The Use of Leap in Herbarium Management and Plant Biodiversity Research

Authors: Knox, Eric B., and Berghe, Edward Vanden

Source: Journal of East African Natural History, 85(1) : 65-79

Published By: Nature Kenya/East African Natural History Society

URL: https://doi.org/10.2982/0012-8317(1996)85[65:TUOLIH]2.0.CO;2

Journal of East African Natural History **85**: 65–79(1996)

THE USE OF LEAP IN HERBARIUM MANAGEMENT AND PLANT BIODIVERSITY RESEARCH

Eric B. Knox[1]
Edward Vanden Berghe[2]
National Museums of Kenya
P.O. Box 40658, Nairobi, Kenya

The List of East African Plants (LEAP) is an electronic database of the species, subspecies, and varieties of vascular plants (angiosperms, gymnosperms, ferns, and fern allies) that are native or naturalised in East Africa (Uganda, Kenya, and Tanzania). The information in LEAP comes primarily from the *Flora of Tropical East Africa* (*FTEA*; published as family fascicles from 1952 to present) and is supplemented by other published sources, such as *The Sedges and Rushes of East Africa* (Haines & Lye, 1983), *Pteridophytes of Tropical East Africa* (Johns, 1991), *Upland Kenya Wild Flowers* (Agnew & Agnew, 1995), and *Kenya Trees, Shrubs and Lianas* (Beentje, 1995), and specimen information from the East African Herbarium (EA) and other sources for those plant families not yet written for FTEA. This information includes taxon names, author citations, life form, origin, altitudinal range, geographic distribution (using FTEA floral division codes), and higher-level classification. Presented below is a brief history of LEAP, the current structure and organisation of LEAP version 1.3, a discussion of potential applications for herbarium management and plant biodiversity research (along with the limitations of such a database) and the work to be conducted in the future.

THE HISTORY OF LEAP

The desire for a species-level database on the plants in East Africa developed during an annual series of regional databasing meetings in East Africa. The work was initiated at Makerere University, Kampala, Uganda, with the entry of information from the published families of FTEA. In 1994, the master version of the database was moved to the East African Herbarium (administered as a department of the National Museums of Kenya, Nairobi) in order to check the initial entries, add the remaining published families, add more detailed information on altitudinal and geographic distribution, and to undertake the more difficult task of compiling lists for the families not yet published.

LEAP VERSION 1.3

LEAP was constructed using a commercially available relational database package (Microsoft

[1] Present address: Department of Biological Sciences, Rutgers University, 101 Warren Street, Newark, NJ 07102 USA

[2] Present address: Brusselmansstraat 14, 1700 Dilbeek, Belgium

Access 2.0) on IBM-compatible computers with Microsoft Windows. The LEAP database consists of five tables:

- The main LEAP Table;
- A Higher Classification Table;
- A Source Table;
- A Synonym Table; and
- The Recipients of the LEAP Table.

The field Family links the main LEAP and Higher Classification Tables (see tables 1 & 2), and a query named Classification Order can be used to view the LEAP data according to the classification order instead of the alphabetic order of the main LEAP Table. The field Source provides cross-references between the LEAP and Source Tables, but these fields are not linked. The Synonym Table lists previously used names that are now regarded as synonyms, along with the currently accepted name listed in LEAP. The Recipients of LEAP Table simply records the distribution of LEAP. The organisational structure of LEAP follows FTEA in two important aspects. First, FTEA uses a phylogenetic sequence for listing the genera in a family and the species in a genus. Second, the three countries of East Africa are subdivided into 21 floral divisions (see recent volumes of FTEA) which are coded and used to report the geographic distribution of each terminal taxon (the species, subspecies, or variety, depending on the extent of taxonomic subdivision). The status of the information in the tables and fields of LEAP is discussed below.

THE MAIN LEAP TABLE

Family

The species of native plants in East Africa are classified in 251 plant families. Of these, 175 families have been published in FTEA and the pteridophytes (31 families) and 32 families of angiosperms are in preparation. The remaining 13 families have not yet been started, but these are mostly small families with only 284 species in the FTEA region. The family entries have been checked against the FTEA list of families for completeness and have been checked against *Vascular Plant Families and Genera* (VPFG, Brummitt, 1992) for correct spelling. The delimitation of plant families and genera and the classification of genera into families still depend on taxonomic judgement. The taxonomic structure of VPFG is based on current thinking at the Royal Botanic Gardens, Kew, UK, but is not universally accepted; it will change as taxonomic research progresses, and there are many points of disagreement between it and the classification used for FTEA. In general, LEAP follows FTEA since this constitutes the principle floristic literature for the region. Notable changes include:

- The genera originally included in the FTEA volume for the Hypericaceae are now listed under the Guttiferae as was discussed in the FTEA treatment for the Guttiferae (Bamps *et al.*, 1978);
- The spelling for the Capparaceae is accepted over that used for FTEA (Capparidaceae);
- The Costaceae have been separated from the Zingiberaceae, as discussed in the FTEA treatment for the Musaceae (Lock, 1993).

Numerous additional changes at generic, specific, and infraspecific levels have been made in accordance with published treatments, and these are indicated in the Source and Synonym Tables.

Table 1. Data structure for the main LEAP table.

Field Name	Contents	Data Type	Size
Record Number	Unique numerical identifier	Counter	
Family	Family name	Text	20
Subfamily	Subfamily name (Leguminosae)	Text	20
Genus Number	Number in FTEA-assigned phylogenetic sequence	Number	
Genus	Genus name	Text	20
Species Number	Number in FTEA-assigned phylogenetic sequence	Number	
Species	Species name	Text	75
Species Author	Author citation for currently accepted species name	Text	50
Subspecies	Subspecies name	Text	25
Subspecies Author	Author citation for currently accepted subspecies name	Text	50
Variety	Variety name	Text	25
Variety Author	Author citation for currently accepted variety name	Text	50
Life Form Code	Tree, Shrub, Herb, or Climber	Text	1
Origin Code	Native or Introduced	Text	1
Lower Alt Limit	Lower altitudinal limit recorded in East Africa	Number	
Upper Alt Limit	Upper altitudinal limit recorded in East Africa	Number	
U	Floral division codes in Uganda	Text	10
K	Floral division codes in Kenya	Text	10
T	Floral division codes in Tanzania	Text	10
Z	Presence or absence on Zanzibar Island	Text	1
P	Presence or absence on Pemba Island	Text	1
Endemic	Marked 'Yes' if known to be endemic to East Africa	Yes/No	
Other Countries	Abbreviations for other countries of taxon's distribution	Text	75
Habitat	Short description of habitat(s)	Text	100
Notes	Miscellaneous comments	Text	100
Source	Source(s) of information	Text	50
Date of Entry	Last date the record was modified	Date	

Table 2. Data structure for the Higher Classification table.

Field Name	Contents	Data Type	Size
Dnum	Number in phylogenetic sequence	Number	
Division	Division Name	Text	15
Cnum	Number in phylogenetic sequence	Number	
Class	Class Name	Text	15
Snum	Number in modified Cronquist/Dahlgren phylogenetic sequence	Number	
Subclass	Subclass Name	Text	15
Onum	Number in modified Cronquist/Dahlgren phylogenetic sequence	Number	
Order	Order Name	Text	15
Fnum	Number in modified Cronquist/Dahlgren phylogenetic sequence	Number	
Family	Family Name	Text	20

Subfamily

The Leguminosae were treated in FTEA as comprising three subfamilies (Mimosoideae, Caesalpinioideae, and Papilionoideae) with the phylogenetic sequence of genera numbered starting with one in each subfamily. The subfamily names of the Leguminosae are included to aid in the retrieval of information within this family, but because it is not a required rank,

genera in other families are not necessarily assigned to subfamilies. In those families that regularly employ subfamilial classification, the omission of this information in LEAP does not negatively affect the retrieval of information.

Genus Number

The phylogenetic sequence numbers for the genera in each of the published FTEA families were entered to check for completeness and to facilitate retrieval of additional information from the FTEA volumes. A phylogeny is better represented by a "tree" than a linear sequence. Sequential numbers may or may not indicate close relationship, and one must have a detailed understanding of the phylogenetic hypothesis for a particular group in order to ascribe any significance to these numbers. In some cases, numbers have been interdigitated to accommodate unnumbered genera (e.g., introduced plants discussed in the introduction but not treated in full). Provisional genus numbers were also used for the FTEA conspectuses for the Compositae and the Cyperaceae, but these will be changed in the final publication of the respective FTEA volumes.

Genus

The genus entries have also been checked against VPFG for correct spelling. In some cases, orthographic variants of genus names used in FTEA are not the preferred standard spelling, and these have been changed in LEAP. In other cases, a genus used in FTEA has been sunk as a synonym of another genus or separate genera have been recognised. Many of these cases represent lumping versus splitting, and although there may be good evidence for a given change, the original names in FTEA remain validly published for those wishing to use them. However, when genera are combined, valid combinations in the aggregate genus may or may not exist, and each species must be checked individually.

Species Numbers

Species numbers, like genus numbers, represent a phylogenetic sequence, and the caveats stated above also apply.

Species

The species entries still need to be checked individually for correct spelling against FTEA (for the published families) or *Index Kewensis* (for the unpublished families). These entries were provisionally checked as additional information and new entries were added, but they have not yet been systematically checked. FTEA has very few spelling errors in specific epithets in the main body of each volume, although spelling errors are more common in the indexes. The CD-ROM version of *Index Kewensis* was electronically scanned, and so, has about a 5% error rate in character reading, and this error rate is particularly high for certain characters or character combinations (e.g., "rn" is often read as "m"). Many species epithets have unusual spellings, but unusual entries in *Index Kewensis* should always be double-checked against the original literature. For species not yet named and designated by a letter, the entry consists of the abbreviation for species (sp.) and then the letter for that species (e.g., sp. A), so that the output reads correctly (e.g., *Vitex* sp. *A*).

Author

The author entries have been standardised and checked against *Authors of Plant Names* (APN; Brummitt & Powell, 1992). The standardised entries listed in APN are commonly different than those used in FTEA, but the purpose of assigning the unique author entries used in APN is to avoid ambiguous author references, particularly for people with common names (e.g.,

Smith) or with names that might be given the same abbreviation (e.g., does Berg. refer to Bergius or Bergeret?). With a few exceptions, the author abbreviations in LEAP are now standardised, but that does not mean that they are necessarily correct. Certain names are more problematic than others, particularly those in which the omission of a period from an abbreviation yields another valid name (e.g., Berg) and names that are commonly confused (e.g., De Wildeman, de Wilde, Willdenow, and Wild). One option for detecting errors in the author entry is to compare the date of publication with the biographical data in APN (dates of birth and death or date of publication). For example, a species name authored by Smith in 1950 obviously cannot be the work of James Edward Smith (the person who has the standardised abbreviation "Sm." and lived from 1760 to 1840), but which of the Smiths alive in 1950 can only be determined by consulting the original literature, perhaps aided by information provided in the biographical work *Taxonomic Literature* (Stafleu & Cowan, 1976–1988). For the published families of FTEA, this work could be conducted on a species-by-species basis by comparing the LEAP author entry against the date of publication of the accepted name (and the basionym, when appropriate). When two people authored a species name, the standardised forms of the two author entries are separated by an ampersand (e.g., Oliv. & Hiern). When three or more people authored a species name, a comma is used to separate the standardised forms of the author entries except for the last author, which is separated only by an ampersand (e.g., Humb., Bonpl. & Kunth). When a species has been transferred from one genus to another, or from one rank to another, the author(s) of the basionym is placed in parentheses preceding the author(s) of the accepted name. When a species name was originally not validly published, but then later validated, the original author(s) name(s) may be included with the original author(s) listed first, then the word ex, then the validating author (e.g., Hochst. ex A.Rich.). In some early volumes of FTEA, the author(s) of the original name was included in square brackets (e.g., [Hochst. ex] A.Rich.) to indicate that the author of the original name is not required according to the International Rules of Botanical Nomenclature. It is acceptable to omit the reference to original author (in this case, Hochstetter), but it is generally preferable to include it. However, the square brackets should no longer be included. Some entries have the abbreviation "ined." following the author name, which indicates that the name has not yet been validly published. For new species, no other name may exist, and this form of citation represents an unsatisfactory (and invalid) temporary arrangement. For species that are being transferred from one genus to another (or from one rank to another), such names are equally unsatisfactory (and invalid), and a valid name exists in the original genus or at the original rank. Entries with "ined." should not be included in output from the database.

Subspecies and Variety

As with species, names of subspecies and varieties need to be checked individually for correct spelling. Unfortunately, for subspecies and varieties in families not yet published for FTEA, *Index Kewensis* did not include infraspecific taxa in the early volumes, and extensive taxonomic research is often required to locate the original literature in order to verify the name. Names at the taxonomic rank forma are not included in the LEAP database. For subspecies and varieties not yet named and designated by a letter, only the letter was included in the entry, because the rank (subsp. or var.) needs to be included in the output for all subspecies and varieties.

Sauthor and Vauthor

The discussion above regarding the authors of species names also applies to authors of subspecies and varieties, with the additional complications regarding verification for taxa not

yet published for FTEA. In cases where a subspecies or variety carries the epithet of the species, or where a variety carries the epithet of its more inclusive subspecies, these entries are termed autonyms (i.e., automatic names), and the author of the name at a higher rank is not repeated [e.g., *Dendrosenecio adnivalis* (Stapf) E.B. Knox subsp. *adnivalis* var. *adnivalis* or *Merremia tridentata* (L.) Hallier f. subsp. *angustifolia* (Jacq.) Ooststr. var. *angustifolia*].

Life Form Code

Entries in this field are single character abbreviations, as follows: C – Climber or Liana; H – Herb; S – Shrub; T – Tree. These entries were not consistently included for the published families, and were generally not included for the unpublished families. If more than one life form was indicated in a publication, then the largest life form was generally entered (e.g., "shrub or small tree" was generally recorded as "T"). Relatively few errors were detected in spot checks of the existing entries.

Origin Code

Entries in this field are single character abbreviations, as follows: N – Native; I – Introduced. These entries were not consistently included for the published families, and were generally not included for the unpublished families. Several errors were detected in spot checks of the existing entries, and therefore all entries need to be rechecked. For many widespread weedy species it is not possible to know the species' place of origin, and these have generally been treated as natives.

Lower Alt Limit and Upper Alt Limit

These entries are in metres. For some taxa in FTEA, the altitudinal distribution was reported simply as "up to 2,400." If the taxon occurs on Zanzibar or Pemba, a lower altitudinal limit was recorded as 0 m. When the information taken from EA only provided a single altitude, this was included as both the lower and upper limit. There are clearly many errors in these data. For example, there are taxa recorded from Zanzibar that have a minimum altitude that is above the highest point on Zanzibar. Correcting these data requires careful research using correctly identified specimens. The altitudinal distribution of marine plants was recorded as 0 to 0 m.

U, K, T, Z, and P

The entries in these fields are the floral division codes for each country used in FTEA. These floral division codes were based on provincial boundaries in 1952, and hence are political rather than biological in nature. However, the use of these codes provides more detailed information than would be gained simply by indicating presence or absence in each country. When location within a country cannot be determined based on specimen information, or when an available conspectus has not yet documented the distribution within a country, an "X" has been used to indicate presence within a given country. There are four divisions in Uganda (U), seven divisions in Kenya (K), and eight divisions in mainland Tanzania (T; Tanganyika). Parentheses are used around a given code number if the distribution is suspected but not known with certainty, and these are used in place of the question marks that were included for some entries in FTEA. Slashes between floral codes have been omitted, and a range of codes separated by a dash has been replaced by a complete listing of the appropriate codes. These codes are entered as a single string (e.g., 134578), but the information can easily be retrieved by parsing or searching within the string, and reporting should use the format in FTEA (e.g., **U**1 or **T**2, 6–8). Entries for Zanzibar (Z) and Pemba (P) are Z and P, respectively, which can be used in reports by simply putting the character in bold font. Mafia

Island is treated as part of T6. The distribution information is fairly complete and accurate for the published FTEA families, although less so for the families published early in the series. Information for the pteridophytes and the Cyperaceae was drawn from the respective conspectuses, but these contain numerous entries without complete distribution information and more work is needed. The distribution information for taxa in the remaining unpublished families was taken from specimens at EA and there are undoubtedly many errors and omissions.

Endemic

This is a Boolean entry (Yes/No) which indicates whether the taxon is not known outside the entries for East Africa. A Yes in this field replaces the statement used in FTEA that a given taxon is "not known elsewhere" or various statements that a taxon is known only from a certain locality or from a certain specimen or gathering. Statements that a taxon is known only from limited material are recorded in the Notes field, as an indication that a taxon may be imperfectly known or may be extremely rare. This entry is relative to the distribution listed in the previous fields (U, K, T, P, and Z) and should not be mistaken as a statement of restricted range.

Other Countries

Distribution of a taxon outside East Africa is indicated with varying amounts of precision, with a listing of entries separated by commas. For countries adjacent to East Africa and those of southern Africa (which have a high degree of floristic affinity with East Africa), abbreviations have been used. For countries in other parts of Africa and elsewhere in the world, names have been written in full, or more often, regions have been recorded (e.g., W. Afr). Sometimes, however, the available information is simply that a taxon is "widespread in tropical Africa," and this is presented in full, but such statements were not consistently entered into LEAP and additional work is needed to check all entries that are not marked as endemics but for which no distribution outside East Africa is indicated. The information in FTEA concerning distribution outside East Africa is not standardised, and no standard protocol was adopted for this field prior to data entry, so the entries are somewhat heterogeneous.

Habitat

FTEA provides a short description of habitat for each taxon, which synthesises information in the collection notes from individual specimens and often reflects the author's knowledge of the plant. Such an overview of habitat is different from the habitat information provided for a single collection. This information still needs to be added to the database regardless of whether some standardised system of habitat codes is also used. Information for the Lobeliaceae has been entered as an example.

Notes

This text field is available for miscellaneous comments on a record.

Source

For information from FTEA, the entry for the Source field simply indicates FTEA and the year of publication, separated by a comma (e.g., FTEA, 1984). Information from the East African Herbarium is designated EA, and information from the Kew Herbarium is designated K. The Apocynaceae are currently being written for FTEA, a specimen database for this family has been prepared, and the source entry is indicated as "FTEA conspectus."

Information on the Cyperaceae is based primarily on *The Sedges and Rushes of East Africa* (Haines & Lye, 1983), with certain modifications based on published work and the FTEA conspectus for this family. Other non-FTEA published sources are indicated (e.g., Jeffrey, 1966, 1967, 1968, 1986, 1988), and the full citation is provided in the Source Table. Where information from various sources has been compiled, the Source field records these multiple sources.

Date of Entry
This entry indicates when the record was completed. Any modification of a given record should be accompanied by a change in the Date of Entry in order to facilitate corrections to the master copy of LEAP.

THE HIGHER CLASSIFICATION TABLE

Higher Taxa
The fields Division, Class, Subclass, Order, and Family contain the names of higher taxa at the indicated rank. There is no clear agreement within the systematic community concerning the rank at which various higher taxa should be placed, so the arrangement presented here is somewhat arbitrary. This arrangement is based primarily on the system for dicotyledons developed by Cronquist (1981, 1988; see also Brummitt, 1992). The system for monocotyledons is taken from Dahlgren *et al.* (1985; see also Brummitt, 1992), but the superorders used in Dahlgren's system have been treated as subclasses to maintain consistency with the Cronquist system. The pteridophytes and gymnosperms have been incorporated into the framework provided by the Cronquist/Dahlgren hybrid system for the angiosperms. Subclass is not a required rank, but subclass names have been entered for the pteridophytes and gymnosperms so that the table does not appear incomplete. The families are those treated in FTEA, and these do not correspond exactly to either the Cronquist or Dahlgren systems.

Taxon Numbers
The fields Dnum, Cnum, Snum, Onum, and Fnum are phylogenetic sequence numbers that have been assigned to the corresponding taxa. These can be used to generate a phylogenetic listing or to otherwise manipulate the main LEAP table. For example, the division numbers (Dnum) can be used to select only the angiosperms, or the class numbers (Cnum) can be used to select only the monocotyledons. The family numbers (Fnum) are those used at EA which are modified from Johns (1991) for the pteridophytes, Hutchinson (1926) for the Dicotyledons, and Dahlgren *et al.* (1985) for the monocotyledons. Use of the EA family numbers improves the practical application of LEAP for herbarium-based research in the region.

OTHER TABLES

The Source Table
This table records the sources of information recorded in LEAP. As additional information from other published sources is added to LEAP, the source table will provide a fairly complete listing of published information on the flora of tropical East Africa that post-dates

the FTEA treatment for each family. Earlier literature is noted in the FTEA volumes, and this information is not repeated in LEAP.

The Synonym Table

This table records changes in taxonomic treatment from the original FTEA volumes or other sources used in LEAP and can be used to determine the currently accepted name based on earlier names that are now regarded as synonyms. Information was included in this table as changes were made to LEAP during its development. The synonym table does not provide a complete cross-referencing of all synonyms of East African vascular plants. For the published FTEA families, much of this information can be retrieved from the indices. Additional entries could be usefully added to this table to improve the cross-referencing, but the priority should be on synonyms that are still commonly used.

The Recipients of LEAP Table

This table provides the names and addresses of recipients of LEAP, and is intended simply as a listing of users. This listing can be used for distributing updated versions of LEAP and for identifying users who are already familiar with the database.

RECORD COUNT

LEAP version 1.3 contains 11,871 species, with the subspecies and varieties bringing the total number of entries (terminal taxa) to 13,140. For the preliminary counts presented below, all information for infraspecific taxa has been summarised to species level.

For 346 species, no country distribution is included in the database. Most of these are species listed in the FTEA conspectuses for the pteridophytes or the Cyperaceae. Other records are derived from old collections with distribution information limited to "East Africa" or are records for taxa known to occur just outside the East African boundaries that might be present in at least one of the three countries. The number of species in East Africa with distribution known to the country level or below is 11,515. Figure 1 is a Venn diagram showing the country-level distribution of species in East Africa. The figure at the centre (2,682) is the number of species that are recorded from all three countries. The three overlapping areas around the centre present the number of species recorded from only two of the three countries in the three pair-wise combinations. The non-overlapping areas present the number of species recorded from only one of the three countries. The total number of species for each country is as follows: Uganda, 4,318; Kenya, 6,302; Tanzania (including Zanzibar and Pemba), 9,105. Uganda possess 37.5% of the East African vascular flora, but only 16.1% of the species in Uganda are not recorded from the other two countries. Kenya possess 54.8% of the East African vascular flora, but only 21.9% of the species in Kenya are not recorded from the other two countries. Tanzania possesses 79.1% of the East African vascular flora, but 42.8% of the species in Tanzania are not recorded from the other two countries.

Endemism is a relative concept. All known species are endemic to the planet earth, and endemism at smaller geographic scales can only be described relative to a stated geographic area. Geopolitical boundaries are inherently arbitrary with respect to biogeographical distributions. Species that are endemic to Mt. Kenya are endemic to Kenya. Species that are endemic to Mt. Elgon are endemic to East Africa, but not endemic to Kenya or Uganda if they grow on both sides of the border. Species that are endemic to the Ruwenzori Mts. are not endemic to East Africa if they grow on the Congo side of the border. The concept of

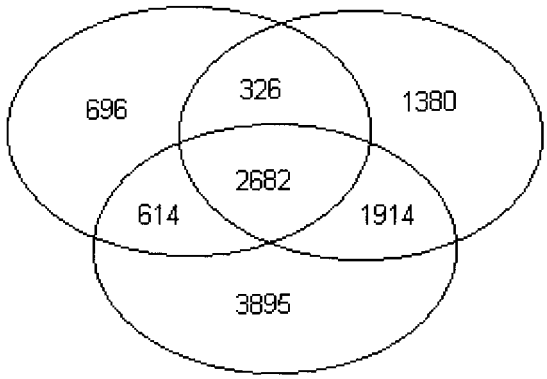

Figure 1. Distribution of 11,507 species in East Africa (K - Kenya; T - Tanzania; U - Uganda).

endemism is also different from the related concepts of abundance and range. A species may be endemic to Tanzania, but common and widespread within Tanzania. In contrast, a species composed solely of a small population of plants growing on the border between Kenya and Ethiopia may be described as rare with restricted range, but it is not endemic relative to the political boundaries of East Africa. With this in mind, it is nonetheless possible to evaluate the country-level patterns of endemism.

Only the published FTEA families provide sufficiently complete distribution information for analysing patterns of endemism, but fortunately, this published work covers two-thirds of the East African vascular flora. Species concepts understandably vary from author to author, but the existing species delimitations provide a general basis for comparison. For species with recognised infraspecific taxa, it is not necessarily obvious whether additional infraspecific taxa occur outside the FTEA region. The published FTEA accounts frequently indicate whether or not additional infraspecific taxa occur outside the region, but this information was not consistently provided and it was not incorporated into LEAP. The subset of LEAP records for published FTEA families that excludes (1) species with infraspecific taxa and (2) records that lack distribution information within East Africa totals 6,559. Of these, 1,845 (28.1%) are thought to be endemic to East Africa, with varying patterns of endemism within and among the three countries (figure 2). A more conservative estimate is obtained by excluding all unnamed species, which eliminates many imperfectly known species, but also eliminates presumably rare species for which sufficient herbarium material is lacking. This reduces the number of East African endemics to 1,595 (24.3%), but has little effect on the distribution of endemics within East Africa. Tanzania possesses a majority of the endemic species in East Africa and these endemics comprise 17.8 - 21.0% of the Tanzanian vascular flora, based on whether the conservative or liberal estimate of endemism is used. Similarly, endemic species comprise 7.9 - 9.7% of the Kenyan vascular flora and only 1.8 - 2.5% of the Ugandan vascular flora. These estimates are not expected to change appreciably as FTEA is completed and research on the East African flora continues. Additional fieldwork in East Africa and the neighbouring countries is expected to reduce the number of perceived endemics due to enlarged understanding of the true distributions, but this will be offset by newly discovered and/or newly described species which will, in general, be rare and/or restricted in range.

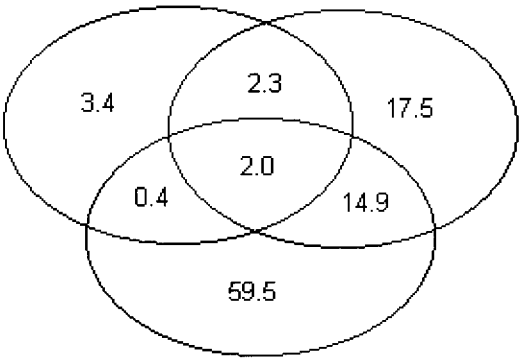

Figure 2. Distribution expressed as percentages for 1,845 endemic species based on the published FTEA families (K - Kenya; T - Tanzania; U - Uganda).

HERBARIUM MANAGEMENT AND PLANT BIODIVERSITY RESEARCH

Herbarium Management

For herbaria that employ a phylogenetic sequence for ordering specimens, indices are used to locate a desired species. LEAP can be used to generate an index to the vascular plant families in East Africa, the genera within each family, and the species within each genus. The family index may be posted in the herbarium or printed as a guide for herbarium users. The genera and species indices are generally placed at appropriate locations in the cupboards and constitute part of the cupboard literature that is useful in herbarium-based research. LEAP only treats the East African vascular flora, but an herbarium may have specimens from outside East Africa, and any index needs to incorporate appropriate information that reflects the holdings of an individual herbarium. This can be simply accomplished by the construction of a table for non-East African specimens that has parallel structure to the LEAP tables. The information from the two parallel tables can be merged to generate an index that treats both the East African and non-East African components of the collection. The question of how information on the non-East African specimens should be structured depends on the organisational arrangement adopted in a particular herbarium.

Regardless of whether an herbarium uses a phylogenetic or an alphabetic sequence for organising its collection, there are additional components of the cupboard literature that can be generated using LEAP. One such component is a distribution sheet that lists the species, subspecies, and varieties (in either phylogenetic or alphabetic order) and presents the known altitudinal range and geographic distribution based on the floral division codes. The dichotomous keys for identification provided in FTEA constitute the preferred method for identifying plants, but herbarium workers routinely encounter situations in which recently discovered species are not included in a key, a specimen is inadequate for identification based on a key, or a suitable key does not exist because the family has not yet been published as part of FTEA. In these cases, a distribution sheet can be used to select the species in a genus which have appropriate altitudinal and geographic distribution relative to the specimen to be identified, and these species can be consulted in an attempt to identify the specimen by matching characteristics with a previously identified specimen.

The database can also be used to generate determination slips. If a person is making a determination for a single specimen, it may be faster and easier to write a determination slip by hand, but curatorial activities requiring numerous determination slips can be greatly facilitated by using pre-printed slips.

Plant Biodiversity Research

LEAP has potential applications in a wide range of research activities, but the limitations of this, or any, database must always be kept in mind. In the analysis of patterns of endemism presented above, it was necessary to exclude incomplete and potentially inaccurate records. The analysis used species as a standard basis of comparison, but evolution is an ongoing process that renders the delimitation of species a matter of taxonomic judgement. LEAP relies on the judgement of the authors of the FTEA volumes and subsequently published literature, but another competent taxonomist might legitimately delimit species, subspecies, and varieties differently in certain cases. It is not possible at this date to provide an exact number of the vascular plant species that are endemic to East Africa, but the analysis does provide a reasonably accurate estimate of the percentage of the flora that is endemic and the pattern of endemism within the three countries. Other analyses may be conducted using LEAP, but the design of the analysis and its interpretation requires an understanding of the structure and limitations of this database.

LEAP can be used as one component in a tripartite approach to ongoing herbarium-based research on the flora of East Africa. Science is not science until the results are published. FTEA constitutes important published literature, but the East African flora is not completely known, and as new specimens are collected and new research is completed, FTEA will gradually become out-of-date. The flora exists in the natural world. Our herbaria contain collections that are more or less representative, but these have known or unknown errors and omissions. The published scientific knowledge represents our understanding based on the collections at the time of publication, and this literature carries its own set of known or unknown errors and omissions. It is unlikely that FTEA will be revised in the foreseeable future, so a need persists for local botanists to publish information that updates our scientific knowledge of the East African flora. By providing a compiled summary of the published information, LEAP can be used as a point of reference for examining current herbarium collections. It may be observed, for example, that recently collected specimens extend the known altitudinal and/or geographical range of a species, particularly for those species treated in early volumes of FTEA, but such observations do not provide a sufficient basis for changing LEAP. These observations need to be published in order to establish this personal knowledge as scientific knowledge. At that point, LEAP can be updated and the research cycle will be complete until new specimens become available. The *Journal of East African Natural History* has published such notes and new records in the past, and it is hoped that local botanists might play an increased role in documenting the regional flora.

The use of LEAP as a point of comparison has additional application for cross-checking a specimen-level database that includes altitude and floral division code. Specimen data from outside the altitudinal and geographic range recorded in LEAP indicates a conflict between the two databases. This conflict might be due to the following potential sources of error: (1) The altitudinal or geographic information in LEAP is inaccurate; (2) The altitudinal or geographic information for the specimen is inaccurate; (3) The altitudinal and geographic information for the specimen is accurate, but the specimen is inaccurately determined; or (4) The specimen information is accurate, LEAP accurately records the published information, and the specimen represents a new altitudinal or geographic record for the species. In the first situation, LEAP simply needs to be corrected. In the second or third situation, the specimen

database simply needs to be corrected. In the fourth situation, the new distribution record needs to be published so that LEAP can be updated. LEAP can similarly be used to check the accuracy of taxon names and authors and to make corrections for species lists to be used in publications.

The ability to generate species lists based on altitudinal, geographical, and taxonomic parameters makes LEAP a useful tool in preparing for fieldwork. For example, by selecting taxa known to occur in K4 at altitudes above 2,000 m, it is possible to generate a preliminary species list of the plants that occur on Mt. Kenya. The floral division K4 also includes the Aberdare Mts., so this list will also include species that occur on the Aberdares but not Mt. Kenya. There may also be some plants that grow on Mt. Kenya which have not yet been recorded in K4, but this is a limitation of our scientific knowledge, not of LEAP *per se*. Once this subset of LEAP records has been selected, it may be sorted taxonomically to facilitate a review of specimens in the herbarium before proceeding into the field. Depending on the purpose of the fieldwork, the list may be sorted again using, say, ascending lower altitudinal limit, which would generate a list of species in the altitudinal progression one might encounter as the mountain is ascended.

LEAP can also be used as a pick-list of taxon names for compiling label information for newly collected specimens. This application eliminates the need to re-type taxon names and author citations and prevents new typing errors from being incorporated into the labels and the associated database of specimens. This pick-list function of LEAP can be applied to many other situations.

Finally, LEAP can be used to record the taxon-level holdings of an herbarium. From this information, it is possible to determine which taxa are not represented in the herbarium and desiderata lists can be compiled. For example, the herbarium at Makerere University may wish to assemble a complete collection of Ugandan plants. If fieldwork is planned in a particular area of Uganda, the appropriate altitudinal and geographical parameters can be used in conjunction with a record of holdings to generate a list of desired species that might be encountered in that area.

The potential applications discussed above are by no means exhaustive. LEAP is a tool that can be put to many uses. As research progresses, LEAP can be updated and refined, thereby increasing its range of application.

TAKING THE LEAP FORWARD

LEAP now provides a useful taxon-level database of the vascular plants of the region, but much can still be done to improve its utility.

For the families not yet published in FTEA, the plant names and authors were checked against *Index Kewensis*. This eliminated many of the transcription and typographic errors that arose as information was culled from the collection at EA. Because these unpublished families have not yet undergone taxonomic revision, there are undoubtedly many names in LEAP that will eventually be treated as synonyms, as well as specimens that are not properly identified and potentially distort the distribution information. The taxa in these families represent the weakest component of LEAP, but this weakness is indicated in the Source field. The librarian at EA recently went through the last thirty years of *Kew Bulletin* and identified many potentially relevant articles, but a more systematic review of the literature is still required. It is beyond the scope of LEAP to resolve all outstanding taxonomic problems in the unpublished families, but LEAP does provide a valuable conspectus based on the EA collection.

LEAP is still incomplete for certain fields, most notably the habitat, life form code, and origin code. The habitat information (taken from FTEA) was not initially included and has only been entered for a few genera in which there has been a specific research interest. The life form and origin codes are mostly present for taxa from the FTEA volumes, but this information was not included for the taxa culled from the EA collection. The information on distribution outside East Africa for taxa culled from EA is also limited by the EA collection. This can be augmented in the future with published information from adjacent regional flora treatments (i.e., *Flora of Somalia*, *Flora of Ethiopia*, *Flore du Rwanda*, *Flore d'Afrique Centrale*, and *Flora Zambesiaca*), as available.

Some of the work that might be done to advance LEAP in the future is as follows:

1. To enter missing information on life-form, native versus introduced status, habitat, and distribution outside East Africa that was not included for certain records during the initial data entry.
2. To enter new information for existing records that includes (i) the date of publication for currently accepted names; (ii) the basionym and the date of publication of the basionym for currently accepted names that were formed as new combinations; and (iii) the publication source for taxa not yet published in FTEA.
3. To enter new records for taxa newly described or revised since publication of the relevant FTEA family.
4. To complete error checking for the spelling of specific epithets, the accuracy of the author citations, and typographic errors in the text entries.
5. To expand the list of synonyms to include most names that are still commonly used but not currently accepted.
6. To develop a set of utility programs that work as mouse-driven menu options so that users can retrieve certain kinds of information from LEAP without requiring knowledge of how to operate database programs like Microsoft Access.

LEAP version 1.3 is not the finished product, but we hope that it is useful in its presently imperfect form. A copy of LEAP may be obtained from either the Botanist in Charge of the East African Herbarium, or the Biodiversity Database Co-ordinator, Centre for Biodiversity, National Museums of Kenya, and comments and/or corrections should be sent to either of these people.

ACKNOWLEDGEMENTS

Numerous people have contributed to the development of LEAP through discussion and data entry, most importantly, H. Bakamwesiga, B. Bytebier, P. Ipulet, A. Katende, A. Kingada, J. Omoding, C. Orwa, T. Pearce, and J. Reynolds.

REFERENCES

Agnew, A. D. Q. & S. Agnew (1995). *Upland Kenya Wild Flowers*. Second Edition. East African Natural History Society: Nairobi.

Bamps, P., N. Robson & B. Verdcourt (1978). *Guttiferae*. In R. M. Polhill (ed.) *Flora of Tropical East Africa*. Crown Agents for Overseas Governments and Administrations: London.

Beentje, H. (1995). *Kenya Trees, Shrubs and Lianas*. National Museums of Kenya: Nairobi.

Brummitt, R. K. (1992). *Vascular Plant Families and Genera*. Royal Botanic Gardens, Kew: London.

Brummitt, R. K. & C. E. Powell (1992). *Authors of Plant Names*. Royal Botanic Gardens, Kew: London.

Cronquist, A. (1981). *An Integrated System of Classification of Flowering Plants*. Columbia University Press: New York.

Cronquist, A. (1988). *The Evolution and Classification of Flowering Plants*. Second Edition. The New York Botanical Garden: New York.

Dahlgren, R. M. T., H. T. Clifford & P. F. Yeo (1985). *The Families of the Monocotyledons*. Springer-Verlag: Berlin and New York.

Haines, R. W. & K. A. Lye (1983). *The Sedges and Rushes of East Africa*. East African Natural History Society: Nairobi.

Hutchinson, J. (1926). *Families of Flowering Plants*, Vol. 1. *Dicotyledons*. First Edition. Macmillan: London.

Jeffrey, C. (1966). Notes on *Compositae*: I. The *Cichorieae* in East Tropical Africa. *Kew Bulletin* 18: 427 - 486.

Jeffrey, C. (1967). Notes on *Compositae*: II. The *Mutisieae* in East Tropical Africa. *Kew Bulletin* 21: 177 - 223.

Jeffrey, C. (1968). Notes on *Compositae*: III. The *Cynareae* in East Tropical Africa. *Kew Bulletin* 22: 107 - 140.

Jeffrey, C. (1986). The *Senecioneae* in East Tropical Africa. Notes on *Compositae*: IV *Kew Bulletin* 41: 873 - 943.

Jeffrey, C. (1988). The *Vernonieae* in East Tropical Africa. Notes on *Compositae*: V. *Kew Bulletin* 43: 195 - 277.

Johns, R. J. (1991). *Pteridophytes of Tropical East Africa. A Preliminary Check-list of the Species*. Royal Botanic Gardens, Kew: London.

Lock, J. M. (1993). *Musaceae*. In R. M. Polhill (ed.) *Flora of Tropical East Africa*. Balkema: Rotterdam and Brookfield.

Stafleu, F. A. & R. S. Cowan (1976–1988). *Taxonomic Literature*. Second Edition. Bohn, Scheltema and Holkema, Utrecht, and W. Junk, The Hague and Boston.