FAO/INFOODS food composition database for biodiversity

U. Ruth Charrondière *, Barbara Stadlmayr, Doris Rittenschober, Beatrice Mouillé, Emma Nilsson, Elinor Medhammar, Temesgen Olango, Sandra Eisenwagen, Diedelinde Persijn, Kristy Ebanks, Verena Nowak, Juan Du, Barbara Burlingame

FAO, Rome, Italy

ABSTRACT

Nutrient content can vary as much between different varieties of the same foods, as they do among different foods. Knowledge of varietal differences can therefore mean the difference between nutrient adequacy and inadequacy. The FAO/INFOODS food composition database for biodiversity has been developed with analytical data for foods described at the level of variety, cultivar and breed, and for underutilized and wild foods. It contains 6411 food entries and values for 451 components together with the bibliographic references and other information. The database is in MS Excel format and can be downloaded free-of-charge from the INFOODS website http://www.fao.org/infoods/biodiversity/index_en.stm. It is intended to annually publish new editions, making these data available for national and regional food composition databases. This database could be used to raise the awareness, promote and investigate food biodiversity and help to better estimate nutrient intakes.

© 2012 Food and Agriculture Organization of the United Nations. Published by Elsevier Ltd. All rights reserved.

1. Introduction

Diets low in variety but high in energy may contribute to the escalating problems of obesity, chronic diseases and micronutrient deficiencies. The Food and Agriculture Organization of the United Nations (FAO) actively promotes the conservation and sustainable use of biodiversity for nutrition and agriculture, as a means to increase dietary diversity (FAO, 2005; FAO, 2008a; FAO, 2008b; FAO, 2010a; FAO, 2010b; FAO, 2010c). However, the worldwide trend is towards dietary simplification and a loss of food biodiversity due to reliance on a limited number of varieties of staple and other crops. The number of rice varieties has dropped from thousands to just a few dozen. In Thailand, for example, the number of varieties under cultivation has fallen from more than 16,000 to just 37, and 50% of the rice-growing areas use only two varieties (IRC, 2006; Kennedy & Burlingame, 2003).

The Convention on Biological Diversity’s (CBD) cross-cutting initiative on biodiversity for food and nutrition responds to an emerging global consensus that (1) the Simplification of diets, the growing incidence of chronic diseases related to nutrition-poor, energy-rich diets and the decline in the use of locally available foods are linked; and (2), that biodiversity is the source of many foods and dietary components that can reverse this unhealthy trend (FAO, 2008b).

The identification and monitoring of nutrition indicators for biodiversity is critical for promoting sustainable diets. FAO and partners therefore developed two such indicators: one on food composition (FAO, 2008b) and one on food consumption (FAO, 2010b). Unfortunately, food composition and consumption data on wild, underutilized, indigenous and traditional plant and animal foods are limited and fragmented. This makes it difficult to evaluate the contribution of food biodiversity to nutritional adequacy.

It is well known that environmental factors and processing influence the composition of foods, but it is less known that varietal differences are responsible for very significant variations in nutrient contents (Burlingame, Charrondière, & Mouillé, 2009; Burlingame, Mouillé, & Charrondière, 2009). Indeed, the nutrient contents can vary as much or more between different varieties of the same foods as between different foods. Sweet potato cultivars, for example, can differ in their carotenoid content by a factor of 200 or more. The protein content of rice varieties can range from 5 to 14 g/100 g food (Kennedy & Burlingame, 2003) and the beta-carotene content of bananas can be less than 1 μg/100 g for some cultivars to as high as 8,500 μg/100 g for others (Englberger, Aalbersberg, Fitzgerald, Marks, & Chand, 2003; Englberger, Schierle, Marks, & Fitzgerald, 2003; Englberger et al., 2003). More examples are found in the FAO/INFOODS food composition database on biodiversity (FAO/INFOODS, 2012). Thus, the intake of one variety rather than another can mean the difference between micronutrient deficiency and adequacy (Burlingame, Mouillé, et al., 2009).

In the past, generic food composition data were considered sufficient for most purposes. Today, there is greater awareness...
surrounding the need for food composition studies that take biodi-versity into account. However, compositional data at the variety,
cultivar and breed level are not yet widely available. As a result,
farmers and consumers are often not aware of the high nutrient
values of certain cultivars and, thus do not grow or consume them.
Introducing more compositional data on biodiversity in food com-
position databases will improve the estimates of nutrient intake
and, assessments of dietary adequacy or inadequacy especially
with regard to micronutrients. Poor decisions have been made in
nutrition and health programmes because the micronutrient val-
ues in same national food composition databases did not reflect
the nutrient content of the varieties actually consumed. Nutrient
content should be one of the criteria which agriculture policy mak-
er and practitioners use to ensure that better and more nutritious
crop varieties are available for consumption (Burlingame, 
Charrondière, Dernini, Stadlmayr, & Mondovi, 2011).

The aim of INFOODS (International Network of Food Data Sys-
tems), since its establishment in 1984, is to stimulate and coordinate
efforts to improve the quality and availability of compositional data
globally. INFOODS has established tools. Standards and guidelines
on the generation, compilation and use of food composition data
(FAO/INFOODS, 2012a; Greenfield & Southgate, 2003), including a
distance learning tool: the Food Composition Study Guide
(Charrondière, Burlingame, Berman, & Elmadfa, 2011a; Charr-
ondière, Burlingame, Berman, & Elmadfa, 2011b). An entire module
of the study guide is devoted to biodiversity to enable researchers
and food composition table compilers to generate and publish more
data on food biodiversity, i.e. on varieties, cultivars and breeds.

Several international bodies and committees (FAO, 1995; IRC,
2006) recommend the analysis and compilation of nutrient data at
the taxonomic level below species. As no such database currently
exists to the authors' knowledge, FAO developed the FAO/INFOODS
food composition database for biodiversity. Version 1.0 was
launched in December 2010, updated in 2011 as version 1.1 (FAO/
INFOODS, 2011) and further updated in 2012 as version 2.0 (FAO/
INFOODS, 2012b). Compilers and users can easily access the data
and incorporate them into national food composition databases. To-
gether with food consumption data on food biodiversity, it will then
be possible to evaluate the contribution of biodiversity to nutrient
intakes thus, improving the quality of dietary assessments.

This article describes the FAO/INFOODS food composition data-
base for biodiversity version 2.0 and its development since version
1.0. In the document, the term 'Biodiversity Database' will be used
when talking about the database in general and it will be abbrevi-
ated as BioFoodComp2.0 when indicating version 2.0 specifically.

2. Materials and methods

The Biodiversity Database was prepared using the FAO/INFOODS
compilation tool (Charrondière & Burlingame, 2011c), a simple food
composition database management system in MS Excel. The tool
was adapted for this purpose by adding new fields to the overall
structure to capture additional information. For example, in version
1.0 the following fields were added: ‘country/region’, ‘species/sub-
species’, ‘cultivar/variety/accession name’, ‘season/other’, while the
‘compiler ID’ was added in version 1.1. Version 2.0 was updated by
including more fields for easy sorting and for capturing additional
information which are important for the understanding of the data
(such as ‘comments on data processing/methods’, ‘publication year’
and ‘latest revision’). INFOODS component identifiers, also called
tagnames (Klensin, Feskanich, Lin, Truswell, & Southgate, 1989)
were selected to compile the compositional data in a standardized
manner as they are suitable for Excel. EuroFIR components
(EuroFIR, 2012) could not be used as they would require several
additional cells to obtain the same information as given by a single
INFOODS tagname such as information on method, expression and
unit. This was considered too complicated for an Excel application.

Data were mainly collected and compiled by interns, volunteers
and consultants working at FAO. Comprehensive literature searches
were done on potatoes, fruits, vegetables, roots and tubers, beef,
fish and shellfish, insects and milk of underutilized species. The
compositional data were entered into the Biodiversity Data-
base and review papers were published (Burlingame, Mouillé,
et al., 2009; Medhammer, Wijesinha-Bettoni, Stadlmayr, Nilsson,
Charrondière, & Burlingame, 2011; Olango, Stadlmayr, & Charr-
ondière, 2011). A general search for food biodiversity to report
on the Nutrition Indicators for Biodiversity (FAO, 2008b; FAO,
2010b) yielded several scientific articles with compositional data
from which data were partially compiled into the Biodiversity Data-
base. Additionally, some original data were received through the
INFOODS network and personal communication.

Papers were scanned for general inclusion and exclusion criteria
such as unequivocal food descriptions, missing units, denominators
and missing values when data presented graphically. Some data
needed to be converted to express them as ‘per 100 g edible portion
on fresh weight basis’ (EP) using the FAO/INFOODS guidelines
on conversion among different units, denominators and expressions
(FAO/INFOODS, 2012c). All data were evaluated using the
FAO/INFOODS guidelines for checking food composition data
prior to the publication of a user table/database (FAO/INFOODS,
2012d) and were peer reviewed by FAO staff.

3. Results and discussion

The FAO/INFOODS food composition database for biodiversity is
up to the author’s best knowledge the first global database on bio-
diversity containing solely analytical data on the composition of foods
counting for biodiversity (FAO/INFOODS, 2012a): at least
one compositional value must be reported at variety/cultivar/breed
level for common foods and at species level (or with local name)
for wild and underutilized foods. The data are reported as stated
in the original publication but expressed per 100 g edible portion
on fresh weight basis (EP). However, no values were added through
estimation or calculation. The only exceptions are the changes of
units and denominators (e.g. from dry matter to fresh weight basis)
and the calculation of the sum of fatty acids using fatty acid con-
version factors from the literature, if needed to calculate the fatty
acid content from g/100 g total fatty acids to g/100 g EP. Each com-
positional value is listed together with the bibliographic reference,
the food name in English, the scientific name including cultivar/
variety/breed or accession name and the compiler ID. If available,
information was also entered on local name, country, region, sea-
son, other specification, sample size, and any additional comment
if relevant (e.g. soil conditions, feeding practices). In general,
scientific articles on the composition of foods, including on biodi-
versity, do not identify the exact genetic coding or give reference
to a genetic databank, therefore, this information is not found in
the Biodiversity Database.

BioFoodComp2.0 is accompanied by a documentation file which
contains information on the development of the database, the de-
finition of components, and explanation of which fields and data
were added in version 2.0.

The database is an Microsoft Excel file composed of different
worksheets: copyright, one sheet for each of the 12 food groups,
on the description of codes, a component list and the bibliography.
Due to time restriction, the bibliographic references were docu-
mented (indicated in the BibiOD field in the datasheet of the food
groups while providing the full reference in the ‘bibliography’ worksheet) but no detailed information was collected on value
documentation, sampling or method description.
The data quality index field was not used for several reasons: firstly, INFOODS has not yet developed a comprehensive data evaluation system of analytical data and it seems that the EuroFIR data evaluation scheme (EuroFIR, 2009) is being further refined. Secondly, it was decided that at this stage, given the limited time and funding, it is more important to compile more data than to add a quality index. However, about 30% of the articles were excluded using an internal inclusion and exclusion criteria and the FAO/INFOODS guidelines (FAO/INFOODS, 2012d).

Also some of the data of version 1.1 were excluded as they did not comply with the criteria and guidelines. Examples are data that were previously entered from review articles (i.e. not original analytical data from primary sources), or the sum of proximates were outside the now defined acceptable range of 95–105 g/100 g EP, or because some of the component names were not yet official INFOODS component identifiers.

Many articles had missing values for the unit conversion (e.g. missing water values when data were expressed per dry matter basis), did not define the foods and/or components unambiguously, or included implausible data (e.g. sum of proximates outside the acceptable range of 95–105 g/100 g EP). Authors of the papers were contacted but in most cases they did not provide the missing data. As a result, letters were written to the editors of the journals so that the common problems leading to poor quality papers could be avoided. In addition, FAO/INFOODS will publish a check list for data presentation before their submission in scientific articles on the INFOODS website. It is hoped that in the future, data will be presented in a clearer and unambiguous manner so that more compositional data can be used from the scientific literature.

The compositional data either on variety/cultivar/breed level or on species level for wild and underutilized foods were assigned to 12 food groups. Most of the 6411 foods entries in version 2.0 (see Table 1) are on potatoes (1671), fruits (1635), finfish (1069), insects (514), vegetables (354), milk (273 including from sheep, goat, mare, camel, yak, buffalo, donkey, moose, reindeer and mithun), mammal meat (217), other roots and tubers (199), crustaceans (129), molluscs (106), nuts and seeds (101), cereals (90), legumes (28), reptiles/amphibians (15) miscellaneous foods (5) and eggs (5). No data were so far entered for poultry, herbs and spices. Data collection and compilation are ongoing. For version 1.1 and 2.0, 98% of the data were compiled by FAO staff and 2% came from collaborators working with FAO on the food composition database on ‘Selected foods from West Africa’ (Stadlmayr et al., 2010) and from Brazil (personal communication). Some of the data are on experimental cultivars which do not reach the market. It was however decided to include these data to complete the variation of compounds in foods.

The Biodiversity Database contains 451 components. As shown in Fig. 1, the data entered are mainly on fatty acids (31%), macronutrients and macronutrient fractions (25%) and minerals (20%), followed by phytochemicals and bioactive compounds (8%), vitamins (8%), other components (3%) and toxic trace elements (1%). In version 1.1, the high amount of data were for phytochemicals, mainly from the food group potatoes, which also accounted for the highest number of data entries. In the BioFoodComp2.0, most data are on fatty acids due to the fish data which were added and most of the mineral data were added for fruits (29% of all mineral data entered). Some of the components are on toxic compounds such as heavy metals or on solanine or glycoalkaloids but the amount is limited as there was no specific search done on these compounds. They were in general included when also other compositional data were available in the respective data source.

Table 1 shows the number of data entries per continent. Countries were assigned to continents following the M49 UN classification (United Nations Statistics Division., 2012). Most of the analytical data compiled were for the Americas (1515 for North America; 752 for South America, 166 for Central America), followed by Europe (360), Asia (317), Oceania (248) and Africa (125) and the Antarctic region (2). The high contribution of analytical data from North America was derived mainly from the food group potatoes and reflects different breeding lines.

The Biodiversity Database can be downloaded free-of-charge from the INFOODS website. Version 1.0 was launched in December 2010, version 1.1 was published in September 2011 and version 2.0 in July 2012. From December 2010 to June 2012, 5060 hits were counted for the Biodiversity Database (FAO, unpublished report), and it is expected that with more food entries, a higher coverage of food groups and components, and an increased knowledge of its existence, it will become more used by nutritionists, compilers and other professionals including those in agriculture.

It is intended to publish annual versions of the Biodiversity Database with new data, if funds are available. With additional funding, FAO will finance data collection and compilation through interns, volunteers and consultants or through other collaborators. The Biodiversity Database will then grow rapidly and will hopefully soon cover all food groups and contain significantly more data per food group. Starting with version 1.1, a compiler ID was added to identify the compiler for each food entry and from version 2.0 on, compiler IDs are added, whenever data of a food entry are revised. Each data compiler is acknowledged for the provided data, not only at data level through the ID but also by listing the name in the copyright spreadsheet, if at least 10 food entries are compiled and sufficient documentation is provided (i.e. data together with laboratory report and/or other documentation on sampling and analytical method – or the published paper). Additionally, the sources of funding are also mentioned. When researchers and other data owners provide more of their published or unpublished...
data, the usefulness of the database will increase more rapidly. Additional international recognition of the data provider could be achieved as their name would be associated with the provided data in the Biodiversity Database. This could motivate researchers, especially those of ‘orphan data (e.g. \( n = 1 \))’ to publish their data in the Biodiversity Database.

More scientific articles on the variability of nutrients and phytochemicals are planned to be published which will assist in the communication of the significant nutritional differences among varieties, cultivars and breeds. It is hoped that through the increased evidence base more researchers will investigate food biodiversity (composition and consumption) and the impact of food biodiversity on nutrition and health. It is also expected that through advocacy more data will be provided to FAO, and many more researchers and compilers will use them to update their databases and allow better nutrient intake estimations of their populations. A good example for the use of the Biodiversity Database is the inclusion of its data in the Composition of Selected Foods in West Africa (Stadlmayr et al., 2010) and in the West African Food Composition Table (Stadlmayr et al., 2012).

4. Conclusion

The Biodiversity Database is a global repository of analytical data on food biodiversity of acceptable data quality. It is hoped that in the future, more and better data will be available in the scientific literature to further develop the Biodiversity Database and provide this essential tool for the investigation and promotion of the sustainable use of food biodiversity and for mainstreaming food biodiversity into nutrition projects, programmes and interventions.

Acknowledgement

Acknowledgment for financial contribution: FAO AGNDA, Multi Disciplinary Funds through FAO ESS division, WAHO/ECOWAS and World Agroforestry Centre.

References


Engberger, L., Aalbersberg, W., Ravi, P., Bonnin, E., Marks, G. C., Fitzgerald, M. H., et al. (2003). Further analyses on Micronesian banana, taro, breadfruit and other...


