

Overview of Current Internet Resources for Molecular Pathology

Anthony A. Killeen

Department of Pathology, University of Michigan, Ann Arbor, Michigan

INTRODUCTION

The past few years have witnessed a dramatic expansion of information distribution across the global network of computers known as the Internet. The changes that accompany this development will affect almost every field of study, and certainly molecular pathology will be impacted by the ready availability of information stored in large genetic databases, the electronic publication of journals, and the rapid exchange of discussion between investigators and practitioners through this new forum. The purpose of this article is to review current Internet resources that relate to molecular pathology, and particularly to point out some of the many resources currently available. Where possible, reference is made to sites that actively maintain lists of resources in a given area. The bulk of Internet communications are presently handled by electronic mail, listservs and majordomos, Usenet, or are available through the World Wide Web. This review focuses primarily on these communication resources.

LISTSERVS AND MAJORDOMOS

For most Internet users, electronic mail (e-mail) remains the primary mechanism of electronic communication on the Internet. E-mail allows one to send files from one user to another. Although these files normally contain text, they can also contain audio and video images. As e-mail grew in popularity, users perceived the need to send mail to multiple recipients simultaneously. Anyone who has attempted to maintain a list of electronic addresses for individuals with some common interest who frequently share e-mail will be familiar with the difficulties of compiling and maintaining such a list. If each person in a group had to independently maintain his or her own list, there would be considerable redundancy of effort, and most likely significant errors.

Listservs (short for list servers), and the more recent Majordomo, are software programs that maintain lists of electronic addresses for mail distribution. The principal function of a listserv or majordomo is that it allows a single site to maintain a list of electronic addresses of people who are interested in a particular topic. An e-mail need be sent only to one address, that of the listserv or majordomo, which then automatically distributes the message to all members of the group, who each receive a copy of the e-mail in their mailbox. Most

listservs and majordomos have several additional features that make them remarkably useful tools such as archiving of messages for later review by any member of the group.

Finding Listservs and Majordomos of Interest

Currently, there are ~6,000–8,000 listserv groups in operation. A catalog of listservs arranged by subject is available on the World Wide Web (vide infra) at the following URL:

<http://www.tile.net/tile/listserv/index.html>

The names and addresses of some listservs related to laboratory medicine and pathology are shown in Table 1. There are currently two major listservs of interest to molecular pathology. One, MxDIAG-L is dedicated to molecular diagnostics and has >300 subscribers. The second, HUM-MOLGEN, is dedicated to the general field of human genetics and has >2,000 subscribers.

USENET

Usenet is a collection of discussion groups devoted to numerous topics. A person "posts" an article on some topic that can be read by anyone in the world with a Usenet reader and access to a Usenet feed. In turn, others can post a reply on this topic. A series of posts on a given topic is known as a "thread." Any given newsgroup may contain discussions on multiple threads at one time. Universities and commercial Internet providers commonly supply access to Usenet "feeds" or connections.

Usenet differs substantially from e-mail and listservs in several ways. First, the Usenet audience is almost entirely unknown. A message is not being posted to an individual, or even to the defined membership of a listserv group. Anyone with Usenet access can read a post and reply. Second, the potential reader of a message normally must actively go and look for messages that have been posted to a newsgroup of his or her interest. The messages are not sent to a recipient's

Received December 28, 1995; accepted December 28, 1995.

Address reprint requests to Dr. A.A. Killeen, Dept. of Pathology, University of Michigan, 1301 Catherine St., Ann Arbor, MI 48109-0602.

TABLE 1. Listservs for Pathology and Molecular Pathology

FAMCAN-L	listproc@moose.uvm.edu (Familial Cancer)
HUM-MOLGEN	listserv@nic.surfnet.nl (Human Molecular Genetics)
MxDIAG-L	listserv@health.state.ny.us (Molecular Diagnostics)
MEDLAB-L	listserv@vm.ucs.ualberta.ca (Laboratory Medicine)
NEPHRO-L	listserv@vm.ucs.ualberta.ca (Nephrology and Renal Pathology)
PATHNET	majordomo@ipas.afip.mil (General Pathology)
PATHO-L	listserv@emuvml.cc.emory.edu (General Pathology)
PEDPATH	listproc@u.washington.edu (Pediatric Pathology)
TX-L	listserv@anima.nums.nwu.edu (Transplantation & Immunology)

e-mail address. Third, the open access leads to a tendency of many groups to lose a clear focus. Some groups have elected to have all posts approved by a moderator to avoid this problem. On medically oriented groups one often sees posts from patients, or relatives of patients, who are looking for medical advice. What medico-legal liabilities one might incur in responding to such queries is, to my knowledge, unknown. Finally, it should be mentioned that many listserv groups have elected to automatically forward their posts to a related Usenet group. This has the advantage that people who have a minor interest in a particular topic can follow a discussion at their convenience on Usenet, rather than having messages delivered directly to their mailboxes from the listserv.

Finding Usenet Groups of Interest

Discovering Usenet groups of interest is accomplished in several ways. Services such as "Usenet Info Center Launch Pad" (<http://sunsite.unc.edu/usenet-i/>) and DeJaNews (<http://www.dejanews.com/>) maintain an index of Usenet groups. One can also review the list of all groups available through one's Usenet reader. It should be pointed out that the entire list of Usenet groups contains several thousand groups and that a local provider may not carry all of these. For convenience, Usenet groups are named on a hierarchical basis around major subject headings, known as top-level groups. Two top-level groups of interest to molecular pathology are the "bionet" and "sci" groups. Each of these contains a hierarchy of subgroups. For example, the group `sci.med.pathology` is a subgroup of the "sci" top-level group. Posts from listservs that are sent to Usenet are normally found within the "bit" hierarchy. The name "bit" originates in the word "bitnet," the network on which many listservs operate.

Usenet posts normally consist of simple text; however, posts can include audio and video images that have been appropriately formatted for Usenet distribution. It should be noted that the latter type of files are generally much larger than simple text files and can take longer to upload and download. Whereas most sites keep Usenet posts for only a relatively short period of time, typically a few weeks, several sites archive Usenet posts for longer peri-

ods. One such archival service that can be searched on-line is located at this URL:

<http://www.dejanews.com/dndn.html>

A list of some Usenet groups related to pathology and molecular pathology is shown in Table 2.

Usenet Software

Many universities and commercial Internet providers will supply a Usenet reader and an e-mail program that both run on a mainframe computer to which the user connects. For the user who wishes to install a Usenet reader on his or her own computer, the following programs are popular and can be obtained by FTP from the indicated addresses:

IBM-PC compatible (DOS and Windows):

Trumpet jazz.trumpet.com.au
or
ftp.trumpet.com

Macintosh:

NewsWatcher ftp://ftp.acns.nwu.edu/pub/newswatcher/
Nuntius ftp://ftp.cit.cornell.edu/pub/mac/comm/test/

Table 2. Usenet Groups for Laboratory Medicine and Pathology

General Interest Groups
bionet.diagnostics
sci.med.laboratory
sci.med.pathology
Specialty Interests
bionet.diagnostics.prenatal
bionet.molbio.ageing
bionet.molbio.bio-matrix
bionet.molbio.emblatabank
bionet.molbio.evolution
bionet.molbio.gdb
bionet.molbio.genbank
bionet.molbio.genbank.updates
bionet.molbio.gene-express
bionet.molbio.gene-linkage
bionet.molbio.gene-org
bionet.molbio.genome-program
bionet.molbio.hiv
bionet.molbio.methods-reagents
bionet.molbio.news
bionet.molbio.oncogenes
bit.med.mxdiag-1 (Molecular Diagnostics)
sci.med.aids
sci.med.diseases.cancer
sci.med.diseases.hepatitis
sci.med.immunology
sci.med.informatics
sci.med.prostate.bph
sci.med.prostate.cancer
sci.med.prostate.prostatitis

Netscape 2.0, a World Wide Web browser (vide infra) contains features for sending and receiving e-mail, and accessing Usenet in one package.

A more complete description of available Usenet software can be found at this URL:

<http://www.cis.ohio-state.edu/hypertext/faq/usenet/usenet-software/part1/faq.html>

It is important to remember that in addition to having Usenet software, one also needs access to a news feed that obtains articles posted on Usenet. Questions on how to locate an available news feed should be directed to one's system administrator.

WORLD WIDE WEB

The World Wide Web ("WWW", "W³", or simply the "Web") is a very versatile mechanism of presenting and retrieving data through the Internet. The Web is defined as a distributed, multimedia, hypertext system. The wide distribution of documents on the Web is accomplished via the Internet using a client/server protocol. In other words, a client computer requests that the server computer send it a document in an agreed transfer protocol. Each site that serves documents to the Web is identified by a unique, Uniform Resource Locator (URL), which is used by the client to locate the server. The documents may take the form of plain text, images, audio, or video files, and hence are multimedia. There are several commonly used transfer protocols including hypertext transfer protocol (http), file transfer protocol (ftp), gopher, telnet, and wais.

A key characteristic of the Web is the "hypertext" document. Hypertext refers to documents that contain "links" to other documents. In other words, a hypertext document is not simply a fixed entity, but contains pointers to other documents that can be retrieved simply by clicking on their representation in the first document. In turn, a linked document may contain links to other documents, generating an expanding web of related material that can be accessed by the user by following the links of interest. Thus the reader becomes an active participant in deciding which documents he or she wishes to retrieve. Moreover, it is possible to obtain a response from the reader through his or her Web browser relating to the contents of a document. A common application of this feature is obtaining a mailing address and credit card information by commercial vendors on the Web from their customers.

The ability to present information in various media such as text, video, and audio is clearly very advantageous for the scientific community. In the area of molecular pathology there are several important resources that should be mentioned:

- The World Wide Web Virtual Library contains an index of Web pages devoted to BioSciences (<http://golgi.harvard.edu/>

biopages.html). At the time of writing, there is no entry for molecular pathology; however, there is an excellent collection of links for molecular biology including links to databases, organizations, journals, commercial suppliers, universities, and some individual investigators.

- A listing of molecular biology resources by category is available from the NIH at

<http://molbio.info.nih.gov/molbio/servers.html>

- The National Center for Biotechnology Information (NCBI), a division of the National Library of Medicine, maintains a home page at <http://atlas.nlm.nih.gov:5700/Entrez/index.html>. This site serves several functions including storing information related to molecular biology, biochemistry, and genetics, maintaining databases including the GenBank (the DNA sequence database), and developing and providing computer software to analyze stored data.

- The Genome Database (<http://gdbwww.gdb.org/>) maintains an extensive collection of locus data for gene mapping and information about disease loci through the on-line version of McKusick's Mendelian Inheritance in Man (OMIM). The latter is an invaluable resource on genetic diseases, which is frequently updated.

- A listing of available software for molecular biology is maintained at

<http://www.gdb.org/Dan/softsearch/softsearch.html>

- A list of molecular biology servers is maintained by the U.S. Geological Survey at <http://info.er.usgs.gov/network/science/biology/index.html>

- Several sites have available protocols for common laboratory procedures involving nucleic acids including:

http://dna1.chem.uoknor.edu/protocol_book/protocol_index.html

Finding Web Sites of Interest

From its earliest stages, the Internet has been characterized by an absence of a central governing body. There is therefore no central system of comprehensively indexing information available on the Web. This, coupled with the marked annual increase in Internet servers, means that any survey of resources in a field is incomplete. There are several good search engines for the Web available on the Internet. These compile data by several means including asking users to submit the Uniform Resource Locator (URL) for their home page and by automated, systematic searching of the Internet for documents. Examples of popular search engines are:

http://www2.infoseek.com/	(InfoSeek)
http://www.yahoo.com/	(Yahoo)
http://www.lycos.com/	(Lycos)
http://webcrawler.com/	(WebCrawler)

Within the pathology community, there are several lists of resources maintained by individuals. These include the following:

http://www.pds.med.umich.edu/users/amp/path_resources.html
<http://www.cc.emory.edu/WHSC/medweb.pathology.html>
<http://worldmall.com/erf/sources.htm>

A listing of URLs for many departments of pathology is given in Table 3.

Academic publishing can take place very rapidly by placing papers on Web sites affiliated with official organizations, and a growing number of journals are available on-line. At this time there are no electronic publications devoted exclusively to molecular pathology, although these will certainly appear. An index of scholarly publications available on the Web is maintained at this URL:

<http://info.lib.uh.edu/pr/v6/n1/bail6n1.html>

World Wide Web Browser Software

Several Web browsers are available for most computer platforms. Perhaps the two most widely used on personal

computers are Netscape, Microsoft Explorer, and Mosaic. These are available from <ftp.netscape.com>, <http://www.microsoft.com> and <ftp.ncsa.uiuc.edu>. Increasingly, Web browsers such as Netscape are incorporating e-mail and Usenet modules into one seamless software package. A listing of available browsers and a discussion of their merits is maintained at this URL:

<http://www.boutell.com/faq/>

Operating a Web Server

Serving files on the Web involves several steps. First, the computer that is going to be a file server must be connected to the Internet and regularly available, usually 24 hours a day. This normally implies that the computer is dedicated to this purpose and that it is hardwired to the Internet. Although it is possible to serve Web documents from a Macintosh or IBM-PC computer, it is preferable that a departmental or institutional Web server operate on a more powerful platform such as a Unix-based computer.

Second, files served to the Web have to contain various formatting instructions written in Hypertext Markup Language

TABLE 3. Uniform Resource Locators for Departments of Pathology Home Pages

http://cobweb.aecom.yu.edu/aecom-courses.html (Albert Einstein College of Medicine, Medical School Syllabus)
http://rpiwww.mdacc.tmc.edu/ (M. D. Anderson Cancer Center Levit Radiologic-Pathologic Institute)
http://www.afip.mil (Armed Forces Institute of Pathology)
http://www.bgsu.wfu.edu/ (Bowman Gray, Wake Forest University)
http://biomedcs.biomed.brown.edu/ (Brown University, Division of Biology and Medicine)
http://edcenter.med.cornell.edu/Departments.html (Cornell University Medical College)
http://www.hitchcock.org/pages/path/index.html (Dartmouth-Hitchcock Medical Center)
http://www.emory.edu/PATHOLOGY/RESIDENT/ (Emory University School of Medicine, Information on Residency Program)
http://www.med.jhu.edu/pathology/research/index2.html (Johns Hopkins)
http://www.med.nyu.edu/Path/HomePage.html (New York University)
http://www.lsumc.edu/campus/path/default.htm (LSU Medical Center)
http://130.219.5.192/one.html (Robert Wood Johnson Medical School, Medical Informatics)
http://www.msccc.org/ (Sloan-Kettering Cancer Center)
http://www.mcl.tulane.edu/departments/pathology/pathtop.html (Tulane University)
http://www.path.usuf2.usuhs.mil/ (Uniformed Services University of the Health Sciences)
http://fester.his.path.cam.ac.uk/big/synapse/000p0025.htm (University of Alberta)
http://www-mp.ucdavis.edu/ (University of California, Davis)
http://www.med.ufl.edu/path/homepage.html (University of Florida)
http://www.med.uiuc.edu/PathAtlas/titlePage.html (University of Illinois, Urbana Urbana Atlas of Pathology)
http://www.pds.med.umich.edu/ (University of Michigan)
http://www.borg.labmed.umn.edu/ (University of Minnesota)
http://www.pathology.med.unc.edu/ (University of North Carolina, Chapel Hill, and Medical Examiner's Office)
http://www.med.upenn.edu/~path (University of Pennsylvania School of Medicine)
http://pathology.dental.upenn.edu/ (University of Pennsylvania School of Dental Medicine)
http://path.upmc.edu/ (University of Pittsburgh)
http://wwwminer.lib.rochester.edu/wwwml/Leon/URPLM.html (University of Rochester)
http://hyrax.med.uth.tmc.edu/home.htm (University of Texas, Houston)
http://www-medlib.med.utah.edu/WebPath/webpath.html (University of Utah WebPath Internet Pathology Laboratory for Medical Education)
http://www.pathology.washington.edu (University of Washington)
http://www.biostat.wisc.edu/educ/path/path703.html (University of Wisconsin Server for Medical School Pathology Course)
http://www.path.cam.ac.uk/ (University of Cambridge, U.K.)
gopher://gopher.ncl.ac.uk:70/11/admin/RSU/directory/medicine/pathsci (University of Newcastle, U.K.)
http://www.ccc.nottingham.ac.uk/~mpzjlowe/PathWWW.html (University of Nottingham, U.K.)
http://www.uniud.it/drrm/drrmeng.html (University of Udine, Italy)

(HTML). Although this topic is beyond the scope of this review, a few comments are in order. HTML indicates how a file is to be presented to the user of a Web browser. HTML instructions, or "tags," are not normally seen by the user, but they determine how the user's browser will display such things as the spacing of text, the use of italicized or bold characters, the placement of images, and importantly, the activation of hypertext links and any forms that require input from the browser's user. An excellent primer on HTML is available at:

<http://www.ncsa.uiuc.edu/General/Internet/Htmlprimer.html>

Third, the Web server must run software that is capable of serving documents to the Web. Software for this purpose is available, often as shareware, for various computer platforms. A more complete discussion of establishing a Web server, and a listing of Web server software is available at several sites including the following URLs:

<http://www.arnes.si/books/www-handbook/handbook.long.html>

<http://www.lib.ncsu.edu/staff/morgan/www-and-libraries.html>

Establishing a Web server enables one to make information available in text, audio, and video files. Examples of the kinds of files that a molecular pathology laboratory might consider making available are sample requirements, test menus, pricing and availability, guidelines for test interpretations, information on who to contact to discuss laboratory results, and teaching materials including text, audio, and video files. Most Web server software allows one to select whether information is to be made available to anyone on the Web or only to local users. With the expansion of the Web, these kinds of offerings will become increasingly common in molecular pathology laboratories.