

Non-targeted comparative metabolomics of leaf samples

Leaf tissue of each biological replicate is ground with a mortar and pestle under liquid nitrogen to a fine frozen powder. Frozen powder (60 mg fresh weight \pm 10% tolerance) is extracted with methanol/chloroform (v/v: 70/30). The fraction of polar metabolites is prepared by liquid partitioning into a water/methanol phase. Aliquots of the polar phase are dried and chemically derived by sequential methoxyamination and trimethyl silylation. GC/EI-TOF-MS profiling is performed using an Agilent 6890N24 gas chromatograph with split or splitless injection and electronic pressure control mounted to a Pegasus III TOF mass spectrometer. We will follow the GC/EI-TOF-MS based profiling method reported by Wagner et al. (2003) and Erban et al. (2007). Retention indices are calibrated by the addition of a C10, C12, C15, C18, C19, C22, C28, C32 and C36 n-alkane mixture to each sample according to Strehmel et al. (2008).

Acquired GC/EI-TOF-MS chromatograms are visually controlled, baseline corrected and exported in NetCDF file format using ChromaTOF software. The GC-MS data processing into a standardized numerical data matrix and compound identification is performed using the TagFinder software (Luedemann et al. 2008). Compounds are identified by mass spectral and retention time index matching to the reference collection of the Golm metabolome database (GMD, <http://gmd.mpimp-golm.mpg.de/>; Kopka et al. 2005; Schauer et al. 2005; Hummel et al. 2010) and to the mass spectra of the NIST08 database (<http://www.nist.gov/srd/mslist.htm>). All spectrometry mass features are normalized by sample fresh weight, internal standard and maximum scaled (e.g., Sanchez et al. 2008a). Laboratory and reagent contaminations are tested and eliminated by non-sample control experiments. Metabolite pool sizes are routinely assessed by relative changes expressed as response ratios, that is, x-fold factors (+ or -) in comparison to a control condition (E-). Mass spectra usually reveal metabolites, which are identifiable by standard addition of authentic reference substances, but also yet non-identified mass spectral tags (MSTs) (Desbrosses et al. 2005; Sanchez et al. 2012). To visualize sample classifications and the within-replicate variation, principal component analysis (PCA) is performed on the complete data set without prior knowledge of metabolite identity. Statistical assessments of data and visualizations are performed using the multi-experiment viewer software, MeV (Version 4.6.2; <http://www.tm4.org/mev/>; Saeed et al. 2003, 2006), or the SPSS program. Probabilistic principal component analyses (PPCA) for data reduction and outlier detection is carried out according to Scholz et al. (2005) via the MetaGeneAlyse web application (Version 1.7.1; <http://metagenealyse.mpimp-golm.mpg.de>) without missing value substitution (Sanchez et al. 2008b).

References

- Erban, A., N. Schauer, A.R. Fernie, and J. Kopka. 2007. Nonsupervised construction and application of mass spectral and retention time index libraries from time-of-flight GC-MS metabolite profiles. *Methods in Molecular Biology* 358: 19–38.
- Desbrosses, G.G., J. Kopka, and M.K. Udvardi. 2005. Lotus japonicus metabolic profiling. Development of gas chromatography-mass spectrometry resources for the study of plant-microbe interactions. *Plant Physiology* 137: 1302-1318.

Luedemann, A., K. Strassburg, A. Erban, and J. Kopka. 2008. Tag-finder for the quantitative analysis of gas chromatography-mass spectrometry (GC-MS) based metabolite profiling experiments. *Bioinformatics* 24: 732–737.

Saeed, A.I., V. Sharov, J. White, et al. 2003. TM4: a free, opensource system for microarray data management and analysis. *Biotechniques* 34: 374–378.

Saeed, A.I., N.K. Bhagabati, J.C. Braisted, W. Liang, V. Sharov, E.A. Howe, J. Li, M. Thiagarajan, J.A. White, and J. Quackenbush. 2006. TM4 microarray software suite. *Methods in Enzymology* 411: 134–193.

Sanchez, D.H., H. Redestig, U. Krämer, M.K. Udvardi, and J. Kopka. 2008b. Metabolome-ionome-biomass interactions: what can we learn about salt stress by multiparallel phenotyping? *Plant Signaling & Behavior* 3: 598–600.

Sanchez, D.H., F. Schwabe, A. Erban, M.K. Udvardi, and J. Kopka. 2012. Comparative metabolomics of drought acclimation in model and forage legumes *Plant, Cell and Environment* 35:136–149

Scholz, M., F. Kaplan, C.L. Guy, J. Kopka, and J. Selbig. 2005. Nonlinear PCA: a missing data approach. *Bioinformatics* 21: 3887–3895.

Strehmel N., J. Hummel, A. Erban, K. Strassburg, and J. Kopka. 2008. Retention index thresholds for compound matching in GC-MS metabolite profiling. *Journal of Chromatography B* 871: 182–190.

Wagner, C., M. Sefkow, J. Kopka. 2003. Construction and application of a mass spectral and retention time index database generated from plant GC/EI-TOF-MS metabolite profiles. *Phytochemistry* 62: 887–900.