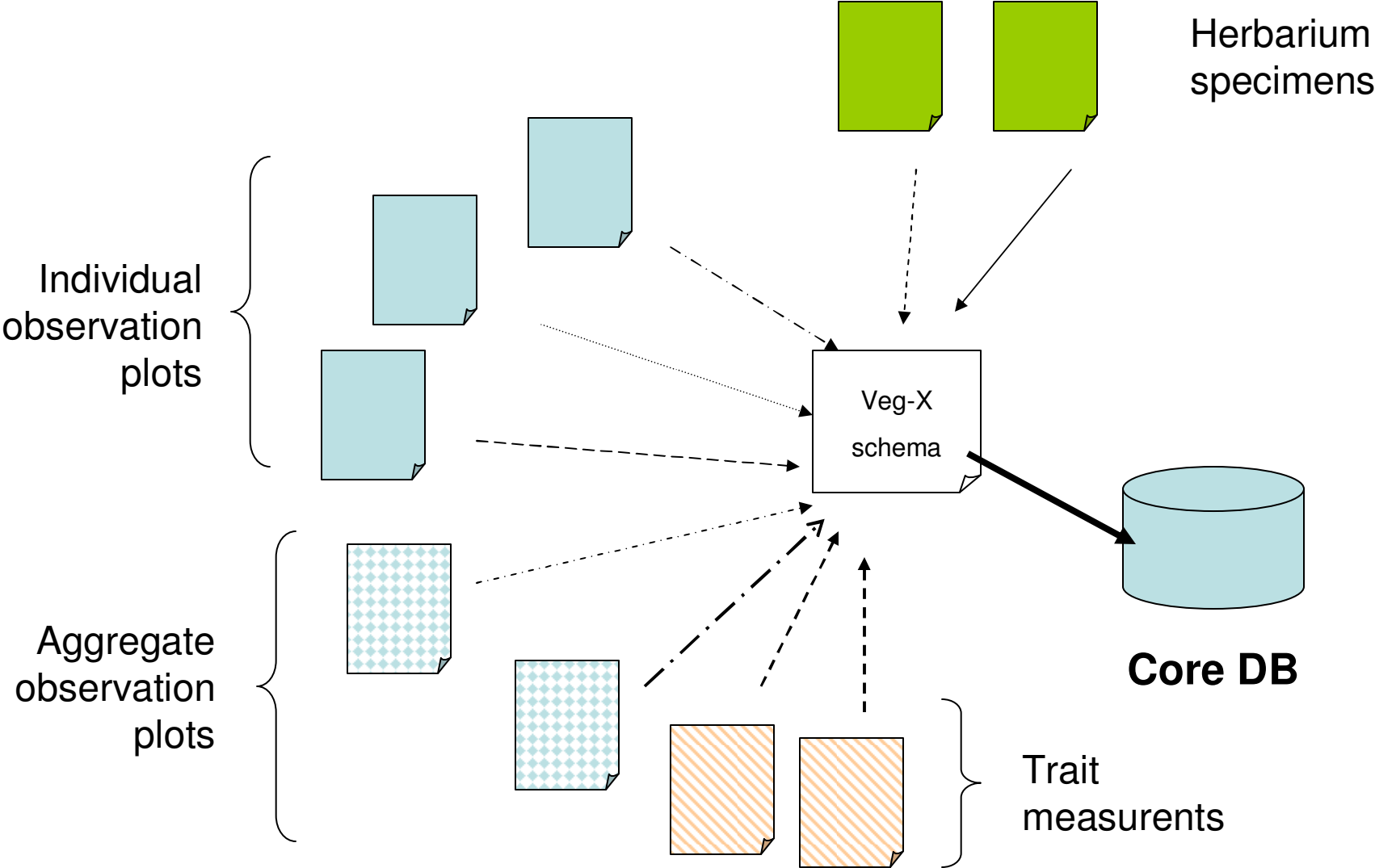


SALVIAS, BIEN:
Biological “megadatasets” and
taxonomic challenge

Brad Boyle
BIEN / University of Arizona



BIEN database



Microsoft Excel - gentry_example.xls

File Edit View Insert Format Tools Data Window Help Adobe PDF

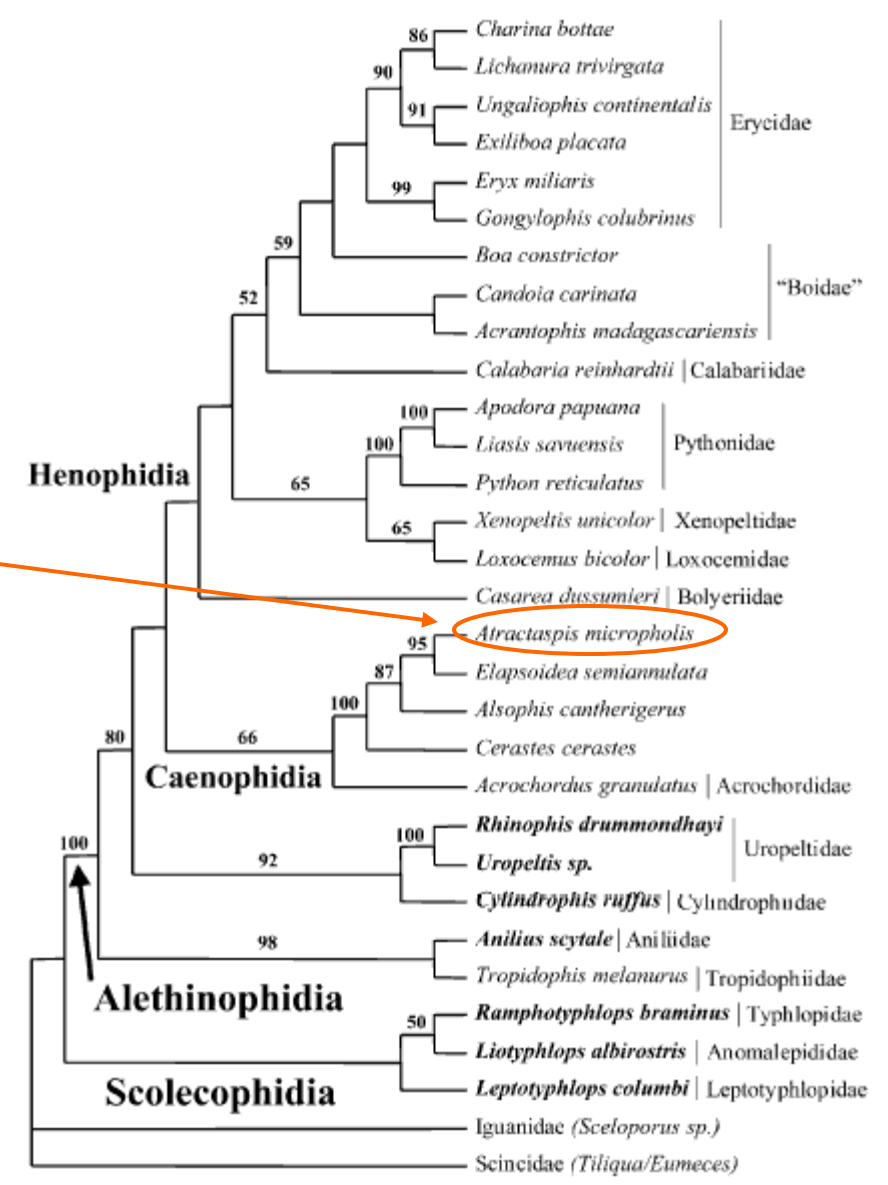
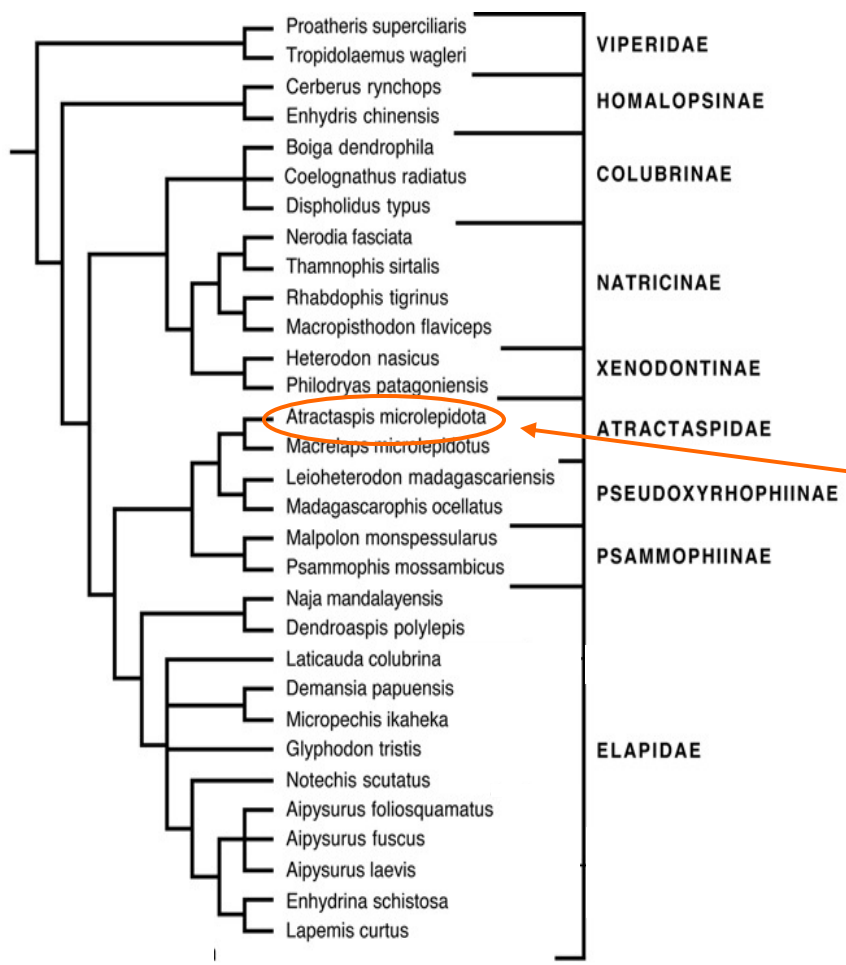
100% Arial 10

	A	C	D	E	F	G	P	Q	R	S
	OBSERVATION ID	plot	subplot	collector_code	coll_lastname	coll_firstname	family	species	habit	no_of_individuals
2	71106	TAMBUPL	1	GENTRY AH	Gentry	Alwyn H.	ANACARDIACEAE	Thyrsoedium M1	T	1
3	71111	TAMBUPL	1	GENTRY AH	Gentry	Alwyn H.	ANNONACEAE	Xylopia M1	T	1
4	71117	TAMBUPL	1				ARECACEAE	Geonoma interrupta	T	1
5	71147	TAMBUPL	1				BIGNONIACEAE	Jacaranda copaia	T	1
6	71160	TAMBUPL	1				CECROPIACEAE	Pourouma cucura	T	1
7	71164	TAMBUPL	1	GENTRY AH	Gentry	Alwyn H.	CECROPIACEAE	Pourouma guianensis	T	1
8	71167	TAMBUPL	1				CECROPIACEAE	Pourouma minor	T	0
9	71167	TAMBUPL	1				CECROPIACEAE	Pourouma minor	T	2
10	71194	TAMBUPL	1	GENTRY AH	Gentry	Alwyn H.	COMBRETACEAE	Combretum M1	L	0
11	71194	TAMBUPL	1	GENTRY AH	Gentry	Alwyn H.	COMBRETACEAE	Combretum M1	L	0
12	71194	TAMBUPL	1	GENTRY AH	Gentry	Alwyn H.	COMBRETACEAE	Combretum M1	L	4
13	71195	TAMBUPL	1	GENTRY AH	Gentry	Alwyn H.	COMBRETACEAE	Combretum M1	L	1
14	71200	TAMBUPL	1	GENTRY AH	Gentry	Alwyn H.	COMBRETACEAE	Thiloo paraguayensis	L	0
15	71200	TAMBUPL	1	GENTRY AH	Gentry	Alwyn H.	COMBRETACEAE	Thiloo paraguayensis	L	0
16	71200	TAMBUPL	1	GENTRY AH	Gentry	Alwyn H.	COMBRETACEAE	Thiloo paraguayensis	L	3
17	71208	TAMBUPL	1	GENTRY AH	Gentry	Alwyn H.	CONNARACEAE	Connarus M1	L	1
18	71210	TAMBUPL	1	GENTRY AH	Gentry	Alwyn H.	CONNARACEAE	Rourea cuspidata	L	1
19	71223	TAMBUPL	1	GENTRY AH	Gentry	Alwyn H.	FABACEAE	Inga chartacea	T	1
20	71230	TAMBUPL	1	GENTRY AH	Gentry	Alwyn H.	FABACEAE	Tachigali polyphylla	T	1
21	71236	TAMBUPL	1				FLACOURTIACEAE	Lindackeria paludosa	T	1
22	71249	TAMBUPL	1	GENTRY AH	Gentry	Alwyn H.	LAURACEAE	Ocotea M1	T	1
23	71259	TAMBUPL	1				LINACEAE	Roucheia punctata	T	1
24	71265	TAMBUPL	1	GENTRY AH	Gentry	Alwyn H.	MALPIGHIACEAE	M1 M1	L	1
25	71267	TAMBUPL	1	GENTRY AH	Gentry	Alwyn H.	MALPIGHIACEAE	M2 M1	L	1
26	71269	TAMBUPL	1	GENTRY AH	Gentry	Alwyn H.	MELASTOMATACEAE	M1 M1	T	1
27	71276	TAMBUPL	1	GENTRY AH	Gentry	Alwyn H.	MELASTOMATACEAE	Miconia serrulata	T	1
28	71284	TAMBUPL	1				MONIMIACEAE	Mollinedia killippii	T	0
29	71284	TAMBUPL	1				MONIMIACEAE	Mollinedia killippii	T	2
30	71292	TAMBUPL	1	GENTRY AH	Gentry	Alwyn H.	MONIMIACEAE	Siparuna bifida	T	1
31	71305	TAMBUPL	1				MYRISTICACEAE	Iryanthera juruensis	T	0
32	71305	TAMBUPL	1				MYRISTICACEAE	Iryanthera juruensis	T	2
33	71321	TAMBUPL	1				MYRISTICACEAE	Virola calophylla	T	1
34	71323	TAMBUPL	1	GENTRY AH	Gentry	Alwyn H.	MYRISTICACEAE	Virola multinervis	T	1
35	71324	TAMBUPL	1	GENTRY AH	Gentry	Alwyn H.	MYRISTICACEAE	Virola sebifera	T	1
36	71331	TAMBUPL	1	GENTRY AH	Gentry	Alwyn H.	MYRSINACEAE	Cybianthus aff resinus	T	1
37	71342	TAMBUPL	1	GENTRY AH	Gentry	Alwyn H.	PIPERACEAE	Piper arboreum cf	T	1
38	71356	TAMBUPL	1				RUBIACEAE	Remijia megistocaula	T	1
39	71365	TAMBUPL	1	GENTRY AH	Gentry	Alwyn H.	SAPOTACEAE	Sarcaulus brasiliensis ssp grasilis	T	1
40	71109	TAMBUPL	10	GENTRY AH	Gentry	Alwyn H.	ANNONACEAE	Guatteria scytophylla	T	1
41	71130	TAMBUPL	10				BIGNONIACEAE	Adenocalymna purpurascens	L	0
42	71130	TAMBUPL	10				BIGNONIACEAE	Adenocalymna purpurascens	L	2

plot_dl

Ready

Start AR... Go... tax... bo... uBi... ph... SQ... C:\... co... SA... Mi...



Taxonomic uncertainty

1. Non-existent names

- Misspellings
- Contamination

2. Synonymy

- Nomenclatural synonyms
- Taxonomic synonyms / concepts

3. Misidentifications, incomplete identifications

Non-existent names

	Genus	Specific epithet	Author
	<i>Acer</i>	<i>rubrum</i>	L.
• Misspellings →	<i>Acer</i>	<i>rubum</i>	L.
• Unpublished names ↗	<i>Clusia</i>	<i>haberi</i>	Hammel ined.
• “Contamination” by:	<i>Acer</i>	<i>rubrum</i>	L.
– annotations of uncertainty →	<i>Acer</i>	<i>cf. rubrum</i>	
– Morphospecies →	<i>Clusia</i>	<i>sp.3</i>	

Common forms of “contamination”

Annotations of
uncertainty, etc.

Miconia cf. *tonduzii*
Palicourea aff. *pallens*
Erythrina biloba –aff.
Randia acuminata??

Morphospecies

Miconia sp.3
Palicourea “fuzzy leaf”
Erythrina M6
Fabaceae gen_indet. sp.6
Indet. Ochoa_6566

Non-existent names

Digitization errors:

- Frame shifts & other import errors
- Character set problems (e.g., accents)

Genus	Specific_ep	Author	Det_date
<i>Acer</i>	<i>rubrum</i>	L.	24/3/05
<i>Clusia</i>	<i>retusa</i>	L.15/4/06	
<i>Ticodendron</i>	<i>incognitum</i>	Gómez-Laurito & L.D. Gómez	3/6/02
<i>Ticodendron</i>	<i>incognitum</i>	G&24mez-Laurito & L.D. G&24mez	5/5/02

Gentry data set

ID	Family	Genus	SpecificEpithet
1	INDET	M1	M1
2	LAURACEAE	M1	M1
3	MELASTOMATAACEAE	MICONIA CF	M6
4	MELASTOMATAACEAE	MICONNIA?	M1
5	POLYGALACEAE	MONINA??	M2
6	MORACEAE	FICUS	CITRIFOLIA CF
7	MORACEAE	PSEUDOLMEDIA CF	M1
8	MORACEAE	SOROCEA	OPIMA
9	SAPOTACEAE	CHRYSOPHYLLUM	ARGENTEUM SSP AURATUM
10	PROTEACEAE	Panopsis?	M1
11	BORAGINACEAE	Cordia	cf. alliodora
12	BORAGINACEAE	Cordia	curasavica
13	BORAGINACEAE	Tournefortia	aff. ternifolia
14	FABACEAE	Erythrina	amazonicum
15	STERCULIACEAE	Dombeya	ankarafantiskae
16	FABACEAE	Leucaena	trichodes

Non-existent names: Herbarium specimens

Total specimens:	1.1 million
Unique species names:	53,052
Published names (legitimate & illegitimate):	44,532
Misspelled names:	9371 (18%)
Specimens with misspelled names:	101,237 (9%)

*New World plant specimens, 34 herbaria, simple match against IPNI and TROPICOS, excluding authors

Non-existent names: Vegetation plots

Total observations (individuals):	41,549
Unique names:	7,367
Published names (legitimate & illegitimate):	5,573
Misspelled names:	1,794 (24%)
Observations with misspelled names:	6,185 (15%)

*Original Gentry transect database, checked against IPNI
and TROPICOS, excluding morphospecies

Synonymy

Centaurium curvistamineum (Wittr.) Abrams (1951)

Centaurium minimum (Howell) Piper (1915)

Centaurium muhlenbergii (Griseb.) Wight ex Piper (1906)

Centaurium muhlenbergii (Griseb.) Wight ex Piper forma *albiflorum*
(Suksd.) St. John (1937)

Centaurium muhlenbergii (Griseb.) Wight ex Piper var. *albiflorum*
Suksd. (1927)

Centaurodes muhlenbergii (Griseb.) Kuntze (1891)

Erythraea curvistaminea Wittr. (1886)

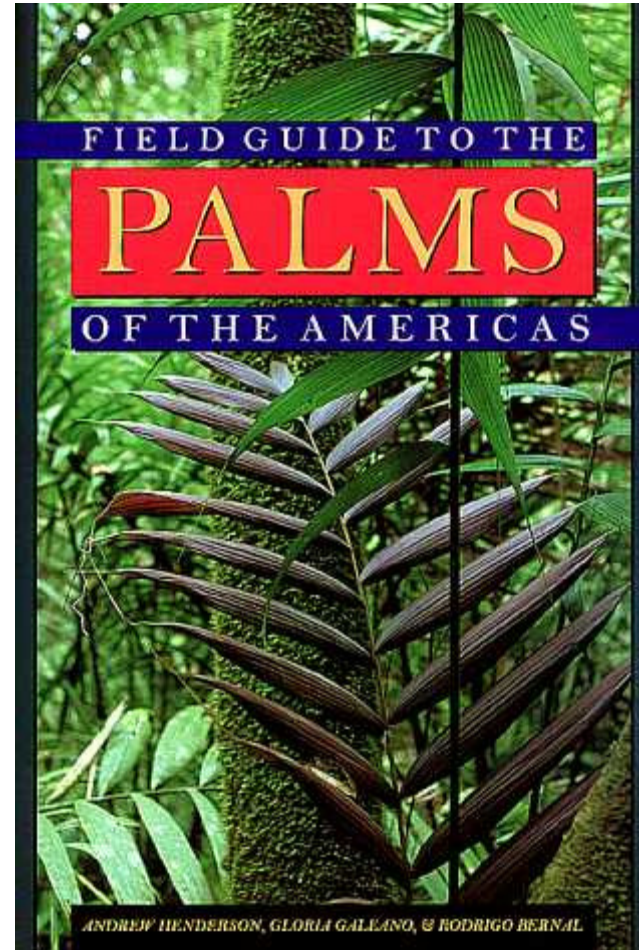
Erythraea minima Howell (1901)

Erythraea muhlenbergii Griseb. (1839)

Synonymy

Total species names:
1903

Number of “accepted”
species recognized:
567



Nomenclatural synonyms

- Based on same type (Homotypic synonyms)
- Change in rank or group
- Same specific epithet

Eschatogramme panamensis (1929)

Polypodium panamense (1954)

Dicranoglossum panamense (1976)

(**Basionym**: earliest name)



Taxonomic synonyms

- Based on different types (Heterotypic synonyms)
- Involves taxonomic judgement (species boundaries)
- Different specific epithets

Example:

Polybotrya caudata Kunze (1834)

Polybotrya costaricensis Brade
(1969)



***Dicranum fuscescens* Sm.**

Flora Britannica, 1804: 27

= *Dicranum congestum* Brid.

[sec. Koperski et al., 2000]

[nomenclatural source]

[heterotypic synonym]

== *Dicranum fuscescens* Sm.

sec. Corley et al. (1981, 1991)

== *Dicranum fuscescens* Sm.

sec. Ludwig et al. (1996)

Ludwig et al. (see there) refer to the concept of Corley et al.

== *Dicranum fuscescens* var. *eu-fuscescens* Mönk.

sec. Mönkemeyer (1927)

< *Dicranum fuscescens* Turner

sec. Frahm and Frey (1992)

Includes *D. flexicaule* (see comments there).

< *Dicranum fuscescens* Turner

sec. Mönkemeyer (1927)

Includes *D. flexicaule* (see comments there).

< *Dicranum fuscescens* Sm.

sec. Smith (1980)

Includes *D. flexicaule* in the type variety (cf. morphological account).

> *Dicranum fuscescens* var. *congestum* (Brid.) Husn.

sec. Smith (1980)

This taxon is evidently a montane growth form of *D. fuscescens*.

>< *Dicranum fuscescens* var. *fuscescens*

sec. Smith (1980)

from: Franz et al, *in press*.

Name 1

{
=
>
<
>
—
}

Name 2

?

Synonymy triage

Automate sorting of synonyms into:

- **simple:** 1:1 equality ($==$), also, old name contained by new name ($<$)
- **complex:** requires inspection by user (e.g., $>$, $><$)

Simple synonym

Synonym

Accepted name

Aberemoa odorata Diels

Duguetia odorata (Diels) J.F. Macbr.

Aberemoa pedunculata Diels

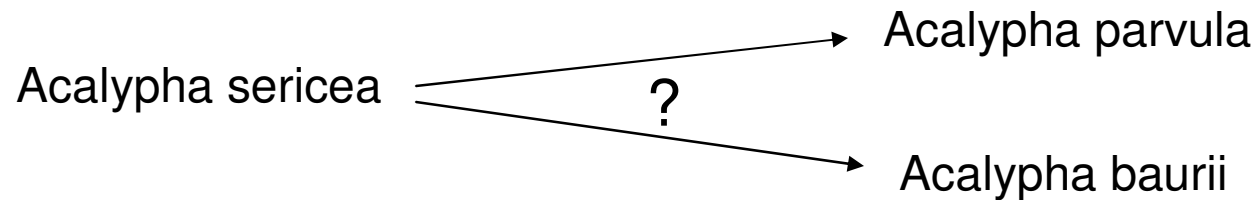
Crematosperma pedunculatum (Diels) R.E. Fr.

Abatia macrostachya Killip

Abatia spicata (Turcz.) Sleumer

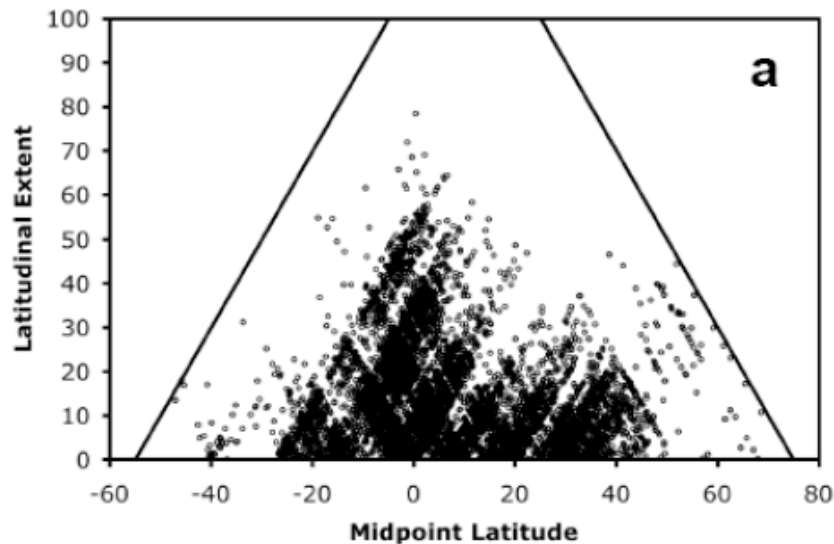
Complex synonyms

Synonym	Accepted name
<i>Acalypha sericea</i> Andersson	<i>Acalypha parvula</i> Hook. f.
<i>Acalypha sericea</i> var. <i>baurii</i> (B.L. Rob. & Greenm.) G.L. Webster	<i>Acalypha baurii</i> B.L. Rob. & Greenm.
<i>Acalypha sericea</i> var. <i>indefessus</i> G.L. Webster	<i>Acalypha baurii</i> B.L. Rob. & Greenm.

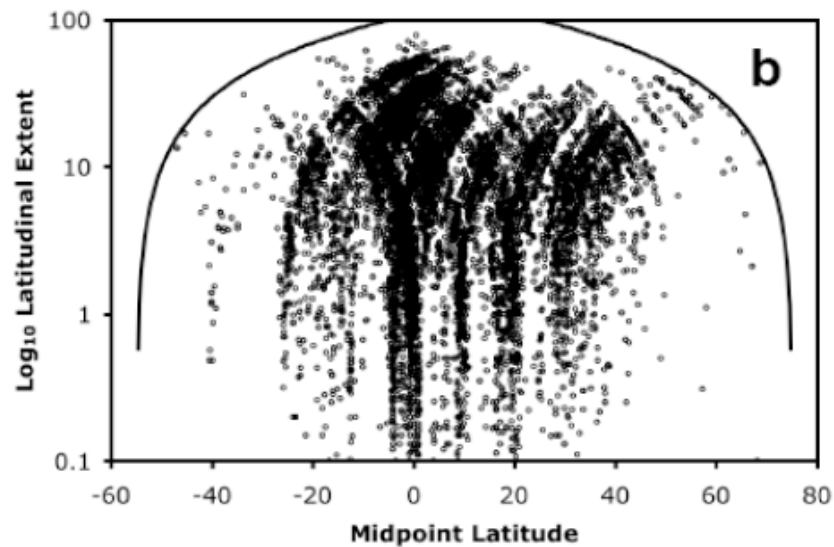


...requires user inspection

One approach to taxonomy...

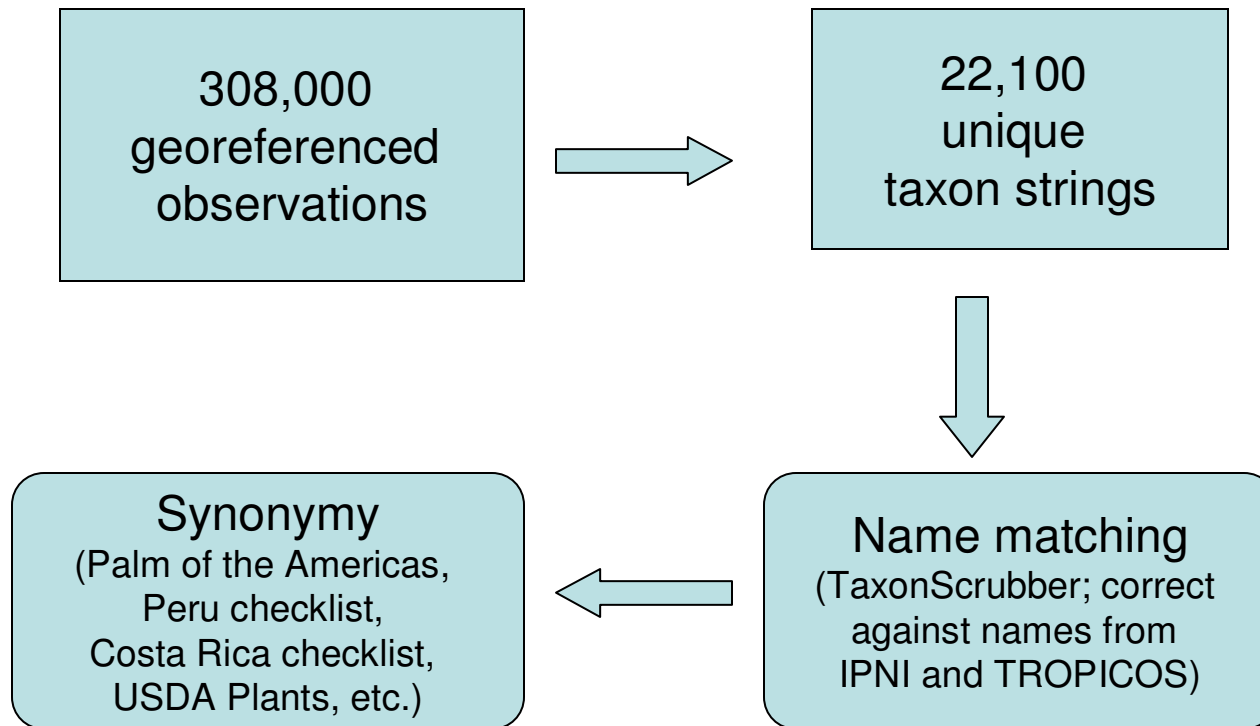


Latitudinal extents of 12,980 species of New World Woody Plants.

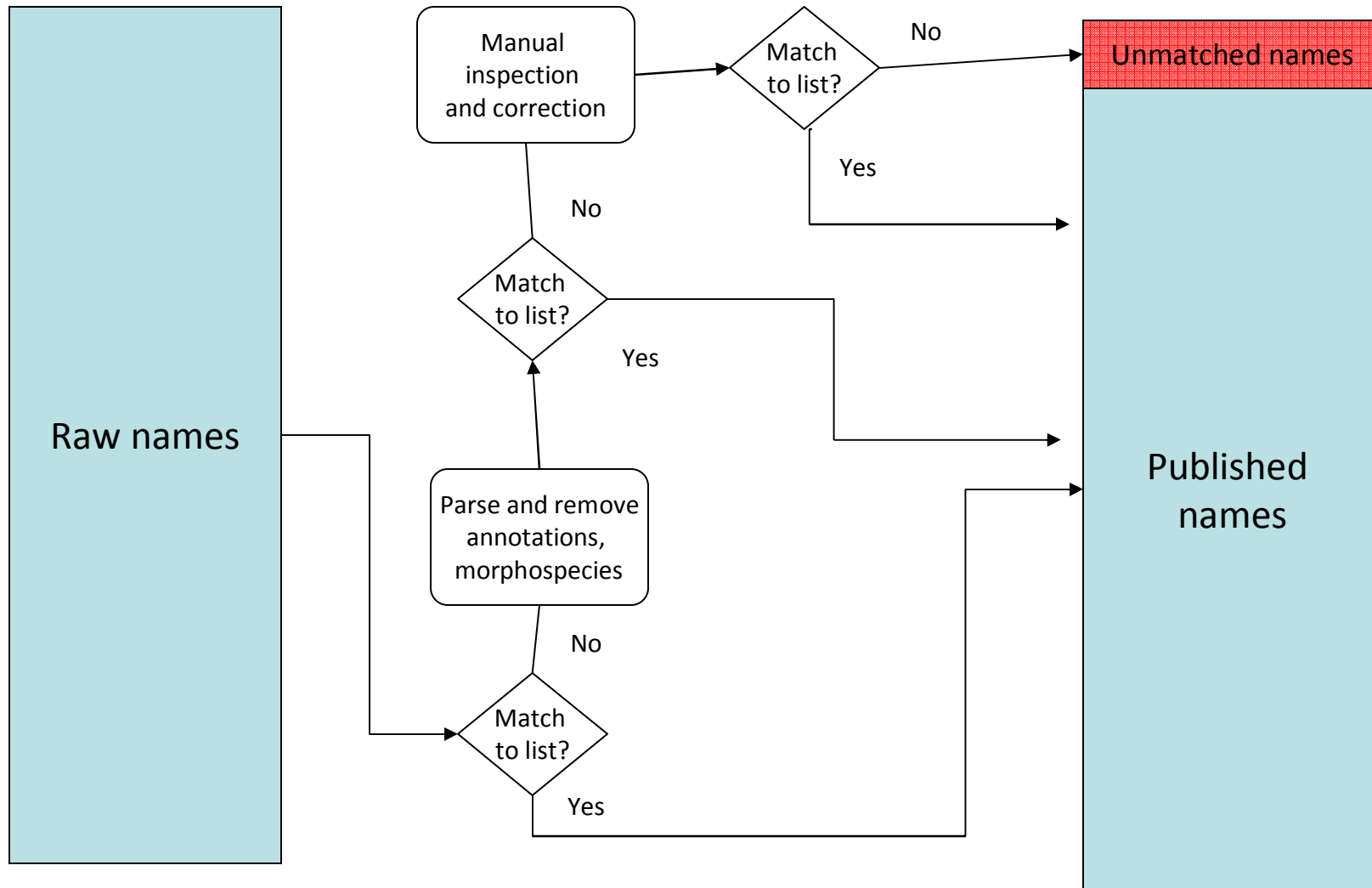


Weiser et al., 2007. GEB 16: 679-688.

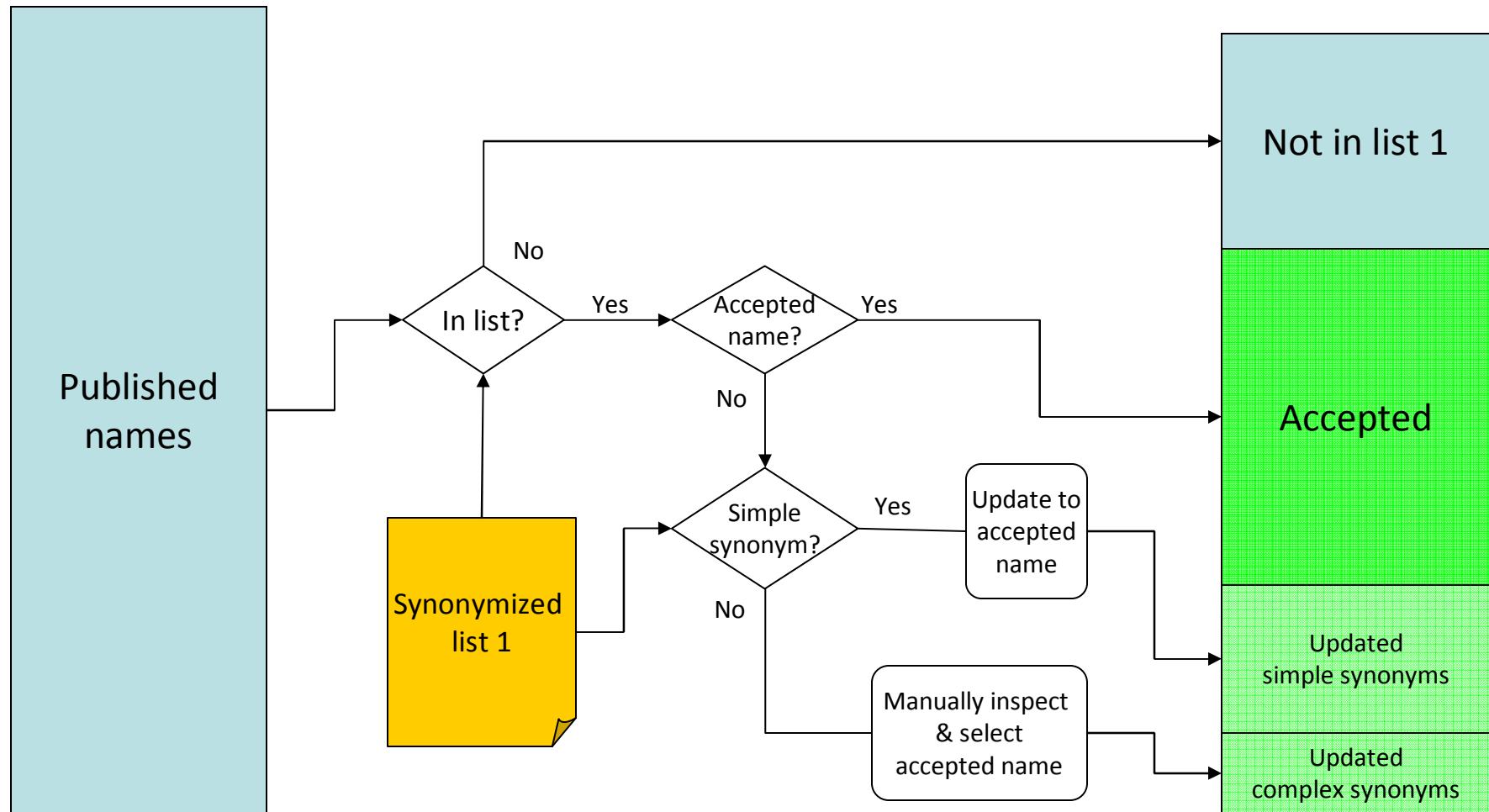
Weiser et al., 2007. Latitudinal Ranges



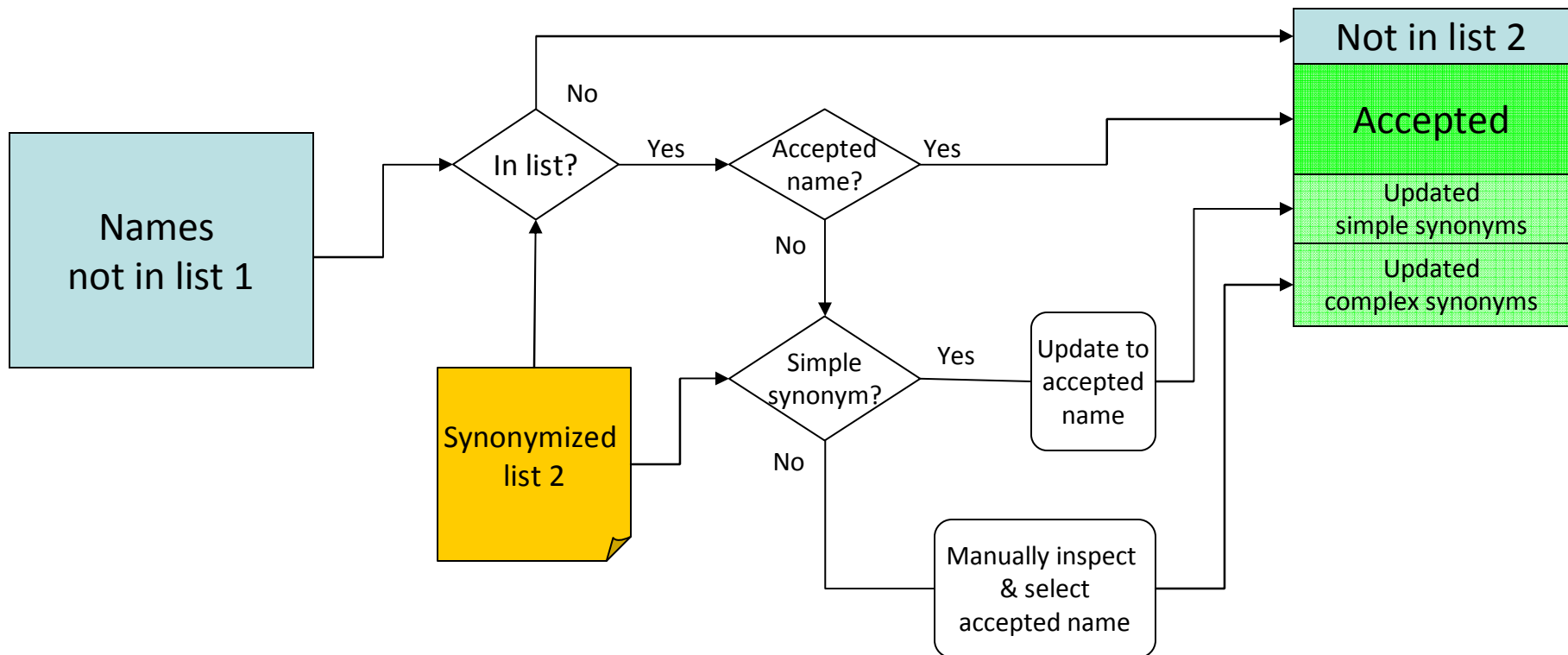
Name matching



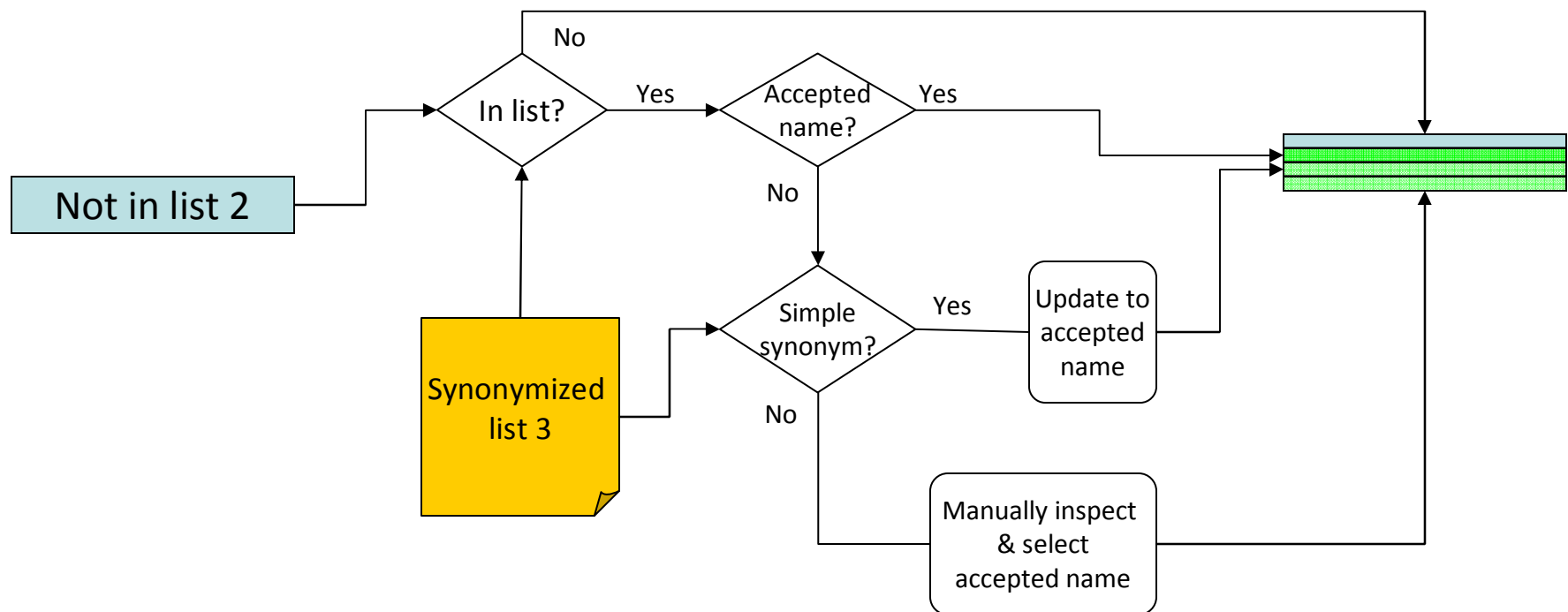
Synonymy round 1



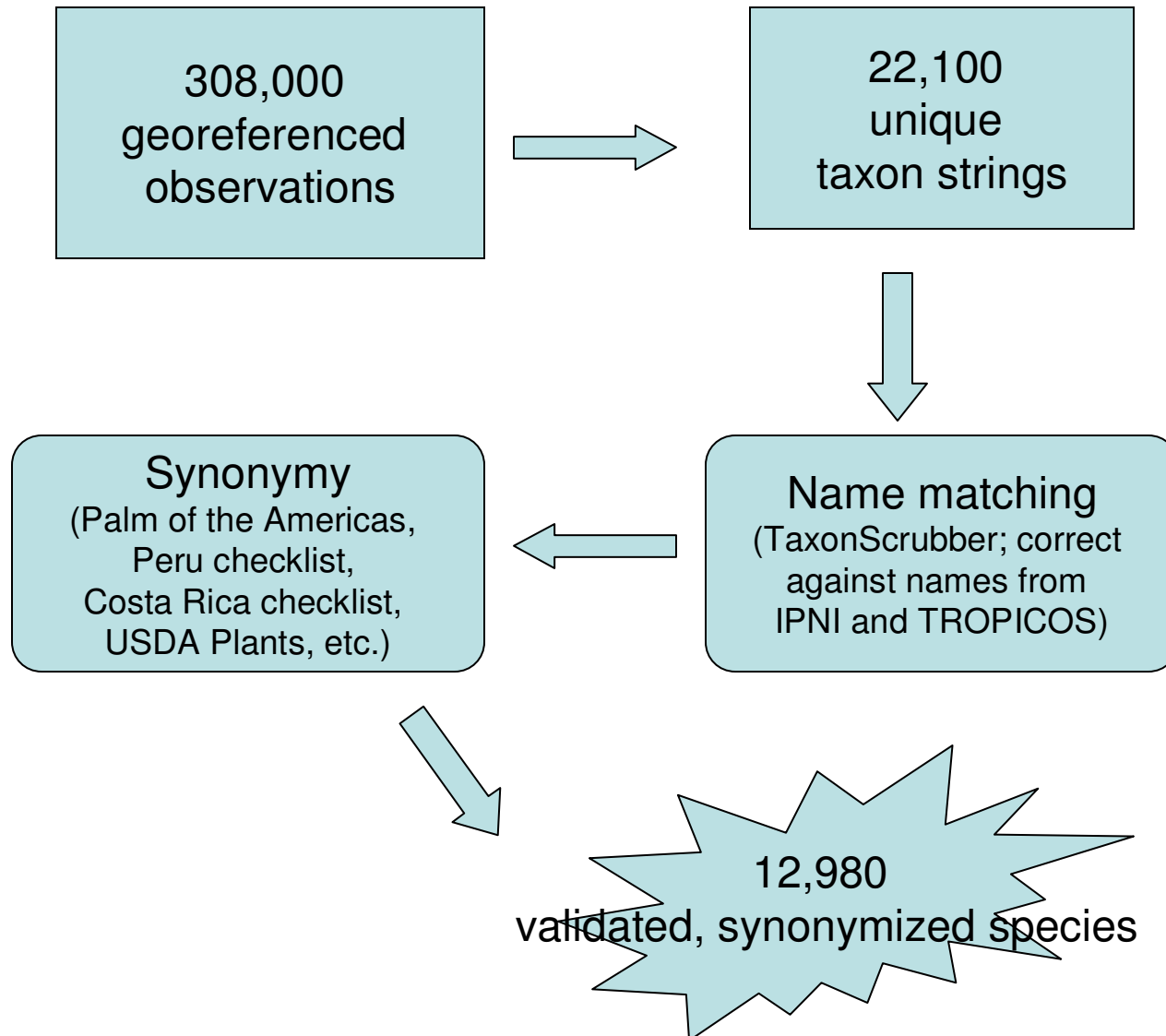
Synonymy round 2



Synonymy round 3



Weiser et al., 2007. Latitudinal Ranges

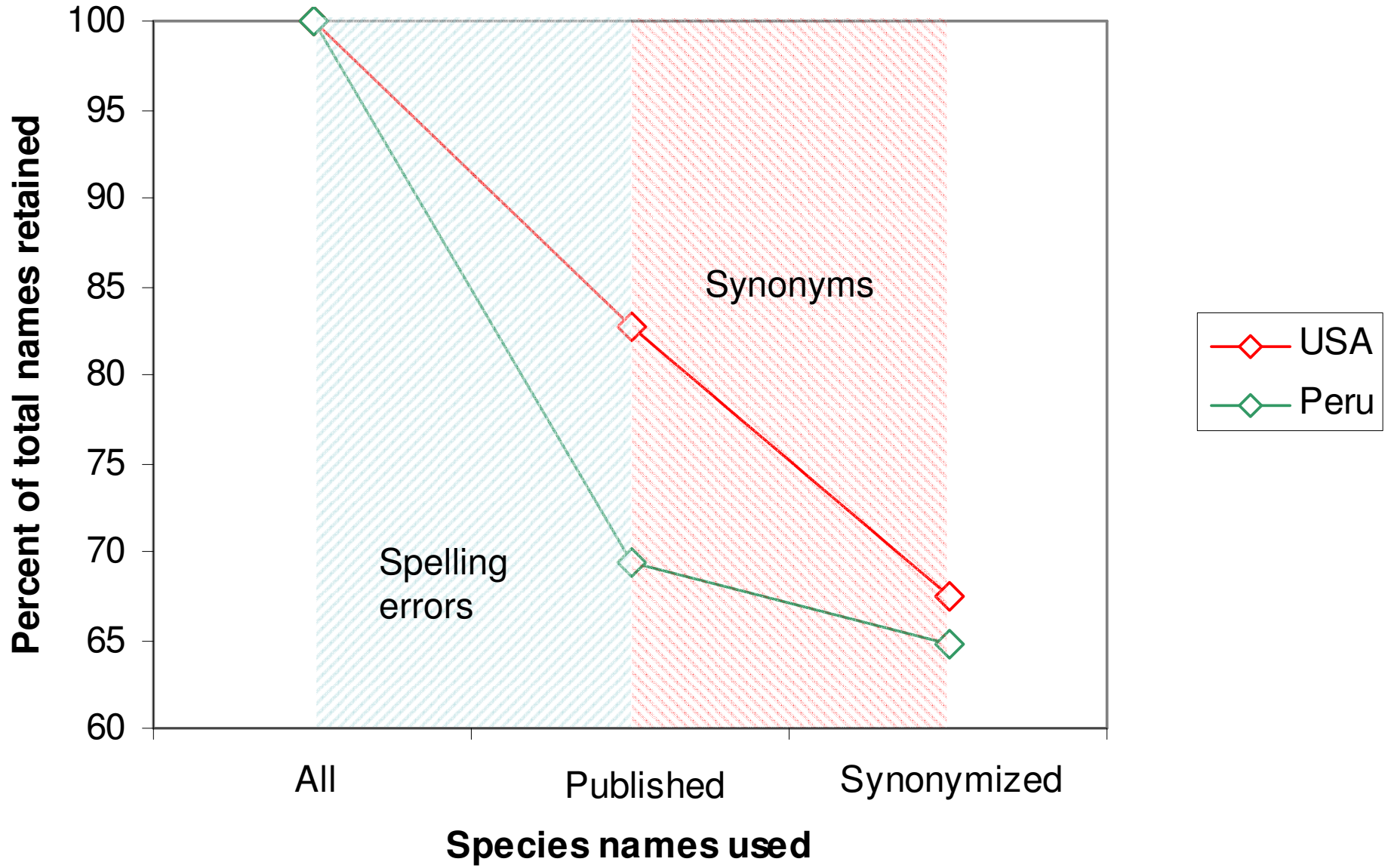


Synonyms: Regional Differences

Herbarium specimens in SALVIAS synonymized using USDA Plants list (USA) and Brako & Zarucchi checklist (Peru)

Region	All names	Published names	Unambiguous synonymized names
USA	9779	8093	6604
Peru	5534	3842	3585

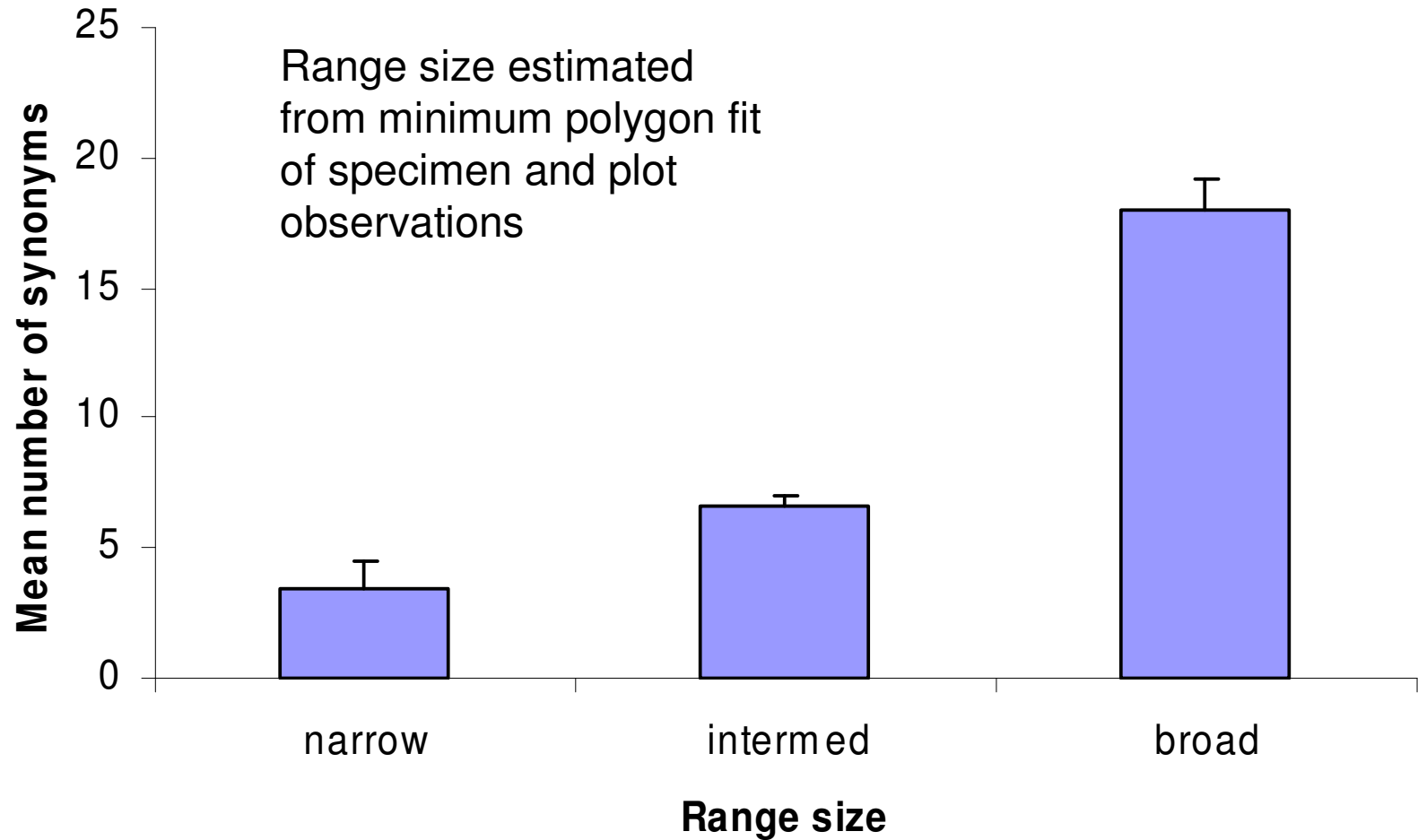
"Loss" of species during taxonomic correction



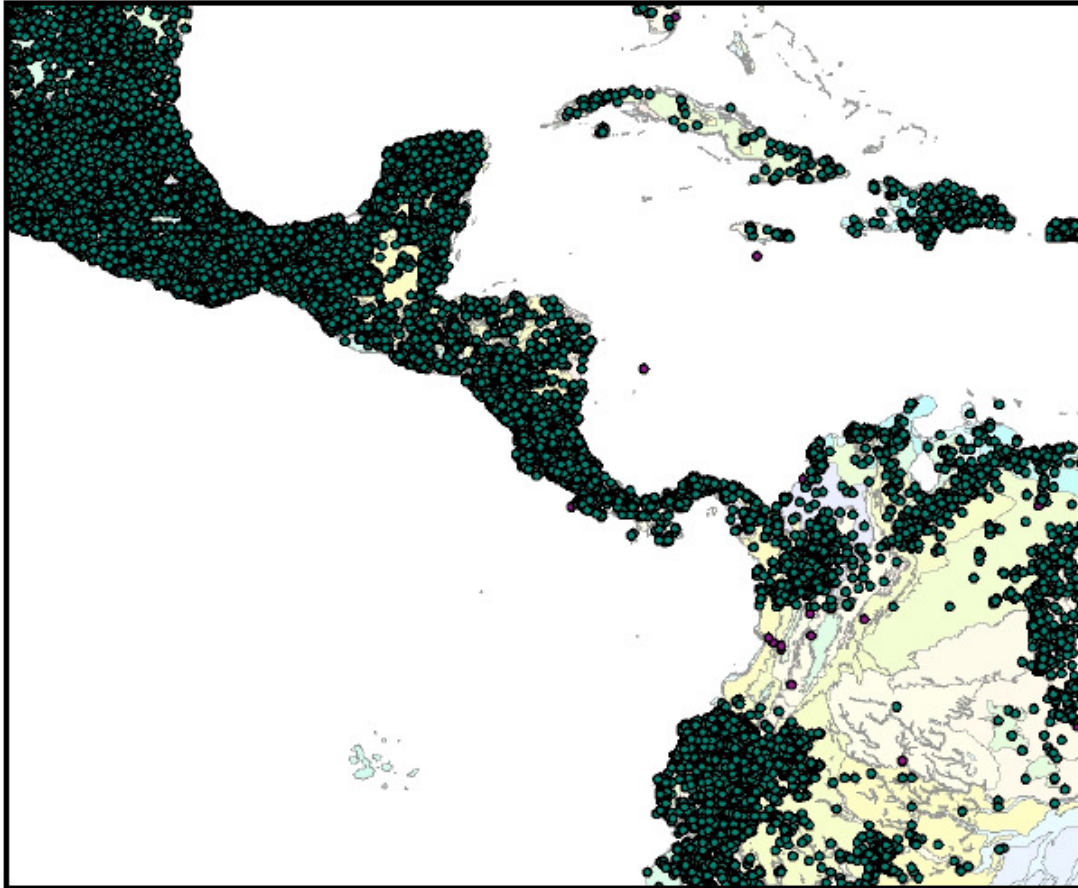
Synonyms per name

	Median
Temperate	6
Tropical	3
Wilcoxon Z	4.22 (P<.0001)

Synonymy and species range



Current BIEN database



Specimens:

7.8 M

Plot observations:

10.7 M

Unique name strings:

~315,000

Total species:

????