Technical Note

Pathology interface for the molecular analysis of tissue by mass spectrometry

Jeremy L. Norris[1], Tina Tsui[1], Danielle B. Gutierrez[1], Richard M. Caprioli[1]

[1]Department of Biochemistry, National Research Resource for Imaging Mass Spectrometry, Mass Spectrometry Research Center, Vanderbilt University School of Medicine, Nashville, Tennessee 37240, USA

E-mail: *Dr. Richard M. Caprioli - r.caprioli@vanderbilt.edu
*Corresponding author

Abstract

Background: Imaging mass spectrometry (IMS) generates molecular images directly from tissue sections to provide better diagnostic insights and expand the capabilities of clinical anatomic pathology. Although IMS technology has matured over recent years, the link between microscopy imaging currently used by pathologists and MS-based molecular imaging has not been established. **Methods:** We adapted the Vanderbilt University Tissue Core workflow for IMS into a web-based system that facilitates remote collaboration. The platform was designed to perform within acceptable web response times for viewing, annotating, and processing high resolution microscopy images. **Results:** We describe a microscopy-driven approach to tissue analysis by IMS. **Conclusion:** The Pathology Interface for Mass Spectrometry is designed to provide clinical access to IMS technology and deliver enhanced diagnostic value.

Key words: Diagnostics, imaging mass spectrometry, molecular imaging

INTRODUCTION

Matrix-assisted laser desorption ionization (MALDI) imaging mass spectrometry (IMS) is a label-free molecular imaging technology that can be used to study a wide array of molecules (e.g., proteins, peptides, lipids, metabolites, and drugs) in their native tissue environment.[1,2] This technology is uniquely suited to the analysis of solid tissues because it provides targeted *in situ* molecular analysis while maintaining tissue morphological information upon which current anatomical pathology standards are constructed. Histology-directed IMS (Histo-IMS) is one modality of IMS that focuses on characterizing analyte distributions in discrete tissue locations, such as tumor versus nontumor regions or cell-type specific areas within a tissue section.[3,4] In contrast to whole specimen imaging, implementation of Histo-IMS improves the specificity and efficiency of the analysis making it possible to measure 100-1,000s of specimens per day as would be required for clinical diagnostic applications.[5]

The molecular specificity and sensitivity of modern MS provide unique capabilities to accurately measure disease markers that have clinical application. In recent years, liquid chromatography ([LC]-MS) has become a routine technology in the clinical laboratory[6,7] and is now responsible for returning millions of tests annually in the US alone.[8] Notwithstanding these enormous strides in the development of clinical MS, few examples of the application of these technologies to the diagnosis of disease in solid tissues have become standard clinical practice. This translational gap is the result of several

factors including (1) increased difficulty and complexity of preparing solid tissues for mass spectrometric analysis and (2) the lack of accessibility to the technology for those individuals with training in pathology and access to specimens. Further, the image manipulation process required to go from marked serial histology sections to pixel coordinates for image acquisition requires considerable time and expertise[9] and is a limiting factor for high-throughput analysis. Current approaches to the design and control of MS-based experiments are insufficient to overcome these problems. Thus, the opportunity exists to develop a user interface for MS suitable for the clinical environment to deliver quality, clinically actionable results to the physician.

RESULTS AND DISCUSSION

We have developed a novel workflow automation web-based software tool – the pathology interface for MS (PIMS) – that overcomes these challenges and facilitates collaboration between anatomical pathologists and analytical scientists through seamless integration of microscopy, sample preparation, and MS analysis [Figure 1]. Importantly, the interface allows biologists and pathologists who are untrained in IMS technology to control the acquisition of MS data. Image annotation performed by the pathologist drives histology-directed sample preparation and IMS analysis,[10-12] making this technology easy to use and ultimately accessible to physicians worldwide.

The core of the software is an image server that permits remote viewing and annotation of high-resolution micrographs by pathologists and biologists anywhere in the world. Other software features include automation tools to permit control of instruments based on annotations and a project management tool to facilitate the transfer of relevant

Figure 1: Pathology interface for mass spectrometry transforms the current IMS approach into one practical for the clinical setting, integrating microscopy, and IMS analyses to deliver molecularly specific, clinically valuable results

sample information. PIMS incorporates a histological examination of the tissue, sample preparation and analysis, and reporting while maintaining the important link between the histology and the molecular measurements that can further guide diagnosis. PIMS facilitates rapid collaboration, reduces experimental error, and increases productivity, saving time and money for the laboratory.

In a normal pathology workflow, tissue samples are sent to the laboratory for analysis. Thin serial sections (approximately 5–10 μm) are prepared for analysis by sectioning on a microtome. One of the serial sections is stained using normal procedures for histological analysis and subsequently imaged using a high-resolution slide scanner, typically at ×20–40 magnification. One of the issues encountered is that high-resolution optical micrographs suitable for interpretation by pathologists are often larger than 30 GB without data compression. Workflows requiring the transfer of these images to a remote collaborator and return of the image after annotation are slow and inefficient, making this process impractical for clinical work that requires rapid testing.

In contrast, the PIMS interface does not necessitate the transfer of a compressed image to capture the annotations needed to execute the required experiment. After tissue sections are imaged, users upload images to the cloud from which they are automatically imported into the PIMS database, eliminating the need to reduce image size and resolution for E-mail transfer of images. Collaborators can open the PIMS interface through any web browser (Chrome, Firefox, Internet Explorer, Safari, etc.,) and therefore, it can be used on PC or Mac desktops, laptops, or tablets with no additional installs required. Images can be composed – viewed, cropped, rotated, and annotated – using the Image Markup Viewer [Figure 2a-c]. The image server supports layered tiling to optimize performance in viewing large, high-resolution (gigapixel) images so that users can efficiently view and annotate tissue areas. Opening images, this size can take hours on a common laptop/ desktop setup, but PIMS accomplishes this task in seconds.

Once the coordinates for each annotated region of interest are recorded, the annotated image is spatially registered with an optical image of the serial unstained tissue section. Currently, registration of the annotated image to the optical image of an unstained tissue section is performed manually. Pixel coordinates of the annotated regions are transferred to the coordinate system of the spotting instrument (e.g., Labcyte Portrait 630) for matrix application to the unstained tissue section and subsequent MS data acquisition [Figure 2d]. There are no internal, software automated controls to verify the accuracy of the overlaid annotations or the position of the resulting matrix spots. However, the relative alignment between the matrix spots and the annotations are checked using overlaid
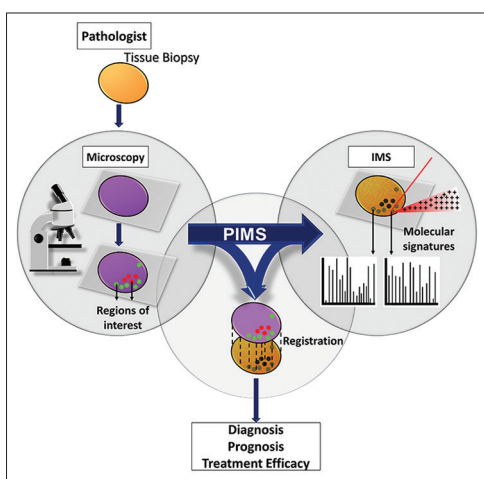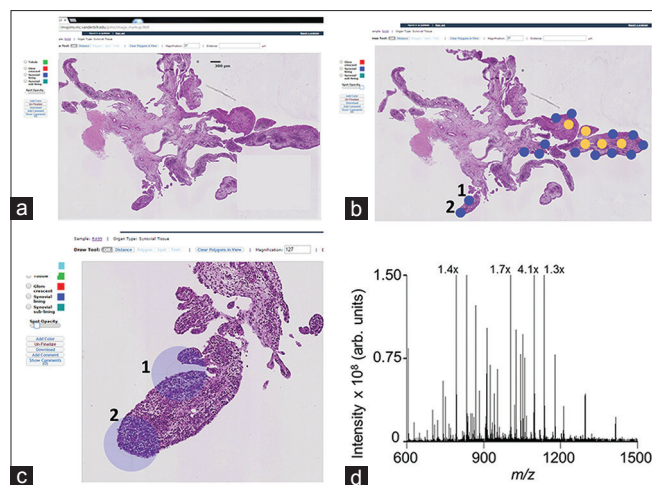
Figure 2: Screenshots from the pathology interface for mass spectrometry interface highlight the software features and workflow. (a) A scanned H and E stained tissue section can be viewed through the image markup screen and examined remotely by a pathologist. (b) The pathologist can then annotate the tissue via a user-defined color palette to distinguish regions of interest, for example, tissue classes (e.g., normal versus disease) or cell types. Annotations can be made as discrete spots (as shown) or regions. (c) Pathology interface for mass spectrometry can zoom to regions of interest. Scale bars and the magnification level are automatically rescaled according to the zoom level. (d) A mass spectrometry spectrum acquired from the spot labeled "1" in panels b and c

optical images by both the analyst and pathologist to ensure accuracy and importantly that the cell population underlying the matrix spot is homogeneous. If the sample preparation is not accurate, the specimen is prepared again for analysis. Development of the registration feature is ongoing, and automated registration of images is being developed for future versions of PIMS.

Within the Image Markup Viewer, many custom features have been specifically designed for the IMS workflow such as a custom color palette to distinguish the classification of annotated regions and cell types, a measurement tool to calculate the distance between sample features, a polygon tool to designate regions of interest, and a text annotation tool for comments. These features are highlighted in a screenshot of the Image Markup Viewer [Figure 2a-c]. The classification of specific features (e.g., cell type and disease state) into a specific color group has been designed to prevent errors in annotation and classification for projects with multiple experiments. For example, if a specific feature is designated by a certain color in the first experiment that color can only represent that feature in subsequent related experiments. In addition, the software guides the proximity of two annotated areas based on achievable spatial resolutions for the project. Annotations are saved as a separate data layer to allow for easy editing, and custom annotated images are generated when a user chooses to save a snapshot of the viewport in the Image Markup Viewer. Once the analysis has been completed,

the annotations can be color coded to represent the classification of the cells of interest in accordance with the MS results.

PIMS also streamlines data management throughout the entire project. For example, in the Vanderbilt University Tissue Core, project information is included such as grant and billing information, project PI, lead project scientist, affiliated scientists, and resource information. Notifications of the project status can be sent through PIMS, thereby expediting the sample preparation process. For example, the project scientist can notify the clinical collaborator that samples were received and enter the planned schedule for sample preparation and analysis into PIMS; the collaborator will be notified and can view the schedule. The collaborator is also notified when samples are ready for viewing and annotation. As soon as the images have been annotated, the project scientist is notified. Furthermore, information managed through PIMS is secured using a single sign-on protocol Central Authentication Server that supports multiple authentication methods and contains five access levels: System administrator, core manager, core user, user, and billing.

PIMS is built on a distributed processing architecture to handle concurrent requests for composing annotated images while maintaining acceptable response times. Supplementary Figure 1 and the accompanying text illustrate the processing architecture and provide detailed descriptions of the technology stack and development environment. Development and improvement of PIMS are ongoing. To provide the latest version, source code will be shared with interested collaborators by request to the corresponding author. In its current implementation, no patient data is stored in the system. A planned further development will involve security measures that meet HIPAA security standards and integration with existing digital pathology platforms so that the molecular classification provided by MS can be integrated into a single clinical report with the histopathological results.

PIMS is currently the central touchstone for collaborators to direct analytical efforts within our laboratory. Currently, there are thirty projects managed through PIMS, and feedback from these collaborators permits continual refinement of the software to ensure a relevant and user-friendly product. Examples of published work that used PIMS include the analysis of racial disparities in Wilms Tumor by IMS[10,11] and tissue profiling and diagnosis of melanocytic skin tumors by IMS.[12] In the latter case, IMS was able to detect spatial, diagnostically useful differences in vimentin and actin expression in spitz nevi and spitzoid malignant melanomas, but neither conventional immunohistochemistry nor automated quantitative analysis of tissue sections could distinguish these differences.[13] These results highlight the power and utility of IMS for diagnostic medicine.

CONCLUSIONS

This report describes a first generation software system for the control of histology-directed MS that was designed and built specifically with pathologists' needs in mind. This control paradigm places the tissue and the visualization of the cell morphology at the heart of the experiment. This approach to the design and execution of the experiment allows for biologists and pathologists, whose expertise is not necessarily this instrumentation, to guide the collection of molecular data from cells of interest in tissues with minimal training. Likewise, collaboration among groups of chemists and pathologists is made easier even across great distances. While the PIMS software described above was implemented for MALDI-IMS, this software platform could be used to guide similar experiments using a number of surface sampling/ionization technologies in MS, such as desorption electrospray ionization. Further, the PIMS technology could be adapted for the remote operation of laser-capture microscopy to facilitate other molecular analysis techniques including genomics, transcript analysis, and LC-based MS. As highly specialized analytical platforms continue to transition to diagnostic applications, software solutions that facilitate access to these technologies will be necessary to connect patients and physicians to important and potentially life-saving results.

Financial Support and Sponsorship

Conflicts of Interest

There are no conflicts of interest.

REFERENCES

1. Caprioli RM, Farmer TB, Gile J. Molecular imaging of biological samples: Localization of peptides and proteins using MALDI-TOF MS. Anal Chem 1997;69:4751-60.
2. Norris JL, Caprioli RM. Analysis of tissue specimens by matrix-assisted laser desorption/ionization imaging mass spectrometry in biological and clinical research. Chem Rev 2013;113:2309-42.
3. Cornett DS, Reyzer ML, Chaurand P, Caprioli RM. MALDI imaging mass spectrometry: Molecular snapshots of biochemical systems. Nat Methods 2007;4:828-33.
4. Taverna D, Boraldi F, De Santis G, Caprioli RM, Quaglino D. Histology-directed and imaging mass spectrometry: An emerging technology in ectopic calcification. Bone 2015;74:83-94.
5. Cornett DS, Mobley JA, Dias EC, Andersson M, Arteaga CL, Sanders ME, *et al.* A novel histology-directed strategy for MALDI-MS tissue profiling that improves throughput and cellular specificity in human breast cancer. Mol Cell Proteomics 2006;5:1975-83.
6. van den Ouweland JM, Kema IP. The role of liquid chromatography-tandem mass spectrometry in the clinical laboratory. J Chromatogr B Analyt Technol Biomed Life Sci 2012;883-884:18-32.
7. Vogeser M, Seger C. A decade of HPLC-MS/MS in the routine clinical laboratory – Goals for further developments. Clin Biochem 2008;41:649-62.
8. Grebe SK, Singh RJ. LC-MS/MS in the clinical laboratory – Where to from here? Clin Biochem Rev 2011;32:5-31.
9. Samsi SS, Krishnamurthy AK, Groseclose M, Caprioli RM, Lozanski G, Gurcan MN. Imaging Mass Spectrometry Analysis for Follicular Lymphoma Grading. Proceedings of the 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2009 September 2-6. Minneapolis, MN: IEEE; 2009.
10. Axt J, Murphy AJ, Seeley EH, Martin CA, Taylor C, Pierce J, *et al.* Race disparities in Wilms tumor incidence and biology. J Surg Res 2011;170:112-9.
11. Libes JM, Seeley EH, Li M, Axt JR, Pierce J, Correa H, *et al.* Race disparities in peptide profiles of North American and Kenyan Wilms tumor specimens. J Am Coll Surg 2014;218:707-20.
12. Lazova R, Seeley EH, Keenan M, Gueorguieva R, Caprioli RM. Imaging mass spectrometry – A new and promising method to differentiate spitz nevi from spitzoid malignant melanomas. Am J Dermatopathol 2012;34:82-90.
13. Alomari AK, Klump V, Neumeister V, Ariyan S, Narayan D, Lazova R. Comparison of the expression of vimentin and actin in spitz nevi and spitzoid malignant melanomas. Am J Dermatopathol 2015;37:46-51.

Supplementary Text

IMPLEMENTATION

System Overview

Pathology interface for mass spectrometry (PIMS) is a web-enabled platform that supports distributed processing of large images to facilitate performance optimization and scaling. The platform integrates multiple open-source technologies, leveraging the strengths of each technology and modularizing services so that the system can be easily enhanced and optimized as new technologies become available. The system consists of seven functional modules: A web application server, an authentication server, a relational database, a persistent message queue, an image server, image processing services, and a browser client. The server-side modules run best with the Linux operating system.

Web Application Server, Authentication, and Database

The web application server integrates the services of the other modules to respond to requests from the browser client. It was implemented with Ruby on Rails[1] and uses Phusion Passenger[2] to integrate with Apache web server.[3] The web application can be configured to use several authentication mechanisms: (1) Database (passwords are encrypted in the database), (2) external direct Lightweight Directory Access Protocol (LDAP), and (3) external single sign-on via Central Authentication Server (CAS) protocol.[4] In our production environment, we used a CAS authentication server. The web application server manages access (secured by role-based rules) to user data (projects, samples, annotations, comments) and other internal data stored in a relational database. The web application supports several relational databases. We use PostgreSQL relational database. To optimize performance, the web application mediates image requests: (1) Requests for existing images that have not been modified on the image server are deferred to the browser cache before going to the image server, (2) requests for images that require composition are submitted to the message queue for processing.

Image Server

We use JPEG2000 as our internal image format for scanned tissue images because it supports multiresolution images (levels) and image segmentation (tiling) within a single image file. Tiling is key to the ability to view high spatial resolution images via the web in an efficient manner. Djatoka image server[5] is used to read scanned images converted to the JPEG2000 format. Djatoka also supplies an additional level of caching, delivering further performance. Djatoka is run within the Apache Tomcat Java container.[6]

Image Processing Services and Message Queue

To enhance throughput and performance, multiple processes manage image conversion and composition. These processes are directed by a persistent message queue serviced by Redis database.[7] A polling service scans for tissue image files to import and then submits the files to the message queue for conversion. The web application server also submits requests to the message queue for image resizing, cropping, and annotation. Multiple server processes consume the queued requests to process the image. ImageMagick[8] (with JPEG2000 enabled using the JasPer library[9] and BigTiff enabled using the BigTIFF library)[10] converts tissue image files to JPEG2000 and is also used to crop, resize, and annotate images for user downloads.
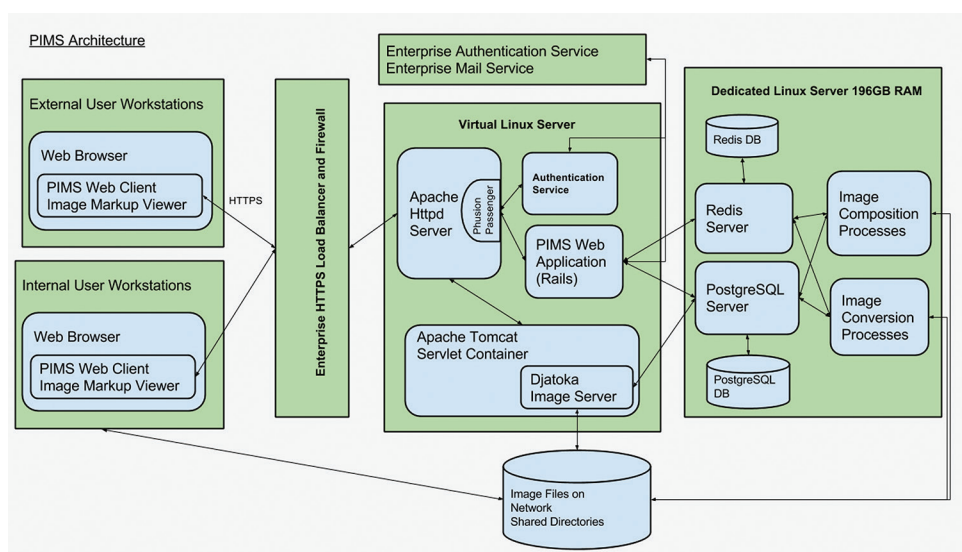
Supplementary Figure 1: Architecture for pathology interface for mass spectrometry software. Pathology interface for mass spectrometry is built on a distributed processing architecture to handle concurrent requests for composing annotated images while maintaining performance and throughput. The system incorporates open-source technologies and is modularized to facilitate future development and optimization

Browser Client

Most of the project and sample management screens use HTML templates facilitated by the Ruby on Rails framework to access and update data in PostgreSQL and Redis databases. The Image Markup Viewer, however, is powered by an internally developed JavaScript library that leverages the scalable vector graphics (SVG) canvas graphics capabilities built into modern browsers supporting HTML5, CSS3, and SVG specifications.

DEVELOPMENT ENVIRONMENT AND PROCESS

PIMS was developed mainly with the RubyMine integrated development environment[11] on Mac OSX, using agile techniques and short development cycles. Testing was done within the development process. Most tests were done manually, emphasizing integration testing over unit testing. The three staging environments included: Development, user acceptance test (UAT), and production. All environments were on the same platform: Linux CentOS 6. Integration testing was done in both development and UAT environments before release into production. Git was used as a source code repository and for version control.[12] Git and shell scripts were used for deployment and administration.

DEPLOYMENT CONFIGURATION

The modular design of the system allows us to distribute processing to enhance performance and make the most of available hardware and enterprise infrastructure in a cost-efficient manner. All of the modules communicate over network via Transmission Control Protocol/Internet Protocol sockets and thus can be deployed to separate servers as long as they can communicate over the network. Our current deployment configuration for production environment consists of (1) a virtualized server running the web application and Djatoka image services behind a load balancing Hypertext Transfer Protocol router, secured with a firewall and (2) a dedicated server with 196GB of RAM running Redis and PostgreSQL databases and the image processing services. If more memory is needed to increase the throughput of image processing, we can add servers and run additional image processing services on those servers. We use enterprise LDAP services for authentication and enterprise secure mail servers for email notifications, which are external to PIMS and provided by the Vanderbilt University Information Technology infrastructure.

REFERENCES

1. Rails. David Heinemeier Hansson; Available from: http://www.rubyonrails.org/. [Last cited on 2016 Mar 10].
2. Phusion Passenger. Phusion BV; c2012-2016. Available from: https://www.phusionpassenger.com/. [Last cited on 2016 Jan 15].
3. Apache HTTP Server Project. The Apache Software Foundation; c1997-2016. Available from: https://www.httpd.apache.org/. [Last cited on 2016 Mar 10].
4. RubyCas Server. Available from: http://rubycas.github.io/. [Last cited on 2016 Mar 10].
5. Source Forge. Djatoka. Slashdot Media; c2016. Available from: http://www.sourceforge.net/projects/djatoka/. [Last cited on 2016 Jan 15].
6. Apache Tomcat. The Apache Software Foundation; c1999-2016. Available from: http:// tomcat.apache.org/. [Last cited on 2016 Jan 15].
7. Redis. Redis Labs. Available from: http://www.redis.io/. [Last cited on 2016 Jan 15].
8. ImageMagick. ImageMagick Studio, LLC; c1999-2016. Available from: http://www.imagemagick.org. [Last cited on 2016 Jan 15].
9. Adams M. The JasPer Project Home Page. Department of Electrical and Computer Engineering, University of Victory. Available from: http://www.ece.uvic.ca/~frodo/jasper/. [Last cited on 2106 Jan 15].
10. LibTIFF – TIFF Library and Utilities. Available from: http://www.remotesensing.org/libtiff/. [Last updated on 2015 Sep 12; Last cited on 2016 Jan 15].
11. RubyMine. JetBrains s.r.o.; c2000-2016. Available from: https://www.jetbrains.com/ruby/. [Last cited on 2016 Jan 15].
12. Git. Available from: https://git-scm.com/. [Last cited on 2016 Mar 10].