Providing access to slide and tissue banks

Robert W. Williams and Glenn D. Rosen

Effective long-distance collaboration is rapidly transforming research. This transformation was initially catalyzed by the almost instantaneous exchange of large data files via e-mail, ftp, and the World Wide Web. However, it is now practical to work together in much more interconnected ways. Sophisticated software, rapid advances in hardware speed and capacity (a terabyte of disk storage now costs as little as $1200), and ubiquitous Internet access can be leveraged to make research resources available in real time to investigators dispersed across continents.

The cost of implementing a "collaboratory" is within the reach of even small groups. The Informatics Center for Neurogenetics is a case in point. The Center consists of research teams located in Memphis, TN; Philadelphia, PA; Buffalo, NY; and Boston, MA. These groups share a common interest in the multiple genetic causes of brain diseases such as dyslexia, autism, Parkinson's, Alzheimer's, and schizophrenia; all of the groups use mice to study these diseases. Genetic studies of this type require group access to massive collections of tissue, particularly histological sections. Mapping a single gene associated with a neurodegenerative disease can require quantitative structural and genetic analyses of hundreds of mice. At this scale, collaboration is a necessity, not an option. The challenge is how to work together effectively to generate and share common resources, minimize duplication of effort, and build tools for distributed data analysis. The greater challenge is to make these resources accessible to an even larger community of hundreds of independent research groups over the Internet.

This paper describes highly effective bioinformatic and software solutions developed over the past three years in the laboratories of the Informatics Center for Neurogenetics. The main focus is on the distribution of images from a large collection of histological slides to anyone with an Internet connection in full color and at resolutions ranging from 25 µm per pixel down to 0.1 µm per pixel. This is beneficial to both large and small clinical and research laboratories with valuable slide collections that they want to curate and distribute electronically. Much can be accomplished in just weeks, with a light box, digital camera, and relational database system such as FileMaker Pro (**Filemaker, Inc.**, Santa Clara, CA) or the MySQL. The Mouse Brain Library (MBL) (www.mbl.org) not only serves static images of slides and sections, but also provides dynamic control of motorized microscopes located in distant laboratories and clinics. This makes it possible to convert any computer system into a virtual Internet microscope, or iScope. The MBL can be implemented on Macintosh OS X (**Apple Computer Corp.**, Cupertino, CA), Windows, and the Linux operating systems. FileMaker Pro is a more suitable introduction for biologists who do not have expertise in database design. Because the program runs on all three major operating systems, it will be used to illustrate the examples throughout this paper.

Mouse Brain Library

The Informatics Center for Neurogenetics is part of an extended network called the Human Brain Project, which is funded jointly by the National Institute of Mental Health, National Institute on Drug Abuse, and National Science Foundation (www.nimh.nih.gov/neuroinformatics/index.cfm). The MBL houses an expanding collection of high-resolution images of histological slides taken from 1800 mice representing over 100 different genetically distinct strains. Each slide contains a systematic sample of sections throughout an individual mouse brain. These slides are cut at different planes of section and are dynamically presented at a variety of resolutions (4–24 µm/pixel). In addition to these images, the MBL contains atlases, magnetic resonance images (MRIs), and databases on the brain structure and genetic makeup of each of these mice, all of which must link together across the four separate laboratories that comprise the MBL. Moreover, these data need to be linked directly to the Web so that an expanding network of researchers interesting in utilizing this resource can access them.

Managing case and image data flow

The record-keeping of animal sources, tissue processing, sectioning, and digital photographs that interest researchers relies on a suite of integrated database files that are maintained on a dual processor Mac G4 server computer that handles colleagues and clients using standard TCP/IP protocols. A key factor in choosing a relational database is cost, development speed, and development control. Small research groups often modify database tables on a weekly basis, and the authors' group needed a software solution over which they had complete control. They considered and experimented with such alternative relational database programs as Access (**Microsoft**, Redmond, WA), Panorama (**ProVUE Development Corp.**, Huntington Beach, CA), the MySQL, and PostgreSQL.

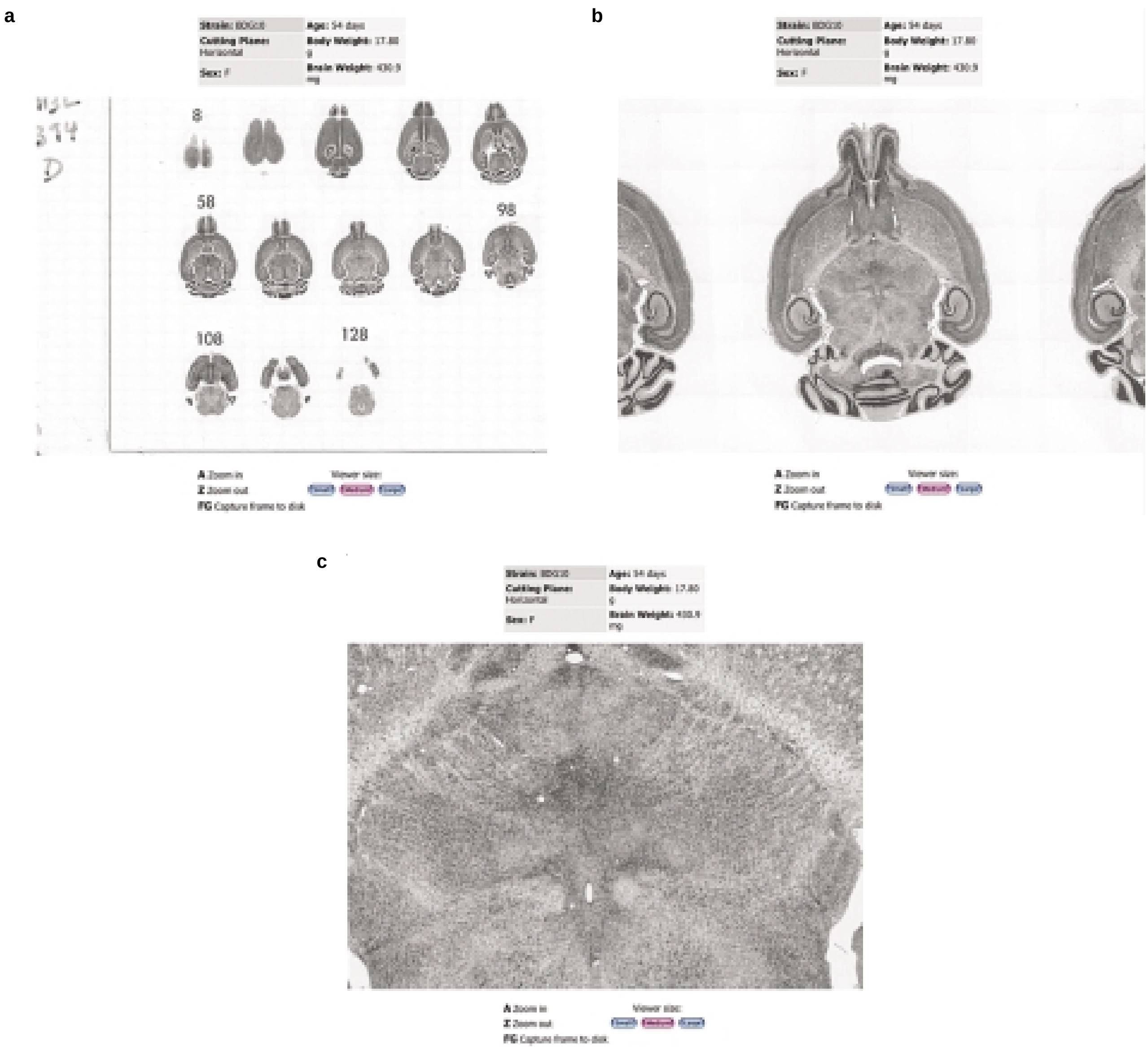

Figure 1 *On-line magnification of high-resolution images of mouse brain slides in the Mouse Brain Library using Zoomify (T) and a Web browser. a) shows the entire slide, b) zooms in on a section in the middle of the slide, and c) shows the highest resolution image of that section (4 µm/pixel). The high-resolution image was obtained using Virtual Slice from* ***Microbrightfield, Inc.***

FileMaker Pro has become the group's primary relational database system, because it is easy to prototype and implement complex and visually self-explanatory tables and relations. The program runs well on the latest Linux, Macintosh, and Windows operating systems. Upgrades have kept pace with technology (most recently, XML support) without sacrificing design clarity. This database interfaces well with **Microsoft** Excel on one side and the Internet on the other.

Data are entered directly into one of several interconnected data tables, all of which reside on a networked server that is backed up daily to a second hard drive. Files are archived on a DVD on a weekly interval. Networked computers (mainly iMac [**Apple**] and Wintel clients) are kept on small carts in the mouse colony and in histology laboratories, and in close proximity to digital photography and microscopy workstations in Tennessee and Boston. Entering data directly into a consolidated set of database files greatly simplifies record-keeping and minimizes data loss and transcription errors. The slide collection is typically photographed with a high-resolution camera system. The current system consists of a PhaseOne H20 (**Phase One Inc.**, Northport, NY) digital camera back with a 4000 × 4000 pixel charge-coupled device (CCD) camera (**Kodak**, Rochester, NY) coupled to a 35-mm-style camera body (K**apture Group Inc.**, Osage Beach, MO) and a PC Micro Nikkor 85-mm f2.8 lens (**Nikon**, Melville, NY). The system can image anything, i.e., a single section, an entire slide, or a large histological specimen, with extremely high resolution and contrast. As images are acquired and processed, links and background information are immediately added to one of several image database files.

Worldwide access to the digital images is provided in three steps that vary in simplicity. Clicking a few boxes in FileMaker can open up a database and images to anyone with a Web browser. Developing a custom FileMaker Web interface takes a bit more finesse and involves simple coding with control vocabulary and scripting language. The prototype www.mbl.org site was built in less than two weeks. Finally, once a working Internet image database is up and running, it is possible, with additional programming and database expertise, to convert a system to a public domain database software such as MySQL or PostgreSQL. The advantage of the latter implementation is that distribution does not require additional licenses. The disadvantage is that such a translation requires much more computer and database expertise to manage. It took a skilled programmer several months to replicate the FileMaker version of the MBL.

The need for digital video

Still images of the slide collection are available at a series of resolutions, but they have limitations. Many users need to see images at high resolutions (0.1–5.0 µm/pixel) to examine cellular structure in detail. Students, clinicians, and researchers may also want to change the focal plane interactively. One solution that is becoming practical is simply to digitize all of the tissue on the slide using 4× to 40× objectives. With a 4× objective, excellent photomosaics can be constructed in under 1 hr for single slides using a digital microscope system from **Microbrightfield Inc.** (Williston, VT) and a simple Web browser plug-in program called Zoomify (**Zoomify, Inc.**, Santa Clara, CA) (See *Figure 1*). A growing collection of remarkably large photomosaics is now available in the MBL. Users can pan to any location and then zoom in as if using a high-quality dissecting microscope. It is also possible to reach the cellular resolution limits, but because even a single slide can occupy close to a terabyte, the coverage must be limited to restricted regions of interest.

High-resolution images

To distribute the highest-resolution images possible from the MBL collection, a suite of programs and Internet interfaces are being developed to provide users access to a set of three remote control light microscopes named Tall, Dark, and Handsome. Each is equipped with a motorized stage, contrast-enhanced optics, and high-power objectives. Images are captured using a DFW-SX900 FireWire camera (**Sony Inc.**, New York, NY) that produces images that match those at the eyepiece (1280 × 960 pixel, 24 bit at 20 Hz) and that also provides a pixel dimension of about 0.1 µm. These iScopes can acquire through-focus stacks of images at locations that are defined by clients over the Internet. The lower-resolution MBL slide images are used as a guide. The iScopes acquire stacks of images and then archive and distribute sliced digital tissue over the Internet at full resolution. Researchers can also view the streaming video output of the camera and move the XYZ to examine different parts of a slide.

Once researchers have acquired and transferred digital tissue from the iScope to their own computer, they need to be able to score, count, and measure cells. Loss of cells is a strong indicator for Parkinson's or Huntington's disease, while changes in cell distribution are typical in diseases such as schizophrenia. The VideoScribbler and ARIA3D software programs (developed at the Informatics Center for Neurogenetics) allow clients to trace directly onto the images to measure and count. ARIA3D is completely generic and the final release will run effectively using Mac OS X, Windows, and Linux.

Conclusion

Tools such as FileMaker Pro, iScope, and ARIA3D are invaluable for enabling the scientists of the MBL, as well as researchers worldwide, to share information, advance their research, and avoid duplicative efforts.

Dr. Williams is a Professor, Department of Anatomy and Neurobiology, Department of Pediatrics, University of Tennessee Health Science Center, 855 Monroe Ave., Memphis, TN 38120 U.S.A.; tel: 901-448-7018; fax: 901-448-7193; e-mail: rwilliam@nb.utmem.edu. Dr. Rosen is an Associate Professor of Neurology (Neuroscience), Department of Neurology, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, MA, U.S.A.