**Abstract**

- Condenser is a novel computational tool for aggregating and merging quantitative xml outputs from the Matrixscience Distiller work package into a common format ready for subsequent bioinformatics investigations.
- Condenser allows the quantitative information extracted from the individual sample files using Mascot Distiller into an SQL database. This allows peptides of the same protein accession number in separate samples to be aggregated to a condensed protein list with accurate protein quantifications.

**Introduction to Condenser**

Isotopic mass tag methods have allowed mass spectrometry-based proteomics to overcome problems with variable analyte ionization efficiency and turn quantitative. The nature of quantitative information in mass spectrometry generally requires a detailed evaluation of profile spectra and chromatograms not necessary for database search algorithms that essentially just require peak lists with m/z values and intensities. The development of such tools has been slow and fragmented, with most tools aimed at specific workflows or restricted to certain instrument vendor platforms.

Subsequently, large parts of the community have had limited access to robust computational and statistical tools for quantitative experiments. Matrix Sciences’ Mascot Distiller support multiple labelling approaches and data formats, while being user friendly, however, an unfortunate drawback of Distiller is the lack of sufficient multidimensional experiment support. Condenser appends this functionality to the Distiller workflow, allowing proteins and peptides originating from sequential MS analyses (gel bands, SCX fractions etc.) to be aggregated into combined lists, with statistical evaluations and re-calculation of protein abundance ratios. Loading data into a local database, Condenser processing the data using tailored queries, which allows significant freedom in generating output tables compatible with downstream applications such as repositories, pathway analyses or Gene Ontology annotation tools. However, the solution presented here requires access to a workstation with Mascot Distiller.

**Conclusion.**

- The SQL-based proteome data handling platform that we here present, solves the lack of multisample statistical analysis support in Matrixscience Mascot server and Distiller workpackage.
- Our solution presented here do require access to a workstation with Matrixscience Mascot Distiller v. 2.3 or 2.4

**Key functions of Condenser**

1. The processed data can be exported in a single step into a tabulated Excel format for easy visualization and data mining.
   - Ratio vs. Intensity plots can visualize data distribution
   - p-value vs. ratio plots can visualize significance at each ratio
   - Density plots can visualize protein distribution

2. Raw data xml files from Distiller are imported, combined and re-refined based on user defined thresholds for the spectral quality values known from Distiller (A). The user has several options for choosing appropriate processing of the combined peptide list, including inclusion/exclusion of shared peptides (B), global normalization (C), outlier removal (D) and choice of significance threshold for t-test statistical calculations (F). In addition, the user can choose to apply a Low-False-Filter (F) to correct for Distiller's tendency to discard stable isotopic label peptide doublets when differences in relative abundance exceed a certain magnitude.

3. Windows 32 and 64 bit GUI

4. The effect of methods used for outlier removal and the effect of the stringency chosen by the user.

5. How did the data get from A to D? Condenser shows the data tables for each step in the user defined processing from the raw data in the imported Distiller units to the final spreadsheet-ready protein table.

6. The user can create a project specific exclusion list to define experimental protein contaminants for subsequent omission.

**References & Acknowledgements**


Acknowledgements: The authors are not affiliated in any way with Matrixscience but appreciate the versatility and transparency of the Distiller workpackage.