UPDATES AND IMPROVEMENTS OF TURKISH PLANTS DATA SERVICE (TÜBİVES)

Yasin Bakış, M. Tekin Babaç, Emel Uslu

Biology Department, Faculty of Science, Abant İzzet Baysal University
14280, Bolu, Turkey
phone: + (90) 3742541000 - 1746, fax: + (90) 2534642, email: bakis_y@ibu.edu.tr
web: http://turkherb.ibu.edu.tr

ABSTRACT

Reaching biological information of the studied taxa became quite simple by introducing biodiversity information databases through internet. TÜBİVES (Turkish Plants Data Service) is the first and the most efficient biodiversity database of the plants in Turkey. In this study, the structural of TÜBİVES has been changed according to needs in new data types such as, chromosomal numbers, vernacular names, uses of plants, IUCN categories for endemics, nomenclature and synonym information of plants distributed in Turkey. The structure of database and querying algorithm have been updated both for TÜRKNOM (Nomenclatural Database of Turkish Vascular Plants) and TÜBİVES.

1 INTRODUCTION

The science of Systematics itself is a vast dimensional database in which enormous data accumulated since Aristotle. But it was a mechanical database whose data are kept in many different sources, such as books, floras, monographs, herbaria, botanical gardens, etc. Therefore, reaching to and collating the data for a given taxon was quite time consuming and expensive process. Fortunately, marvellous progresses in electronic world have brought many advantages and easy solutions to systematics today. One of them is the databases in computer sciences which immediately have become applicable as an invaluable procedure especially in every level of plant systematics today (Bisby, 1984). By means of Internet, the massive development of information systems have been enabled to reach data about any plant in anywhere in the world (Bisby 2000, Bisby & al. 2002, and Babaç 2004), and all these data are, now, at the tip of ours finger at institutes, research stations and homes, too. Nowadays, we can come across with the perfect information systems based on different database technologies, providing tremendous knowledge about plant diversity at earth such as Species 2000 (Bisby 2003), Integrated Taxonomic Information System (ITIS 2006), wild California plants (CallFlora 2006) in the USA, Vascular Families of Genera database in Royal Botanic Gardens of Kew (RBGK 2006), Flora Europaea Database in Royal Botanic Garden of Edinburgh (RBGE 2006), Environmental Resources Information Network in Australia (ERIN 2006), TROPICOS in the Missouri Botanical Garden, USA (TROPICOS 2006), Global Biodiversity Information Facility (GBIF 2006), Turkish Plants Data Service (TÜBİVES 2006) and ANTHOS in Spain (Castroviejo et al. 2006)

TÜBİVES (Turkish Plants Data Service) is a first the data service of plant diversity in Turkey (Babaç, 2004), based on the TUBVET (A Database of the Turkish Plants) that has been created in 1992 (Babaç & al. 1995). Since 1998 it has been working on the Internet in Turkish. In 1996, a database management program has been developed for the project of the plant specimens kept in 23 Turkish herbaria (Erik & al. 1996). At the beginning of 1999, TURKHERB (Central Database of Turkish Herbaria) was put into service on Internet from Abant İzzet Baysal University in Bolu as a central database (Babaç & al. 1996). Within next five years, most of the herbaria have chosen to create their own databases using TURKHERB management program as a pioneering system - some are over Internet.

The last project is BIOCES (Biodiversity of Turkey) that is going to be completed at the end of 2007. This information system contains many different relational data tables belonging to Turkish micro and macro fungus, algae, lichens, bryophytes, and animals (both invertebrates and vertebrates).

In this paper, new information about the TÜBİVES is mentioned as the version of 2.0. In accordance with new data required, its structure has been changed, and some new data fields and new relational data tables relevant with Turkish plants are added and also the data which are already warehoused are updated.

1.1 Structure of the Database

The database structure of TÜBİVES 1.0 has been changed completely. This new version is designed as a database with centre of taxonomic information. The database consists of 10 tables which are (Fig. 1) chromosome_counts, chromosome_specimens, general taxon information (taxon_information), geographic_distributions, IUCN information (iucn), plant usage (plant_use), bibliographic references (references), nomenclatural information (TÜRKNOM), taxon list (taxa), and local names (vernacular_names). The primary table is taxa to which the rest of the tables are linked. There are also linkages between some of the secondary tables, such as chromosome_counts and chromosome_specimens (Fig. 1). The first fields of all tables are designed as id fields standing for record numbers. Some of the tables which are containing information from certain ref-

978-1-4673-4431-9/11//$31.00 ©2012 IEEE

HIBIT May 02-05, 2011 Izmir, Turkey

erence are also structured as the last fields linking references table.

1.1.1 The New Data Tables and Data Fields Added In TÜBİVES 2.0

The Taxa Table – In the taxa table, name and rank fields contain taxon names and taxonomic ranks, respectively. This table is established somewhat in hierarchical order. In a record in taxa table, tax_id stands for the taxon number referencing the taxon and parent_tax_id can be used to constract relationships between taxonomic categories. In this case, tax_id of the upper taxon entered the parent_tax_id of all lower taxa. All tax_id fields of the remained tables are linked to tax_id field of the taxa table (Fig. 1).

Chromosome Data Tables And Their Data Fields – The data were collected from different sources such as, Flora of Turkey (Davis & al. 1986, and Güner & al. 2000).

Chromosome data tables (chromosome_specimens and chromosome_counts) are found in the database as a complete module with taxa table (Fig. 1). This module provides taxonomic, reference, specimens, chromosome count information of which is placed in different data tables. Besides retrieving information from the module above, the user can also search for specific data among these tables. While entering chromosome numbers, related chromosome specimen should be indicated in the chromosome_specimens table. Chromosome specimen record also needs related reference to be indicated in the references table.

Among chromosome data there might be more than one record for each taxon. The reason for that either specimens or authors could be different. On the other hand, two different chromosome numbers of the same taxon had also been reported into different publications, stated by the word of "or". Such chromosome numbers were entered the database as separate records. If there is any information about the number of "B" chromosome of the same taxon, it is also entered the data table as a separate field. For some taxa, the approximate chromosome numbers published are entered as "circa" field. Unknown chromosome numbers of some taxa are left blank in the data table.

IUCN information – The data have been collated from the Turkey red book of plants (Ekim & al. 2000). As a starting point, IUCN categories are entered data tables for Turkish endemic taxa. The worldwide distributed taxa which span in the limited areas in Turkey, and the taxa distributed only in Turkey, and neighboring countries have been planned to be included in data tables in near future. The table has linked taxa and references tables by tax_id and reference fields respectively.

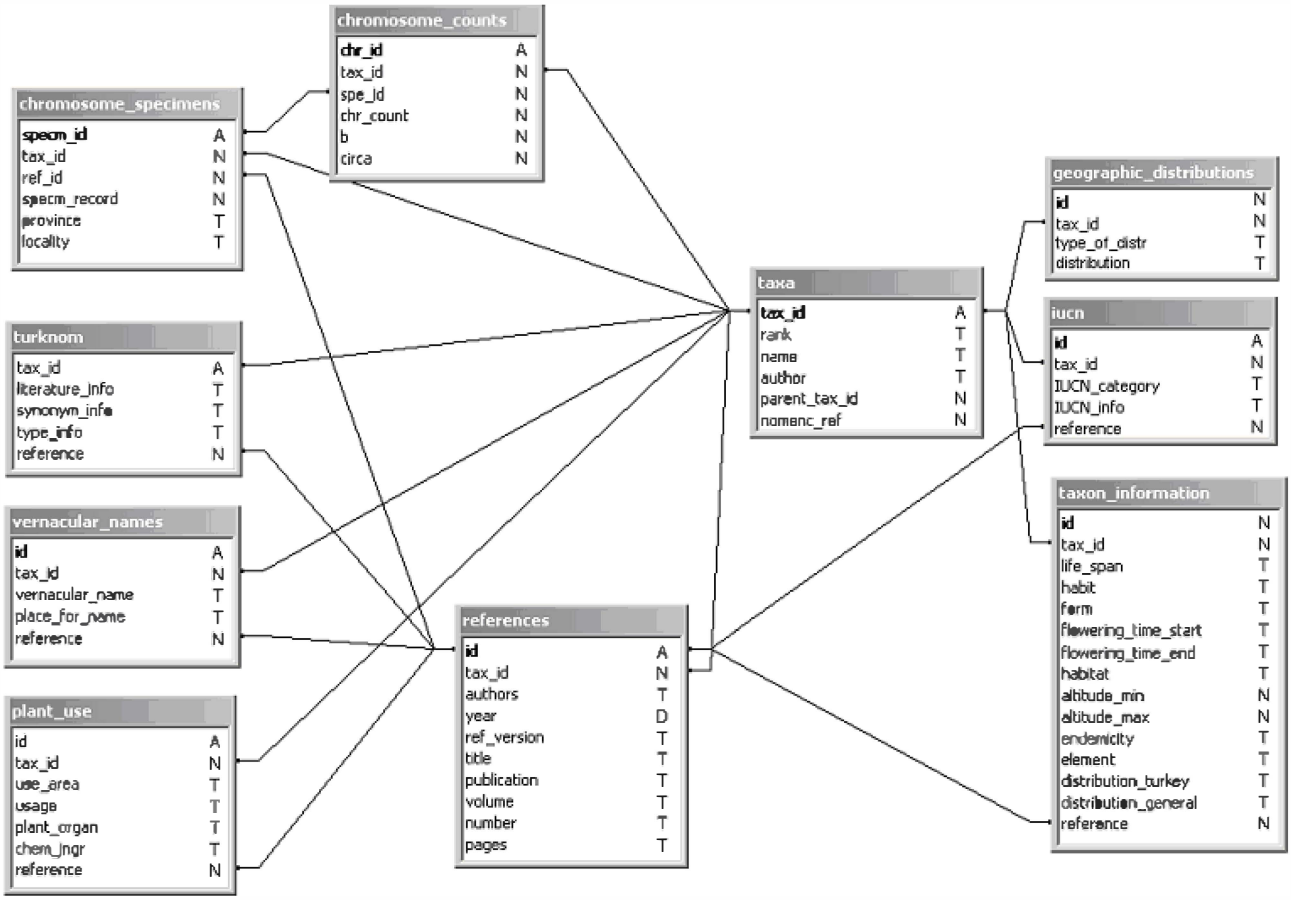

Figure 1: General database structure and types of data fields in TÜBİVES 2.0. T = text; N = number; D = date; A = auto number.

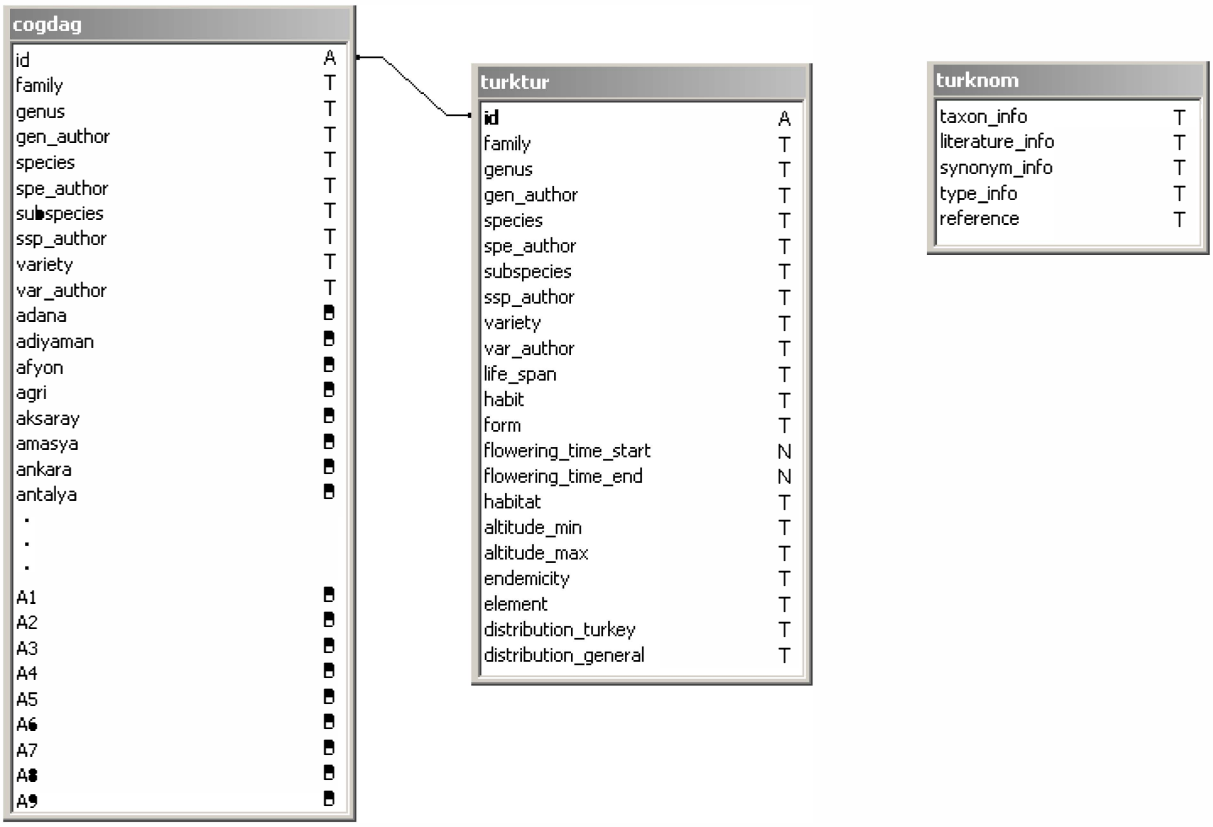

Figure 2: General database structures and the types of data fields in TÜBİVES 1.0 and TÜRKNOM . T = text; N = number; D = date; B = binary. Since the cogdag table in TÜBİVES has 120 fields, the figure shows the summary of the original cogdag table.

The Vernacular Names and Usages of Plants – Local names of Turkish plants have been kept in database in vernacular_names table, with their names, areas, and source information. The table is linked to the taxa and references tables by tax_id and reference fields respectively. The data for this aim collected mainly from Baytop (1984 & 1994). If a taxon has more than one vernacular names, they are entered the table as separate records.

In table plant_uses, fields of use_area, usage, plant_organ and chemical ingredients (chem_ingr) were linked references table. The data also have been recorded with tax_id by linking taxa table.

The Table of References – The table of references provides bibliographic information for any record in relational tables in the database. These tables are chromosome_specimens, taxa, vernacular_names, and plant_use. Among these tables, a specimen record needs to be declared as a reference record. Therefore, chromosome_specimens table is obligatory for references. On the other hand, the user can search for data among tables independently. The nomenclatural literature data is also stored in this table as well as literature of both vernacular_names and plant_use tables.

1.2 The Integration of TÜRKNOM

TÜRKNOM (Nomenclatural Database of the Vascular Plants in Turkey) was created as a project of TÜBİTAK (The Scientific and Technological Research Council of Turkey) in 2001 to serve the synonyms, types and bibliographic data of related taxa for nomenclatural purposes. The structure of database is a two dimensional flat table with fields taxa, literature, synonyms, type and source. The original structure of the TÜRKNOM was preserved except the fields of taxa and source, the taxa field is changed to tax_id field which includes taxonomic identification number and providing link to taxa table. The source field is changed to reference field which also provides link to references table by containing reference identification number (Fig. 1). The organization of the data in the fields has been designed as in Turkish Flora.

The Integration Tables From Version 1.0 to 2.0 – The original structure of the tables were kept except the fields including taxonomic information. These fields were replaced by tax_id field which is containing the taxonomic identification number and providing link to taxa table.

The Tables of Geographical Distribution – This table is originated from the old version of TÜBİVES. It is adapted to the new version by changing the table structure from multifield table with binary data (Fig. 2) to single field with multistate data. The table consists of fields; record number, taxon identification number, type of distribution and geographical distribution. The type of distribution field contains three data types which are grid system of Turkish Flora (Davis 1965), Turkish Vilayets, and both of them. In this case, the distribu-

tion field contains either codes of grid system or names of Vilayets or both grid system and names of Vilayet, such as A1, İstanbul, and A1-İstanbul, respectively (Fig.3).

General Taxon Information – The fields in the taxon_information table are id (record number), tax_id, life_span, habit, form, flowering_time_start (start of flowering time), flowering_time_end (end of flowering time), habitat, altitude_min (minimum altidute), altitude_max (maximum altitude), endemism, element (geographical element), distribution_turkey (distribution in Turkey), distribution_general (general distribution of taxa on the world), reference (reference link).

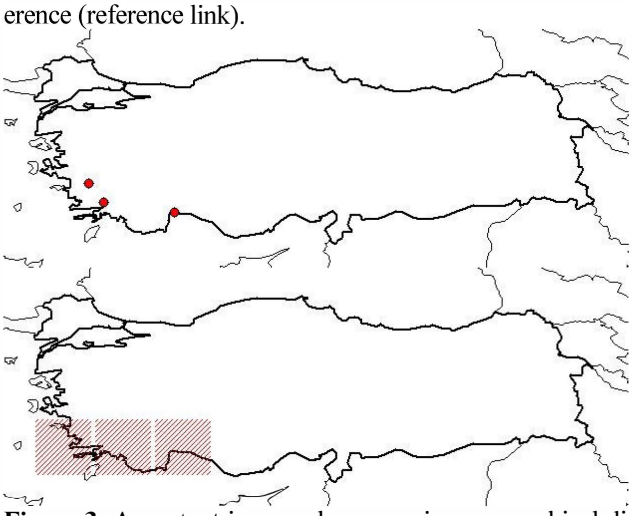

Figure 3: An output image when querying geographical distribution of *Quercus aucheri* in Vilayets Aydın, Muğla, and Antalya (top) and C1, C2 and C3 grid squares belong to the Davis's grid system (bottom).

1.3 Data Update

In the old version of database, there were four relational data tables that are, FAMTAB (Family Table), GENTAB (Genus Table), TURTAB (Species Table) and COGTAB (Geographic Table).

The data about the taxa spanned in Turkey were entered all the tables according to the order that was followed in The Flora of Turkey And East Aegean Islands (Davis 1965-1980, Davis & al. 1981, Özhatay & al. 1999, Güner & al. 2000, and Özhatay & Kültür 2006) which is similar to old version. Although the data about the endemics and new reported taxa present merely in the Aegean Islands which were not used. Furthermore the data about taxa mentioned either in Turkey and Aegean Islands are entered the database but the taxa are not accepted as endemic for Turkey.

According to the results of the last revisions and floristic studies, the new combinations and separations of the taxa are executed in the tables one by one, and their data were re-edited. Especially, the taxa identified wrongly during the preparation of the Flora of Turkey were revised and they were discarded from the tables.

After the publication of the Flora of Turkey, the numbers of Turkish Vilayets (province) were increased from 67 to 71 (Aksaray, Bartın, Karaman, and Kırıkkale) firstly and to 81 (Ardahan, Batman, Bayburt, Düzce, Iğdır, Şırnak, Karabük, Kilis, Osmaniye and Yalova) later by the government within last 10 years. The data about geographical distributions of the taxa were entered according to the 71 Turkish Vilayets and grid squares (Davis 1963) in TÜBİVES 1.0. By the statistical analyses, It was found that distributions of taxa on the basis of Vilayets is one of the most retrieved data from TÜBİVES (Fig. 1). Thus, geographical distributions of the taxa were reconstructed according to 81 Vilayets in version 2.0. These reconstructions have been executed using the old and new governmental maps (Tanoğlu 1961, Map1 1977, Map2 1992, Map3 1998, and Map 2004). In this process, localities of the taxa were fixed on the maps and then they were transferred into borders of new Vilayets one by one.

Thus, 97% of the localities in the new Vilayets were exactly determined but only 33 localities have not been found and they were left inside of their old Vilayets' border.

All data in new modules have been entered via internet by several scientists using "filling data" sub-programs developed for this aim. Therefore, the data in the tables will also be updated automatically by on-line processes. The data from the older version is transformed and then imported into the database by some programming.

The strategy in the searching mechanism in new version has been established to be suitable for a relational database model. Thus, a user will have the capability to search data by querying from different tables.

One way of creating a query is searching from primary table through outer ones. An example for this query is retrieving the vilayet distributions of a taxon. In this situation, searching mechanism looks for the selected taxon from taxa table. By using relationships between taxa and geographic_distributions tables, selects the geographical distributions from distribution field which are having the same number in tax_id field (Fig. 1). The reverse query – from outer tables to primary table – is also possible in a similar way.

The queries through database can also be between outer layer tables. As an example, a query can be established to find out the chromosome numbers according to related vernacular name. In this query example; user asks for chromosome number data from chromosome module, by entering a vernacular name of a plant that will be checked from vernacular names table.

2 CONCLUSION

The old version of database has been used since 1992. The new version of database created according to the statistical information of database usage. The new version gives two important advantages. Firstly, the information that contains are upgraded and also new fields are added to the old and new tables. Among the new tables, relevant with chromosome information would be helpful for cytogeneticists and cytotaxonomists. The IUCN categories of the Turkish endemic plants and their synonym information are served over Internet for the first time. These two data tables would also be quite beneficial for taxonomists and especially for conservation biologists. On the other hand, the last new tables about local Turkish names and plant usages are thought to be very informative for the ethno botanists and publics as a reference

point. Therefore, this database meets the connection between science and public.

Secondly, the structure has been changed for the aim of easy and fast data retrieving according to the statistics of old version. Search event and data extracting from database is much more independent and flexible which is designed user friendly.

Inclusion of digitalized endemic type specimens in herbaria into the service is currently progressing. There are also some developments in the area of taxon images including photographic images of live specimens and hand-drawings.

The categories recently added to IUCN lists will also be included in the database as soon as possible.

Geographic distribution in current version is mainly based on grid squares and administrative areas. Since there are comprehensive developments in the area of geographic information systems, TÜBİVES workgroup offers scientists to collect distribution data in the form of GPS or global coordinates.

In Flora of Turkey (Davis 1965 – 1985, Davis & al. 1989, Güner & al. 2000 and Özhatay & Kültür 2006), the grid square system is probably not very useful for the citation of taxa having no coordinate values. The squares are too large, such as two degrees of latitude and longitude for the citation of taxa distributing in Turkey. Thus, a new grid square system for those taxa having no coordinates should be smaller than Davis' grid square system. Each square – submitted by Davis (Davis 1963) – is more meaningful to be divided into 10 x 10 sub-squares.

3 REFERENCES

[1] Babaç, M. T., Güner, A., Doğan, M., Düzenli A., & Şahin, A. 1995: TUBVET: A data base of the Turkish plants. – Pp. 371-383 in: Öztürk, M. A. & Seçmen, Ö. (ed.) Plant Life in Southwest and Central Asia 1. – İzmir.

[2] Babaç, M. T., Bilgin, C. C., Usta, E. 1996: TURKHERB: Türkiye herbaryumları merkezi veri tabanı – Abstrakt, 13. Ulusal Biyoloji Kongresi, İstanbul, 18-21.6.1996.

[3] Babaç, M. T. 2004: Possibilities of an information system about plants of south west Asia: with particular reference to Turkish Plants Data Service (TÜBİVES) – Turk. J. Bot. 28: 19-127.

[4] Baytop, T. 1994: Türkçe Bitki Adları Sözlüğü – Ankara.

[5] Baytop, T. 1984: Türkiyede Bitkiler ile Tedavi (Geçmişte ve Bugün). – İstanbul.

[6] Bisby, F. A. 1984: Information Services in Taxonomy – Pp. 17-33 in: Allkin, R. & Bisby, F. A. (ed.), Databases in Systematics. – London and Orlando.

[7] Bisby, F. A. 2000: The quite revolution: biodiversity informatics and the Internet. – Science 289: 2309-2313.

[8] Bisby, F. A. 2003: Species 2000: an architecture and strategy for creating the Catalogue of Life. – Abstracts, Proceedings of the fourth biennial meeting of the Systematics Association, Dublin, Ireland, 18-23.8.2003-Dublin.

[9] Bisby, F. A., Shimura, J., Ruggiero, M., Edwards, J. and Haeuser, C. 2002: Taxonomy, at the click of a mouse (Nature Correspondence). – Nature 418: 367.

[10] CallFlora 2006: Wild California plants. – Published on the Internet http://www.calflora.org

[11] Castroviejo, A., Aedo, C. & Medina, L. 2006: Management of floristic information on the Internet: the Anthos Solution. Willdenowia 36: 127-136.

[12] Davis, P. H. 1963-85: Flora of Turkey and the East Aegean Islands 1-9. – Edinburgh.

[13] Davis, P. H., Mill, R. R. & Kit Tan 1986: Flora of Turkey and the East Aegean Islands Suppl. 1. – Edinburgh.

[14] Ekim, T., Koyuncu, M., Vural, M., Duman, H., Aytaç, Z. & Adıgüzel, N. 2000: Red Data Book of Turkish Plants – Ankara.

[15] Erik, S, Babaç, M. T. & Bilgin, C. C. 1996: HUB ve FUH herbaryumlarında veri tabanı geliştirmesi üzerine bir araştırma. – Hacettepe Fen ve Müh. Bil. Der. 17: 81-103.

[16] ERIN 2006: Environmental Resources Information Network. Published on the Internet http://www.deh.gov.au/erin

[17] GBIF 2006: The Global Biodiversity Information Facility. – Published on the Internet http://www.europe.gbif.net:80/portal/index.jsp

[18] Güner, A., Özhatay, N., Ekim, T. & Başer, K. H. C. 2000: Flora of Turkey and the East Aegean Islands Suppl. 2. – Edinburgh.

[19] ITIS 2006: Integrated Taxonomic Information System. – Published on the Internet http:// www.itis.gov

[20] Map 1977: Yeni Türkiye Atlası – M. S. B. Harita Genel Müdürlüğü, 1/500.000 ölçek – Ankara.

[21] Map 1992: Büyük Dünya Atlası Arkın Kitabevi – Ankara.

[22] Map 1998: Türkiye Mülki İdare Bölümleri Haritası, 1/800.000 ölçek- Ankara.

[23] Map 2004: Türkiye Karayolları Haritası – Ankara.

[24] Özhatay, N., Kültür, S. & Aksoy, N. 1999: Check-list of additional taxa to the suppliment Flora of Turkey II. – Turk. J. Bot. 23: 151-169.

[25] Özhatay, N. & Kültür, S. 2006: Check-list of additional taxa to the supplement Flora of Turkey III. Turk. J. Bot. 30: 281-316.

[26] RBGE 2006: Flora Europaea Database in Royal Botanic Garden of Edinburgh. – Published on the Internet http://www.rbge.org.uk/rbge/web/search/index.jsp

[27] RBGK 2006: Vascular Families of Genera database in Royal Botanic Gardens of Kew. – Published on the Internet http://www.rbgkew.org.uk/data/vascplnt.html

[28] Tanoğlu, A., Erinç, S. & Tümertekin E. 1961:Türkiye Atlası – İstanbul Üniversitesi Edebiyat Fakültesi Yayınları No:93.

[29] TROPICOS 2006: Missouri Botanical Garden's Vast (VAScular Tropicos) nomenclature database and associated authority files. – Published on the Internet http://mobot.mobot.org/W3T/Search/vast.html

[30] TÜBİVES 2006: The Data Service of the Plants in Turkey. – Published on the Internet http://www.tubitak.gov.tr/TÜBİVES