

A public repository for mass spectrometry imaging data

Andreas Römpp · Rui Wang · Juan Pablo Albar · Andrea Urbani ·
Henning Hermjakob · Bernhard Spengler · Juan Antonio Vizcaino

Published online: 27 December 2014

© The Author(s) 2014. This article is published with open access at Springerlink.com

Introduction

Mass spectrometry (MS) imaging is a very active field of research, and has seen impressive progress in recent years [1, 2]. The number of groups that are working on this topic is constantly increasing. However, the field is still very heterogeneous in terms of applied instrumentation and data processing methods. In addition, complex datasets are reduced to a set of two-dimensional “images,” which inevitably results in information loss. This simplified graphical representation also strongly depends on processing options such as color scale, intensity normalization, and spatial interpolation. Consequently, experimental data are presented in very diverse

ways, and published results can therefore be difficult to evaluate and compare. With a growing number of published studies, the issue of standardization and quality control of MS imaging data is becoming more important. This is a natural process for any new field that is maturing. The MS-based proteomics community has been facing similar issues in the last decade, and this discipline is therefore discussed as a “role model” herein. Since its inception in 2002, the Proteomics Standards Initiative (<http://www.psdev.info>) has driven the development of a number of minimum reporting guidelines (called “minimum information about a proteomics experiment” documents) [3] and several standard data formats for the different data types relevant in proteomics. For example, for raw and processed MS data, the data standard is called mzML [4].

In addition, several data repositories were established about 10 years ago to address the demand for storage and availability of MS data in the public domain [5–9]. A big step forward in this area has been the establishment of the ProteomeXchange (PX; <http://www.proteomexchange.org/>) consortium [10], led by the PRIDE [9] and PeptideAtlas [8] resources. The overall aim of PX is to provide a common framework and infrastructure for the cooperation of proteomics resources by defining and implementing consistent, harmonized, user-friendly data deposition and exchange procedures among the members. Thanks to the guidelines promoted by several scientific journals and funding agencies, and the general perception that sharing data is good scientific practice, the culture in the proteomics community has evolved toward data deposition as part of the publication process.

In analogy to these activities in the MS proteomics field, similar mechanisms have been discussed and to some extent already implemented in the MS imaging community in recent years. A common data format for MS imaging—imzML—has been established [11]. This format is being used more and more, and the number of available tools is constantly growing (see <http://www.imzml.org> for more details). Reporting guidelines have been discussed for several years, and a first suggestion of those is included in this topical collection [12].

Nevertheless, owing to the lack of suitable resources, a missing element so far has been the possibility to make MS imaging

Juan Pablo Albar died on 19 July 2014.

Published in the topical collection *Mass Spectrometry Imaging* with guest editors Andreas Römpp and Uwe Karst.

Electronic supplementary material The online version of this article (doi:10.1007/s00216-014-8357-8) contains supplementary material, which is available to authorized users.

A. Römpp (✉) · B. Spengler
Institute of Inorganic and Analytical Chemistry, Justus Liebig
University, Schubertstrasse 60, 35392 Giessen, Germany
e-mail: andreas.roempp@anorg.chemie.uni-giessen.de

R. Wang · H. Hermjakob · J. A. Vizcaino
European Molecular Biology Laboratory, European Bioinformatics
Institute, Wellcome Trust Genome Campus, Hinxton,
Cambridge CB10 1SD, UK

J. P. Albar
Centro Nacional de Biotecnología/CSIC, UAM Campus
Cantoblanco, Darwin 3, 28049 Madrid, Spain

A. Urbani
Department of Experimental Medicine and Surgery, University of
Rome “Tor Vergata”, 00133 Rome, Italy

A. Urbani
IRCCS – Fondazione Santa Lucia, Via del Fosso di Fiorano 65,
00143 Rome, Italy

datasets available in the public domain. Earlier attempts to develop a data repository were abandoned mainly because of the large size of MS imaging datasets. However, nowadays very large datasets (i.e., file size on the order of a few terabytes) can also be generated in MS-based proteomic and metabolomic studies, and can be submitted to established repositories.

From a purely technical point of view, the infrastructure available in existing MS repositories is also suited for MS imaging data. Therefore, the missing step is to define and adopt a submission procedure in order to be compatible with MS-imaging-specific parameters. Here we describe the newly implemented way of submitting MS imaging data to PX via the PRIDE database. We also describe how to retrieve these data and to reproduce the MS images.

Experimental procedures/submission process

Procedure for submitting datasets

For MS proteomics, there are two different PRIDE/PX submission modes: “complete” and “partial.” For both types of submission, a set of common metadata (agreed by all PX partners) and MS raw data are mandatory for each dataset. The difference is in the way the processed results are provided. After a “complete” submission has been performed it is possible for the repository to directly connect the processed peptide/protein identification results with the mass spectra. This can be achieved if the processed results are available in an open data format supported by the repository. The alternative is to perform a “partial” submission, and in this case, the connection between the spectra and the identification results cannot be done in a straightforward way. For “partial” submissions, the processed results are not available in a format supported by the repository. Instead, the corresponding analysis software’s output files (in heterogeneous formats) are made available for download [13].

The submission procedure has been adapted for MS imaging data using the “partial” submission mechanism. The PX Submission Tool [10] is the main tool used to perform the submissions. It is developed in Java and provides a user-friendly graphical user interface for performing the actual data submission, through a series of stages: (1) select all the files needed for the submission; (2) interactively group related files; (3) ensure the minimum level of metadata annotation; and (4) transfer the files via Aspera (<http://asperasoft.com/>) or FTP. Aspera can perform much faster transfers than FTP, resulting in a very convenient way of transferring potentially very large datasets.

In addition to the PX Submission Tool, datasets containing a high number of files can also be submitted using a command-line-based alternative [13].

An exemplary dataset has been used to demonstrate the submission process. The data describe a matrix-assisted laser desorption ionization imaging experiment of mouse urinary

bladder tissue sections acquired with high mass accuracy at a pixel size of 10 μm which was reported previously [14]. Experimental details are described in the electronic supplementary material (the results published originally are shown in Fig. 2, panels A and B). The dataset was deposited in PRIDE/PX (accession number PXD001283).

The submission process is described below; screenshots of all submission steps are provided in Figs. S1–S10. For a detailed description of the submission process, a tutorial for MS proteomics data is available [13]. Here, we will mainly focus on the changes made to accommodate the MS imaging datasets.

Before starting, users must first register at PRIDE (<http://www.ebi.ac.uk/pride/archive/register>). The default assumption is that all of the files belonging to one study or manuscript will be uploaded at once and handled as a unit, although there is some flexibility in the process. After the PX Submission Tool has been launched, in step 1 the type of submission must be indicated. In this case, “partial” must always be chosen.

In step 2, metadata are provided to describe the overall study, such as title, description, sample processing and data processing protocols, keywords, and experiment type (Fig. 1a). Most of them can be provided as simple text. A new term called “MS imaging” (MS: 1002521) has been created in the PSI-MS controlled vocabulary [15] to include this new experiment type in the submission process. The sample processing protocol should include all details on sample preparation, data acquisition, and instrumentation used. A crucial value that needs to be provided in any case is the pixel size that was used to acquire the reported MS image. A general suggestion for parameters that should be reported for MS imaging experiments has been proposed recently [12]. The data processing protocol is of particular importance for reproducing the MS images shown in the publication. A critical parameter is the bin size of the MS images, i.e., the mass range that was used to generate the image (e.g., ± 0.1 Da of the theoretical mass of the imaged compound). Other important details are the information about (pixelwise) intensity normalization as well as the type and level of spatial interpolation (and color map applied).

Actual data selection is performed in step 3 (“Add Files”) (Fig. 1b). These files include mass spectral data (labeled as “RAW”), metadata information about the images (labeled as “MS_IMAGE_DATA”), and an optical image as a reference (labeled as “OPTICAL_IMAGE”). The new file tags “MS_IMAGE_DATA” and “OPTICAL_IMAGE” have been incorporated to support MS imaging data. The files submitted for the example dataset are summarized in Table 1. The main specific points to consider for this step are as follows:

1. It is mandatory to provide the MS raw data (labeled as “RAW”). It is recommended to submit MS imaging data in imzML format as it offers the most flexible options for viewing, but proprietary data formats are also accepted.

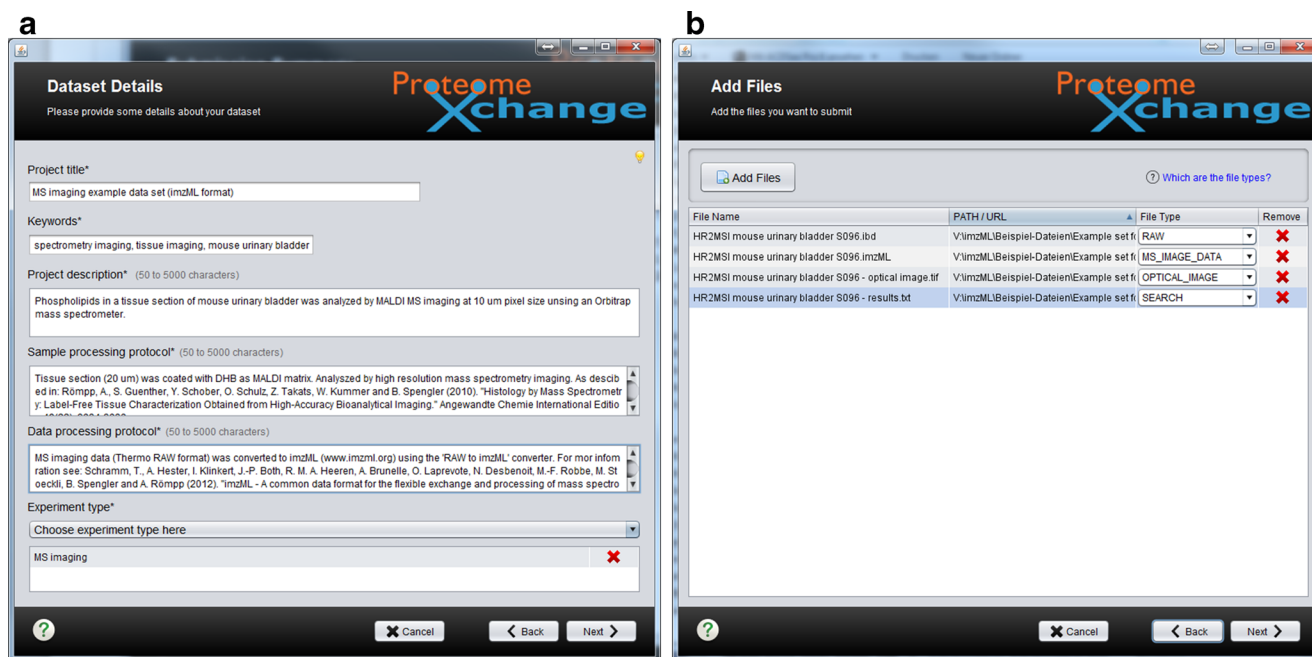


Fig. 1 Screenshots of the ProteomeXchange submission tool: **a** step 2 “Dataset Details” and **b** step 3 “Add Files”

The default submission process has been modified, and it now includes the possibility to submit two different mass-spectral-related files for one dataset, as required for several MS imaging data formats (e.g., imzML and Analyze). The mass spectral data file (*.ibd file for imzML or *.img file in Analyze format) must be labeled as “RAW.” The file that contains metadata (such as pixel dimensions and additional information) must be labeled as “MS_IMAGE_DATA” (e.g., *.imzml file for imzML or *.hdr file for Analyze). If an “ibd” file (imzML format) is submitted as “RAW,” an “MS_IMAGE_DATA” file (*.imzml) is mandatory. However, in the case of “RAW” proprietary formats that consist of only one file, an “MS_IMAGE_DATA” file is not required.

- In addition, PRIDE requires a mandatory “SEARCH” file for “partial” submissions, which corresponds to the processed results. This file typically contains a list of identified peptides and proteins, and related metadata. There is currently no strict definition for the format of this mandatory file, but it should contain a list of m/z values, names of (tentatively) identified compounds, and additional

information that was used to generate the MS images in the published work. The “SEARCH” file for the example dataset is given in the electronic supplementary material. These data should be available from the manuscript, but including them in a concise way in the data submission facilitates the reproduction of the published MS images.

We acknowledge that the file tag “SEARCH” might not be the best option to describe these data, but this file is still mandatory for consistency with the overall PX submission framework. Another more specific file tag might be added in the future if this turns out to be necessary.

- Since MS imaging data contain spatial information, the data submission also supports the inclusion of an optical image (“OPTICAL_IMAGE”) of the measured sample, which can allow validation and/or interpretation. The “OPTICAL_IMAGE” file could contain a photograph of the imaged sample or an adjacent section that shows comparable spatial features. Native samples, classic histological techniques (hematoxylin and eosin, toluidine), or immunohistochemistry staining (antibody staining) can be provided for this purpose.

Table 1 Submitted data files in the example dataset (accession number PXD001283)

File name	File type	Comment
HR2MSI mouse urinary bladder S096.ibd	RAW	Binary part of imzML data
HR2MSI mouse urinary bladder S096.imzml	MS_IMAGE_DATA	XML part of imzML data
HR2MSI mouse urinary bladder S096 - optical image.tiff	OPTICAL_IMAGE	Optical image of measured sample (tissue section)
HR2MSI mouse urinary bladder S096 - results.txt	SEARCH	Information on MS images which have been presented in the corresponding manuscript (e.g., m/z values)

The rest of the submission steps are identical to the MS proteomics “partial” submission process [13] (see Figs. S1–S10). An updated tutorial is available at <http://www.proteomexchange.org/submission>. Briefly, in step 4 (“Mapping Files”), the relationships between the different files can be captured. Each “RAW” file requires at least one “SEARCH” file. In addition, for MS imaging data, “RAW” files and “MS_IMAGE_DATA” files need to be connected. Step 5 is devoted to the annotation of biological and technical metadata. Information about species, tissues, and instrumentation is mandatory. In step 6, contact details for the principal investigator need to be provided. Step 7 is devoted to additional details. For instance, there it is possible to provide a PubMed identifier if the corresponding manuscript has already been published at submission time (as is the case in the example dataset). Step 8 (“Summary Screen”) provides an overview of all the files and file mappings, for the submitters to perform a final review before the actual file upload occurs. Step 9 is the final step where the files are uploaded to PRIDE. When the transfer is finished, the submitter will receive a confirmation e-mail. The actual time required to upload a dataset depends on the size of the dataset and the bandwidth available. In our example, it took approximately 1 h to prepare the data and collect the information for the submission, and the actual transfer of 830 MB data from Giessen to Cambridge, UK, took less than 5 min.

The data submission is then processed by the PRIDE team, and a “PXD identifier” is issued for each dataset. The submitter also receives by e-mail a username and password to allow private access to the data.

Retrieve and display data from the repository

All submitted datasets are private by default. Each dataset becomes publicly available on acceptance or publication of the corresponding manuscript, or when the submitters tell PRIDE to do so. All public PX datasets (including those in PRIDE) are accessible via the portal ProteomeCentral (<http://proteomecentral.proteomexchange.org/>). There it is possible to search for PX datasets, independently from the receiving repository (Fig. S11). However, although the datasets are private, reviewers/editors need to go to PRIDE (<http://www.ebi.ac.uk/pride>) and log in using the username and password provided by the authors. Detailed information about how to access private datasets is provided in the electronic supplementary material.

The imzML data can be downloaded and viewed in freely available software tools such as MSiReader [16], Datacube Explorer [17], and OmniSpect [18]. Alternative commercial tools include Quantinetix [19] and MALDIVision [20]. An updated list of available tools is available at <http://www.imzml.org>.

Panel C in Fig. 2 shows the imzML dataset as downloaded from PRIDE (accession number PXD001283) and displayed in the open source program MSiReader on a local computer (a screenshot that includes the user interface is shown in Fig. S12). The MS image shows features basically identical to those in the figure included in the original publication (Fig. 2, panel A). Alternatively, individual masses can be displayed as a grayscale image as shown for the example of m/z 798.5410 in Fig. 2, panel D (a corresponding MS image was included in the supporting online information of the original publication [14]). This MS image represents the raw data without any modifications such as interpolation or normalization. All necessary parameters to generate these MS images (bin width, color map, normalization, and interpolation options) are provided in the PX submission. This example demonstrates how MS imaging data used in publications can be reproduced with freely available and vendor-independent software. The reader can now access and search fully functional MS datasets, instead of evaluating the graphical representation of selected mass spectra and predefined MS images.

Discussion and outlook

In this article we have provided an overview of the newly implemented process for submitting MS imaging datasets to PX via PRIDE. The procedure allows the evaluation of data during the manuscript reviewing process. In the MS proteomics field, there is a growing trend toward public data reuse which is triggering the assessment, reuse, comparative analysis, and extraction of new findings from already published data. As an example, one very prominent case of reuse of PX datasets occurred in the elaboration of one of the recently published “drafts” of the human proteome [21].

Likewise, submitted datasets of MS imaging data could be used as input for bioinformatics reanalysis by other groups. One particular type of data reuse, popular already in other disciplines, is to analyze data coming from a large number of publications/datasets in a combined way. This is a so-called meta-analysis study, which could also be applicable to MS imaging.

A possible next step would be to establish procedures for making possible PX “complete” submissions for MS imaging data. For this it would be necessary to agree on a standard procedure that can make online/direct data visualization possible and, if applicable, make possible the direct connection between the mass spectra and the compounds reported in a given study. The submitted exemplary dataset represents the first fully functional (independent of vendor software) MS imaging dataset that is publically available and that can be used freely for subsequent reprocessing and interpretation (the original publication [14] should be referenced in this case).

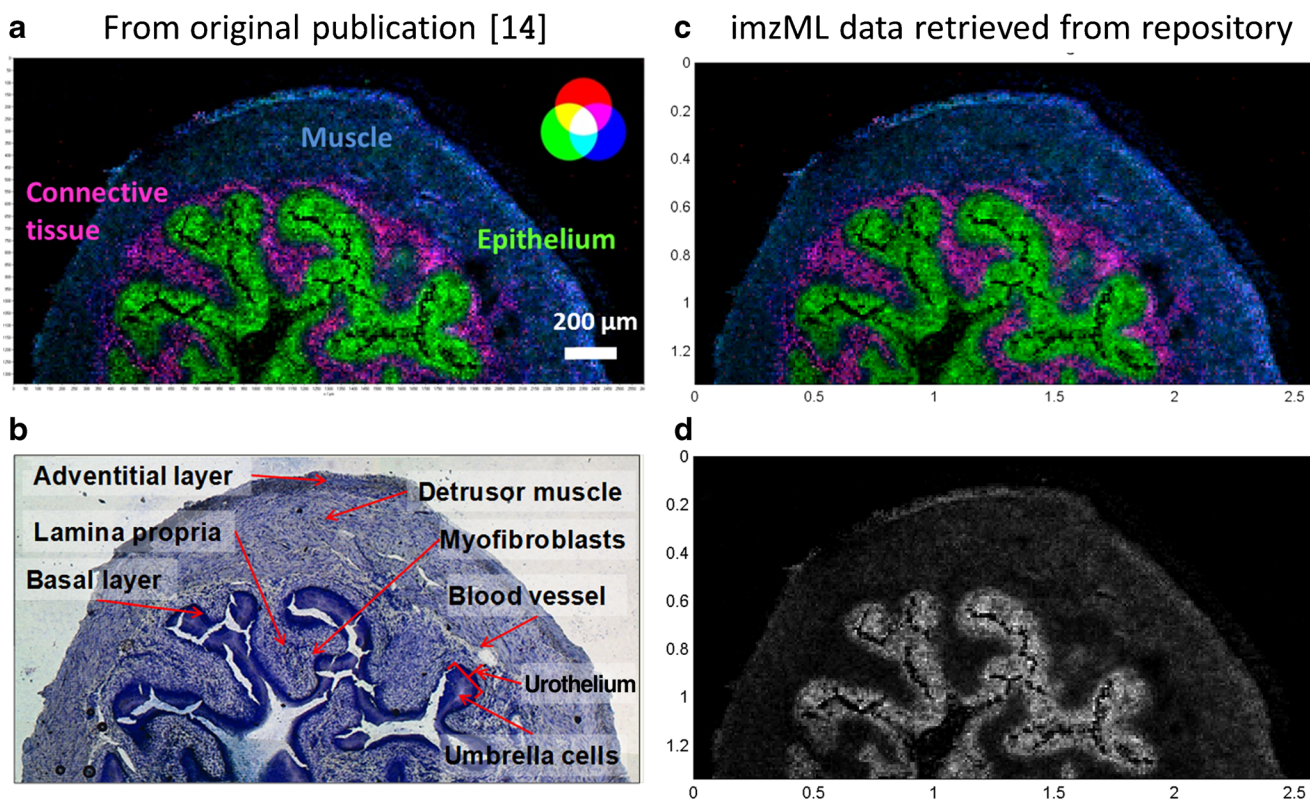


Fig. 2 Submitted dataset as displayed in the original publication [14]: mouse urinary bladder (A, B). A overlay of selected ion images for m/z 741.5307 (blue, muscle tissue, sphingomyelin 34:1), m/z 798.5410 (green, urothelium, phosphatidylcholine 34:1), and m/z 743.5482 (red, lamina propria); bin size $\Delta m/z=0.01$, 10- μm pixel size, 260×134 pixels. B optical image of measured section stained with toluidine. Mass

spectrometry images regenerated from the imzML dataset as retrieved from ProteomeXchange and displayed in the open source program MSiReader (C, D). C overlay of selected ion images corresponding to A. D selected ion image of m/z 798.5410 (phosphatidylcholine 34:1). (A, B modified from the original figure in Römpp et al. Copyright 2010 Wiley-VCH Verlag GmbH & Co. KGaA)

For example, this dataset could be used as a reference and test dataset for developers of data processing software, i.e., to evaluate procedures for mass recalibration or spatial segmentation. We have been asked to provide data for such activities in the past, which triggered a discussion about how to share and exchange data and what the conditions for their reuse would be. Now interested parties can directly download the data and use the data as necessary. This submitted example also acts as a template for the submission of larger datasets, since the procedure and requirements are identical. Currently, there is no upper limit on the size of the submitted datasets. MS datasets with sizes up to a few terabytes have been submitted successfully to PX.

Our main objective here has been to show that a data repository is now available for MS imaging data. A relatively small dataset was chosen in order to facilitate data retrieval and processing, but the mechanism is also applicable for larger-scale data. We are not proposing this process as the only and default procedure for publication of MS imaging data in the immediate future, but rather see this as a demonstration that there are no longer any technical obstacles to a public repository for MS imaging data. The availability of

such informatics tools represents an important step toward establishing MS imaging as a routine and reliable method in biomedical research. For instance, it may well open the route to fully integrate MS imaging within large-scale international research programs such as the Human Proteome Project.

The proposed approach allows public dissemination of MS imaging data based on existing public investment and infrastructure, avoiding significant extra costs for the development of a new, dedicated resource. The current setup allows the community to evaluate and test this submission process. We welcome any feedback since it could be used to improve and facilitate the submission process in the future.

Acknowledgments J.A.V. is supported by the Wellcome Trust (grant number WT101477MA) and the EU Seventh Framework Programme grants ProteomeXchange (grant number 260558) and PRIME-XS (grant number 262067). R.W. is supported by the Biotechnology and Biological Sciences Research Council Quantitative Proteomics grant (reference BB/I00095X/1). We acknowledge the COST Action BM1104 Mass Spectrometry Imaging: New Tools for Healthcare Research, which provided the framework for the initial discussion between A.R., J.P.A., and A.U. We especially acknowledge the contribution made to this work by Juan

Pablo Albar (1953–2014). His untimely death has deprived us of both a valued friend and colleague.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Römpp A, Spengler B (2013) Mass spectrometry imaging with high resolution in mass and space. *Histochem Cell Biol* 139(6):759–783
- Chughtai K, Heeren RMA (2010) Mass spectrometric imaging for biomedical tissue analysis. *Chem Rev* 110(5):3237–3277
- Taylor CF, Binz P-A, Aebersold R, Affolter M, Barkovich R, Deutsch EW, Horn DM, Huhmer A, Kussmann M, Lilley K, Macht M, Mann M, Muller D, Neubert TA, Nickson J, Patterson SD, Raso R, Resing K, Seymour SL, Tsugita A, Xenarios I, Zeng R, Julian RK (2008) Guidelines for reporting the use of mass spectrometry in proteomics. *Nat Biotechnol* 26(8):860–861
- Martens L, Chambers M, Sturm M, Kessner D, Levander F, Shofstahl J, Tang WH, Römpp A, Neumann S, Pizarro AD, Montecchi-Palazzi L, Tasman N, Coleman M, Reisinger F, Souda P, Hermjakob H, Binz P-A, Deutsch EW (2011) mzML—a community standard for mass spectrometry data. *Mol Cell Proteomics*. doi:10.1074/mcp.R1110.000133
- Perez-Riverol Y, Alpi E, Wang R, Hermjakob H, Vizcaino JA (2014) Making proteomics data accessible and reusable: current state of proteomics databases and repositories. *Proteomics*. doi:10.1002/pmic.201400302
- Haug K, Salek RM, Conesa P, Hastings J, de Matos P, Rijnbeek M, Mahendrakar T, Williams M, Neumann S, Rocca-Serra P, Maguire E, González-Beltrán A, Sansone S-A, Griffin JL, Steinbeck C (2013) MetaboLights—an open-access general-purpose repository for metabolomics studies and associated meta-data. *Nucleic Acids Res* 41(D1):D781–D786
- Craig R, Cortens JP, Beavis RC (2004) Open source system for analyzing, validating, and storing protein identification data. *J Proteome Res* 3(6):1234–1242
- Farrar T, Deutsch EW, Omenn GS, Sun Z, Watts JD, Yamamoto T, Shteynberg D, Harris MM, Moritz RL (2014) State of the human proteome in 2013 as viewed through PeptideAtlas: comparing the kidney, urine, and plasma proteomes for the biology- and disease-driven Human Proteome Project. *J Proteome Res* 13(1):60–75
- Vizcaino JA, Cote RG, Csordas A, Dianes JA, Fabregat A, Foster JM, Griss J, Alpi E, Birim M, Contell J, O'Kelly G, Schoenegger A, Ovelheiro D, Perez-Riverol Y, Reisinger F, Rios D, Wang R, Hermjakob H (2013) The PRoteomics IDentifications (PRIDE) database and associated tools: status in 2013. *Nucleic Acids Res* 41(D1):D1063–D1069
- Vizcaino JA, Deutsch EW, Wang R, Csordas A, Reisinger F, Rios D, Dianes JA, Sun Z, Farrar T, Bandeira N, Binz P-A, Xenarios I, Eisenacher M, Mayer G, Gatto L, Campos A, Chalkley RJ, Kraus H-J, Albar JP, Martinez-Bartolome S, Apweiler R, Omenn GS, Martens L, Jones AR, Hermjakob H (2014) ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat Biotechnol* 32(3):223–226
- Schramm T, Hester A, Klinkert I, Both J-P, Heeren RMA, Brunelle A, Laprèvote O, Desbenoit N, Robbe M-F, Stoeckli M, Spengler B, Römpp A (2012) imzML—a common data format for the flexible exchange and processing of mass spectrometry imaging data. *J Proteome* 75(16):5106–5110
- McDonnell LA, Römpp A, Balluff B, Heeren RMA, Albar JP, Andren P, Corthals G, Walch A, Stoeckli M (2014). Discussion point: reporting guidelines for mass spectrometry imaging. *Anal Bioanal Chem*. doi:10.1007/s00216-014-8322-6
- Ternent T, Csordas A, Qi D, Gomez-Baena G, Beynon RJ, Jones AR, Hermjakob H, Vizcaino JA (2014) How to submit MS proteomics data to ProteomeXchange via the PRIDE database. *Proteomics*. doi:10.1002/pmic.201400120
- Römpp A, Guenther S, Schober Y, Schulz O, Takats Z, Kummer W, Spengler B (2010) Histology by mass spectrometry: label-free tissue characterization obtained from high-accuracy bioanalytical imaging. *Angew Chem Int Ed* 49(22):3834–3838
- Mayer G, Montecchi-Palazzi L, Ovelheiro D, Jones AR, Binz PA, Deutsch EW, Chambers M, Kallhardt M, Levander F, Shofstahl J, Orchard S, Vizcaino JA, Hermjakob H, Stephan C, Meyer HE, Eisenacher M (2013) The HUPO proteomics standards initiative-mass spectrometry controlled vocabulary. Database. doi:10.1093/database/bat009
- Robichaud G, Garrard K, Barry J, Muddiman D (2013) MSiReader: an open-source interface to view and analyze high resolving power MS imaging files on Matlab platform. *J Am Soc Mass Spectrom* 24(5):718–721
- Klinkert I, Chughtai K, Ellis SR, Heeren RMA (2014) Methods for full resolution data exploration and visualization for large 2D and 3D mass spectrometry imaging datasets. *Int J Mass Spectrom* 362(1):40–47
- Parry RM, Galhena A, Gamage C, Bennett R, Wang M, Fernández F (2013) OmniSpect: an open MATLAB-based tool for visualization and analysis of matrix-assisted laser desorption/ionization and desorption electrospray ionization mass spectrometry images. *J Am Soc Mass Spectrom* 24(4):646–649
- Imabiotech (2012) Quantinetix MALDI imaging software. <http://www.imabiotech.com/Quantinetix-TM-Maldi-Imaging.html?lang=en>. Accessed 8 Jul 2012
- PREMIERBiosoft (2012) MALDIVision. <http://www.premierbiosoft.com/maldi-tissue-imaging/index.html>. Accessed 3 Jul 2012
- Wilhelm M, Schlegl J, Hahne H, Moghaddas Gholami A, Lieberenz M, Savitski MM, Ziegler E, Butzmann L, Gessulat S, Marx H, Mathieson T, Lemeer S, Schnatbaum K, Reimer U, Wenschuh H, Mollenhauer M, Slotta-Huspenina J, Boese JH, Bantscheff M, Gerstmair A, Faerber F, Kuster B (2014) Mass-spectrometry-based draft of the human proteome. *Nature* 509(7502):582–587



Andreas Römpp is a group leader and lecturer (*Privatdozent*) at the Institute of Inorganic and Analytical Chemistry of Justus Liebig University in Giessen, Germany. His work has always been focused on method development for high-resolution mass spectrometry. In recent years, mass spectrometry imaging has become his main area of interest. He is the coordinator of the common data format for mass spectrometry imaging—imzML. He is also actively involved in COST Action

BM1104 as a management committee member and workgroup co-chair.



Rui Wang works as a senior software engineer in the Proteomics Services Team at the European Bioinformatics Institute of the European Molecular Biology Laboratory. He has special interests in big data, data visualization, data analysis, enterprise application architecture, and bioinformatics. Previously, he has worked as a software engineer at a pharmaceutical start-up and at IBM.



Henning Hermjakob leads the Proteomics Services Team at the European Bioinformatics Institute, providing a broad portfolio of resources for systems biology, ranging from protein expression (PRIDE) via molecular interactions (IntAct) and curated pathways (Reactome) to systems biology models (BioModels) at the highest level of abstraction. As a founding member and co-chair of the Human Proteome Organization Proteomics Standards Initiative, and as a senior editor of the

journal *Proteomics*, he contributes to the standardization of data representation in proteomics.



Juan Pablo Albar obtained a degree in chemistry from Complutense University of Madrid, where he obtained his Ph.D. degree in 1981. After several years working in the private sector, he started his scientific activity at the National Centre for Biotechnology (CSIC), becoming one of the pioneers of proteomics in Spain. He was Head of the CSIC Proteomics Facility, Biomolecular and Bioinformatics Resources Platform Coordinator, a member of the Human Proteome

Organization Council, and a member of the Chromosome-Based Human Proteome Project Executive Committee. His work has been published in more than 160 scientific articles and was a source of inspiration for many scientists in the field. He passed away in July 2014 at the age of 61.



Bernhard Spengler is Full Professor of Analytical Chemistry at Justus Liebig University in Gießen, Germany. He contributed to the development of matrix-assisted laser desorption/ionization mass spectrometry in the 1980s, developed the postsource decay technique, and in 1994 introduced the matrix-assisted laser desorption/ionization imaging method. He and his group have worked in the field of mass spectrometry imaging for many years, in several fields of atmospheric

pressure in situ mass spectrometry techniques, and in aerosol analysis.



Andrea Urbani is Head of the Proteomics and Metabonomics Laboratory at IRCCS – Fondazione Santa Lucia, Rome. He is Associate Professor in Clinical Biochemistry and Molecular Biology, and Faculty Chair at the University of Rome “Tor Vergata.” Since 2009, he has been organizing the National Congress of the Italian Proteomics Association (<http://www.itpa.it>) as President of the Italian Proteomics Association for 2009–2015, and he will be

European Proteomics Association President for 2015–2017. Since he was awarded his Ph.D. degree in 1998, he has authored over 100 articles in the field of proteomics and metabonomics.



Juan Antonio Vizcaino is the project leader of the PRIDE database of mass spectrometry proteomics data at the European Bioinformatics Institute of the European Molecular Biology Laboratory (Cambridge, UK). He has also managed the ProteomeXchange consortium of proteomics resources since its inception in 2011. He has degrees in pharmacy and biochemistry and a Ph.D. degree in molecular biology from the University of Salamanca (Spain). In the last few years, he has been

involved in the development of PRIDE and its related tools and several data standards (e.g., mzIdentML and mzTab), as part of his contribution to the Human Proteome Organization Proteomics Standards Initiative.