

Providing web servers and training in Bioinformatics: 2010 update on the Bioinformatics Links Directory

Michelle D. Brazas, Joseph T. Yamada and B. F. Francis Ouellette*

Ontario Institute for Cancer Research, 101 College Street, Suite 800, Toronto, Ontario, Canada M5G 0A3

Received May 29, 2010; Accepted May 31, 2010

ABSTRACT

The Links Directory at [Bioinformatics.ca](http://bioinformatics.ca) continues its collaboration with *Nucleic Acids Research* to jointly publish and compile a freely accessible, online collection of tools, databases and resource materials for bioinformatics and molecular biology research. The July 2010 Web Server issue of *Nucleic Acids Research* adds an additional 115 web server tools and 7 updates to the directory at http://bioinformatics.ca/links_directory/, bringing the total number of servers listed close to an impressive 1500 links. The Bioinformatics Links Directory represents an excellent community resource for locating bioinformatic tools and databases to aid one's research, and in this context bioinformatic education needs and initiatives are discussed. A complete list of all links featured in this *Nucleic Acids Research* 2010 Web Server issue can be accessed online at http://bioinformatics.ca/links_directory/narweb2010/. The 2010 update of the Bioinformatics Links Directory, which includes the Web Server list and summaries, is also available online at the *Nucleic Acids Research* website, <http://nar.oxfordjournals.org/>.

COMMENTARY

The annual *Nucleic Acids Research* Web Server issue presented here, along with the annual Database issue also published by *NAR*, continues to be an invaluable resource for the scientific community. The 2010 Web Server issue highlights the latest web servers and open access bioinformatic tools available online to guide and enable research in any number of life science domains. The complete listing of URLs cited in the 2010 Web Server issue can be accessed online at the *Nucleic*

Acids Research website, <http://nar.oxfordjournals.org/>, as well as at http://bioinformatics.ca/links_directory/narweb2010/.

In partnership with *Nucleic Acids Research* since 2005, the Bioinformatics Links Directory (http://bioinformatics.ca/links_directory/) has collected and organized all of the Web Server issue's published links in its comprehensive public repository (1–5). The Directory is organized by biological subject with subcategories of common tasks relevant to each subject listed. All entries in the Directory contain a short description of the tool's function, as well as the accompanying PubMed citation and web server URL. Such information facilitates easy browsing of tools relevant for a particular biological subject, as well keyword searches to locate tools that suit a user's research needs.

This year's Web Server issue introduces an additional 115 web servers, plus 7 server updates (Table 1). The 2010 *Nucleic Acids Research* Web Server update brings the total number of servers and tools listed in the Bioinformatics Links Directory close to 1500 unique links. Since community input continually adds new links and non-functional links are removed, the list of web servers in Table 1 and online is in constant flux. The up-to-date complete listings accessible through the Bioinformatics Links Directory, including the *Nucleic Acids Research* 2010 web servers, can be accessed online at http://bioinformatics.ca/links_directory/.

As this list of bioinformatic tools, web servers and databases expands alongside new research technologies, data types and research ideas, there is a growing need for training and educational tools to enable and empower a wider and more varied audience of users.

The Bioinformatics Links Directory is only useful if accompanied by user guidelines and training assistance. The challenge then is to provide this training in a manner that addresses both the diversity of research arenas requiring bioinformatics training, and the increasing complexity of data sets and research questions.

*To whom correspondence should be addressed. Tel: +1 416 673 8511; Fax: +1 416 977 7446; Email: francis@oicr.on.ca

Table 1. Historical summary (2006–10) of the number of web servers listed in each subcategory of the Bioinformatics Links Directory

Name	2006	2007	2008	2009	2010 ^a
Computer Related					
Bio-* Programming Tools	20	20	20	20	18
C/C++	3	3	3	3	2
Databases	2	2	2	3	5
Java	4	4	4	4	4
Linux/Unix	12	11	11	11	10
PERL	5	5	5	5	5
PHP	3	1	1	1	1
Statistics	9	9	9	9	9
Web development	6	2	2	2	2
Web Services	6	7	7	10	17
Workflows					3
DNA					
Annotations	38	56	57	62	71
Gene Prediction	32	33	34	37	37
Mapping and Assembly	14	15	15	17	20
Phylogeny Reconstruction	37	43	46	49	53
Structure and Sequence Feature Detection	118	142	145	150	161
Sequence Polymorphisms	32	39	41	42	44
Sequence Retrieval and Submission	26	30	32	32	30
Tools For the Bench	55	63	65	71	76
Utilities	19	20	23	24	25
Education					
Bioinformatics Related News Sources	9	9	9	9	9
Community	19	24	23	23	18
Courses, Programs and Workshops	5	5	5	5	4
Directories and Portals	15	15	15	15	18
General	15	14	14	14	14
Tutorials and Directed Learning Resources	9	9	9	9	10
Expression					
cDNA, EST, SAGE	29	36	44	48	49
Gene Regulation	96	119	120	128	138
Transcript Expression and Microarrays	75	89	101	108	121
Protein Expression	8	9	17	22	23
Splicing	16	19	19	21	22
Networks			8	12	16
Gene Set Analysis				11	24
Human Genome					
Annotations	31	37	38	39	46
Ethics	7	8	8	8	6
Genomics	4	3	3	10	19
Health and Disease	14	19	23	27	29
Other Resources	25	29	29	29	31
Sequence Polymorphisms	25	33	36	38	46
Literature					
Goldmines	6	6	6	6	5
Open Access Resources	2	2	2	3	4
Search Tools	10	12	12	13	14
Text Mining	11	15	22	30	31
Model Organisms					
Fish	11	11	11	11	11
Fly	16	17	17	17	21
General Resources	23	27	28	29	32
Microbes	31	38	45	53	60
Mouse and Rat	32	35	35	36	42
Other Organisms	18	21	21	21	22
Other Vertebrates	10	10	10	10	11
Plants	16	19	21	25	28
Worm	9	9	9	9	10
Yeast	15	18	18	18	21
Other Molecules					
Carbohydrates	6	6	6	6	7
Metabolites		3	4	7	12
Small Molecules	3	6	6	9	13
Compounds			2	6	12
Protein					
2-D Structure Prediction	51	58	60	63	65
3-D Structural Features	53	70	75	85	100
3-D Structure Comparison	35	45	50	59	71

(continued)

Table 1. Continued

Name	2006	2007	2008	2009	2010 ^a
3-D Structure Prediction	48	59	60	70	83
3-D Structure Retrieval, Viewing	45	51	52	56	58
Biochemical Features	37	40	41	46	46
Do-it-all Tools for Protein	8	8	13	14	15
Domains and Motifs	86	112	115	121	124
Annotation and Function	35	44	47	53	57
Interactions, Pathways, Enzymes	66	88	94	107	125
Localization and Targeting	30	38	38	39	41
Molecular Dynamics and Docking	19	21	27	34	40
Phylogeny Reconstruction	36	44	45	53	54
Presentation and Format	13	14	14	14	14
Protein Expression	8	8	8	10	10
Proteomics	25	27	33	37	39
Sequence Data	7	8	9	10	10
Sequence Comparison			7	14	17
Sequence Features	25	31	33	38	46
Sequence Retrieval	27	29	29	31	29
RNA					
Functional RNAs	14	19	26	32	37
General Resources	10	10	10	10	11
Motifs	19	21	22	23	25
Sequence Retrieval	11	10	11	11	9
Structure Prediction, Visualization, and Design	38	47	54	58	62
Sequence Comparison					
Alignment Editing and Visualization	20	21	21	23	25
Analysis of Aligned Sequences	43	59	60	62	64
Comparative Genomics	26	33	35	37	48
Multiple Sequence Alignments	38	50	56	57	65
Other Alignment Tools	11	11	11	11	12
Pairwise Sequence Alignments	22	23	26	33	35
Similarity Searching	31	47	47	49	50

^aA complete listing of all URLs listed in the *Nucleic Acids Research* 2010 Web Server Issue can be accessed online at: http://bioinformatics.ca/links_directory/narweb2010

Below, we discuss some education ideas and initiatives that take aim at these challenges.

Brief tutorials and case examples to accompany new applications

Most new applications and bioinformatic tools are posted with supporting readme documentation that contains useful tips on how to navigate within and use a given tool. However, such documentation is often unread, and usually does not contain practical information on how the tool may be applied in research. Brief tutorials or case examples offer a mechanism to enhance the uptake of new computational tools by showing potential users how to perform simple to advanced analyses with the tool or how to use a tool for a given research problem. Good examples of applications with tutorials include many of the tools available through the National Center for Biotechnology Information (NCBI; <http://www.ncbi.nlm.nih.gov/guide/training-tutorials/>), the UCSC Genome Browser (<http://genome.ucsc.edu/training.html>) and Ensembl (<http://uswest.ensembl.org/info/website/tutorials/index.html>) to name a few.

Roadshow training programs at conferences in diverse fields

With the increasing number and diversity of research fields needing specific computational applications to address

their research problem, there are an increasing number of researchers realizing the need for some level of bioinformatic skills training. Conferences offer an ideal venue for hosting a discipline-specific bioinformatics training program, since a large number of researchers from a particular scientific field are already gathered. The International Society for Computational Biology (ISCB) Education Outreach Task Force has recognized this opportunity to empower potential users in applications relevant to their research and has initiated traveling road shows to accompany field-specific conferences with the aim to offer three to five workshops per year. Similarly other larger institutions offer traveling road shows for any of their applications [e.g. NCBI, European Bioinformatics Institute (EBI), etc.], although these are typically on a cost recovery basis.

Open access to collections of online training programs

Online lists or repositories of available educational training materials similarly enable users and potential new users to access and acquire computational skills in a particular application or computational work flow at their own pace and to suit their own research needs. Such lists are an invaluable source of information, and are often the starting point for many researchers. The Bioinformatics Links Directory (http://bioinformatics.ca/links_directory/) maintains an Education category and lists numerous education-related resources including

'Courses, Programs and Workshops' and 'Tutorials and Directed Learning' resources. Several of the links located in this directory redirect the user to open access, full content bioinformatics training workshops (e.g. Canadian Bioinformatics Workshops, <http://bioinformatics.ca>) and self-directed tutorials (e.g. OpenHelix, (<http://www.openhelix.com>)).

Effort must be made to continually improve education and training in bioinformatics, to keep pace with the rapid development of novel computational applications and the research communities' needs. The Bioinformatics Links Directory welcomes new entries both to its 'Education' section and other informatic resource sections. Suggestions for new links or updates and corrections to existing links may be submitted through email directly to links@bioinformatics.ca.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the efforts of *Nucleic Acids Research* and the researchers and developers worldwide who invest considerable effort into ensuring that their research is freely accessible to all. The Bioinformatics Links Directory is a community resource built on this commitment to the spirit of open access. In particular, the authors would like to acknowledge all of the contributors to the Bioinformatics Links Directory for their valuable input and suggestions for improvements to the directory; these individuals are listed on the

Acknowledgements page at http://bioinformatics.ca/links_directory/acknowledgements/.

FUNDING

Ontario Institute for Cancer Research through funding provided by the Government of Ontario. Funding for open access charge: The Open Access publication charge for this paper has been waived by the Oxford University Press, in recognition of the work on behalf of the *Journal*.

Conflict of interest statement. None declared.

REFERENCES

1. Brazas, M.D., Fox, J.A., Brown, T., McMillan, S. and Ouellette, B.F. (2008) Keeping pace with the data: 2008 update on the Bioinformatics Links Directory. *Nucleic Acids Res.*, **36**, W2–W4.
2. Brazas, M.D., Yamada, J.T. and Ouellette, B.F. (2009) Evolution in bioinformatic resources: 2009 update on the Bioinformatics Links Directory. *Nucleic Acids Res.*, **37**, W3–W5.
3. Fox, J.A., Butland, S.L., McMillan, S., Campbell, G. and Ouellette, B.F. (2005) The Bioinformatics Links Directory: a compilation of molecular biology web servers. *Nucleic Acids Res.*, **33**, W3–W24.
4. Fox, J.A., McMillan, S. and Ouellette, B.F. (2006) A compilation of molecular biology web servers: 2006 update on the Bioinformatics Links Directory. *Nucleic Acids Res.*, **34**, W3–W5.
5. Fox, J.A., McMillan, S. and Ouellette, B.F. (2007) Conducting research on the web: 2007 update for the bioinformatics links directory. *Nucleic Acids Res.*, **35**, W3–W5.