

## **Extraction, Derivatization, and Analysis of Seed Metabolites Using GC-TOF-MS**

The extraction method is modified from the method previously described by (Roessner et al., 2000). Approximately a bulk of approximately 70-100 seeds (30mg) seeds were homogenized in 2 ml tubes with 2 iron balls (2.5mm), pre-cooled in liquid nitrogen. For the homogenization the micro-dismembrator (Sartorius) is used at 1500 rpm. 700µl methanol/chloroform (4:3) was added together with the standard (0.2mg/ml ribitol) and mixed thoroughly. After 10 minutes sonication, 200µl MQ was added to the mixture followed by vortexing and centrifuging (5 mins 13,500rpm). Methanol phase was collected in a glass vial. 500µl methanol/chloroform was added to the remaining organic phase and kept on ice for 10 minutes. 200µl MQ was added followed by vortexing and centrifuging (5 mins 13,500rpm). Again, the methanol phase was collected and mixed with the other collected phase. 100µl was dried overnight in a speedvac (35°C Savant SPD121). The GC-TOF-MS method was previously described by (Carreno-Quintero et al., 2012) with some minor modifications. Detector voltage was set at 1600V. Raw data was processed using the chromaTOF software 2.0 (leco instruments) and further processed using the Metalign software (Lommen, 2009), to extract and align the mass signals. A signal-to-noise ratio of 2 was used. The output was further processed by the Metalign Output Transformer (METOT; Plant Research International, Wageningen) and the mass signals that were present in less than 3 RILs were discarded. Out of all the mass signals, centrotypes are formed using the MSclust program (Tikunov et al., 2011). This resulted in 160 unique centrotypes (representative masses). The mass spectra of these centrotypes were used for identification by matching to an in-house constructed library, the NIST05 (National Institute of Standards and Technology, Gaithersburg, MD, USA; <http://www.nist.gov/srd/mslist.htm>) and Golm libraries (<http://csbdb.mpimp-golm.mpg.de/csbdb/gmd/gmd.html>). This identification is based on spectra similarity and comparison with retention indices calculated by using a 3rd order polynomial function (Strehmel et al., 2008).