# MoTo DB: A Metabolic Database for Tomato

An organism's metabolome is the cataloguing of the entire set of metabolites found in the organism. It is dynamic and thus just a snapshot of the processes going on in the organism at one specific time point. When combined with other ''omics'' data, a more complete understanding of the organism emerges (Fiehn et al., 2000). However, several technical challenges to obtaining a metabolomic profile of an organism exist, as no single technique is capable of detecting all of the metabolites. Thus, a variety of analytical techniques must be used due to the large variety of compounds that make up a metabolome. Common metabolite databases need to be established to facilitate the identification of the compounds. In 2006, Moco et al. published an article in *Plant Physiology* titled ''A Liquid Chromatography-Mass Spectrometry-Based Metabolome Database for Tomato'' describing a new database. The authors validated the database by comparing tomato (*Solanum lycopersicum*) fruit peel and flesh tissues.

## BACKGROUND

The diverse chemical nature of metabolites necessitates the use of multiple analytical techniques to obtain a profile of the compounds present in a plant. Liquid chromatography (LC; in particular, reverse phase is suitable for detection of semipolar metabolites) and gas chromatography (GC; analysis of aromatics and other volatile compounds as well as those that can be chemically volatilized) provide a way to distinguish between compounds biochemically, but many compounds have overlapping detection signals and are not able to be conclusively identified. The combining of these separation techniques with a second analytical technique, such as mass spectrometry (MS), which, depending on the machine, allows mass estimation. Use of NMR provides the most detailed information on a compound's chemical structure, allowing an absolute estimation. The identification of the metabolites is accompanied by comparing spectral data with known, authenticated compounds. GC-MS is highly reproducible, and chromatographic traces and mass fragmentation patterns can readily be compared. However, this is not the case with LC-MS, since the spectra can be instrument dependent.

There are still many challenges to obtaining a complete metabolic profile of a plant. The plant metabolome (indeed, any metabolome) is not complete. There are still many metabolites that have not been identified, and of those that have, many need to be positively identified, as some are actually products of the isolation process and not present in the biological sample. As mentioned above, publicly available databases need to be developed and added to. A challenge to this is the wide range of analytical techniques, with different resolutions, used to isolate metabolites. This is especially true for LC-MS data. Moco et al. (2006) have started to address this challenge in the creation of an LC-MS database of tomato metabolites.

## WHAT WAS SHOWN

The Metabolome Tomato Database (MoTo DB) is an open-access metabolome database for tomato fruit. The database was developed with LC-MS data and primarily contains semipolar metabolites. MoTo DB also contains metabolite information from the literature and includes a variety of tomato fruit extracts from cultivated, wild, and transgenic tomato varieties. As this is LC-MS, a highly reproducible system for large-scale detection and identification needed to be first developed. The authors utilized a system of reverse-phase LC coupled to quadrupole time-of-flight mass spectrophotometry (QTOF-MS) and photodiode array detection (PDA).

Use of fruits from 96 different tomato cultivars in different stages of ripening ensured a representative fruit sample. The selected cultivars produced red- or orange-colored beef-, round-, or cherry-type fruits along with some purple-skinned fruits for anthocyanin analysis. Peels were used since they contain the highest amount of flavonoids, an important class of natural products (secondary metabolites) in tomato fruit. To extract the mass signals and alignment of chromatograms, the software metAlign (Vorst et al., 2005; De Vos et al., 2007) was used. Reproducibility of the LC system was determined by comparing retention time shifts from chromatograms of tomato fruits analyzed over a period of 2 years with a very small variation between the retention times. Both positive and negative ionization modes were tested for the tomato extracts, and the results were compared.

To search the database for a compound of interest, the mass is entered, either as a range or an accurate value, along with the ionization mode. The mass accuracy can also be set to accommodate data from a variety of detectors (low or high accuracy). The results pane displays the projected formula, a link to any PubChem identifiers, and CAS number as well as a link to the retention time, calculated accurate mass, MS/MS fragments, UV/visible spectral information, and alternative names for the compound (if any) and literature references. All of the ''hits'' in the database have been validated by experimental evidence and represent either a compound previously found in the literature or a novel, vetted compound. Unknown or

www.plantphysiol.org/cgi/doi/10.1104/pp.109.900308

novel compounds that have not been experimentally confirmed have not been included in the database.

The database was tested with peel and flesh tissues from red, ripe tomato fruits analyzed by LC-PDA-electrospray ionization (ESI)-QTOF-MS in both positive and negative ionization modes. Many unknown metabolites still exist in the flesh. The results confirmed earlier findings that flavonoids are found in the peel and not at detectable levels in the flesh.

## THE IMPACT

The areas of food science and nutrition research have much to gain from the emerging field of metabolomics. The ability to monitor food components, quality, and authenticity are some of the applications of this technology (Wishart, 2008). Using both biochemical and metabolomic approaches, Capanoglu et al. (2008) examined changes in the metabolite profile of tomatoes during processing. Tomatoes and tomato products are an important part of many diets, providing both vitamins and antioxidants. In addition to consumption of fresh fruits, tomatoes are typically processed and cooked, potentially leading to a degradation of the vitamins and antioxidants. This degradation was investigated in this study by analyzing samples collected during processing from five independent tomato paste productions over a 2-year period. An LC-QTOF-MS-based system was used for the metabolomics aspect, and the MoTo DB (Moco et al., 2006) was used to identify the mass signals. As expected, each processing step changes the metabolic profile of the product, and most significant change occurred when the seeds and skins were removed and the fruit-breaking step. These finding prompted the authors to suggest a separate extraction of the seeds and skins and using this to supplement the final product, tomato paste, so it will have a similar metabolite profile as the unpulped fruits. An alternative approach would be to make changes to the pulping process in order to promote a more efficient extraction of nutritive compounds.

## CONCLUSION

Metabolomics has the potential for many applications, from a comparison between an ecotype or species to examination of the effects of a stress upon a plant and, as mentioned above, to examining the effects of processing on the nutritional value of foods. This is still an emerging field with many challenges to be met. For metabolic data to be most useful and informative, it needs to be combined with other functional genomic data (Bino et al., 2004). A start in this direction is the creation of high-quality, online databases that are publicly available, such as MoTo DB (Moco et al., 2006).

## LITERATURE CITED

**Bino RJ, Hall RD, Fiehn O, Kopka J, Saito K, Draper J, Nikolau BJ, Mendes P, Roessner-Tunali U, Beale MH, et al** (2004) Potential of metabolomics as a functional genomics tool. Trends Plant Sci **9:** 418–425

**Capanoglu E, Beekwilder J, Boyacioglu D, Hall R, de Vos R** (2008) Changes in antioxidant and metabolite profiles during production of tomato paste. J Agric Food Chem **56:** 964–973

**De Vos RCH, Moco S, Lommen A, Keurentjes JJB, Bino RJ, Hall RD** (2007) Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. Nat Protoc **2:** 778–791

**Fiehn O, Kopka J, Dormann P, Altmann T, Trethewey RN, Willmitzer L** (2000) Metabolite profiling for plant functional genomics. Nat Biotechnol **18:** 1157–1161

**Moco S, Bino RJ, Vorst O, Verhoeven HA, de Groot J, van Beek TA, Vervoort J, de Vos CHR** (2006) A liquid chromatography-mass spectrometry-based metabolome database for tomato. Plant Physiol **141:** 1205–1218

**Vorst O, de Vos CHR, Lommen A, Staps RV, Visser RGF, Bino RJ, Hall RD** (2005) A non-directed approach to the differential analysis of multiple LC-MS derived metabolic profiles. Metabolomics **1:** 169–180

**Wishart DS** (2008) Metabolomics: applications to food science and nutrition research. Trends Food Sci Technol **19:** 482–493

**Aleel K. Grennan**
**University of Illinois**
**Urbana, IL 61801**