Journal of Integrative Bioinformatics, 10(1):226, 2013 http://journal.imbio.de/

Editorial

Ralf Hofestädt[1,*], Benjamin Kormeier[1], Matthias Lange[2], Falk Schreiber[2,3], Björn Sommer[1], and Stephan Weise[2]

[1]Bio-/Medical Informatics Department, Bielefeld University, Universitätsstraße 25,
33615 Bielefeld, Germany, `http://www.uni-bielefeld.de/`

[2]Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), OT Gatersleben,
Corrensstr. 3, 06466 Stadt Seeland, Germany, `http://www.ipk-gatersleben.de/`

[3]Martin Luther University Halle-Wittenberg, Institute of Computer Science,
Von-Seckendorff-Platz 1, 06120 Halle Germany, `http://www.uni-halle.de/`

1 JIBtools: a Strategy to Reduce the Bioinformatics Analysis Gap

The aim of the Journal of Integrative Bioinformatics (JIB) is to provide a free-of-charge open access and peer-reviewed platform for original research articles in all aspects of integrative bioinformatics. Over the last ten years, the editorial board accompanied the process and trends in integrative bioinformatics and perceives a strong tendency towards a glut of tools and databases. As a consequence, a strategic decision was made to reduce the increasing *Bioinformatics Analysis Gap* by adding an additional focus to JIB as a platform for tool-related publications.

1.1 Data as Blessing and Curse

The progress in molecular biology, ranging from experimental data acquisition on individual genes and proteins, over post-genomics technologies, such as RNA-seq, to phenotyping, proteomics, systems biology and integrative bioinformatics aims to capture the big picture of entire biological systems [1]. As a consequence of this revolution, the amount of data in the life sciences has exploded. The wave of new technologies, for example in genomics, is enabling data to be generated at unprecedented scales [2]. As of February 2013, NCBI Genbank provides access to more than 162 million sequences and PubMed comprises over 22 million citations for biomedical literature from MEDLINE, life science journals, and online books. The number of publicly available databases recently passed the high water mark of 1,512 [3]. This data deluge must now be harnessed and exploited.

Another aspect is the continuous development in information procurement, preparation and processing. Over the past years, information processing techniques evolved from library research and individual data archives to web-based systems using intercontinental high speed network links for an ad-hoc data exchange, cloud computing and distributed databases. This continuous and ongoing shift is attended by the use of database management systems (DBMS) which are applied to the management of increasingly complex data structures and voluminous content [4].

---

[*] To whom correspondence should be addressed. Email: ralf.hofestaedt@uni-bielefeld.de

doi:10.2390/biecoll-jib-2013-226

The key concepts in bioinformatics with regard to data handling are a consistent classification and unambiguous definition of the modeled biological objects in the databases, the raising use of ontologies connected with methods of knowledge processing, as well as information extraction and data mining [5, 6].

The first step to face this challenge is a raising need to find, extract, merge and synthesize information from multiple, disparate sources [7]. In particular, the convergence of biology, computer science and information technology will accelerate this multidisciplinary endeavor. The araising infrastructural requirements are summarized in [8]:

1. On demand access and retrieval of the most up-to-date biological data and the ability to perform complex queries across multiple heterogeneous databases to find the most relevant information.
2. Access to the best-of-breed analytical tools and algorithms for extraction of useful information from the massive volume and diversity of biological data.
3. A robust information integration infrastructure that connects various computational steps involving database queries, computational algorithms and application software.

A consequence of this development is the rising need for integration tools, such as SRS [9] or Entrez [10]. Regarding the literature, most of the publications in this research area still use specific workflows which integrate user-relevant data and analysis tools as specific pipeline solutions. Finally, the data warehouse approach has become very successful in this field of research [11, 12, 13, 14].

1.2 Heaven and Hell of Bioinformatics Tool Glut

Another key task in bioinformatics is data analysis. In this context, hundreds of new tools are published every year. In the domain of web services alone, 102 new publications were reviewed [15] for the annual NAR Web Server Issue. Nowadays, the problem is that no one has an overview of existing applications and new tools. We would like to call this the *Bioinformatics Analysis Gap*. In our case, analysis can be a simple statistical approach extending to complex simulations. Many of the analysis tools, which are available via the WWW, are based on database solutions. It is the idea of our integrative bioinformatics activity to focus on these kinds of tasks: molecular and life science databases, data integration, as well as data mining and analysis. Therefore, in 2003 we started the annual International Symposium on Integrative Bioinformatics. Based on this conference, we established the online Journal of Integrative Bioinformatics (JIB) in the year 2004. After nearly ten years of publishing JIB, the integration aspect becomes more and more important. Today, topics like systems biology and systems medicine focus exactly on this field of research. The main problem is that research is distributed over the entire world. Accordingly, the bioinformatics analysis gap is accompanied by an information gap. Besides the emerging of many new tools, many redundant tools are being developed. This problem could be reduced by a central collection of applications. Taking a look at the history of JIB, we have to mention that it has not been possible to close this gap until now. Therefore, we decided to modify the structure of JIB from the onset of 2013. In the future, JIB will place a stronger focus on the publication of tools (or information systems) or on new versions of already existing tools, which are based on integrative methods.

1.3 JIBtools – JIB's Strategy towards a Tool Ark

The JIB will be extended by JIBtools - the Journal of Integrative Bioinformatics tool lists. In the form of a collaborative curated website, a well-structured list of the most important integrative bioinformatics tools will be maintained. Obviously it will not be possible to give a complete list of all existing applications, but it is intended to provide a link to the most important available tools. First, we will start with a selected number of tools from the field of integrative bioinformatics. We would like to invite and motivate all bioinformatics software developers to contribute new tools.

Now, we introduce this new strategy and hope that we can help to reduce the Bioinformatics Analysis Gap by opening JIBtools. Because integrative bioinformatics is a very diverse and interdisciplinary field, each topic will be presented by an expert, who will be selected as an editor responsible for his/her category. The editors are invited to write a review to be published in JIB. Each tool will be described by an abstract, a citation (or a JIB link) and a web link to the tool.

We will invite each representative of a listed tool to send an invited JIB paper for the description and documentation of this tool. If a new tool is published in JIB, this tool will be automatically attached to JIBtools. It is our hope that this new strategy will finally help to systematically construct a well-categorized tool list of high value.

Starting in 2014, an annual JIBtools Issue will be published in JIB, containing an overview of all categories and updates, authored by the editors of JIBtools. Accompanied by this annual issue, editors of each tool list are responsible for checking listed tools for availability at least once a year to prevent the listing of outdated tools.

Everyone is invited to send a scientific paper describing an integrative bioinformatics tool to the Journal of Integrative Bioinformatics. Furthermore, we would like to ask you to be an editor of your area of research by sending your review paper to the Journal of Integrative Bioinformatics. Finally, JIBtools will have its own editorial board.

JIBtools is available at `http://journal.imbio.de/JIBtools`.

References

[1] H. Kitano. Systems Biology: A Brief Overview. *Science*, 295:1662–1664, 2002.

[2] E. Schadt, M. Linderman, J. Sorenson, L. Lee and G. Nolan. Computational solutions to large-scale data management and analysis. *Nature Reviews Genetics*, 11(9):647–657, 2010.

[3] X. M. Fernández-Suárez and M. Y. Galperin. The 2013 Nucleic Acids Research Database Issue and the online Molecular Biology Database Collection. *Nucleic Acids Research*, 41(D1):D1–D7, 2013.

[4] S. M. Stephens, J. Y. Chen, M. G. Davidson, S. Thomas and B. M. Trute. Oracle Database 10g: a platform for BLAST search and Regular Expression pattern matching in life sciences. *Nucleid Acids Research*, 33(suppl_1):D675–679, 2005.

[5] D. S. Roos. Bioinformatics–Trying to Swim in a Sea of Data. *Science*, 291(5507):1260–1261, 2001.

[6] L. Stein. The case for cloud computing in genome informatics. *Genome Biology*, 11(5):207, 2010.

[7] M. Krallinger, A. Valencia and L. Hirschman. Linking genes to literature: text mining, information extraction, and retrieval applications for biology. *Genome Biology*, 9(Suppl 2):S8, 2008.

[8] Z. Lacroix and T. Critchlow. *Bioinformatics: Managing Scientific Data*. Morgan Kaufmann Publishers, 2003.

[9] E. M. Zdobnov, R. Lopez, R. Apweiler and T. Etzold. The EBI SRS server – recent developments. *Bioinformatics*, 18(2):368–373, 2002.

[10] G. D. Schuler, J. A. Epstein, H. Ohkawa and J. A. Kans. Entrez: Molecular biology database and retrieval system. In R. F. Doolittle (editor), *Computer Methods for Macromolecular Sequence Analysis*, volume 266 of *Methods in Enzymology*, pages 141–162. Academic Press, 1996.

[11] S. Shah, Y. Huang, T. Xu, M. Yuen, J. Ling and B. F. F. Ouellette. Atlas - a data warehouse for integrative bioinformatics. *BMC Bioinformatics*, 6(1):34, 2005.

[12] T. Lee, Y. Pouliot, V. Wagner, P. Gupta, D. Stringer-Calvert, J. Tenenbaum and P. Karp. BioWarehouse: a bioinformatics database warehouse toolkit. *BMC Bioinformatics*, 7(1):170, 2006.

[13] D. Smedley, S. Haider, B. Ballester, R. Holland, D. London, G. Thorisson and A. Kasprzyk. BioMart - biological queries made easy. *BMC Genomics*, 10(1):22, 2009.

[14] A. Kasprzyk. Biomart: driving a paradigm change in biological data management. *Database*, 2011, 2011.

[15] G. Benson. Editorial. *Nucleic Acids Research*, 40(W1):W1–W2, 2012.