INTRODUCTION

Medical librarians can expedite biomedical research by broadening the scope of their instruction curricula to include courses on the use of network visualization tools. With the dramatic increase in the number of databases, data sets, tools, and software being used to store, retrieve, and analyze proteomic, genomic, and metabolomic data, researchers are being forced to navigate an ever-complex information environment. Medical librarians can accelerate data analysis by learning about, conducting comparative analyses for, and providing instruction on these resources. Many libraries, such as the University of Florida’s Health Science Center Libraries and Washington University’s Bernard Becker Medical Library, offer instruction on biomedical databases and tools such as those provided by the National Center for Biotechnology Information (NCBI) [1, 2]. Instruction on network visualization tools is not as commonly provided by libraries.

Network visualization is a method of data analysis that uses nodes to represent a data point, such as a gene, and edges (lines connecting nodes) to represent a relationship between two data points, such as an interaction between two genes (Figure 1, online only). Nodes and edges can have attributes associated with them that provide additional information, such as the chromosome on which a gene is located. Network visualization is an effective method for representing biologic relationships by succinctly highlighting properties and trends in complex systems. Networks also allow the integration of multiple different kinds of data such as gene expression, ontologies, and protein structures that have traditionally been stored in their own repositories [3]. Network analysis and network modeling techniques are being applied to biological networks in order to provide new hypotheses for biological systems, and as a result, a large number of tools for visualizing networks are constantly being developed, used, and reviewed [4].

Teaching biomedical resources such as network visualization tools, in addition to those developed by NCBI, is an area where medical librarians can make a significant impact. While resources like OpenHelix...
provide some instruction on these types of resources, these sources are limited, providing an important opportunity for libraries to expand the scope of their instruction curricula and to use instruction methods that allow broad participation by users. This paper discusses the evolution of a successful training class on a free and useful network visualization tool and the methods used to reach a wide-ranging audience.

METHODS

Two librarians at the University of Michigan have provided instruction on a network visualization tool called Cytoscape (www.cytoscape.org) since 2009. Cytoscape is an open source software tool used to visualize and analyze biological networks and pathways. The Cytoscape core software is developed by a consortium but has a plug-in architecture that allows anyone to develop components that meet their specific needs. Plug-ins provide additional functionality to Cytoscape, such as importing data from specific external databases, determining portions of the network that cluster together, finding the shortest path between two nodes, and more. Plug-ins can be developed by anyone with programming knowledge, using the Cytoscape Java code library. The librarians have focused on Cytoscape because University of Michigan researchers have developed plug-ins for it, and it is freely available for anyone to use regardless of their affiliation. Cytoscape can be used to create any type of network including social and physical networks, although the University of Michigan librarians have focused on gene and protein interaction networks. The two librarians teaching Cytoscape both have a biological sciences education, one with an undergraduate degree and the other with a master’s degree. The librarians have developed and delivered different types of instructional sessions for Cytoscape over the course of several years.

The librarians saw the opportunity for teaching classes on network visualization software to help fill researchers’ burgeoning need to find patterns in large data sets. As noted above, unlike certain subscription resources, finding training courses on these types of resources can be challenging. Cytoscape’s website has several tutorials [5] in addition to a user manual [6], but these do not have the advantages of an in-person training session, where participants can have their questions answered immediately and engage in collaborative learning. The librarians, therefore, developed a hands-on workshop for people to attend.

Stage 1: initial development

The first librarian-led Cytoscape training session using originally developed materials was offered at the University of Michigan in July 2009. To make the training more valuable to biomedical researchers, the tool was taught in a valid biological context. A biological case was presented and used throughout the instructional session in order to demonstrate Cytoscape functionality. Further course development led the librarians to a late 2009 article, summarizing two articles discussing three genes newly implicated in late onset Alzheimer’s disease [7]. These three genes—CLU, PICALM, and CR1—became the focus for the Cytoscape workshop. These novel gene associations for a well-researched inherited disease presented an interesting biological case study that allowed attendees of the training session to further investigate the new relationships between these genes that were previously considered unrelated.

The initial workshops covered the core features of Cytoscape, such as how to select nodes and edges or change visualization layouts (hierarchical, circular, etc.), and introduced several plug-ins: Michigan Molecular Interactions (MiMI), Enhanced Search, Shortest Path, and MCODE. The functionality of these plug-ins ranged from importing gene interaction data to isolating clusters of highly interconnected regions within a network (Table 1, online only). These plug-ins were chosen because they met at least one of the following criteria: were developed at the University of Michigan, had quality documentation available, or were frequently downloaded.

Stage 2: integrating changes

As course feedback was received, three more plug-ins were integrated into the Cytoscape workshop: MetScape, BINGO, and Agilent Literature Search (Table 1, online only). The addition of these plug-ins expanded on the type of data covered in the workshop. For example, the MetScape plug-in focuses on small molecules and compounds, in addition to genes. Adding these plug-ins required extending the training session period to two hours from one and a half. Presenting a longer training session did not seem to deter researchers from attending. Registration was full for the first two-hour long session, held in February 2011, for the classroom size of twenty-three people, with a wait list of ten (Figure 2).

Because of attendees’ interest in using their own data, a new handout was provided at the extended workshop that illustrated how to import data into Cytoscape. The librarians determined that demonstrating how to import researchers’ own data during the training was too time consuming. As a result, a handout walking users through importing network and attribute data was developed and distributed to session attendees. Thus, attendees could still get the necessary information for importing their data, but additional workshop time did not have to be designated to the task.

Stage 3: remote Cytoscape session

As the hands-on training sessions became more popular, the librarians realized they were being constrained by the physical training space. The training
The room only had twenty-three computers and was geographically located in a room that was difficult for registrants to find. Although hands-on training sessions are often preferred, there are many reasons why people are unable to attend them, such as having insufficient funds to travel to the training location or having time constraints. A remote training session is one solution to allow people to receive meaningful instruction. Blyth et al. found that the amount of money and time saved for many participants of the online faculty development workshops that they offered made this form of training worthwhile, even without in-person interaction [8].

When an individual outside of the state expressed interest in attending the Cytoscape workshop, the librarians decided a remote training session via webcast might be a valuable instructional format for reaching a broader audience, even though this type of training would lack the in-person component of the class.

The librarians developed a remote session that was first offered in March 2011 (Figure 2). The advantage of the additional format was to expand the audience, both within the University of Michigan and beyond the physical bounds of the school. Adobe Connect software, which was already licensed centrally by the university, was used to allow attendees to remotely stream the session from their own computers. The librarians had free access to a seminar room that had the existing technology hardware infrastructure and personnel support for Adobe Connect.

Attendees were asked to register to attend this remote session to allow the librarians to gather contact information for pre-training and post-training information and follow up. Before the day of the session, a librarian sent registrants a link to a web page containing the link for viewing the live session, along with course handouts. Because hands-on time was difficult to accommodate for a remote session, the session length was curtailed to one hour. A text chat box, staffed by a second librarian, was available for attendees to ask questions. The session was recorded to allow other interested individuals to view it in the future. As a result, the recording has been sent to several people outside of the University of Michigan who have expressed interest in Cytoscape training and has been posted on the Training web page of the National Resource for Network Biology [9]. The session recording also provided a point of need training format for researchers who were unavailable during the scheduled live remote session but could view the recording at a time that was convenient and useful for them.

**OUTCOMES**

Each stage of development of this workshop on a network visualization tool expanded the scope of the course. By adding content and offering instruction in different formats, the librarians have been able to reach a broad audience. Class attendees have included faculty, staff, researchers, graduate students, and postdoctoral fellows. Viewers of the remote sessions were not only from locations within the state of Michigan, but also various states throughout the contiguous United States, Puerto Rico, and Europe.

Based on class evaluations, the majority of the hands-on attendees felt they left the training with an understanding of how and why to use Cytoscape, showing a successful transfer of knowledge. Several respondents commented that they liked learning about the plug-ins. Due to the popularity of the Cytoscape webinar, a second one was held in September 2011. A follow-up survey was sent to session attendees (19 attendees and 2 site coordinators) the week after the first remote training session and immediately following the second training. Of the 17 people who responded, 16 people (94.12%) thought the webinar format was a useful method for learning about Cytoscape. In addition, 16 attendees (94.12%) said that the text chat box was
an appropriate way to handle questions. The continued interest in these training sessions, evidenced by the registration numbers (Figure 2), shows the value and importance of researchers receiving training on network visualization software, such as Cytoscape.

CONCLUSIONS

Stage 4: training the trainer

In the future, the librarians hope to teach a continuing education course for librarians on Cytoscape and VisANT, another free, open-source software tool used to visualize and analyze biological networks and pathways. The value of training the trainer would be to create a pool of instructors to further reach a larger audience of researchers who may benefit from these types of analytical tools.

Although there was a substantial learning curve for the librarians to initially learn the software, the number of training attendees and questions received about Cytoscape proved that the investment was valuable. The large registration numbers for the Cytoscape classes demonstrate the interest in and need for learning about network visualization software. These trainings have successfully introduced researchers to Cytoscape basics so that they can use it for their own work and have resulted in meaningful post-training interactions between the library and researchers. In the future, the librarians would like to expand on their Cytoscape trainings by introducing additional plug-ins and offering a session focused on importing personal data.

Librarians can take advantage of the opportunity to fill a training gap presented by the ever increasing number of tools being used for network visualization and analysis of large data sets. This case study of an institution’s development of a course on Cytoscape demonstrates how successful library instruction on these types of tools can be for researchers.

REFERENCES

1. Tennant MR. Bioinformatics librarian: meeting the information needs of genetics and bioinformatics researchers. Ref Serv Rev. 2005;33(1):12–9.
2. Wang L, Lipsey K, Murray C, Prendergrast N, Schoening P. The bioinformatics program at Washington University’s Bernard Becker Medical Library. Med Ref Serv Q. 2007 Summer;26(2):87–98.
3. Pavlopoulos GA, Wegener A, Schneider R. A survey of visualization tools for biological network analysis. BioData Min. 2008 Nov;1:12.
4. Suderman M, Hallet M. Tools for visually exploring biological networks. Bioinformatics. 2007 Oct;23(20):2651–9.
5. Cytoscape Consortium. Cytoscape user documentation [Internet]. The Consortium [cited 30 Jan 2012]. <http://www.cytoscape.org/documentation_users.html>.
6. Cytoscape Consortium. Cytoscape user manual [Internet]. The Consortium; 2006 [cited 30 Jan 2012]. <http://www.cytoscape.org/manual/Cytoscape2_8Manual.pdf>.
7. Van Es MA, van den Berg LH. Alzheimer’s disease beyond APOE. Nat Genet. 2009 Oct;41(10):1047–8.
8. Blyth RD, May MK, Rainbolt JG. Effective live online faculty development workshops: one model. EDUCAUSE Q. 2006;29(4):33–40.
9. National Resource for Network Biology. NRNB: training [Internet]. [cited 30 Jan 2012]. <http://www.nrnb.org/training.html>.

AUTHORS’ AFFILIATIONS

Marci D. Brandenburg, MS, MSI, mbradenb@umich.edu, Bioinformationist; Jean Song, MSI, jeansong@umich.edu, Research and Informatics Coordinator; Taubman Health Sciences Library, University of Michigan, 1135 East Catherine Street, Ann Arbor, MI 48109

Received February 2012; accepted April 2012