



Biodiversity in Southern Africa

Vol. 2

Patterns and Processes at Regional Scale

SPONSORED BY THE



Federal Ministry
of Education
and Research



© University of Hamburg 2010

All rights reserved

Klaus Hess Publishers

www.k-hess-verlag.de

ISBN all volumes: 978-3-933117-44-1 (Germany), 978-99916-57-30-1 (Namibia)

ISBN this volume: 978-3-933117-46-5 (Germany), 978-99916-57-32-5 (Namibia)

Printed in Germany

Suggestion for citations:

Volume:

Schmiedel, U., Jürgens, N. (2010) (eds.): Biodiversity in southern Africa **2**: Patterns and processes at regional scale. Göttingen & Windhoek: Klaus Hess Publishers.

Article (example):

Petersen, A., Gröngröft, A., Mills, A., Miehlich, G. (2010): Soils along the BIOTA transect. – In: Schmiedel, U., Jürgens, N. (eds.): Biodiversity in southern Africa **2**: Patterns and processes at regional scale: 84–92. Göttingen & Windhoek: Klaus Hess Publishers.

Corrections brought to our attention will be published at the following location: <http://www.biota-africa.org/biotabook/>

Cover photograph: Giraffes on the game farm Omatako Ranch (Observatory S04 Toggekry) in the Namibian Thornbush Savanna.

Photo: Jürgen Deckert, Berlin/Germany.

Cover Design: Ria Henning

Article III.8.6

– Author's copy –

Please cite this article as follows:

Muche, G., Hillmann, T., Suwald, A., Jürgens, N. (2010): Data access and availability: BIOTA Data Facility. – In: Schmiedel, U., Jürgens, N. [Eds.]: *Biodiversity in southern Africa. Volume 2: Patterns and processes at regional scale*: pp. 337–342, Klaus Hess Publishers, Göttingen & Windhoek.

Data access and availability: BIOTA Data Facility

GERHARD MUCHE *; THOMAS HILLMANN, ANDRZEJ SUWALD & NORBERT JÜRGENS

Summary: The BioMonitoring Data Facility is an infrastructure that serves all the regional projects within BIOTA AFRICA, i.e. BIOTA Maroc, BIOTA East Africa, BIOTA West Africa, and BIOTA Southern Africa. BIOTA AFRICA aims to provide robust scientific information about the current status, recent change and future development of biodiversity in the major African biomes. The present activities of the BioMonitoring Data Facility focusses on a continuous effort to enable inventorying and archiving of all Observatory-relevant BIOTA data, the presentation of data and information via the internet, and the facilitation of web communication between all BIOTA partners.

Since the start of the BIOTA AFRICA project in the year 2000, a wide array of digital data has been gathered relating to over 50 subprojects, work-packages and core topics of BIOTA and its associated projects. Due to the different subjects and key questions within the various disciplines, this data is of relatively heterogeneous structure. For establishing a BIOTA AFRICA data archive, the data files in their different formats were collected and archived in a central data pool at the Biocentre Klein Flottbek, University of Hamburg.

The internet offers an easy way to identify and access the data that BIOTA AFRICA supplies. Direct download from the website www.biota-africa.org is offered for a wide spectrum of open access data. More complex data requests have to be processed upon request. The conditions of data access are fixed in the Data Sharing Protocol of BIOTA AFRICA.

Introduction

Modern global change science provides various types of outputs such as scientific publications, recommendations for practitioners, and tools for stakeholders. In the third phase of BIOTA Southern Africa, the following work layers were distinguished: A Drivers of Change, B Monitoring of Change, C Land Use, Impact & Value, D Process Analysis, E Interventions & Restoration, F Transformation & Capacity Development (see www.biota-africa.org, Part I). All these work layers produce quite different types of data and they should all be made available to a broad range of users (from land managers to policy makers). Amongst the different elements, a very strong focus of BIOTA AFRICA has been

on long-term observation. The observations result in data that robustly document environmental changes and they are also the basis for empirical testing of hypotheses. Such data are often collected to answer a particular research question. At the same time they are a reference of the state of biodiversity at a specific point in time. Therefore, these data need to be continuously archived, technically maintained, understandable, utilisable and made available for unpredictable future research needs. Safeguarding the long-term availability of the BIOTA AFRICA data is one of the main objectives of the BIOTA Data Facility or BioMonitoring Data Facility.

The Data Facility attends to data management of the regional projects BIOTA Southern Africa, BIOTA West Africa,

BIOTA East Africa, and BIOTA Maroc. This data management includes the online presentation of working groups and their achievements, the processing of the data according to professional standards, the supply of datasets for interested persons or institutions, and archiving of data for future use. The BIOTA Data Facility, together with the African host countries, presently also aims to facilitate the hosting, maintenance and provision of data beyond the funding period of BIOTA AFRICA.

One of the main tasks of the Data Facility is the design and maintenance of the BIOTA AFRICA website. There are several arguments for the importance of this task. The internet is an important platform for disseminating information regarding projects and their achievements, so that potential users can learn about the available databases. Websites also facilitate a) communication between the data provider and data user; and b) availability of data packages at anytime from anywhere and for everyone. The availability of data via websites has much more impact than delivery on request via e-mail, for instance (for similar experiences with another research project see Göttlicher & Bendix 2004).

The BIOTA AFRICA website provides information on the goals, approaches and structure of the project. Furthermore, each of the many working groups and subprojects supplied information on their specific expertise, their research and research findings. The BIOTA website is accessible at www.biota-africa.org.

Data availability and access

Users of the BIOTA AFRICA website can gain access to data in various ways:

Open access via the website: a fundamental organisational aspect within

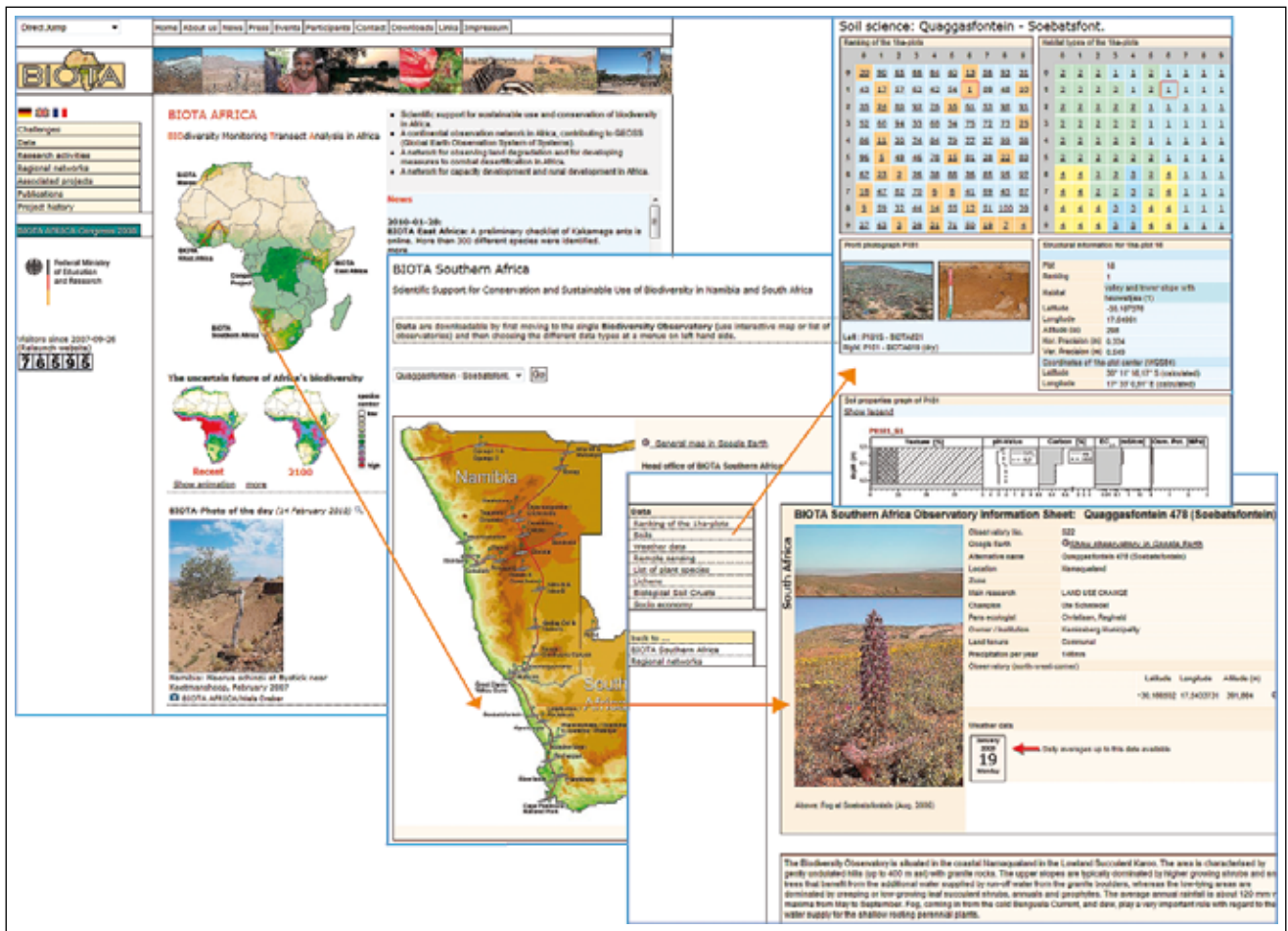


Fig. 1: Website of BIOTA AFRICA—three clicks from homepage to data access.

BIOTA Southern Africa is the arrangement of research sites along transects. Transects and BIOTA Biodiversity Observatories (hereafter called BIOTA Observatories) are shown on the homepage of the website on an interactive map and by clicking on the symbol of a BIOTA Observatory, the user reaches the page of that specific Observatory. Some texts and figures are shown on a compact information sheet. A menu on the left hand side offers more complex datasets. Each BIOTA Observatory may show a slightly different menu because data availability may differ between the Observatories. The data relates to topics such as weather, vegetation, soil, lichens, crusts, insects and socio-economy. In this way the user can easily find what type of data (i.e. topic or discipline) is available for a particular site. Alternatively, a user may be interested in a particular topic or discipline and want to find all data relating to it along the BIOTA transect. If the user is

interested in soil data for instance, he or she will find general information about soil science research and from there go to all BIOTA Observatories where the available data are displayed. A user can conveniently move to the information on any of the Observatory sheets with a few mouse clicks (Fig. 1).

Another means of accessing data is via the **metadatabase**. Metadata describes the contents (types of data, e.g. vegetation data), format (e.g. BIOTABase), and availability (e.g. open access) of data. The open access datasets, as well as datasets which are only available on request in order to protect PhD and other projects in progress, are listed within the metadatabase.

Besides metadata and open access information, published articles and books are also data which is generated by the BIOTA AFRICA initiative. The **List of Publications** is an online database of all publications that have emerged from BIOTA AFRICA. The details of the pub-

lications are searchable with flexible search functions. Abstracts are provided for downloading, and often the entire publication is available in pdf-format (see Fig. 2).

The BIOTA AFRICA staff and team members are listed in an online database named **List of Participants**. This makes it easy for users to contact a data provider or any expert on the BIOTA AFRICA network.

The BIOTA AFRICA website was initially only hosted in Hamburg/Germany, but a mirror has been established in Windhoek, Namibia (<http://www.nbri.org.na/BIOTA/>). This **Namibian data mirror** has been installed for two reasons. Firstly, the German-hosted website responded relatively slowly when accessed from Namibia because of the limited infrastructure there. The website response has improved significantly since the National Botanical Research Institute of Namibia (NBRI) in Windhoek


BIOTA AFRICA																																																																																											
	Total	in preparation					submitted					in review					in press					2010					2009					2008					2007					2006					2005					2004					2003					2002					2001					2000					1999					1998					unspec. year				
		in preparation	submitted	in review	in press	2010	2009	2008	2007	2006	2005	2004	2003	2002	2001	2000	1999	1998	unspec. year																																																																								
Publications in Journals (peer-reviewed)	459	28	36	7	19	3	25	54	51	70	55	49	34	13	10	1	3	1	0																																																																								
Publications in Journals (not peer-reviewed)	127	1	3	0	2	0	4	13	8	13	16	31	11	12	9	4	0	0	0																																																																								
Contributions in Books	82	1	0	0	4	0	7	8	10	10	11	9	5	7	7	1	0	2	0																																																																								
Conference Proceedings	101	0	4	1	0	0	11	18	9	13	12	17	7	5	4	0	0	0	0																																																																								
Monographs	16	1	0	0	1	0	2	3	4	0	0	1	0	3	0	1	0	0	0																																																																								
PhD Thesis	63	14	0	0	1	0	3	10	10	8	9	6	2	0	0	0	0	0	0																																																																								
Diploma Thesis	121	0	0	0	0	1	4	8	14	8	20	25	28	10	5	0	0	0	0																																																																								
DEA: Diplôme d'étude approfondie	14	1	0	0	0	0	2	2	1	0	2	2	1	3	0	0	0	0	0																																																																								
Master Thesis	34	0	0	0	0	0	3	6	6	2	5	5	2	3	2	0	0	0	0																																																																								
Bachelor Thesis	11	0	0	0	0	0	0	5	3	0	1	1	1	0	0	0	0	0	0																																																																								
Internet Publications	55	0	0	0	0	0	0	3	5	2	15	2	22	3	3	0	0	0	0																																																																								
Software Products	6	0	0	0	0	0	0	2	0	2	2	0	0	0	0	0	0	0	0																																																																								
Total	1089	46	43	8	27	4	61	132	121	128	148	147	112	56	43	6	4	3	0																																																																								
Posters	242	0	0	0	0	0	8	37	8	27	73	21	48	12	10	0	0	0	0																																																																								
Conference Contributions Abstracts	473	1	0	0	2	1	22	61	49	48	80	62	80	34	31	2	0	0	0																																																																								
Grand total	1804	47	43	8	29	5	91	230	178	203	301	230	238	102	84	8	4	3	0																																																																								

Fig. 2: List of Publications—statistics as of February 2010.

began hosting the Namibian data mirror. Secondly, the Namibian data mirror supports the notion of ownership of African data by Africans.

To facilitate the exchange, presentation and analysis of data, the datasets are stored as digital files. Special software has been designed to read and process these files. In addition, descriptions of the data are provided so that the user can understand the components of the data. Simple datasets often consist of a single file which can be managed with commonly available software packages, and are filled with content which can be interpreted by any user. With complex datasets, special software is necessary and detailed documentation of the content of tables and single fields is indispensable. This important information regarding datasets, their formats and documentations is provided in the metadatabase (see: www.biota-africa.org > Data > Metadata). There is also information provided on data availability and limitations on accessibility. The fact sheets in Part II provide an overview of the data collected at each BIOTA Observatory.

The rules for data access were formulated jointly by scientists of the participating countries within a Data Sharing

Protocol which defines the conditions for providers and users of data produced by the BIOTA AFRICA initiative (Electronic Appendix). The spirit of the BIOTA AFRICA **Data Sharing Protocol** is based on the Open Access paradigm that is documented, for example, in the Berliner Declaration (Gruss 2003). However there are several obvious reasons limiting immediate delivery of data to everybody. For example, based on national law, geographical distribution data of endangered and protected species may not be published, and these datasets are supplied exclusively for scientific use. Another reason for limitations is that some datasets are restricted to those people who have collected the data and who have the right to publish their findings and scientific results first. After a few years, such data become openly accessible.

Products of the BIOTA Data Facility

Within the framework of the BIOTA AFRICA project, the need for software that enabled data storage and processing emerged. Field work observations are typically recorded on standardised data sheets

and these records then have to be digitised so that they can be analysed with the help of computer software. At the beginning of the project it was impossible to find an adequate software package which met the requirements for the ambitious scientific approach. Such software should be able to handle (a) time series, (b) data from nested plots of the standardised BIOTA Observatories, (c) undescribed or unidentified species, and (d) data on the abiotic environment (e.g. topography, geology, soil). To fulfil all these requirements the IT team of BIOTA AFRICA undertook the development of a software package named BIOTABase which met the needs for the storage, administration and analysis of the BIOTA Southern Africa vegetation and ecological data (see Mucbe & Finckh 2009). BIOTABase can be downloaded free of charge from the BIOTA AFRICA Website (www.biota-africa.org > Downloads > BIOTABase), so that everybody can share and store BIOTA data in the original BIOTABase format. The software may also be used for building up new databases, independent of BIOTA AFRICA.

BIOTABase enables a user to store time series data from the same plot, which emerges from the annual surveys



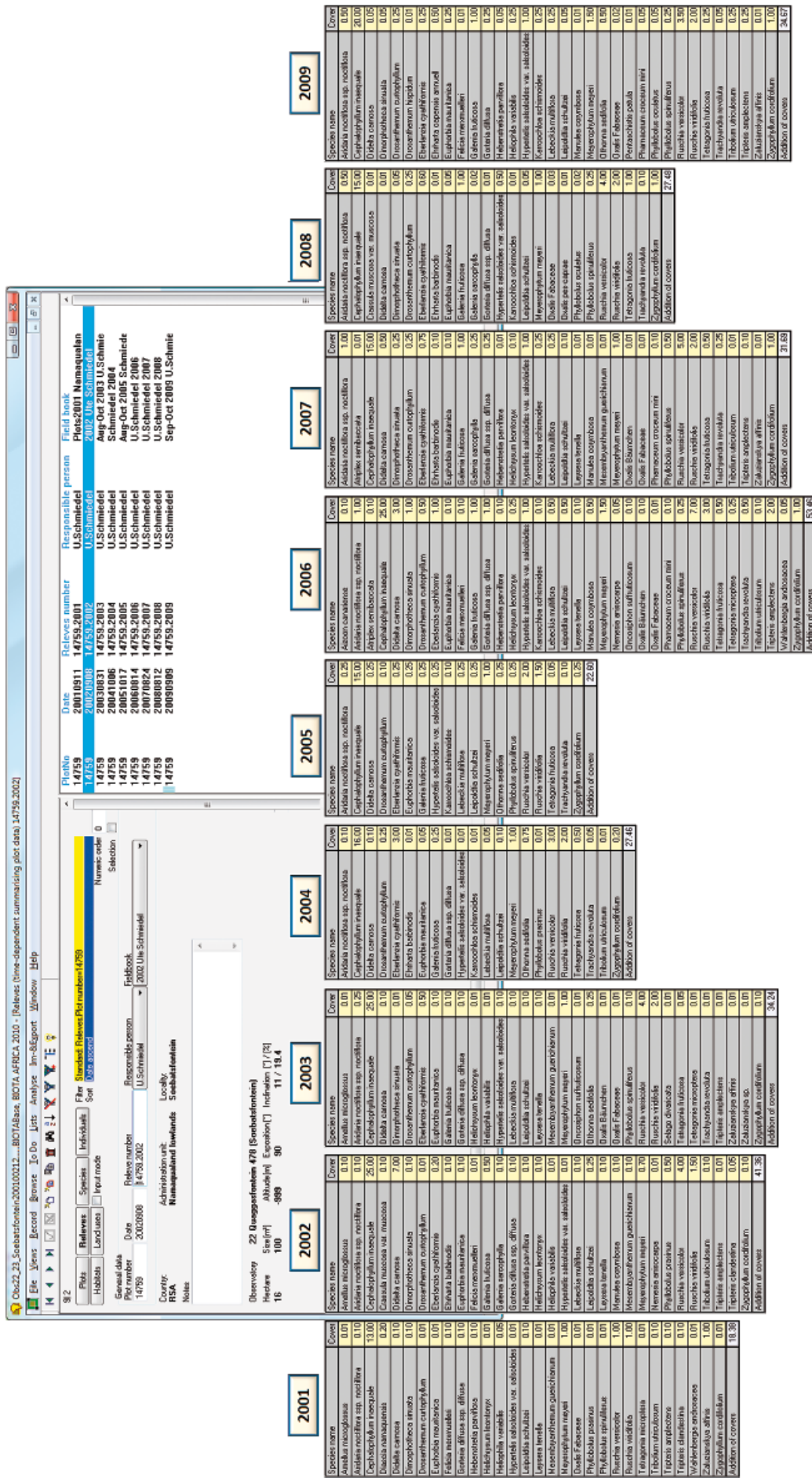


Fig. 3: BIOTABase: relevés of BIOTA Observatory S22 Soebatsfontein, hectare 16, rank 1, size 100 m², time series of species occurrence.

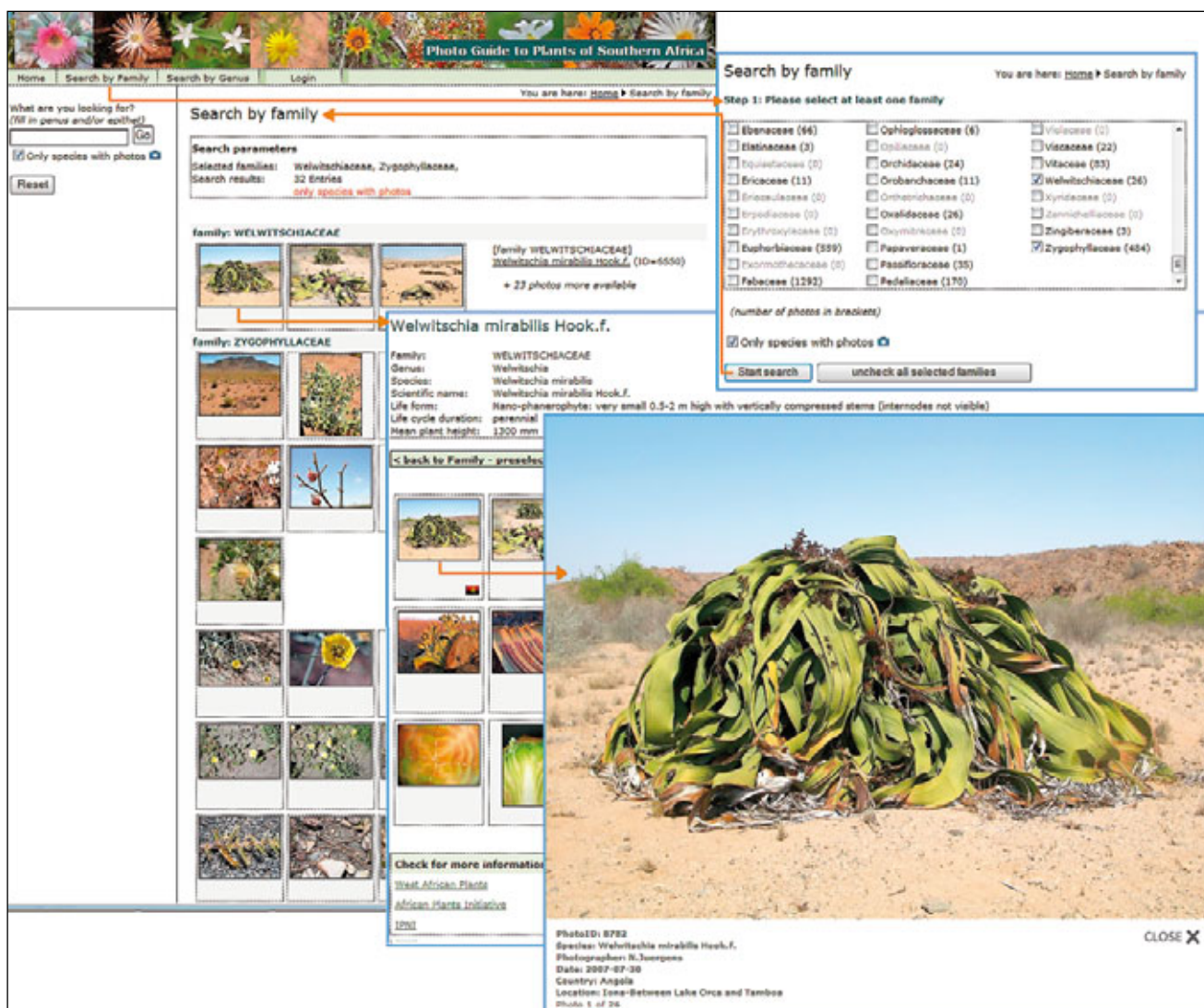


Fig. 4: Photo Guide to Plants of Southern Africa.

according to the standards of the BIOTA Observatories (Part I). Occurrence (presence-absence), species quantity measures (abundance, cover) as well as locality and date of occurrence are recorded. Thus, spatial and temporal patterns of species occurrence from all BIOTA Observatories along the 2,000 kilometres of the southern Africa transect can be analysed.

Independent of the observational and collection data, valuable reference data covering the nine year period of the BIOTA AFRICA project has been compiled. Reference data includes a) taxa (**Taxa Reference List**), b) geomorphology (**Reference List of Geomorphological Terms**) and c) landscape characteristics (> 25 km x 25 km, SOTER = Global and National Soils and Terrain Digital Databases, FAO; van Engelen & Wen 1995, Tempel 2002).

Another product—named BIOTA-Collections—has been designed at the requests of several users. BIOTACollections, the compact version of BIOTABase, has the advantage of storing simple collection data and observational data in a practical flat structure. It can also be downloaded from the BIOTA AFRICA website free of charge.

BIOTABase and BIOTACollections are offline tools for the flexible creation and use of datasets. Databases are subject to constant changes during the compilation of new datasets, as new records have to be added or changed in the process of data cleaning. Such activities should be carried out offline, even if the data are intended for public use. The taxa reference list of the research area has been made available for the public and linked to another nice feature: **The Photo Guide to**

Plants of Southern Africa (see: www.southernafricanplants.net, see Fig. 4).

African plants have evolved an overwhelming diversity of forms and functions in many different habitats. Due to the limited number of published plant field guides, species identification is still a large problem, especially when large numbers of specimens need to be identified within a limited time. Besides using published identification keys and identification based on herbarium records, the screening of a sufficient number of good photographs can be an efficient method of plant identification for many taxonomic groups. The development of a regional online photo guide for southern Africa is aimed at enhancing the identification of African plant species. The guide presents plant photographs in phylogenetic order, thus allowing rapid visual

comparisons. A number of tools allow the users to scan taxonomic groups such as genera and families. Online accessibility enables non-scientific users with limited access to taxonomic literature to gain knowledge on the regional flora and to identify plant taxa. The website also allows users to identify unknown species by entering plant characteristics such as dominant flower colour, life form, and growth form into a multi-access key.

The Guide to southern African Plants includes a comprehensive and rapidly growing representation of all known plant species of the southern African region. At present, Namibia and the western parts of South Africa are best represented, while a growing number of photos from Angola, Botswana and Zambia are also included. The Guide to southern African Plants has received contributions (photographs, identification data, quality control checks) from the National Botanical Research Institute of Namibia (NBRI), the Biocentre Klein Flottbek, the Botanical Garden and Herbarium of the University of Hamburg, and independent botanists from Namibia and South Africa. Contributions from the South African National Biodiversity In-

stitute (SANBI), the Harry Oppenheimer Okavango Research Centre (HOORC), and from Angolan colleagues at Luanda and Lubango, as well as the Research Institute Senckenberg in Frankfurt are also to be included.

A second photo guide has been developed in a similar manner: **The Photo Guide to Plants of Southern Morocco** (see: <http://plantsofsouthernmorocco.biota-africa.net>) aims at presenting the diversity of vascular plant species which occur in the transition zone between the High Atlas Mountains and the Sahara Desert. Morocco is a hotspot of plant diversity in the western Mediterranean region, and the High Atlas Mountains constitute an important centre of endemism in Morocco. The Photo Guide to Plants of Southern Morocco received images from the Institut Agronomique et Vétérinaire Hassan II in Morocco, the BioCentre Klein Flottbek, and the Botanical Garden and Herbarium of the University of Hamburg.

The **West African Photo Guide**, published in 2008 by the colleagues of BIOTA West Africa, inspired these initiatives (www.westafricanplants.senckenberg.de; Brunken et al. 2008). The three online photo guides provide free access to a large number of African plant species.

Acknowledgements

The authors' general acknowledgements to the organisations and institutions, which supported this work are provided in Volume 1.

References

- Brunken, U., Schmidt, M., Dressler, S., Janssen, T., Thiombiano, A., Zizka, G. (2008): www.westafricanplants.senckenberg.de. An image-based identification tool for West African Plants. – *Taxon* **57**: 1027–1028.
- Engelen, V.W.P. van, Wen, T.T. (eds.) (1995): Global and national soils and terrain digital databases (SOTER). Procedures manual. Revised ed. – *World Soil Resources Reports* **74**: i–ix, 1–126.
- Göttlicher, D., Bendix, J. (2004): Eine modulare Multi-User Datenbank für eine ökologische Forschergruppe mit heterogenem Datenbestand. – *Zeitschrift für Agrar-Informatik* **12**: 95–103.
- Gruss, P. (2003): Berlin declaration on open access to knowledge in the sciences and humanities. – <http://oa.mpg.de/openaccess-berlin/berlindeclaration.html> [04.08.2009].
- Muche, G., Finckh, M. (2009): BIOTABase short manual. – Hamburg: Biocentre Klein Flottbek, University of Hamburg. <http://www.biota-africa.org/downloads/biotabase/BIOTABaseManual.pdf>
- Tempel, P. (2002): SOTER global and national soils and terrain digital databases, database structure v3. – Wageningen: ISRIC. <http://www.isric.org/ISRIC/webdocs/docs/DatabaseStructureM1.pdf>