Introduction
It has recently been shown that targeted analysis of data independently acquired (DIA) peptide fragment ion spectra (SWATH-MS) has the potential to combine the scale of shotgun proteomics with the analytical accuracy of MRM [1]. We have developed the software Spectronaut for analysis of high content datasets which are typically produced in DIA mode. Spectronaut adapts algorithms of mProphet [2] providing fast and efficient processing of DIA data and accurate error statistics.

Methods
The software Spectronaut is programmed in C#. Spectronaut processes DIA data using a targeting strategy similar to data analysis of MRM. In addition to mProphet scoring [2], Spectronaut uses retention time prediction based on iRT [3], mass accuracy, and isotopic distribution of fragment ions. iRT based retention time prediction allows slicing of the data into small bits. This provides high performance of processing and increases the specificity because it effectively blinds out distracting signals. False discovery rates are determined using an adapted decoy model similar to the model used by mProphet.

Total yeast digests were measured on a TripleTOF 5600 mass spectrometer (ABSciex, Toronto) coupled to an Eksigent Ultra-NanoLC System.

Figure 1. Screenshot of Spectronaut in the review perspective. Peptides targeted in a yeast sample are visualized analogously to MRM data. Ion traces are extracted from fragment ion spectra using a narrow extraction window according to the resolution of the mass spectrometer. Scores, such as the intensity correlation score shown in the lower right panel, are calculated similarly to mProphet [2].

Figure 2. Four ion traces corresponding to four fragment ions of the triply charged peptide GLVWEGSDLDEEGIR targeted in a yeast sample are shown. The time range corresponding to the complete LC-gradient is displayed. In the zoom-in one can see the targeted peptide in the retention time window from 116 to 126 minutes as predicted using the IRT of the peptide [3].

Figure 3. Maximal memory consumption and computation time needed for the targeted analysis of 1982 proteins, 11594 peptides or 114'862 transitions in one SWATH-MS run as a function of the retention time extraction window. The center of the extraction window is predicted using iRT which was empirically pre-determined for all peptides. Memory consumption and computation time decrease roughly linear with the size of the extraction window.

Figure 4. Number of peptides identified as a function of the retention time extraction window. Prediction of the extraction window center is based on pre-determined iRTs or on an adapted SSRCalc algorithm. iRT shows a better performance than SSRCalc over the complete range of window sizes. The optimal extraction window size in this data set is 10 minutes for iRT and 80 minutes for the adapted SSRCalc. Windows smaller than the error range of prediction result in decreasing performance because some peptides elute outside the extraction window.

References