Description of Additional Supplementary Files

File: Supplementary Data 1
Description: Taxa selected in step 1 for the construction of the sub-databases.

File: Supplementary Data 2
Description: Results of database searches for each step of the pipeline.

File: Supplementary Data 3
Description: List of identified genus and their relative abundances based on the number of total and specific TSMs, and number of specific peptides.

File: Supplementary Data 4
Description: List of protein groups identified in Step 3.

File: Supplementary Data 5
Description: List of GOslim annotations and their relative abundances.

File: Supplementary Data 6
Description: List of genus associated with biosynthetic process and cellular nitrogen compound metabolic process

File: Supplementary Data 7
Description: Univariate non-parametric statistics performed on the discriminant features from sPLS-DA.