



Drug discovery in the era of Facebook— new tools for scientific networking

David S. Bailey and Edward D. Zanders

BioVillage Ltd, St John's Innovation Centre, Cambridge CB4 0WS, United Kingdom

Social networking is beginning to make an impact on the drug discovery process. While bioinformatics and chemoinformatics underpin research at a scientific level, rapid communication between individual researchers across continents now allows the global exchange of ideas, tools and technologies. Networking at this level of speed and reach is quite a recent phenomenon. It facilitates the development of common interests, accelerates technology transfer and increases cooperative and competitive behaviour. In this review, we critically evaluate different web based networking approaches as effective resources for the drug discovery scientist. We also ask whether social networking sites will evolve into serious and credible resources for the drug discovery community.

The issues confronting the pharmaceutical and biotechnology industries are well known and much discussed. The slow introduction of new medicines into the marketplace and the loss of blockbuster revenues caused by the introduction of generic products have always been of primary concern. The consequence of these issues and other global trends is clear to all whose livelihoods depend upon employment in the drug discovery industry. Recent restructuring of large pharmaceutical companies and the precarious nature of investment in biotechnology are leading to significant job losses in the G8 economies. Many of these jobs are now being outsourced to high-growth economies such as China and India, including a recent trend towards outsourcing discovery research in medicinal chemistry and biology. This globalisation is being facilitated by the removal of communication barriers that result from different time zones and other geographical factors.

It is instructive to look at how these new communication tools are being used by individual researchers in industry and academia. To be successful, they must increase the interchange of ideas between scientists and assist in the optimal deployment of human resources worldwide. To be effective, they must help to increase the rate of discovery and development of new medicines.

The purpose of this article is to review current trends in scientific networking, particularly in relation to pharmaceutical research by individual chemists and biologists. We examine a range of net-

working sites and highlight features that have the potential to benefit individual scientists and their collective research efforts.

We define a networking website as one where individuals have an online presence that is visible to all who have access to the site. This presence may range from a name on a list of contacts to an 'online conversation' on message boards or blogs. Networking is possible either by linking one individual to another enabling one-to-one contact or by being made aware of groups and organisations that are working on similar problems, either as competitors or potential collaborators.

We consider four main categories of interactive networking sites with potential utility in drug discovery research:

- General social networking sites
- Business networking sites
- General science and/or medicine networking sites
- Drug-discovery-specific networking sites

There are also sites that exploit video technology for scientific and business information exchange, for example web seminars (webinars) and animations of processes to demonstrate experimental techniques, but these are not the primary focus of this review.

Quantitative analysis of community interactions in science

The mathematical tools of Network Theory have been applied to many different problems, from the structure of the Internet to biological communities at the population and molecular level

Corresponding author: Zanders, E.D. (ed.zanders@biovillage.co.uk)

[1,2]. A quantitative theory of human community behaviour has yet to be formulated, but an increasing number of studies have addressed the dynamics of human networks using such diverse data sets as scientific publications and even the teams required to stage Broadway musicals [3].

One aim of such work is to identify the features that contribute to productive scientific networks. Historically, scientific networks started as loose configurations centred on individual scientists, such as Robert Hooke and Robert Boyle—early members of the Royal Society [4]. Increasing specialisation in the 20th century has led to significant partnerships, such as Crick and Watson at the LMB, while the advent of large-scale collaborations such as the Human Genome Project has significantly increased the network size [5].

Drug discovery research has followed a similar path, from individuals such as Ehrlich and Von Baeyer, through small groups such as Fleming, Florey and Chain, to multinational drug development programmes comprising hundreds of individual researchers. An issue of *Nature* published today contains a similar number of Articles and Letters to one published in 1950, but about four times as many authors, as observed by Greene [6]. This supports the idea that large-scale team collaborations are becoming the norm, as does an extensive analysis of nearly 20 million publications and patents by Wuchty *et al.* (see Ref. [7], with commentary in Ref. [8]). This even applies in Mathematics where team collaborations have historically been seen as less important than in experimental science.

There are a number of possible reasons why larger teams may function more effectively (i.e. in research output, publications, among others). One of these is their increased efficiency in terms of time, money, shared technologies and expertise. As team members become dispersed, however, the difficulty of effective communication increases. While physically visiting these locations was made considerably easier through the introduction of mass air travel via the jumbo jet from the 1970s onwards, 30 years later we see the Internet and increasing computer power underpinning the second revolution in communication that is now underway.

Internet networks

Historical background

Portable computers and high-speed Internet connections are taken for granted today, and it is hard to believe how primitive these tools were a couple of decades ago. Laboratory managers were obliged to buy different types of computers to cater for each software platform (e.g. Microsoft Windows™, Apple Macintosh™ and Unix), a situation that has now been remedied with improved interoperability.

As the drug discovery industry was embracing new lab technologies, such as combinatorial chemistry and genomics in the 1990s, the idea of networked communities (eScience) was still fanciful, partly because of poor Internet connections and patchy penetration of the World Wide Web into countries with a modest science base. Although email communication was adopted rapidly, visual communication tools like PowerPoint presentation software were only embraced later in the decade. This also applied to the electronic distribution of images, which increased after the introduction of cheap digital cameras. Finally, and very importantly,

finding the location of relevant information on the Internet through search engines has improved dramatically, to the point where 'to Google' is now part of the modern lexicon.

The evolving Web

The potential uses of the Internet for scientific networking were recognised at a time before the communications infrastructure was able to turn visions into reality [9–11].

Now that computer and Internet technology allows greater speed and global accessibility the resulting improvements in networking and communication have given rise to new terminologies and concepts [12]. Prominent among these is the concept of the Internet currently evolving from Web 1.0 to Web 2.0 (with a Web 3.0. the 'semantic web' on the horizon).

Web 1.0 is the well-established Internet of static web pages and downloadable content ('pull' technology). Web 2.0, however, allows direct user control ('push' technology) of content on remote web pages. The Web 2.0 concept involves 'Harnessing the Collective Intelligence' of a community (<http://www.oreilly-net.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>). The tools developed to achieve this have spawned a new vocabulary of blogs, wikis, folksonomies, RSS feeds and mashups.

Blogs (web logs) are websites hosted by individuals or organisations that contain regularly updated personal commentary, descriptions of events, diary entries, and so on. Because others may contribute interactively in the form of online conversations, blogs are very powerful networking tools. Sometimes the content is delivered as audio in the form of podcasts, or video (vlogging).

Wikis (from the Hawaiian word for 'fast') are information resources that can be edited by the online community. The most famous example is Wikipedia (<http://en.wikipedia.org/>), an online encyclopaedia that is continually evolving as content is added, removed or modified.

Folksonomies (folk taxonomies) are essentially information tags for searching the Web that have been submitted manually by the online community according to individual preferences. This differs from the more controlled vocabularies used by search engine companies and online literature resources to find information on the Internet.

RSS feeds (Real Simple Syndication) are a form of aggregator that automatically pulls together updated information, such as news headlines or recent publications, and delivers the content directly to users as a single display on the web page.

Mashups are applications that combine data from more than one source to create a feature such as a generalised world map overlaid with specific information. An example of this is the community map for the Drug Design Resource (DDR), described later. These technologies are evolving rapidly and can be found within an increasing number of Internet networking sites.

We now review the four different categories of a site listed in the introduction from the perspective of the scientist who is an active member of the drug discovery community.

General social networking

Few Internet users are unaware of the social networking sites Facebook and MySpace, if only because of some high profile issues of fraud and abuse connected with open and transparent sharing of personal data. Facebook started as a college network and was

therefore mostly used by the 18–25 year age group. Unrestricted access has allowed its growth to over 60 million users worldwide. With this and related sites, the age demographic is changing, with an upward shift towards the 35 year and older population.

Publications dealing with the Facebook phenomenon are beginning to appear in the scientific literature (see *Journal of Computer-Mediated Communication* for a useful introduction; <http://jcmc.indiana.edu/vol13/issue1/boyd.ellison.html>). We are not aware of any specific literature on social networking sites and drug discovery, but Cain [13] has reviewed this area from the perspective of pharmacy education. Although social networking sites appear to lack rigour and literacy and are not subject to peer review they have many positive features. For example, they give scientists the opportunity to form networks regardless of their professional status. Those who participate and are willing to contribute to the network can benefit greatly by making new contacts almost instantly. Traditional means of networking, such as telephoning, sending emails or attending conferences, are still vital but, in the authors' opinion, can be much less efficient. The contacts established through social networking sites are also more likely to be productive because each member has voluntarily signed up to the same community and operates with the same rules of engagement.

To assess the potential utility of social networking sites for the drug discovery scientist, we have searched three major sites – Facebook, MySpace and Bebo – for community groups with an interest in drug discovery and development. While not exhaustive, this search reveals the presence of a small but serious drug discovery community on Facebook (Table 1). By contrast, only two groups feature on MySpace and there is zero activity on Bebo. Interestingly, the latter two sites do, however, display advertisements from suppliers of drug discovery tools and services, presumably in the hope that relevant groups will arise in the future.

In order to see how these drug discovery sites operate and to assess their usefulness, we examined the 'Bioinformatics' group on Facebook. This group has the largest membership of those surveyed (currently 1242). The site consists of a series of online bulletin boards entitled 'Posted Items', 'Discussion Board' and 'The Wall' along with a link to the group members. The material posted (consisting of job advertisements, queries about where to publish, how to find some specific technical information, highlighted resources and general musings about bioinformatics) seems to be spread fairly randomly over the different titles. An examination of members' photographs reveals few, if any, who appear older than their mid-40s. These scientists share information about jobs, resources and technical questions enthusiastically, and the informality of participation overcomes many of the inhibitions traditionally encountered within more hierarchical professional groups.

Is there growth potential for drug discovery groups on social networking sites? This may depend partly on the age demographic of scientists, because users seem to be mostly students and young investigators. Perhaps older scientists are less inclined to visit Facebook and similar sites, simply because they have little interest in the social aspects of networking. Younger scientists, by contrast, particularly those still in training, will continue to use these sites because the informality of Internet networking is a fundamental part of their lives. These lives are becoming ever more complex, so

TABLE 1

Network groups relevant to drug discovery on Facebook

Search term	Total number of groups	Number of relevant groups
(a)		
Drug discovery	48	11
Drug development	168	11
Medicinal chemistry	16	7
Genomics	19	11
Proteomics	15	12
Group name	Number in group	
(b)		
Bioinformatics	1260	
Pharmaceutical and medicinal chemistry	124	
Computational chemists	115	
Chemoinformatics	77	
Computer aided drug design and development	23	
QSAR	17	
Pharmacogenomics	3	
Biotechnology Networking Ltd.	392	
American Association of Pharmaceutical Scientists	93	
UK Pharma Industry Group	43	

(a) Survey using general search terms. (b) Specific groups by name and number of members (as of June 2008). The groups identified on Facebook can be classified into specific interest groups, recruiting vehicles, publicity for conferences, academic and industrial institutions. Not surprisingly, there is a strong representation of scientific disciplines related to computing (e.g. bioinformatics).

tools that allow them to participate in the social and professional scenes at the same time might be eagerly adopted.

Business networking

Traditionally, most business networking is conducted face-to-face at specialised conferences, such as those organised by Cambridge Healthtech Institute (CHI) and Techvision (Table 2). In some cases, this has been enhanced by the establishment of regional business networks, such as US-LBE and ERBI.

More recently, new scientific conferencing tools, such as the podcast and webinars, have emerged. Conference organisers such as CHI have embraced this aspect of the Web to deliver authoritative presentations to a global audience.

Social networking sites can, in principle, also be used to generate productive business relationships, but their unconstrained nature makes them less suitable for the world of business.

The power of the Internet as a networking tool has not been lost on the general business community; as a result there has been a proliferation of sites, such as LinkedIn, Ryze and Xing, which allow the establishment of personal networks in a more structured environment than the social sites. These are useful for building connections with specific individuals such as former work colleagues. From a drug discoverer's perspective, these are probably most useful for establishing relationships that allow new job or business opportunities to arise rather than group discussions on technical issues of interest to specialist groups.

General scientific and medical networking

Access to information underpins scientific enterprise. Early databases such as Index Medicus were only available in printed form, and quickly became outdated. Web technology has revolutionised

TABLE 2

Networking sites by category

	URLs
Social networking	
Bebo	http://www.bebo.com/
Facebook	www.facebook.com
MySpace	http://www.myspace.com
Plaxo	http://www.plaxo.com/
Windows live spaces	http://home.services.spaces.live.com
Business networking	
LinkedIn	http://www.linkedin.com/
Ryze	www.ryze.com
Xing	http://www.xing.com/
Zoominfo	http://www.zoominfo.com/
Scientific/medical networking	
Authoratory	
	http://www.authoratory.com/index.htm
Biomedexperts	http://www.biomedexperts.com/
Bioscreencast	http://bioscreencast.com/
Community of Science	http://www.cos.com/
Connotea	http://www.connotea.org/
ERBI	http://www.erbi.co.uk/
ExpertMapper	http://www.expertmapper.com/index.html
Faculty of 1000 Biology	
	http://www.f1000biology.com/home
Gopubmed	http://www.gopubmed.org/
Innocentive	http://www.innocentive.com/
KD Net Knowledge Discovery	http://www.kdnet.org/kdnet/control/about_kdnet
Life Science Executive Exchange	http://www.lifescienceexec.com/
Nature Network	http://network.nature.com/
Scintilla	http://scintilla.nature.com/
Sermo	www.sermo.com
ScienceRoll	http://scienceroll.com/
Drug discovery networking	
Australian Protease Network	
BlueObelisk.org	http://www.protease.net.au/index.php http://blueobelisk.sourceforge.net/wiki/Main_Page
CCL.Net	http://www.ccl.net/
Chemical Genomics Centre	http://www.cgc.mpg.de/
Chemspider	http://www.chemspider.com/
Collaborative Drug Discovery	http://www.collaborativedrug.com/
FBLD 2008	http://fbld2008.com/forum/
Kinase.com	http://kinase.com/
Protein Kinase Resource	http://www.kinasenet.org
Protein Phosphatase Research	http://www.phosphatase.net/
PubChem	http://pubchem.ncbi.nlm.nih.gov/
QSAR World	www.qsarworld.com
The Chemoinformatics and QSAR society	http://www.ndsu.edu/qsar_soc/index.htm
The Drug Design Resource	http://www.drugdesignresource.com
Conferences, podcasts and webinars	
Cambridge Healthtech Institute	http://www.healthtech.com/
BC Life Sciences	http://www.ibclifesciences.com/
Informa Life Sciences	http://www.iir-events.com/IIR-conf/LifeSciences/
Nature Publishing Group	http://www.nature.com/podcast
Techvision	http://www.techvision.com

information access and, today, Internet databases such as PubMed have become essential resources.

The Internet revolution has also provided a powerful impetus to scientific publishers to distribute journal articles directly to readers

via their computers. The more recent emergence of publisher-sponsored scientific blogs and other networking tools indicates how seriously Internet networking is being taken by major science publishers (see Nature Network, Scintilla and the clinical networking site Sermo for examples). These groups are more disciplined and focused than their equivalents on social (and some business) networking sites. All operate on similar principles and share features in common (e.g. message boards, posting of personal details, among others). Differences arise in their content, particularly literature citation.

Publication databases, such as PubMed, are often the first port of call for those wishing to get more information about a particular scientist, including details of their affiliation, expertise and any collaborators. New sites are beginning to exploit this by using Web 2.0 technology to reveal collaboration networks and expert opinion, for example Gopubmed, Connotea, Community of Science and Faculty of 1000 Biology. Web communities have also been established to address specific scientific topics, for example Alzheimer's disease [14] or proteomics [15]. Scientific applications of Web 2.0 (eScience, Science 2.0 [16]) will provide yet more networking opportunities for scientists from all backgrounds.

Networking in drug discovery

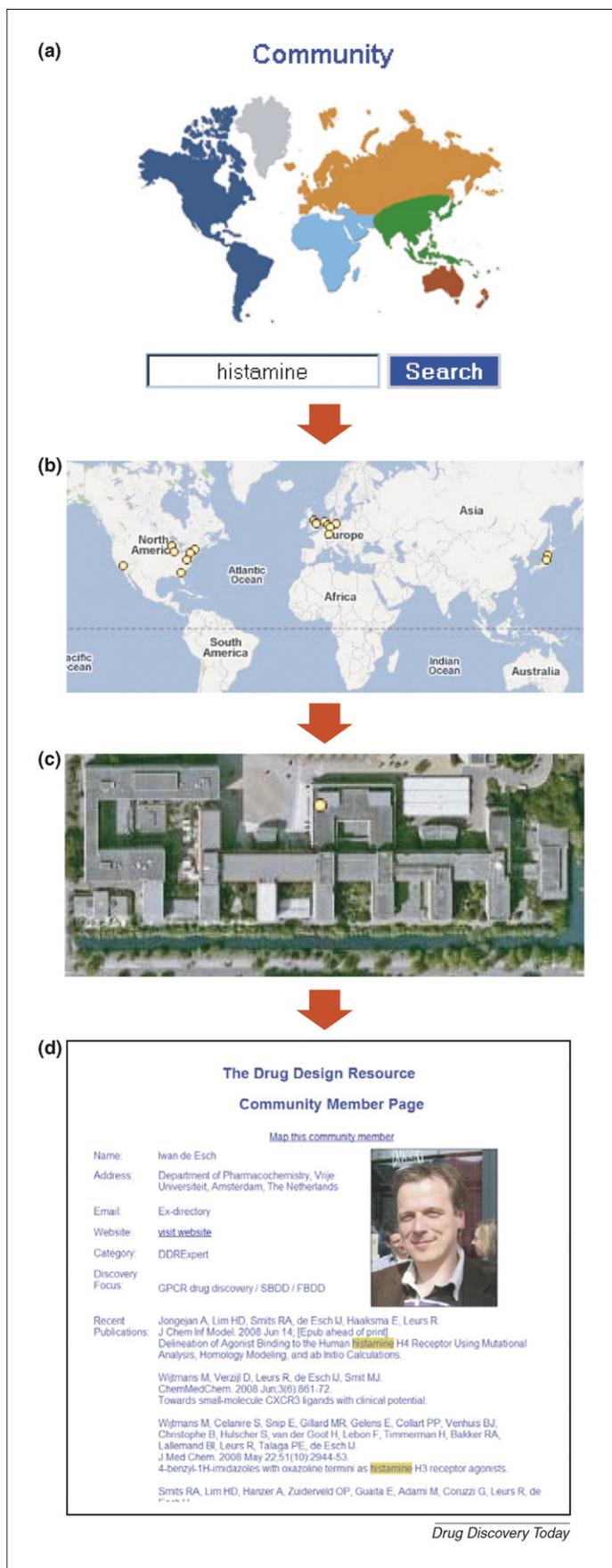
We now review some of the networking sites that have been created specifically for drug discovery scientists. A representative list is shown in Table 2, along with the social, business and general science networking sites described earlier.

Features vary between sites, some offering blogs and online discussion groups to draw out opinions from the community. Others are repositories of information, partly extracted from published resources or provided by the online community.

A broad group of scientific disciplines, including chemistry, biology, medicine and informatics, are required for drug discovery. Expert communities centred on these disciplines can share valuable information and experience on topics of mutual interest. In the case of drug design, the users could be specialists in chemoinformatics, bioinformatics, medicinal chemistry and/or protein structure.

The DDR is an example of such a specialist networking site, employing traditional information resources enhanced by Web 2.0 technologies. The latter includes RSS feeds of the latest drug design literature as well as map-based mashups of the community. Traditional information resources, including Web 1.0, define the drug discovery community as a static list of names, addresses, research interests, among others. Modern web technology has the potential to transform such lists into a community network where geographical proximities and wider scientific interactions can be readily visualised.

We have used the DDR as an example to illustrate these points in Figure 1, where the results of searching the drug design community using the search term 'histamine' are shown. Antagonists of histamine that interact with H1 or H2 receptors have previously proven to be highly effective drugs, and there is an active community in this area working on further receptors. The geographical spread of labs and individuals working on histamine drug design is shown in panel B; if the user zooms into a particular location (in this example a part of Europe) information can be displayed about key individuals and their scientific activities. This can include a list of drug discovery targets, recent publications and patents, contact details, personal blogs, and so on.



It is legitimate to ask what is different about this approach to information management compared with, for example, a conventional search engine or publication database. The main difference is the ability of the Web 2.0 resource to access, process and display relevant information very rapidly with a minimal number of clicks on the computer mouse. We have performed conventional searches for 'histamine' and 'drug design' using Google and PubMed. The former search engine returns over 91 000 hits that, although full of information, are unstructured and time consuming to work through. The PubMed search reveals 107 publications, which can be further processed by assigning them to specific authors and laboratories, thereby defining the community, albeit with difficulty. The Web 2.0 approach to the same area of histamine drug design instantly reveals the geographical distribution of key contributors to this field and eliminates the mountain of irrelevant information that consumes time that can be spent more productively.

Other resources (see Table 2) use the database model to focus on specific gene families (e.g. kinases, phosphatases and proteases) or data produced through specific technologies such as protein mass spectrometry [17] or small molecule crystallography [18]. A comprehensive list of biological databases is published annually in a special issue of *Nucleic Acids Research* [19]. These biologically orientated sites are paralleled by similar ones devoted to chemistry, such as Chemspider, PubChem and BlueObelisk.org. This topic has been reviewed recently by Antony Williams [20].

Finally, information on specific aspects of drug discovery (or indeed many other subjects) can be found in wikis such as Wikipedia. More specialised wikis exist, for example, the WISDOM Wiki (<http://wiki.healthgrid.org/Wisdom>) devoted to the *in silico* docking of compounds onto protein targets for malaria and influenza. This example of a distributed computing approach to drug discovery is yet another way in which Web 2.0 technology can connect a worldwide network of computers to extract valuable information.

Time saver or time waster?

The number of available resources is expanding, so a site should provide a means to an end, not be an end in itself (despite a growing impression to the contrary).

Peer review, although not without its flaws, is still considered to be the 'gold standard' of scientific quality control. Is this being undermined by science blogs and discussion groups? Peer reviewers have expertise in the area under consideration and are generally able to filter out information that will mislead other scientists. This is of course in contrast to easy access web based discussion groups where everyone has their say. Here, comments from the broader community of scientists may be of considerable value, but there is a clear danger of misinterpretation by those without expertise in the relevant field. These contrasting points have been highlighted recently in *Nature Geoscience* [21,22]. It is hard to avoid the conclusion that some editorial control of online community dialogue is required albeit administered with a light

FIGURE 1

Visualisation of the drug design community with an interest in histamine and its receptors. **(a)** Search term entry. **(b)** Geographical distribution of drug design scientists working on histamine. **(c)** European laboratory of individual scientist with expertise in drug design to histamine receptors. **(d)** Personal web page of a scientist showing contact details, drug targets of interest and publications.

hand; this is particularly important in the case of posting details of community membership on the website. If these details are uploaded without review there is a danger of abuse; we have noticed fictitious and highly inappropriate address and occupation details that were posted on a respectable chemoinformatics website because there was no human intervention in the hosting process. One of the attractions of community websites, blogs and wikis is their democratic open-to-all nature, but some control over content may be advisable if they want to be taken seriously by researchers.

More than a telephone directory

Community networking sites should be more than a passive list of members and their contact details. They should have features that encourage the scientist to visit frequently, despite the information overload that afflicts us all. One way of doing this is to provide a material reward or simply the challenge of solving a problem. A good example is the Innocentive website originally developed by Eli Lilly and Company. Contributors post specific technical problems and invite solutions from experts within the community. These problems were originally related to drug discovery, but the site has now expanded to cover a range of scientific, engineering and business disciplines.

Another way to encourage participation is to make accessing information as straightforward as possible. Search functions are extremely useful for rapidly accessing literature, conference and community data, but they must focus down on relevant information as quickly as possible, as noted above in the earlier search for individuals and laboratories working on 'histamine' in the DDR. Given the global expansion of laboratories, and the movement of scientists between them, a regularly updated resource is invaluable for keeping track of competitors or potential collaborators.

Conclusion

The world of scientific communication and networking is changing rapidly through the introduction of new web-based tools (Web 2.0). This evolutionary process mirrors its biological counterpart, in that the sites that survive for more than a few months or years will be those that provide value rather than distraction. We have given some examples of networking sites that we feel are relevant to drug discovery. The common theme is interactive participation, whether by signing up to a social or professional online network, contributing to a blog or discussion board, or creating and/or editing a wiki.

It has always been recognised that personal interaction is vital for seeding and developing scientific ideas. It is no coincidence therefore that enlightened laboratory planners deliberately create social areas to encourage the exchange of ideas over a meal or a drink. Web 2.0 technologies simply expand the number of people that can participate in these conversations, whether they are deep discussions or merely superficial badinage. Although the use of these networking tools is currently biased towards younger drug discovery scientists, we believe that this will change rapidly as their utility in essential professional activities such as job hunting, opinion seeking, data sharing and collaboration becomes apparent. There will be no alternative but to participate! As drug discovery scientists who started their careers before the Internet was born, we feel optimistic about the future of eScience and Internet-catalysed scientific networking and shall follow future developments with great interest.

Acknowledgements

We would like to thank the referees of this paper for their helpful comments.

References

- Newman, M.E.J. (2002) The structure and function of networks. *Comput. Phys. Commun.* 147, 40–45
- Jeong, H.B. *et al.* (2000) The large-scale organization of metabolic networks. *Nature* 407, 651–654
- Guimerà, R. *et al.* (2005) Team assembly mechanisms determine collaboration network structure and team performance. *Science* 308, 697–702
- Hartley, H., ed. (1960) *The Royal Society: Its Origins and Founders*, Royal Society, London
- Barabasi, A-L. (2005) Network theory—the emergence of the creative enterprise. *Science* 308, 639–641
- Greene, M. (2007) The demise of the lone author. *Nature* 450, 1165
- Wuchty, S. *et al.* (2007) The increasing dominance of teams in the production of knowledge. *Science* 316, 1036–1039
- Valderas, J.M. *et al.* (2007) Why do team-authored papers get cited more? *Science* 317, 1496–1498
- Smith, U.R. (1993) Networking on the Internet. *Trends Biochem. Sci.* 18, 398–399
- Bingham, C. (1998) Peer review on the Internet: a better class of conversation. *Lancet* 351, S10–S14
- Washer, P. (2002) Professional networking using computer-mediated communication. *Br. J. Nurs.* 11, 1215–1218
- Berners-Lee, T. *et al.* (2006) Computer science. Creating a science of the Web. *Science* 313, 769–771
- Cain, J. (2008) Online social networking issues within academia and pharmacy education. *Am. J. Pharm. Educ.* 72, 1–7
- Clark, T. and Kinoshita, J. (2007) Alzforum and SWAN: the present and future of scientific web communities. *Brief Bioinform.* 8, 163–171
- Smith, A.K. *et al.* (2007) LinkHub: a Semantic Web system that facilitates cross-database queries and information retrieval in proteomics. *BMC Bioinform.* 8 (Suppl. 3), S5
- Shneiderman, B. (2008) Science 2.0. *Science* 319, 1349–1350
- Murray, K.K. (2007) Mass spectrometry and Web 2.0. *J. Mass Spectrom.* 42, 1263–1271
- Murray-Rust, P. (2008) Chemistry for everyone. *Nature* 451, 648–651
- Database Issue (2008) *Nucleic Acids Research* 36
- Williams, A.J. (2008) Internet-based tools for communication and collaboration in chemistry. *Drug Discov. Today* 13, 502–506
- Schmidt, G. (2008) To blog or not to blog? *Nat. Geosci.* 1, 208
- Allen, M. (2008) Minority report. *Nat. Geosci.* 1, 209