ELSEVIER

Journal of Food Composition and Analysis 17 (2004) 423–438

JOURNAL OF FOOD COMPOSITION AND ANALYSIS

www.elsevier.com/locate/jfca

Original Article

# Development of the International Life Sciences Institute Crop Composition Database ☆

William P. Ridley[a], Raymond D. Shillito[b], Isabelle Coats[c,1], Henry-York Steiner[d], Mary Shawgo[e], Amy Phillips[f], Patricia Dussold[g], Lucyna Kurtyka[h,*]

[a] *Monsanto Company, 800 North Lindbergh Blvd., St. Louis, MO 63167, USA*
[b] *Bayer CropScience, 2 T. W. Alexander Drive, Research Triangle Park, NC 27709-2014, USA*
[c] *Pioneer/Dupont, 7250 NW 62nd Ave., Johnston, IA 50131, USA*
[d] *Syngenta, 3054, Cornwallis Road, Research Triangle Park, NC 27709-2014, USA*
[e] *Renessen LLC, 3000 Lakeside Drive, Suite 300 South, Bannockburn, IL 60015, USA*
[f] *Dow AgroSciences LLC, 9330 Zionsville Road, Indianapolis, IN 46268, USA*
[g] *Digital Consulting and Software Services, 12655 Olive Blvd., Suite 315, St. Louis, MO 63146, USA*
[h] *International Life Sciences Institute (ILSI), One Thomas Circle, NW 9th Floor, Washington, DC 20005, USA*

Received 24 September 2003; received in revised form 27 February 2004; accepted 5 March 2004

**Abstract**

In May 2003, the International Life Sciences Institute released an online comprehensive Crop Composition Database (www.cropcomposition.org) that provides up-to-date information on the natural variability in composition of conventional crops. The database is a compilation of data on the nutrients, anti-nutrients, and secondary metabolites for maize and soybean samples obtained from controlled field trials, in multiple world-wide locations over a 6-year period. The analyses of the samples were conducted using validated analytical methods with reference citations and, in most cases, under the guidelines of Environmental Protection Agency Good Laboratory Practices. Version 1.0 of the database contains more than 53,000 data points that may be searched and accessed based upon user-selected queries. The database complements existing food and nutrient databases and should be of interest to research and regulatory scientists in many areas such as plant biology, food science, and animal nutrition.


*Keywords:* International Life Sciences Institute; ILSI; Crop composition; Nutrients; Maize; Soybean; Database

☆ The data in this paper comes from version 1.0 of the database. Version 2.0 was released on 5 April 2004 and was not available when this article was submitted for publication. Version 2.0 contains all the data in version 1.0 plus additional data points.

*Corresponding author. Tel.: +1-202-659-0074, fax: +1-202-659-3859.
*E-mail address:* lkurtyka@ilsi.org (L. Kurtyka).
[1] Current address: Proteia Consulting Services, 41 Calle Sol, San Clemente, California 92672, USA.


doi:10.1016/j.jfca.2004.03.006

## 1. Introduction

Food and feed crop composition studies are considered an essential part of the safety assessment of new crop varieties, including those developed through biotechnology. Information obtained from such studies is used to assess similarities and differences in important nutrients and anti-nutrients. Information on the composition of crop materials has traditionally relied on publicly available data sources that were 20 or more years old (Ensminger et al., 1990; Jugenheimer, 1976; Watson, 1982). In some cases the analytical methods used to generate the crop data are not known and/or their performance parameters are not available. As a result, there was a clear need for a single, easily accessible and up-to-date source of crop composition information.

The requirements for data quality of food nutrients including representative sampling, accurate and precise methodology and data compilation, and their importance for users of the data have been described (Southgate, 2002). These principles have been incorporated into the United States Department of Agriculture (USDA) Nutrient Database (Holden et al., 2002), which is widely used by nutritionists and other health professionals for making essential dietary decisions related to public health. The International Network of Food Data Systems (INFOODS) was established under the auspices of the United Nations University (Burlingame, 1996) to foster the development of quality food composition data and its dissemination on a global basis. A database that provided nutrient and anti-nutrient data on the raw agricultural commodities such as maize grain and soybean seed—the raw materials for food and feed products—would complement both the USDA Nutrient Database and INFOODS. In addition, accurate information on the composition of crop materials and their variability, along with production year and location, would be important for many food commodities whose processing and desired properties such as flavor and texture depend on composition parameters.

The appearance of the first plant agricultural biotechnology product, the FlavrSavr™ tomato in 1994 (Martineau, 2001) followed by the introduction of herbicide-tolerant and insect-protected soybean and maize products in 1995–1996 signaled the potential for this new technology to provide products that could increase the productivity and decrease the environmental impact of agricultural practices (Ridley, 2004). An essential aspect of the development of these new products is the assessment of compositional equivalence compared to conventional crop varieties in a process known as “substantial equivalence” (Organization for European Cooperation and Development (OECD), 1993; World Health Organization, 1995; Food and Agriculture Organization of the United Nations, 1996). The evaluation of substantial equivalence requires the availability of comprehensive, up-to-date composition data to determine if the new biotechnology product falls within the natural variability of a crop. OECD has published a series of Consensus Documents summarizing the existing data on the composition of crops such as soybean, maize, and canola (OECD, 2002). These documents are an excellent resource, but are limited by the availability of published composition data and are not easily updated to reflect current data or technologies.

Making an electronic, easily accessible world-wide compilation of crop composition data available for use by research and regulatory scientists was in alignment with the mission of the International Life Sciences Institute (ILSI) to further the understanding of scientific issues relating to nutrition, food safety, toxicology, risk assessment, and the environment. An ILSI International Food Biotechnology Committee’s Task Force composed of representatives from six agricultural biotechnology companies who agreed to share their data on crop composition directed the development of the database. These data were obtained from samples generated in controlled field trials at various global locations and analyzed in most cases under the strict documentation guidelines of Good Laboratory Practices (GLP) (United States Environmental Protection Agency, 1989). A proposal was developed in the fall of 2000 and accepted by the ILSI International Food Biotechnology Committee in February 2001. In May 2003, the ILSI Crop Composition Database (www.cropcomposition.org) was launched. This report describes the rationale, structure, and data used for constructing the ILSI Crop Composition Database.

## 2. Materials and methods

### 2.1. Structure of the database

There are a number of different strategies that have been used in the construction of food databases: (a) monohierarchical classification (b) faceted description using standard vocabularies and (c) faceted description using free text (Schlotke et al., 2000). All of these systems require the definition of terms and a clear description of the relationships among the terms. Table 1 contains a list of the key terms or ''descriptors'' used in the construction of the ILSI Crop Composition Database. The central term *numeric value* is the foundation upon which the other terms are dependent. All other descriptors and sub-descriptors in Table 1 provide an increased clarity for *numeric value* including *unit*s, *analyte*, *crop*, and *tissue* that are essential for the definition of *numeric value*. Additional terms including *analysis method*, *country*, and *seed* provide additional information that reflects the origin and source of the numeric values and help define the fitness of the data for particular users.

Table 2 contains a list of the analyte categories, analytes, primary units of measure, and crop types for Version 1.0 of the database. The selection of the categories was based on an assessment of the properties of the analytes contained within that category. For example, the selection of analytes to be included in the category *amino acids* was unambiguous and reflected the output of a standard amino acid analytical method (e.g., glutamic acid represents both glutamic acid and glutamine since the acidic conditions used to hydrolyze proteins converts glutamine to glutamic acid). The category *bioactive* was not as transparent and resulted from an evaluation of the known

Table 1
Data descriptors and sub-descriptors for ILSI Crop Composition Database, version 1.0

| Description | Sub-descriptor | | |
|---|---|---|---|
| Analyte category | Analyte | | |
| Analysis method | Method reference | Method code | |
| Numeric value | Units | | |
| Country | State | | |
| Crop | Source | Type | Year |
| Seed | Vendor | Variety | |
| Tissue | Forage | Grain | Seed |

Table 2
List of analyte categories and analytes for ILSI Crop Composition Database, version 1.0

| Analyte category | Analyte | Primary unit of measure[a] | Crop types |
|---|---|---|---|
| Amino acid | Alanine | mg/g FW | Maize, soy |
| | Arginine | mg/g FW | Maize, soy |
| | Aspartic acid | mg/g FW | Maize, soy |
| | Cystine/cysteine | mg/g FW | Maize, soy |
| | Glutamic acid | mg/g FW | Maize, soy |
| | Glycine | mg/g FW | Maize, soy |
| | Histidine | mg/g FW | Maize, soy |
| | Isoleucine | mg/g FW | Maize, soy |
| | Leucine | mg/g FW | Maize, soy |
| | Lysine | mg/g FW | Maize, soy |
| | Methionine | mg/g FW | Maize, soy |
| | Phenylalanine | mg/g FW | Maize, soy |
| | Proline | mg/g FW | Maize, soy |
| | Serine | mg/g FW | Maize, soy |
| | Threonine | mg/g FW | Maize, soy |
| | Tryptophan | mg/g FW | Maize, soy |
| | Tyrosine | mg/g FW | Maize, soy |
| | Valine | mg/g FW | Maize, soy |
| Bioactive | Lectins | HU/mg FW | Soy |
| | Phytic acid | % FW | Maize, soy |
| | Raffinose | % FW | Maize, soy |
| | Stachyose | % FW | Soy |
| | Total diadzein | ppm FW | Soy |
| | Total genistein | ppm FW | Soy |
| | Total glycitein | ppm FW | Soy |
| | Total isoflavones | ppm FW | Soy |
| | Trypsin inhibitor | TIU/mg FW | Maize, soy |
| Carbohydrates | Starch | % FW | Maize |
| Fatty acid | 8:0 Caprylic | % Total FA | Maize, soy |
| | 10:0 Capric | % Total FA | Maize, soy |
| | 11:0 | % Total FA | Maize |
| | 12:0 Lauric | % Total FA | Maize, soy |
| | 14:0 Myristic | % Total FA | Maize, soy |
| | 14:1 Myristoleic | % Total FA | Maize, soy |
| | 15:0 Pentadecanoic | % Total FA | Maize, soy |
| | 15:1 Pentadecenoic | % Total FA | Maize, soy |
| | 16:0 Palmitic | % Total FA | Maize, soy |
| | 16:1 Palmitoleic | % Total FA | Maize, soy |
| | 16:2 | % Total FA | Maize |
| | 16:3 | % Total FA | Maize |
| | 16:4 | % Total FA | Maize |
| | 17:0 Heptadecanoic | % Total FA | Maize, soy |
| | 17:1 Heptadecenoic | % Total FA | Maize, soy |
| | 18:0 Stearic | % Total FA | Maize, soy |

Table 2 (*continued*)

| Analyte category | Analyte | Primary unit of measure[a] | Crop types |
|---|---|---|---|
| | 18:1 Oleic | % Total FA | Maize, soy |
| | 18:2 Linoleic | % Total FA | Maize, soy |
| | 18:3 Gamma linolenic | % Total FA | Maize, soy |
| | 18:3 Linolenic | % Total FA | Maize, soy |
| | 18:4 | % Total FA | Maize |
| | 19:0 | % Total FA | Maize |
| | 20:0 Arachidic | % Total FA | Maize, soy |
| | 20:1 Eicosenoic | % Total FA | Maize, soy |
| | 20:2 Eicosadienoic | % Total FA | Maize, soy |
| | 20:3 Eicosatrienoic | % Total FA | Maize, soy |
| | 20:4 Arachidonic | % Total FA | Maize, soy |
| | 20:5 | % Total FA | Maize |
| | 21:0 | % Total FA | Maize |
| | 21:5 | % Total FA | Maize |
| | 22:0 Behenic | % Total FA | Maize, soy |
| | 22:1 | % Total FA | Maize |
| | 22:2 | % Total FA | Maize |
| | 22:3 | % Total FA | Maize |
| | 22:4 | % Total FA | Maize |
| | 22:5 | % Total FA | Maize |
| | 22:6 | % Total FA | Maize |
| | 24:0 | % Total FA | Maize |
| | 24:1 | % Total FA | Maize |
| Fiber | Acid detergent fiber | % FW | Maize, soy |
| | Crude fiber | % FW | Maize, soy |
| | Neutral detergent fiber | % FW | Maize, soy |
| Mineral | Cadmium | ppm FW | Maize |
| | Calcium | ppm FW | Maize, soy |
| | Chloride | ppm FW | Maize |
| | Copper | ppm FW | Maize |
| | Iron | ppm FW | Maize, soy |
| | Magnesium | ppm FW | Maize, soy |
| | Manganese | ppm FW | Maize |
| | Phosphorus | ppm FW | Maize, soy |
| | Potassium | ppm FW | Maize, soy |
| | Selenium | ppm FW | Maize |
| | Sodium | ppm FW | Maize |
| | Sulfur | ppm FW | Maize |
| | Zinc | ppm FW | Maize |
| Other metabolites | Ferulic acid | ppm FW | Maize |
| | Furfural | ppm FW | Maize |
| | Inositol | μg/g FW | Maize |
| | *p*-Coumaric acid | ppm FW | Maize |
| Proximate | Ash | % FW | Maize, soy |
| | Carbohydrates by calculation | % FW | Maize, soy |

Table 2 (*continued*)

| Analyte category | Analyte | Primary unit of measure[a] | Crop types |
|---|---|---|---|
| | Crude fat | % FW | Maize |
| | Crude protein | % FW | Maize, soy |
| | Moisture | % FW | Maize, soy |
| | Total fat | % FW | Maize, soy |
| Vitamin | Beta-carotene | mg/100 g FW | Maize |
| | Beta-tocopherol | mg/100 g FW | Maize |
| | Delta-tocopherol | mg/100 g FW | Maize |
| | Folic acid | mg/100 g FW | Maize, soy |
| | Gamma-tocopherol | mg/100 g FW | Maize |
| | Total tocopherols | mg/100 g FW | Maize |
| | B1 (Thiamin) | mg/100 g FW | Maize, soy |
| | B2 (Riboflavin) | mg/100 g FW | Maize, soy |
| | E | mg/g FW | Maize, soy |

[a] For the units of measure used in the table: FW = fresh weight, HU = hemagglutinating unit as defined by the method reference in the database, ppm = parts per million, TIU = trypsin inhibitor unit as defined by the method reference in the database, FA = fatty acids.

biological properties of the analytes represented in this category, which includes lectins, phytic acid, isoflavones, and other molecules with important physiological activity. Fiber has been frequently included under proximates. However, there are many different means for categorizing fiber such as acid detergent fiber, neutral detergent fiber, and total dietary fiber that warranted a separate category for this analyte. The selection of the primary unit of measure for an analyte was based on common practice and also the desire to minimize the number of digits after the decimal point in the resulting values. All analytes were not associated with all crop types. Isoflavones and lectins, e.g., were only associated with soybeans. A glossary of terms and listing of analyte categories with associated units was also included in the database website.

## 2.2. Source and handling of samples

The quality of the data in a database is directly associated with the quality of the samples which are analyzed. The representative character of the samples, as well as the manner in which they are processed and stored, impacts sample quality. The samples analyzed for the ILSI database were all derived from controlled field trials under the direction of a production plan. Records were maintained to document growing conditions, incidence of insect and disease infestation, and even inadvertent crop damage due to vandalism and wildlife crop destruction. Each maize grain or soybean seed sample represents a composite sample that was collected from representative plants throughout a single plot. This sampling procedure was similar to the common practice employed for pesticide residue trials. In some cases, the sample may have been collected from a large number of plants (e.g., a bulk sample from a large plot). In these cases care was taken to obtain a representative sample from the bulk by employing appropriate sampling methods. Forage samples represented a composite sample from a minimum of two whole plants.

The grain samples used to generate the analytical data were analyzed within 12 months of harvest since grain samples used as food and feed are commonly stored for up to 12 months or even longer before use. While raw grain may have been stored at ambient temperature for several months, any grain that was ground was always stored frozen. Green tissue forage samples were either frozen immediately after harvest or oven-dried and then stored frozen until they were analyzed, typically within 12 months of harvest.

### 2.3. Methods of analysis

The important relationship between the numeric value for a specific food or crop component and the analytical method used to measure that component has been recognized by database developers (Holden et al., 2002; Schlotke et al., 2000). All data points in the ILSI database were associated with a referenced method that was identified by a method code. The method code refers to a citation that can be included with the data output at the discretion of the user. The Environmental Protection Agency (EPA) GLP model was used as the standard for defining acceptable analytical procedures and documentation. The analyses were thus only conducted using validated methods in either accredited/certified laboratories or laboratories experienced with specific analytical methodology. Quality control checks were included with the analytical runs using certified or historically verified standards. In addition, the laboratory and equipment records and data were archived for at least 5 years after submission to the database and in most cases these supporting data are retained in company archives for the lifetime of the products for which the data were collected.

No analytical data in the database were based on averaging of values for multiple samples. Each data point was derived from the analysis of a single composite sample. Therefore, a dataset consisted of a series of data points for a number of different analytes obtained from a single sample.

### 2.4. Data submission

Data were submitted to the database as a comma-delimited (.csv) or tab-delimited (.txt) file with each analyte value given in one of the pre-defined units of measure. In addition to the numeric value, the data files contained the sub-descriptor information outlined in Table 1 that defined the data point in greater detail and serves as the basis for user-directed searches of the database. For example, a reference for the analytical method, the year and production site for the crop, the seed vendor, and variety/hybrid were part of the data file. The format for data submission was standardized to speed the compilation process and to ensure that the data were incorporated correctly. The submission requirements were included in a manual that was available to all data sources (data submitters).

An outline of the flow of data from the data source to the database is shown in Fig. 1. An important step in the compilation of the data files is provided by the *Curator*. The Curator is an information technology professional employed by ILSI who serves to control the entry and integrity of the data as well as maintaining the confidentiality of the data submitted. The Curator has password-protected access to the database and is the only individual who can upload data files and add or delete information in the database.

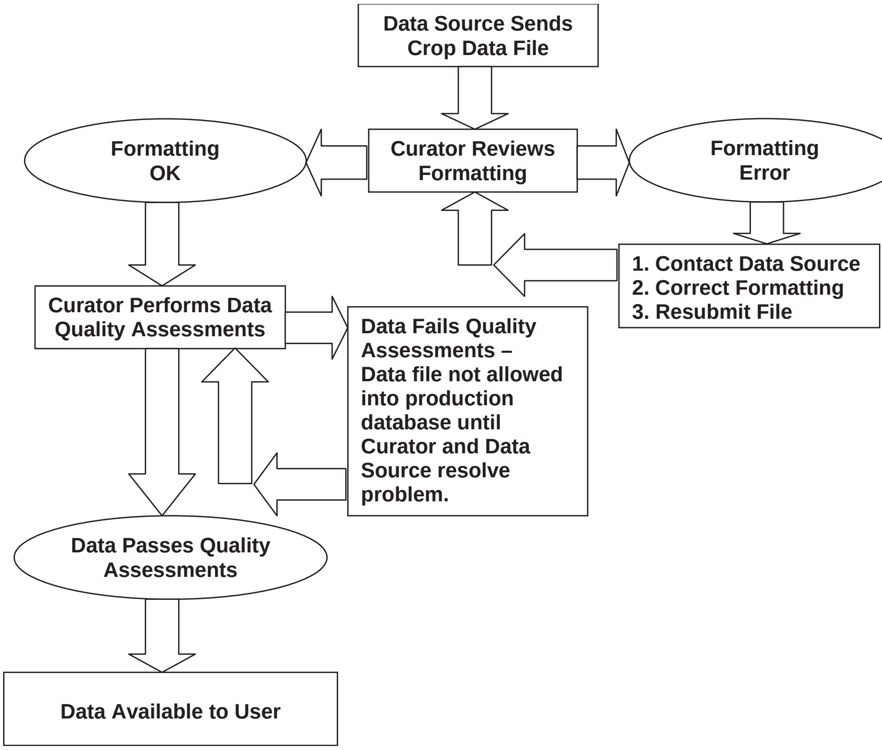


Fig. 1. Compilation of data for the ILSI Crop Composition Database.

The Curator reviewed the format of files and if errors were identified the files were returned to the data source for correction. As part of the data quality assessment, the Curator screened the data against a set of control values for each analyte that were used to examine the data for potential keying or entry errors. These control values were derived from an evaluation of the range of values for each analyte included in the initial 1000 datasets submitted to the database. If a value was found to fall outside the control range, the data source was asked to reconfirm the accuracy of the value before it was included in the database. As another means of data quality assessment, select datasets were retrieved from the database after import and returned to the appropriate data source for review and comparison against the original data file.

### *2.5. System documentation and testing*

Documentation of the requirements and development of the code for the database was included in a series of formal specification documents. The testing of performance was also detailed in written documents of the system test plan, system testing summary, development testing, production testing, and quality control summary. Additional written documents to support the system prior to release were (a) system requirements for the Host Provider, (b) user manual, (c) Curator manual, (d) crop data submission requirements, and (e) system upgrades protocol.

### 2.6. System description

Version 1.0 of the ILSI Crop Composition Database is an Oracle database, version 9.2. The application required one dedicated server for web hosting and one dedicated server for the Oracle database, both housed on one physical server. There were two web interfaces, a user interface and a password-protected Curator interface, written in Perl with CGI and HTML.

## 3. Results and discussion

### 3.1. Search capabilities

The ILSI Crop Composition Database was structured to allow the user to select attributes of the datasets and data points and then search the database for the information of interest. The database interface prepares a search query that will select all the datasets or assays indicated by the search criteria. Once the search criteria are selected, the user can select the attributes and analytes of interest within the datasets returned by the search. The user may or may not choose to display all the analytes contained within the search since the display selection is a separate operation.

Six screens within the program as specified below implement the search selection and refinement process:

1. First screen allows the user to select the dataset attributes to be included in the search as well as the analyte category for further refinement (see Fig. 2).
2. Second screen allows the user to define the numeric and method criteria for each analyte contained in the analyte category(ies) selected on the first screen.
3. Third screen confirms the search criteria selected on the first two screens.
4. Fourth screen allows the user to choose data fields to be displayed on the results page.
5. Fifth screen allows the user to select the analytes to be displayed within the samples returned by the search query.
6. Sixth screen is the results display page (see Fig. 3).

### 3.2. Output of data

As shown in Fig. 3, the data output from the ILSI database is summarized with the minimum, maximum, and average values, as well as the number of individual measurements that were used to derive the summary values. The output can be comprehensive, e.g., contain all the proximates for all years and all locations, or may be more refined and focus on a single amino acid in a particular geography in a single year. Data can be expressed in fresh weight or dry weight units. Method citations keyed by a method code plus a list of varieties/hybrids and vendors can also be provided separately.

### 3.3. Scope of maize and soybean data in version 1.0

Table 3 contains a summary of the maize and soybean data included in version 1.0 of the ILSI Crop Composition Database. The data cover 6 years and multiple locations in North and South

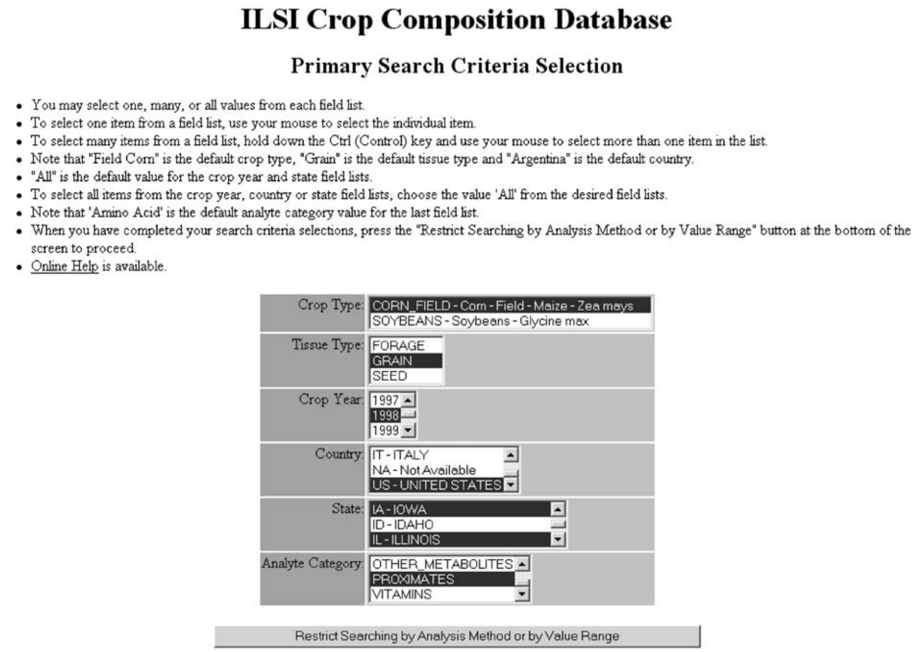


Fig. 2. Selection of criteria for ILSI Crop Composition Database search.

America and the European Union. The number of analytes, datasets, and data points varies with the crop but the scope of the data is indicated by a total of 102 analytes, 1446 datasets and 53,221 data points.

Tables 4–6 contain comparisons of the data in the ILSI database to the USDA Nutrient Database and the OECD Consensus Documents. In general, the values are in agreement with a few exceptions. The calcium values for both maize and soybeans are lower than the values reported in the USDA Nutrient Database as shown in Tables 4 and 5. These differences could be due to differences in methodology or differences in sample source and scope. The levels of riboflavin in maize reported in Table 5 and the levels of glycitein in soybeans reported in Table 6 for the OECD Consensus Document have a broader range than those of the ILSI database and could reflect differences in methodology, sampling, or scope of the data.

### 3.4. Applications of database

Two applications of the data in the ILSI database are shown in Figs. 4 and 5. Fig. 4 shows bar graphs of the % fat, % protein, and % moisture for maize expressed as % fresh weight in the United States for the years 1998–2000. It was generated by importing the data from the database into a data spreadsheet such as Excel® and then plotting the results. Fig. 5 is a graph of all the data (651 data points) in version 1.0 for % protein in maize expressed on a dry weight basis. The resulting graph is a clear indication of the natural variability of this analyte due to the range and interaction of genotype with the variable global environmental conditions.

# ILSI Crop Composition Database

## Search Results

Disclaimer: The database is provided "as is" and without warranty of any kind, whether express or implied. ILSI and its member companies expressly disclaim implied warranties of merchantability, fitness for a particular purpose, and noninfringement. In no event shall ILSI or any of its member companies be liable to database users or any third party in any way, including, without limitation, for direct, indirect, consequential, incidental, reliance, or special damages. Because it is not feasible to provide in this database statistical analyses for search results derived from all combinations of selection criteria, the responsibility for any statistical analyses and interpretation of results rests with the user.

Note: The number of significant figures displayed in the table reflects the actual result submitted, and does not imply a specific accuracy for the analytical result. Please consult the individual method references for accuracy and precision information on each method. Further information on accuracy, repeatability and uncertainty estimates for analytical methods can be found in Thompson et al., Pure Appl. Chem., Vol. 74, No. 5, pp. 835–855, 2002, and at http://www.measurementuncertainty.org/ .

Generated: Tue Aug 19 10:01:13 2003

### Selection Criteria

| | |
|---|---|
| Crop Type: | Corn - Field - Maize - Zea mays |
| Tissue Type: | Grain |
| Crop Year: | 1998 |
| Country: | UNITED STATES |
| State: | IOWA, ILLINOIS |

### Analyte Values - Minimum, Maximum and Average

**Note that Below LOQ values are not included in the table below.**

| Analyte | Minimum | Maximum | Average | n | Units |
|---|---|---|---|---|---|
| Fatty_Acids - 18:0 Stearic | 1.74 | 2.56 | 2.11 | 7 | % Total FA |
| Proximates - Total Fat | 2.770 | 3.710 | 3.443 | 7 | % FW |

### Search Results - Data Values

| Analyte | Value | Units | State | Analysis Method Code |
|---|---|---|---|---|
| Fatty_Acids - 18:0 Stearic | 2.17 | % Total FA | ILLINOIS | A0033 |
| Fatty_Acids - 18:0 Stearic | 1.98 | % Total FA | ILLINOIS | A0033 |
| Fatty_Acids - 18:0 Stearic | 1.74 | % Total FA | IOWA | A0033 |
| Fatty_Acids - 18:0 Stearic | 2.45 | % Total FA | IOWA | A0033 |
| Fatty_Acids - 18:0 Stearic | 1.82 | % Total FA | ILLINOIS | A0033 |
| Fatty_Acids - 18:0 Stearic | 2.56 | % Total FA | IOWA | A0033 |
| Fatty_Acids - 18:0 Stearic | 2.05 | % Total FA | IOWA | A0033 |
| Proximates - Total Fat | 3.550 | % FW | ILLINOIS | A0005 |
| Proximates - Total Fat | 3.660 | % FW | IOWA | A0005 |
| Proximates - Total Fat | 3.360 | % FW | ILLINOIS | A0005 |
| Proximates - Total Fat | 3.600 | % FW | ILLINOIS | A0005 |
| Proximates - Total Fat | 2.770 | % FW | IOWA | A0005 |
| Proximates - Total Fat | 3.450 | % FW | IOWA | A0005 |
| Proximates - Total Fat | 3.710 | % FW | IOWA | A0005 |

### Analysis Methods

| Code | Analysis Method |
|---|---|
| A0005 | AOAC Official Method 960.39 (1995) |
| A0033 | AOCS Official Method Ce 1-62 (1981) |

New Search

<< Back to Select Specific Analytes

Fig. 3. Example of data output from ILSI Crop Composition Database search.

Table 3
Scope of the data in ILSI Crop Composition Database, version 1.0

| Attribute | Maize | Soybean | Total unique attributes |
|---|---|---|---|
| Years | 1995, 1997–2001 | 1998–2001 | 6 |
| Location | | | |
| North America | United States (12) | United States (6) | 14[a] |
| South America | Argentina | Argentina (3), Brazil (5) | 8 |
| European Union | France, Germany, Hungary, Spain, Italy | | 5 |
| Number of analytes | 96 | 66 | 102[a] |
| Number of datasets | 1137 | 309 | 1446 |
| Number of datapoints | 41,329 | 11,892 | 53,221 |

[a] In some cases the location and analytes for maize and soybean were the same and therefore the total unique attributes is less than the sum of the maize and soybean individual attributes.

Table 4
Comparison of maize and soybean data acquired from ILSI Crop Composition Database, version 1.0 and USDA Nutrient Databases

| Grain/seed | Maize grain | | Soybean seed | |
|---|---|---|---|---|
| Attribute[a] | ILSI | USDA[b] | ILSI | USDA[b] |
| Moisture (% FW) | 11.0 | 10.37 | 9.1 | 8.54 |
| Protein (% FW) | 8.95 | 9.42 | 35.71 | 36.49 |
| Calcium (ppm FW) | 42.6 | 70 | 1941 | 2770 |
| Phosphorus (ppm FW) | 2928 | 2100 | 6381 | 7040 |
| Thiamin (mg/100 g FW) | 0.32 | 0.39 | 0.175 | 0.874 |
| Leucine (mg/g FW) | 11.71 | 11.55 | 27.04 | 29.72 |

[a] Values are expressed on a fresh weight (FW) or "as is" basis for both the ILSI and USDA data.
[b] Data from USDA Nutrient Database (United States Department of Agriculture, Agricultural Research Service, 1999).

## 4. Conclusions

The ILSI Crop Composition Database (www.cropcomposition.org) currently contains a broad scope of compositional data for maize and soybeans that spans 6 years and multiple world-wide geographies. The data are structured in a relational format that permits electronic retrieval of information based upon the interests of the user. The data were generated from samples produced in controlled field trials using validated methodology and the detailed documentation based on the concept of EPA GLP guidelines. It is anticipated that the data will be of importance to a broad range of disciplines including researchers in plant, food, and animal sciences who are interested in understanding the natural variability that results from the interaction of genotype with environment. The lasting value of the database will depend upon periodic updates, the

Table 5
Comparison of representative maize data from ILSI Crop Composition Database, version 1.0 and OECD Consensus Document

| Grain component[a] | OECD range[b] | ILSI range | ILSI average |
|---|---|---|---|
| Proximates (% DW) | | | |
| Protein | 6–12.7 | 6.15–15.01 | 10.06 |
| Fat | 3.1–5.8 | 1.74–5.56 | 3.53 |
| Ash | 1.1–3.9 | 0.62–6.28 | 1.45 |
| Carbohydrate | 82.2–82.9 | 77.4–89.5 | 85.0 |
| Fiber (% DW) | | | |
| NDF | 8.3–11.9 | 5.59–22.64 | 10.99 |
| ADF | 3.0–4.3 | 1.82–11.34 | 3.78 |
| Minerals (mg/100 g DW) | | | |
| Potassium | 320–720 | 283.5–527.6 | 383.5 |
| Calcium | 3–100 | 2.16–20.84 | 4.79 |
| Phosphorus | 234–750 | 208.1–433.1 | 328.3 |
| Magnesium | 82–100 | 78.8–156.0 | 117.3 |
| Amino acids (% DW) | | | |
| Methionine | 0.10–0.46 | 0.130–0.306 | 0.199 |
| Cysteine | 0.08–0.32 | 0.148–0.316 | 0.213 |
| Lysine | 0.05–0.55 | 0.236–0.403 | 0.305 |
| Tryptophan | 0.04–0.13 | 0.0355–0.0900 | 0.0595 |
| Leucine | 0.79–2.41 | 0.642–2.174 | 1.314 |
| Phenylalanine | 0.29–0.64 | 0.263–0.830 | 0.516 |
| Fatty acids (% total FA[c]) | | | |
| 16:0 Palmitic acid | 11.4 | 8.51–15.60 | 10.98 |
| 18:0 Stearic acid | 1.9 | 1.02–2.76 | 1.78 |
| 18:1 Oleic acid | 25.3 | 18.6–39.2 | 25.6 |
| 18:2 Linoleic acid | 60.7 | 46.9–65.6 | 58.1 |
| Vitamins (mg/kg DW) | | | |
| Thiamin | 2.3–8.6 | 1.0–5.0 | 3.0 |
| Riboflavin | 0.25–5.6 | 0.70–1.93 | 1.12 |
| Other (% DW) | | | |
| Phytic acid | 0.45–1.0 | 0.290–1.287 | 0.753 |
| Ferulic acid | 0.02–0.03 | 0.174–0.373 | 0.250 |
| *p*-Coumaric acid | 0.003–0.03 | 0.011–0.058 | 0.027 |

[a] For the units of measure: DW = dry weight, FA = fatty acids.
[b] Data from OECD Consensus Document for maize (Organization of European Cooperation and Development (OECD), 2002).
[c] Values from OECD Consensus Document are expressed as % total fatty acids in refined maize oil.

Table 6
Comparison of representative soybean data from ILSI Crop Composition Database, version 1.0 and OECD Consensus Document

| Grain component[a] | OECD range[b] | ILSI range | ILSI average |
|---|---|---|---|
| Amino acids (% DW)[c] | | | |
| Methionine | 0.5–0.67 | 0.431–0.681 | 0.547 |
| Cysteine | 0.45–0.67 | 0.370–0.808 | 0.584 |
| Lysine | 2.5–2.66 | 2.29–2.84 | 2.54 |
| Tryptophan | 0.51–0.67 | 0.356–0.502 | 0.434 |
| Leucine | 2.2–4.0 | 2.59–3.39 | 3.00 |
| Phenylalanine | 1.6–2.08 | 1.63–2.24 | 1.96 |
| Histidine | 1.0–1.22 | 0.878–1.18 | 1.04 |
| Isoleucine | 1.76–1.98 | 1.56–2.04 | 1.80 |
| Threonine | 1.4–1.89 | 1.25–1.62 | 1.44 |
| Valine | 1.5–2.44 | 1.63–2.20 | 1.90 |
| Isoflavones (mg/100 g FW) | | | |
| Diadzein | 20.2–206 | 5.41–211 | 80.4 |
| Genistein | 31.5–268 | 13.0–244 | 90.8 |
| Glycitein | 10.9–107 | 1.37–28.8 | 14.9 |
| Other bioactives (g/100 g FW) | | | |
| Raffinose | 0.1–0.9 | 0.180–0.611 | 0.320 |
| Stachyose | 1.4–4.1 | 1.07–3.28 | 1.97 |
| Phytic acid[d] | 1–2.74 | 0.634–1.96 | 1.12 |

[a] For the units of measure: DW = dry weight and FW = fresh weight.
[b] Data from OECD Consensus Document for soybean (Organization of European Cooperation and Development(OECD), 2001).
[c] The amino acid values were derived from soybean seed samples with protein minimum, maximum and mean values of 33.2%, 45.5% and 39.3% DW, respectively.
[d] Phytic acid is expressed as g/100 g dry weight in OECD Consensus Document.

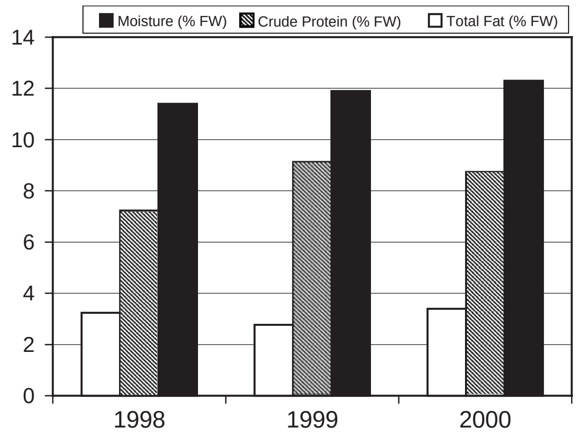


Fig. 4. Maize proximate data from the ILSI Crop Composition Database, version 1.0 for the United States during 1998–2000. For the units of measure: FW = fresh weight.

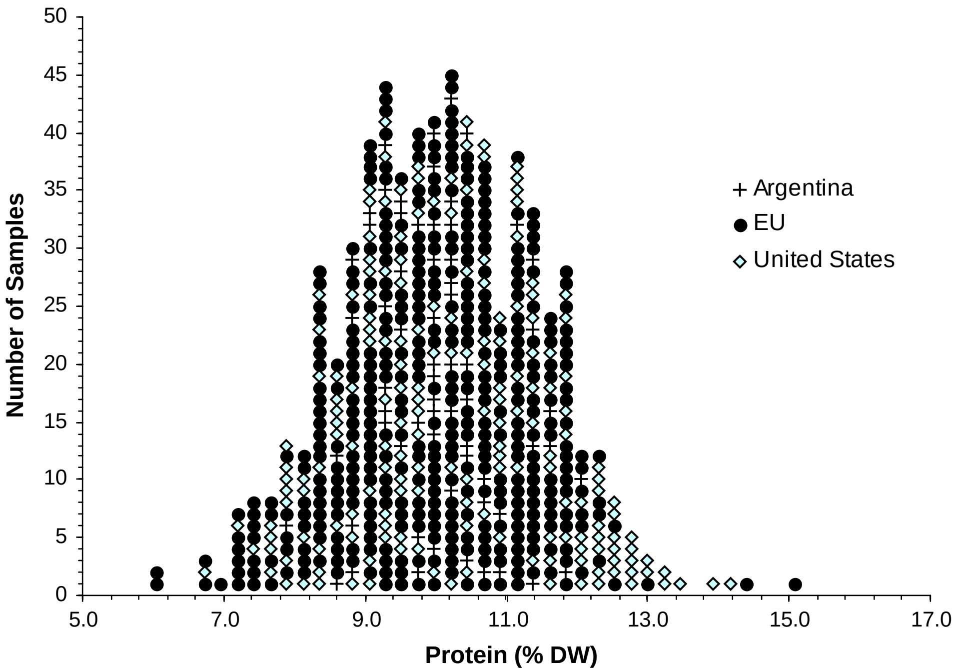


Fig. 5. Distribution of protein values for maize in ILSI Crop Composition Database, version 1.0. The figure represents 651 data points with a minimum, maximum, and mean values of 6.15, 15.01 and 10.06 protein % DW, respectively (see Table 5); DW = dry weight.

addition of other crops such as cotton, canola, and rice and the continued development of the system based upon user needs.

## Acknowledgements

The authors thank Keith Steward and Michael S. Bennett for their help and guidance with the design and development of the software and hardware for the database, Michael King for his work in drafting the initial code and Daniel Suffian for his contributions to defining the Curator's role for the database. The help of Susan Riordan in assembling figures and Melinda McCann in reviewing and editing the text for the screens and formulae for conversion and scaling of units is also gratefully acknowledged.

## References

Burlingame, B., 1996. Development of food composition database management systems: the New Zealand experience. Food Chemistry 57, 127–131.

Ensminger, M.E., Oldfield, J.E., Heinemann, W.W., 1990. Feeds and Nutrition 2nd Edition. The Ensminger Publishing Company, Clovis, CA.

Food and Agriculture Organization of the United Nations, 1996. Biotechnology and food safety. Report of a joint FAO/WHO consultation. In: Food and Nutrition Paper 61, Food and Agriculture Organization of the United Nations, Rome, Italy.

Holden, J.M., Bhagwat, S.A., Patterson, K.Y., 2002. Development of a multi-nutrient data quality evaluation system. Journal of Food Composition and Analysis 15, 339–348.

Jugenheimer, R.W., 1976. Corn: Improvements, Seed Production and Uses. Wiley, New York.

Martineau, B., 2001. First Fruit: The Creation of the Flavr Savr™ Tomato and the Birth of Biotech Food. McGraw-Hill, New York, NY.

Organization for Economic Cooperation and Development (OECD), 1993. Safety Evaluation of Foods Produced by Modern Biotechnology: Concept and Principles. Organization for Economic Cooperation and Development, Paris, France.

Organization of European Cooperation and Development (OECD), 2001. Consensus Document on Compositional Considerations for New Varieties of Soybean: Key Food and Feed Nutrients and Anti-Nutrients. Organization of European Cooperation and Development, Series on the Safety of Novel Foods and Feeds, No. 2, Paris, France.

Organization of European Cooperation and Development (OECD), 2002. Consensus Document on Compositional Considerations for New Varieties of Maize (*Zea mays*): Key Food and Feed Nutrients, Anti-Nutrients and Secondary Plant Metabolites. Organization of European Cooperation and Development, Series on the Safety of Novel Foods and Feeds, No. 6, Paris, France.

Ridley, W.P., 2004. Introduction to agricultural biotechnology: challenges and prospects. In: Bhalgat, M.K., Ridley, W.P., Felsot, A.S., Seiber, J.N. (Eds.), Agricultural Biotechnology: Challenges and Prospects ACS Symposium Series 866. American Chemical Society, Washington, DC, pp. 3–18 (Chapter 1).

Schlotke, F., Becker, W., Ireland, J., Moller, A., Ovaskainen, M.-L., Monspart, J., Unwin, I., 2000. EUROFOODS recommendations for food composition database management and data interchange. Journal of Food Composition and Analysis 13, 709–744.

Southgate, D.A.T., 2002. Data quality in sampling, analysis and compilation. Journal of Food Composition and Analysis 15, 507–513.

United States Department of Agriculture, Agricultural Research Service, 1999. USDA Nutrient Database for Standard Reference, Release 13. Retrieved July 7, 2003 from the Nutrient Data Laboratory Home Page on the World Wide Web, http://www.nal.usda.gov/fnic/foodcomp.

United States Environmental Protection Agency, 1989. Federal Insecticide, Fungicide and Rodenticide Act (FIFRA): Good Laboratory Practice Standards, 40 CFR Part 160. Federal Register, Vol. 54, August 17, pp. 34052–34074.

Watson, S.A., 1982. Corn: amazing maize. General properties. In: Wolff, I.A. (Ed.), CRC Handbook of Processing and Utilization in Agriculture Plant Products, Vol. II (Part 1). CRC Press, Boca Raton, FL.

World Health Organization, 1995. Application of the principles of substantial equivalence to the safety evaluation of foods and food components from plants derived by modern biotechnology. In: Report of WHO Workshop WHO/FNU/FOS/95.1, World Health Organization, Geneva, Switzerland.