



Dr. Bertrand Matthäus

## Editorial

### The database *Seed Oil Fatty Acids* (SOFA) is back on the Internet!

In the 21st century the procurement of information and especially the processing of data is a huge challenge. In science, an inconceivable amount of new information is generated every day, but scientists have more and more difficulties to navigate the plethora of data with the aim to extract the relevant information. Databases are useful tools for collecting information on specific topics and for searching for appropriate information. Such databases are available offline, paper-based or on CD, searchable on an individual computer or online on the Internet. With the help of such databases combined with efficient query engines it is possible to navigate huge volumes of information.

In the plant kingdom about 380 000 different plant species exist on earth. Plant seeds of different species represent an enormous genetic potential in terms of the lipidomic profile from which only a small part is already known. The great variability in the composition of fatty acids, tocopherols, TAGs, phospholipids, sphingolipids and sterols is the basis for different applications in a huge number of scientific disciplines. Examples are the oleochemical industry, which is searching for raw materials with a high proportion of only one specific fatty acid [1, 2], plant breeders, who have to know more about the enzyme systems of plants to which the fatty acid composition can give information [3, 4] or plant chemotaxonomy and phylogenetic botany who want to clarify questions regarding hierarchy and evolutionary relationship of plant species [5, 6].

Since April 2012 the database *Seed Oil Fatty Acids* (SOFA) from the Max Rubner Institut (MRI) has been online again free of charge on the Internet at <http://sofa.mri.bund.de>. After the database had to be removed from the internet in 2008, it has been now replaced by a new database system programmed with the financial support of the German Federal Ministry of Food, Agriculture and Consumer Protection provided by the project sponsor Agency of Renewable Resources (FNR) (FKZ: 08NR144) under the supervision of the MRI.

In comparison to other available databases such as the bibliographical data collection *Nouveau Dictionnaire de Huiles Vegetables – Compositions en acides gras* or the collection of the American Oil Chemists' Society within the *Official Methods and Recommended Practices* with information about the physical and chemical data of about 150 commonly known vegetable oils, the SOFA database contains data about the lipidomic profile of seeds from wild plants which have been collected from the appropriate pharmaceutical, botanical and chemical literature over a period of more than 40 years. Today, SOFA contains about 18 000 tables with about 130 000 individual data of more than 7000 plant species. Additionally to the plant name (genus and species) and the family name, the tables contain at least the information about the oil content, but in most cases also the fatty acid composition is given together with the bibliographical reference from where the information was taken. Many of the tables also contain information about the tocopherol, sterol and TAG composition as well as data about some physical and chemical characteristics such as density, refraction index or unsaponifiables.

Together with the query algorithms SOFA is a comprehensive source of information about the lipidomic profile of plant seeds. It is possible to search for *fatty acids*, *sterols*, *tocopherols* and *TAGs*, *parameters* and *literature*, and in the *Expert form* the user can combine the search for *plants*, *literature* together with *molecules* (fatty acids, tocopherols, sterols or TAGs) and *parameters* (oil content, refraction index, etc.) by Boolean operators.

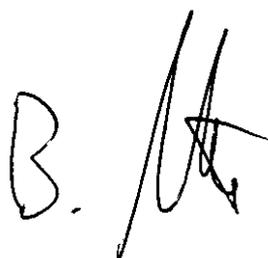
The strongest tool provided by the database is the so-called delta-notation, which allows describing structural elements of fatty acids by a number of alpha-numeric characters. Thus it is possible to use information about the number of carbon atoms and the number, the position and the configuration of double bounds in the fatty acid chain, but also the position of

chemically interesting structure elements such as OH– [OH–], keto– [–O=], cyclopentene [cy], cyclopropane [cpa] and allene [allene] groups or triple bonds [a] to make the search very variable and to allow an almost limitless range of enquiries.

The publication of new data is going on very fast and the information from relevant literature [7–9] will be added by the staff of the MRI to keep the database up-to-date.

Especially the variable search algorithms and the delta-notation make the new database SOFA a very useful and vigorous tool for all lipid scientists who need information about the occurrence of specific fatty acids. I think that scientists from many different disciplines such as chemistry, food chemistry, biochemistry, botany, plant breeding, biology or plant genetics and nutritionists from various institutions will use the database. SOFA will make their work easier and more successful, because without SOFA the huge amount of information about the fatty acid composition of plant seeds would have to be collected from the literature, which is a very time-consuming and labour-intensive task.

I wish all users a successful retrieval of the needed information in the new SOFA database.



Bertrand Matthäus  
Max Rubner-Institut,  
Federal Research Institute for Nutrition and Food,  
Department for Lipid Research, Detmold, Germany

## References

- [1] Meier, M., The oleochemical feedstock wish list. *Eur. J. Lipid Sci. Technol.* 2011, 113, 1297–1298.
- [2] Carlsson, A. S., Lindberg Yilmaz, J., Green, A. G., Stymne, S., Hofvander, P., Replacing fossil oil with fresh oil – with what and for what? *Eur. J. Lipid Sci. Technol.* 2011, 113, 812–813.
- [3] Abbadi, A., Leckband, G., Rapeseed breeding for oil content, quality, and sustainability. *Eur. J. Lipid Sci. Technol.* 2011, 113, 1198–1206.
- [4] Cahoon, E. B., Kinney, A. J., The production of vegetable oils with novel properties: Using genomic tools to probe and manipulate plant fatty acid metabolism. *Eur. J. Lipid Sci. Technol.* 2005, 107, 239–243.
- [5] Aitzetmüller, K., Agonandra seed oil – An excellent source of a truly multifunctional fatty acid. *Fett/Lipid* 1998, 100, 21–23.
- [6] Spitzer, V., Screening analysis of unknown seed oils. *Eur. J. Lipid Sci. Technol.* 1999, 101, 2–19.
- [7] Ali, S., Ahmet, U., Zeki, K., Fatty acid and tocopherol constituents of the seed oil extracted from 21 grape varieties (*Vitis* spp.). *J. Sci. Food Agric.* 2012, 92, 1982–1987.
- [8] Hussain, I., Khan, S., Khan, M. I., Rehman, I. U., Ahmed, M., Investigation of fatty acid composition of ammi majus seed oil by gas chromatography mass spectrometry. *J. Chin. Chem. Soc.* 2012, 59, 655–658.
- [9] Manzoor, M., Anwar, F., Ashraf, M., Alkharfy, K. M., Physico-chemical characteristics of seed oils extracted from different apricot (*Prunus armeniaca* L.) varieties from Pakistan. *Grasas y Aceites.* 2012, 63, 193–201.