TECHNICAL BRIEF

Visualization of LC-MS/MS proteomics data in MaxQuant

Stefka Tyanova[1], Tikira Temu[1], Arthur Carlson[1], Pavel Sinitcyn[1], Matthias Mann[2] and Juergen Cox[1]

[1] Max-Planck-Institute of Biochemistry, Computational Systems Biochemistry, Martinsried, Germany
[2] Max-Planck-Institute of Biochemistry, Proteomics and Signal Transduction, Martinsried, Germany

Modern software platforms enable the analysis of shotgun proteomics data in an automated fashion resulting in high quality identification and quantification results. Additional understanding of the underlying data can be gained with the help of advanced visualization tools that allow for easy navigation through large LC-MS/MS datasets potentially consisting of terabytes of raw data. The updated MaxQuant version has a map navigation component that steers the users through mass and retention time-dependent mass spectrometric signals. It can be used to monitor a peptide feature used in label-free quantification over many LC-MS runs and visualize it with advanced 3D graphic models. An expert annotation system aids the interpretation of the MS/MS spectra used for the identification of these peptide features.

Received: September 19, 2014
Revised: December 12, 2014
Accepted: January 28, 2015

Keywords:
Bioinformatics / Mass spectrometry / LC-MS/MS / Visualization

MaxQuant is a widely used software platform for the analysis of shotgun proteomics data [1, 2]. From version 1.5 on, MaxQuant has been distributed with a unified data exploration component, termed the Viewer, which enables integrated visual inspection of the mass spectrometric raw data together with results from the identification and quantification pipeline. It provides navigation through arbitrarily large datasets by efficient indexing of the underlying data structures and on-demand loading of the required raw data. MaxQuant, including the integrated Viewer described here, is programmed in C# using the .NET framework 4.5. MaxQuant uses the Windows Ribbon Framework to achieve easy navigation and quick access to the graphical user interface. It is freely available and can be downloaded from www.maxquant.org. MaxQuant can read raw data in native vendor formats from Thermo Fisher Scientific, Bruker Daltonics, and AB/Sciex as well as the open mzML format. Concise information on "Getting started," including software requirements, trouble shooting, and a test dataset, is available online (http://141.61.102.17/maxquant_doku/doku.php?id=maxquant:viewer).

Correspondence: Dr. Juergen Cox, Max-Planck-Institute of Biochemistry, Computational Systems Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany
E-mail: cox@biochem.mpg.de
Fax: +49-89-8578-2219

Figure 1A displays the fully redesigned graphical user interface of MaxQuant. Organization of main components in tabs allows for easy navigation in the main level control of MaxQuant. Here, we are interested in the "Viewer" tab. Below the tab selector is the command ribbon, which hosts multiple buttons and other control elements acting on the main visual components beneath it. The main display is subdivided into four parts. The upper left component is the map view, which displays the mass spectrometric color-coded intensities, typically at the MS1 level, in the m/z-retention time plane. Peak boundaries are displayed in different coloring schemes, which indicate the grouping of peaks into isotope patterns or of isotope patterns into labeling pairs or triplets. For instance, the peak shapes in Fig. 1A are colored according to the isotope pattern to which they belong. The blue indicator rectangle encloses a peptide whose 3D peak boundaries are shown in blue. There are eight isotopic peaks found for this peptide in this label-free sample. The flat red rectangle indicates the region in the m/z-retention time plane in which ions were collected for the fragmentation spectrum that was recorded in order to identify the peptide colored in blue. By visual inspection one can conclude that no major cofragmentation of other peptide species is expected in this case. The proportion of cofragmented ions is also

Colour Online: See the article online to view Figs. 1 and 2 in colour.

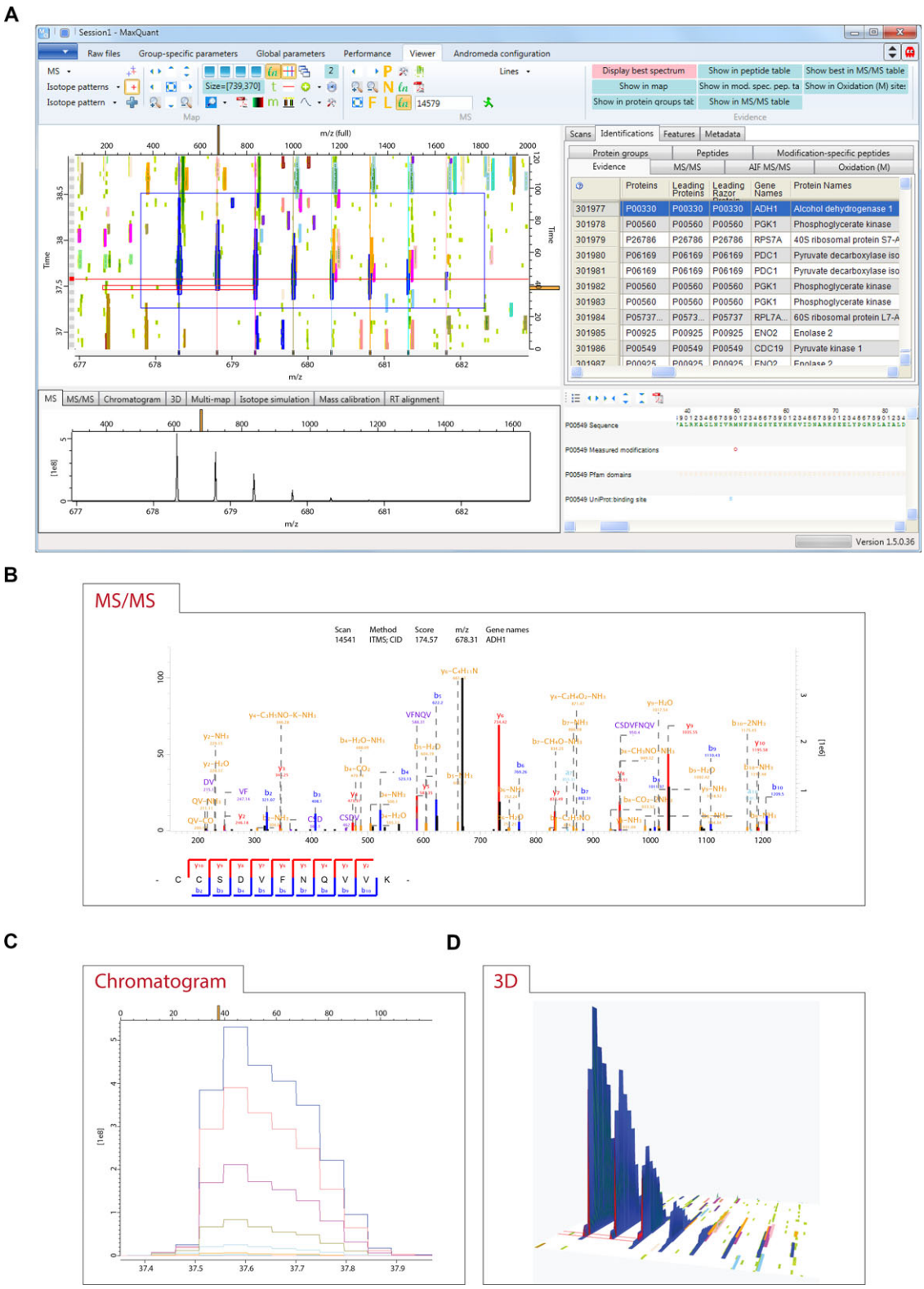

Figure 1. Overview of the MaxQuant Viewer tab. (A) Upon selection of a feature of interest from the evidence table, the MS intensities of the feature (color coded by isotope pattern in blue) are displayed in the *m/z*-retention time map as indicated by the blue rectangle. Underneath the map the MS mode of the feature view is shown. Various modes of the feature view are represented: (B) MS/MS of the selected feature with advanced annotation enabled and display of the peptide sequence; (C) chromatogram with mass traces enabled; (D) 3D view of the isotope peaks.

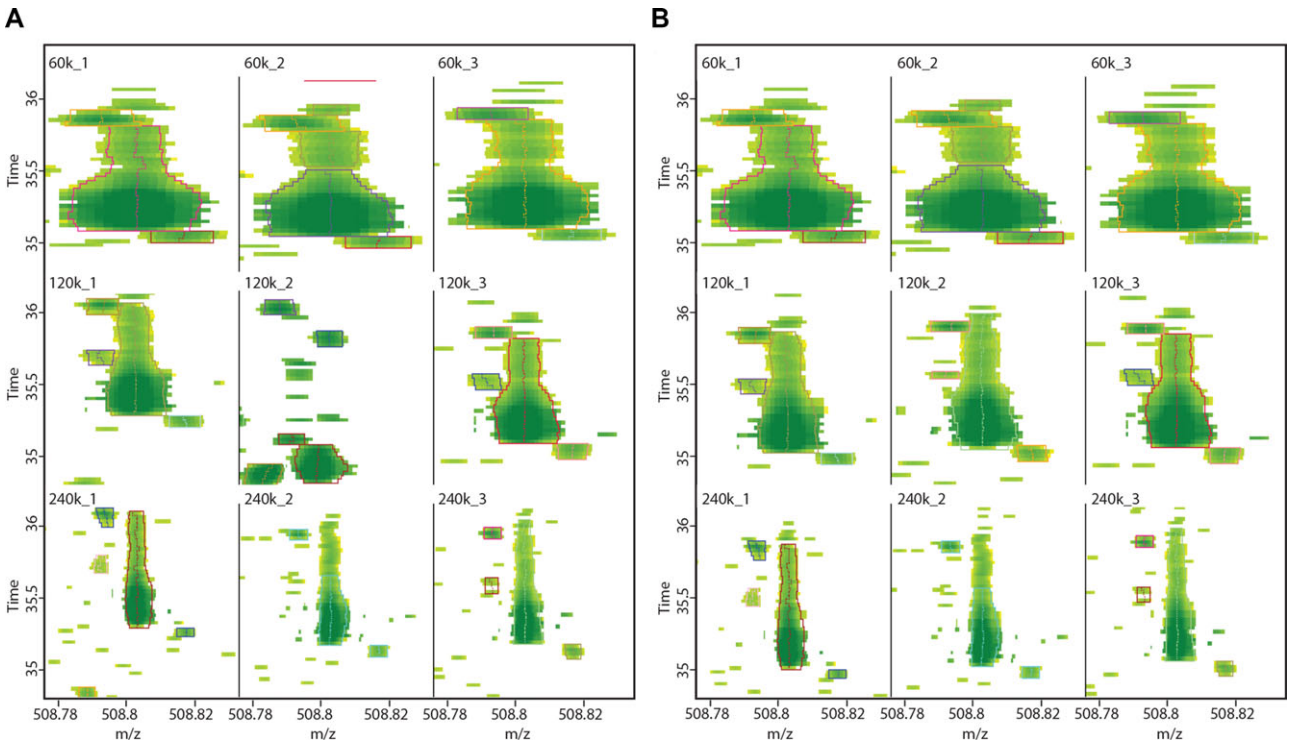

Figure 2. Multi-map view. A view of the same feature in different experiments: three mass resolutions with three replicates for each resolution are shown: (A) without retention time alignment and (B) with alignment.

automatically quantified in MaxQuant for every fragmentation event [3]. The map view can be applied as well to MS/MS data that is continuous in time, as is produced, for instance, during data independent acquisition. For "all ion fragmentation" data [4] MaxQuant can determine 3D peaks as well, which will be displayed in the Viewer in the same way as MS1 level peaks. The map view is fast and responsive because it reads the minimal amount of data from the raw file that is needed for determining the colors of the pixels in the specific zoom level.

In the lower left corner is a subordinate tab document that hosts several detailed displays relating to the content in the current zoom window of the map display. In Fig. 1A the MS spectrum display is visible, which shows the mass spectrum at the horizontal cross-section as indicated by the horizontal red line in the map view. The display can rapidly be moved through consecutive scans with the help of the up-down arrow keys or positioned on a specific retention time by clicking on the left border of the map display. This is useful, for instance, for inspecting the isotopic intensity distribution at the maximum of the elution profile of a peptide.

Another display item is the annotated MS/MS spectrum shown in Fig. 1B that identifies the peptide whose MS1 features are displayed in Fig. 1A. Annotations can include only the main series fragments plus water and ammonia losses, or alternatively an extensive set of peak annotations generated by an expert system [5], which includes internal fragments and a more extensive set of potential neutral losses. The annotated peptide sequence is displayed as well, indicating which main series peaks are identified along the sequence.

The tab "Chromatogram" (Fig. 1C) shows either the time dependence of the total ion current in the given retention time zoom window or similarly multiple intensity profiles in selected mass channels. For that purpose the m/z values are selected in the lower border of the map view. They can also be set automatically to all m/z values occurring in an isotope pattern or in a label pair or triplet, which is particularly useful for inspection of retention time shifts between different labeled versions of a peptide, which can be an issue of concern for particular labeling techniques [6].

Figure 1D shows the 3D visualization of the m/z-retention time area that is indicated by the blue rectangle in Fig. 1A. The 3D rendering is done with the 3D graphics capabilities of Windows Presentation Foundation, which is part of the .NET Framework 4.5 and therefore readily available on every windows computer. For comparison, other displays in this part of the software include an isotope peak simulation display where for a peptide molecule with known elemental composition the intensity profile is calculated with a desired resolution for the purpose of comparison to measured isotopic envelopes. Also the retention time calibration curves resulting from the nonlinear retention time alignment algorithm in MaxQuant can be displayed here at the current zoom levels.

In the upper right area all MaxQuant output tables can be browsed. The information content is the same as in the tab-separated output tables that can be found in the folder "combined\txt" in the processed MaxQuant project. These tables contain all results regarding identification and quantification of peptides and proteins where each table corresponds to a different organization level of the data. The most important of these tables are the protein groups, peptides, modification-specific peptides, evidence, and MS/MS tables. All these tables are interconnected among each other as well as with the displays on the left side. For instance all MS/MS spectra identifying a particular protein of interest can be selected in the corresponding table. The annotated MS/MS spectrum as well as the surroundings of the precursor in the MS1 plane are continuously updated when selecting a row in the MS/MS table. Publication-ready annotated MS/MS spectra can be exported in vector graphics format for any subset of spectra of interest by a single button click.

The lower right area of the main display contains the protein sequence view. For each protein group, sequences of all members are displayed in which the identified peptides are indicated and color-coded according to their uniqueness regarding their occurrence in the protein database. Additional annotation tracks can be displayed along the sequence. As a particularly interesting example, one can display here the PTMs identified in the MaxQuant project in one track and show in another the already known PTMs from a central repository, such as PhosphoSitePlus [7] or MaxQB [8]. A multitude of different sequence-specific annotations as derived from UniProt [9] as well as the Pfam domain structure [10] can be displayed as tracks.

Finally, a very useful feature is the multi-map view (Fig. 2). Here one can monitor a selection region in the *m/z*-retention time plane across many LC-MS runs. In Fig. 2A this can be seen for a particular peak that is viewed over nine different runs. These runs are triplicate groups of measurements of the same biological sample at three different mass resolutions. It is apparent that the retention time of the peak center varies appreciably between the runs, and in the central panel the peak of interest is not even present in the view area. The main map as well as the small maps in the multi-map view can also be displayed with the recalibrated retention time as the vertical axis that is shown in Fig. 2B. Here, the peak occurs at the same retention time in all nine samples. The multi-map view can verify that a particular feature used by the MaxLFQ algorithm for label-free quantification [11] is well aligned across the LC-MS runs involved and that it fulfils the criteria for the feature matching algorithm.

In summary, the Viewer component of MaxQuant has been thoroughly updated and now fulfils the demands of rich content visualization of high resolution proteomics data.

The authors have declared no conflict of interest.

References

[1] Cox, J., Mann, M., MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 2008, *26*, 1367–1372.

[2] Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. A. et al., Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* 2011, *10*, 1794–1805.

[3] Michalski, A., Cox, J., Mann, M., More than 100000 detectable peptide species elute in single shotgun proteomics runs but the majority is inaccessible to data-dependent LC-MS/MS. *J. Proteome Res.* 2011, *10*, 1785–1793.

[4] Geiger, T., Cox, J., Mann, M., Proteomics on an Orbitrap benchtop mass spectrometer using all-ion fragmentation. *Mol. Cell. Proteomics* 2010, *9*, 2252–2261.

[5] Neuhauser, N., Michalski, A., Cox, J., Mann, M., Expert system for computer-assisted annotation of MS/MS spectra. *Mol. Cell. Proteomics* 2012, *11*, 1500–1509.

[6] Boersema, P. J., Raijmakers, R., Lemeer, S., Mohammed, S., Heck, A. J., Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nat. Protocols* 2009, *4*, 484–494.

[7] Hornbeck, P. V., Kornhauser, J. M., Tkachev, S., Zhang, B. et al., PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res.* 2012, *40*, D261–D270.

[8] Schaab, C., Geiger, T., Stoehr, G., Cox, J., Mann, M., Analysis of high accuracy, quantitative proteomics data in the MaxQB database. *Mol. Cell. Proteomics* 2012, *11*, M111 014068.

[9] UniProt Consortium, Activities at the universal protein resource (UniProt). *Nucleic Acids Res.* 2014, *42*, D191–D198.

[10] Finn, R. D., Bateman, A., Clements, J., Coggill, P. et al., Pfam: the protein families database. *Nucleic Acids Res.* 2014, *42*, D222–D230.

[11] Cox, J., Hein, M. Y., Luber, C. A., Paron, I. et al., Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell. Proteomics* 2014, *13*, 2513–2526.

