



OPINION ARTICLE

Matchmaking in Bioinformatics [version 1; referees: 2 approved]

Ewy Mathé ^{1*}, Ben Busby ^{2*}, Helen Piontkivska ^{3*}, Team of Developers

¹Department of Biomedical Informatics, College of Medicine, The Ohio State University, Columbus, OH, 43210, USA

²National Center for Biotechnology Information (NCBI), National Library of Medicine, National Institutes of Health, Bethesda, MD, 20894, USA

³Department of Biological Sciences and School of Biomedical Sciences, Kent State University, Kent, OH, 44242, USA

* Equal contributors

v1 **First published:** 09 Feb 2018, 7(ISCB Comm J):171 (doi: 10.12688/f1000research.13705.1)
Latest published: 09 Feb 2018, 7(ISCB Comm J):171 (doi: 10.12688/f1000research.13705.1)

Abstract

Ever return from a meeting feeling elated by all those exciting talks, yet unsure how all those presented glamorous and/or exciting tools can be useful in your research? Or do you have a great piece of software you want to share, yet only a handful of people visited your poster? We have all been there, and that is why we organized the Matchmaking for Computational and Experimental Biologists Session at the latest ISCB/GLBIO'2017 meeting in Chicago (May 15-17, 2017). The session exemplifies a novel approach, mimicking "matchmaking", to encouraging communication, making connections and fostering collaborations between computational and non-computational biologists. More specifically, the session facilitates face-to-face communication between researchers with similar or differing research interests, which we feel are critical for promoting productive discussions and collaborations. To accomplish this, three short scheduled talks were delivered, focusing on RNA-seq, integration of clinical and genomic data, and chromatin accessibility analyses. Next, small-table developer-led discussions, modeled after speed-dating, enabled each developer (including the speakers) to introduce a specific tool and to engage potential users or other developers around the table. Notably, we asked the audience whether any other tool developers would want to showcase their tool and we thus added four developers as moderators of these small-table discussions. Given the positive feedback from the tool developers, we feel that this type of session is an effective approach for promoting valuable scientific discussion, and is particularly helpful in the context of conferences where the number of participants and activities could hamper such interactions.

Open Peer Review

Referee Status:

	Invited Referees	
	1	2
version 1 published 09 Feb 2018	 report	 report

- Robert M. Blumenthal**, University of Toledo College of Medicine and Life Sciences, USA
- Guenter Tusch**, Grand Valley State University, USA

Discuss this article

Comments (1)



This article is included in the **International Society for Computational Biology Community Journal** gateway.



This article is included in the **Bioinformatics Education and Training Collection** collection.

Corresponding author: Ewy Mathé (ewy.mathe@osumc.edu)

Author roles: **Mathé E:** Conceptualization, Project Administration, Writing – Original Draft Preparation, Writing – Review & Editing; **Busby B:** Conceptualization, Project Administration, Writing – Original Draft Preparation, Writing – Review & Editing; **Piontkivska H:** Conceptualization, Methodology, Writing – Original Draft Preparation;

Competing interests: No competing interests were disclosed.

How to cite this article: Mathé E, Busby B, Piontkivska H and Team of Developers. **Matchmaking in Bioinformatics [version 1; referees: 2 approved]** *F1000Research* 2018, **7**(ISCB Comm J):171 (doi: [10.12688/f1000research.13705.1](https://doi.org/10.12688/f1000research.13705.1))

Copyright: © 2018 Mathé E *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution Licence](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Grant information: Ben Busby's work on this project was supported by the Intramural Research Program of the National Institutes of Health (NIH)/National Library of Medicine (NLM)/NCBI.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

First published: 09 Feb 2018, **7**(ISCB Comm J):171 (doi: [10.12688/f1000research.13705.1](https://doi.org/10.12688/f1000research.13705.1))

Introduction

Informal, face-to-face communication between participants is a vital piece of a scientific conference, just as important, if not more important, as formal activities such as keynote addresses and formal talk sessions (Saunders *et al.*, 2009). However, as the number of attendees grows, coupled with multiple research plenary sessions that often run concurrently (a regular feature of conferences in bioinformatics and other fields), the time available for individual contact with conference participants drops dramatically. Further, for new attendees, it can be difficult to navigate abstracts, posters, and talks to figure out the key people to engage with. While social media interactions via Twitter and other similar social media platforms (Biospace, 2009; Saunders *et al.*, 2009; Tachibana, 2014), or dedicated online communities (Budd *et al.*, 2015) have their own role in facilitating conversations, face-to-face conversations remain invaluable (Budd *et al.*, 2015; Fuller *et al.*, 2013).

Even for those of us who conduct most of our interactions online, face-to-face interactions can solidify relationships, spur novel ideas and research directions, further promote collaborations, and speed up project implementations. Moreover, it is critical for tool developers to carefully assess the utility (e.g., is their tool addressing an unmet need?) and usability (e.g., how streamlined and simple to use is the tool?) of their software. In the open source community especially, these aspects often tend to be overlooked or there are not enough resources to implement them (Al-Geel *et al.*, 2015). To assess utility and usability, developers need to establish a network of potential users, and need to get direct input from those users, including whether the software is sufficiently user-friendly to enable the user to focus on hypothesis- generation and testing in lieu of tool tweaking (Kumar & Dudley, 2007). These interactions can be key in addressing specific needs and/or offering a vision and/or a wish-list for further development (e.g., addition of new features).

For users, another source for finding tools of interest is via formal publication (peer-reviewed). However, this avenue is relatively slow, and is occasionally inefficient and/or insufficient in reaching a broader audience. Pre-peer-reviewed venues, e.g. bioRxiv, Figshare (Huang & Lapp, 2013), Zenodo, are trying to address this gap. Nonetheless, often the user's needs are not well articulated (or even formalized), and that's where face-to-face discussions can be much more helpful.

Developing novel tools that are usable to the wider community

While many tools are being developed, a relatively fewer number are routinely used by the larger biological and medical community. In fact, the average lifespan of an open-source Bioinformatics software is often relatively short, frequently limited by the transient nature of work contracts of developers, many of whom are post-docs or graduate students (Ahmed *et al.*, 2014). Through literature mining, a recent study reported that many database and software resources are mentioned only in the Bioinformatics literature, while only a fraction of the tools are mentioned in the biological and medical literature (Duck *et al.*, 2016). Specifically, only 5% of the resources account for 47% of

total usage and over 70% of the resources are only mentioned once in the literature (Duck *et al.*, 2016). This striking bias suggests that while the Bioinformatics community promotes development of novel software, the biological and medical communities only access a fraction of what is available. It is quite reasonable to think that these latter communities only access software that are intuitive and usable, and that perhaps usability could trump accuracy of analyses performed (Huang & Lapp, 2013; Pavelin *et al.*, 2012).

Of note, two broad approaches could be undertaken when developing Bioinformatics software. First, developers can develop a tool that solves a known issue in the field (e.g. RNA-seq analysis, omics integration), and then can seek users and data to test their approach and software. With this approach, it may be difficult for their tool to have visibility outside the Bioinformatics community, since 1) it is less likely that non-computational users are aware of your tool, and/or 2) your tool may not be user-friendly to non-computational users, and/or 3) your tool may not be readily adaptable to answer specific biological questions, or to accommodate a specific dataset format. With the surge in volume and variety of data types in high-dimensional biological data, adaptability is becoming more and more of a challenge. For example, a novel tool that integrates high-throughput omics data that is collected in the same samples may not be readily adaptable to data that is collected in different samples. Second, developers can develop Bioinformatics solutions that try to answer a specific biological or biomedical question, and can then broaden the utility of the tool by developing an associated software. Because the emphasis is on the biology, the resources and time available to generalize the software to other datasets are oftentimes lacking. This often results in a gap between a goal of developing a user-friendly software and 'on the ground' availability of low-level computational infrastructure (which is frequently scripting based) (Kumar & Dudley, 2007). We believe that this gap could be narrowed by further communication between biologists, computational biologists, clinicians, and users.

Importantly, it is worth noting that developers of widely adopted tools have often formally assessed utility and usability, enabling them to broadly disseminate their software. Guidelines for adopting a user-centered design when developing software have been formally assessed (Ahmed *et al.*, 2014; Pavelin *et al.*, 2012), and if applied, could yield highly usable software and could facilitate novel scientific discoveries. These formal assessments typically require face-to-face meetings between developers and users, and require developers to understand what problems need to be addressed, and how users will interact with the software. While taking these aspects into consideration prior to developing software can be lengthy, the resulting software will surely be useful and used by a wider community. Creating useful software can also provide a lot of job satisfaction to developers.

Reproducibility and software in biomedical research

Creating sustainable computational solutions can have a strong, positive impact on reproducibility of analysis results. With the recent rising concerns in reproducibility of scientific research

(Clark, 2017; Editorial, 2016), it is critically important to ensure that the analysis of large biological datasets is reproducible. More often than not, it is difficult to reproduce graphs and results in publications, and this is largely due to incomplete methods (e.g. parameters missing for statistical methods used, manual curation of results, etc.), and the use of in-house scripts or software. Methods for increasing computational reproducibility include reporting code and documentation used, and automating research analyses (Piccolo & Frampton, 2016). Computational frameworks, including but not limited to Taverna (Hull *et al.*, 2006; Wolstencroft *et al.*, 2013), Galaxy (Goecks *et al.*, 2010) and R markdown (Baumer & Udwin, 2015; Baumer *et al.*, 2014), facilitate reproducibility and oftentimes create reports that record all parameters used during the analysis. In addition to usability, developers can thus take into account the importance of reproducibility and in talking with users, better understand which parameters and analysis information needs to be reported.

ISCB/GLBIO'2017 conference

Hosted by the University of Illinois at Chicago, International Society for Computational Biology affiliate meeting, Great Lakes Bioinformatics Conference (ISCB/GLBIO'2017), has attracted a record 347 registered participants, including ~60% graduate students and post-docs with a broad range of computational and experimental expertise. First convened in 2006 as the Ohio Collaborative Conferences on Bioinformatics (OCCBIO), since 2010 joining forces with ISCB, over the years GLBIO has established itself as an ideal conference for showcasing the latest developments in analysis approaches and tools that span many different fields, and is a venue that attracts both computational and bench scientists. As we are all aware though, communication between computational and bench scientists can be challenging, particularly during the initial introduction stages when the overlap in mutual interests is not clear, and the match-making session that we ran is a first attempt at promoting such communication.

As Dr. Funmi Olopade (University of Chicago) mentioned in her keynote speech, clinicians, basic researchers, and computational biologists must better communicate to advance research. This sentiment is generally shared in the biological sciences, yet each field has its own language and culture. Encouraging communication across different fields via a common theme (e.g. RNA-seq analysis, chromatin accessibility analysis, etc.) is precisely what our matchmaking session aimed to accomplish.

Matchmaking for Computational and Experimental Biologists Session

The Matchmaking Session (Matchmaking@GLBIO session, #GenoMatch, #CompMatchBio) attracted over 40 participants, including 9 tool developers. The session, held at 8 am on the first day of the conference, kicked off with three short introductory talks, followed by multiple rounds of 4–5 minutes long small-table discussions led by individual tool developers, and then

open discussion. Short (10 minutes each) introductory talks by Drs. Ben Busby (NCBI), James Chen (OSU) and Ewy Mathé (OSU) covered available NCBI tools for RNA-seq analyses, approaches in integration of clinical and genomic data, and chromatin accessibility analyses, respectively. The purpose of these talks was to introduce broad topics that pose current, relevant topics and challenges in computational biology, and to present developers that are working on tools to address these challenges.

Next, small-table developer-led discussions were modeled after speed-dating. In each round, participants joined a table, listened to the developer's pitch, asked questions, discovered common interests, exchanged contact information, and then moved on to the next table. Because these small-table discussions were timed (4–5 minutes each), each participant had an opportunity to visit all the tables. At the end of "speed-dating" small-table discussions, participants still had 30–45 minutes available for further discussion. At this point, most users had identified developers that were presenting tools useful to them, and thus had the opportunity to discuss their own data needs in more detail.

Tools and representatives of tool developing teams (developers)

When planning the session, three main themes for tools were considered: analysis of RNA-seq, chromatin accessibility, and omics/multi-dimensional integration. A total 5 representatives of tool developer teams (Ben Busby, James Chen, Ewy Mathé, Arunima Srivastava, and Rick Farouni) were pre-registered for the session. However, at the start of the session, we asked whether other developers were interested in sharing their tool and, thus, were able to include 4 more developers. This near doubling of presenter-participants with a last minute change shows the level of interest that already exists in the community for sharing their tools. Table 1 lists all tools that were presented, with relevant reference information.

Feedback from presenters

As a follow-up to the session, developers were asked about their experiences afterwards, whether they had the sufficient opportunity to discuss their tools with potential users, and whether the subsequent interactions have occurred during the remainder of the conference. The majority of developers have found the session to be quite useful, in part due to the opportunity to network with many potential users, during the session or afterwards. Having time constraints for the matchmaking rounds have also allowed the session participants to quickly determine whether or not they were interested in learning about a specific tool in depth, and if the latter, move on to another tool.

Of note, the 5-minutes rounds were sufficiently long to accommodate exchange of contact information for subsequent follow-up, which occurred later during the conference functions and/or after the conference was over. The primary aim of the session was to provide face-to-face interactions between users and

Table 1. Tools highlighted by developers during the matchmaking session. Each developer had a chance to showcase their tool and to further discuss its usage with potential collaborators during the “speed-dating” small-table discussions.

Tool name	Presenters	Publication/Website
Clust: Optimized consensus clustering of one or more heterogeneous gene expression datasets (e.g. Microarrays and RNASeq)	Basel Abu-Jamous and Steven Kelly	https://github.com/BaselAbujamous/clust
ProcessDriver: Tools that computes copy-number based cancer drivers and associated dysregulated biological processes GSEPD: An R package to compute differentially expressed genes, enriched GO terms and projection-based clustering of samples	Serdar Bozdog	B. Baur and S. Bozdog. <i>ProcessDriver: A computational pipeline to identify copy number drivers and associated disrupted biological processes in cancer</i> . Genomics, 2017, 109(3–4): 233–240. https://github.com/brittanybaur/ProcessDriver
RNA-seq resources at NCBI	Ben Busby	https://www.ncbi.nlm.nih.gov/guide/dna-rna/
MatchTX: An automated learning system for patient cohort matching using high-dimensional genomic data	James Chen	www.match-tx.com
Kover: A machine learning tool to learn interpretable models of phenotypes from k-mer data	Alexandre Drouin	https://github.com/aldro61/kover Drouin, A., Giguère, S., Déraspe, M., Marchand, M., Tyers, M., Loo, V. G., ... & Corbeil, J. (2016). <i>Predictive computational phenotyping and biomarker discovery using reference-free genome comparisons</i> . BMC genomics, 17(1), 754.
ALTRE: workflow for defining ALTERed Regulatory Elements using chromatin accessibility data	Rick Farouni	https://github.com/mathelab/altre Baskin E., Farouni R. , Mathé E.A. <i>ALTRE: workflow for defining ALTERed Regulatory Elements using chromatin accessibility data</i> . Bioinformatics 2017; 33 (5): 740–742.
IntLIM: Integration of metabolomics and gene expression data	Ewy Mathé	https://github.com/mathelab/intlim
SeqclusterViz: Small RNASeq visualization	Lorena Pantano	https://github.com/lpantano/seqclusterViz https://f1000research.com/posters/6-673
OSUMO: Multi-Omic data utilization and patient stratification	Arunima Srivastava	https://github.com/osumo/

developers, and to provide ample opportunities for contact information exchange. Per feedback we received afterwards, this aim appeared to be successfully accomplished.

Future matchmaking sessions

We plan to build up and expand on our successful experiment during GLBIO’2017, to offer similar matchmaking sessions at other ISCB venues, such as ISMB in Chicago in 2018, and GLBIO in 2019 Madison, WI. We have already run an informal session at the ISCB DC-RSG summer workshop in College Park, Maryland (July 12, 2017) with lightweight planning, enormous popularity, and a very positive response.

In the future, to broaden participation and improve participants’ experience, presenters/developers will be given the opportunity to prepare and present 1-2 slides about their tools at the beginning

similar to ‘flash talks’. This format will help developers to find other developers interested in solving similar problems. In our first matchmaking session, developers had little time to interact with each other during the session. In the future this flash talks-format could replace the broad, introductory topic-focused talks given at the beginning of the matchmaking session. Notably, though, these flash talks will not replace the small-table matchmaking portion of the session, which we believe is critical to foster communication between users and developers.

Lastly, it is important to note that this session was scheduled at 8 am at the start of the conference. While we had anticipated lower participation due to this scheduling (assuming that a number of participants would chose to come in later on the first day to avoid traveling the Sunday prior to the start of the conference), the timing of the session turned out to be advantageous. Indeed,

having a discussion-promoting, interactive session as a start of the conference is a great way to engage participants and “break the ice” for subsequent interactions during the conference. Further, it provides ample time for attendees to find each other later during the conference and formalize potential collaborations.

Conclusions

The short-talk/“speed-dating” format provided a platform in which participants could learn about as many tools as possible in a short period of time, while making valuable connections across fields. Given the fast moving pace of Bioinformatics and the rapid advances across clinical/experimental biology fields, it is critical to keep the communication lines open between the communities. Our matchmaking session opened these communication lines by facilitating informal face-to-face interactions.

Data availability

All data underlying the results are available as part of the article and no additional source data are required.

Competing interests

No competing interests were disclosed.

Grant information

Ben Busby’s work on this project was supported by the Intramural Research Program of the National Institutes of Health (NIH)/National Library of Medicine (NLM)/NCBI.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements

We would like to thank all the co-organizers and GLBIO participants for their contributions to the success of our session, and Belinda Hanson and Dr. Tandy Warnow for their help in developing the session. We would also like to thank developers that presented at the Matchmaking for Computational and Experimental Biologists Session, including Basel Abu-Jamous, Steven Kelly, Serdar Bozdog, James Chen, Alexandre Drouin, Rick Farouni, Lorena Pantano, and Arunima Srivastava.

References

- Ahmed Z, Zeeshan S, Dandekar T: **Developing sustainable software solutions for bioinformatics by the “Butterfly” paradigm [version 1; referees: 2 approved with reservations]**. *F1000Res*. 2014; **3**: 71.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Al-Ageel N, Al-Wabil A, Badr G, *et al.*: **Human Factors in the Design and Evaluation of Bioinformatics Tools**. *Procedia Manufacturing*. 2015; **3**: 2003–2010.
[Publisher Full Text](#)
- Baumer B, Cetinkaya-Rundel M, Bray A, *et al.*: **R Markdown: Integrating a reproducible analysis tool into introductory statistics**. arXiv preprint arXiv:1402.1894. 2014.
[Reference Source](#)
- Baumer B, Udwin D: **R markdown**. *Wiley Interdisciplinary Reviews: Computational Statistics*. 2015; **7**(3): 167–177.
[Publisher Full Text](#)
- Biospace: **Why Social Networking Is Important for a Bioinformatics Developer**. 2009; Retrieved on August 16, 2017.
[Reference Source](#)
- Budd A, Corpas M, Brazas MD, *et al.*: **A quick guide for building a successful bioinformatics community**. *PLoS Comput Biol*. 2015; **11**(2): e1003972.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Clark TD: **Science, lies and video-taped experiments**. *Nature*. 2017; **542**(7640):139.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Duck G, Nenadic G, Filannino M, *et al.*: **A Survey of Bioinformatics Database and Software Usage through Mining the Literature**. *PLoS One*. 2016; **11**(6): e0157989.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Editorial: **Reality check on reproducibility**. *Nature*. 2016; **533**(7604): 437.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Fuller JC, Khoueiry P, Dinkel H, *et al.*: **Biggest challenges in bioinformatics**. *EMBO Rep*. 2013; **14**(4): 302–304.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Goecks J, Nekruteno A, Taylor J, *et al.*: **Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences**. *Genome Biol*. 2010; **11**(8): R86.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Huang D, Lapp H: **Software Engineering as Instrumentation for the Long Tail of Scientific Software**. *Figshare*. 2013.
[Publisher Full Text](#)
- Hull D, Wolstencroft K, Stevens R, *et al.*: **Taverna: a tool for building and running workflows of services**. *Nucleic Acids Res*. 2006; **34**(Web Server issue): W729–732.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Kumar S, Dudley J: **Bioinformatics software for biologists in the genomics era**. *Bioinformatics*. 2007; **23**(14): 1713–7.
[PubMed Abstract](#) | [Publisher Full Text](#)
- Pavelin K, Cham JA, de Matos P, *et al.*: **Bioinformatics meets user-centred design: a perspective**. *PLoS Comput Biol*. 2012; **8**(7): e1002554.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Piccolo SR, Frampton MB: **Tools and techniques for computational reproducibility**. *Gigascience*. 2016; **5**(1): 30.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Saunders N, Beltrão P, Jensen L, *et al.*: **Microblogging the ISMB: a new approach to conference reporting**. *PLoS Comput Biol*. 2009; **5**(1): e1000263.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Tachibana C: **A scientist’s guide to social media**. *Science*. 2014; **343**(6174): 1032–1035.
[Publisher Full Text](#)
- Wolstencroft K, Haines R, Fellows D, *et al.*: **The Taverna workflow suite: designing and executing workflows of Web Services on the desktop, web or in the cloud**. *Nucleic Acids Res*. 2013; **41**(Web Server issue): W557–561.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Open Peer Review

Current Referee Status:  

Version 1

Referee Report 27 March 2018

doi:[10.5256/f1000research.14887.r30752](https://doi.org/10.5256/f1000research.14887.r30752)



Guenter Tusch

Grand Valley State University, Allendale, MI, USA

The authors discuss a unique experimental session that they initiated at the ISCB/GLBIO'2017 meeting in Chicago (May 15-17, 2017) in form of an opinion article. Based on the model of speed dating they teamed up interested parties with developers of bioinformatics software in order to connect those developers with potential users. The paper consists of roughly four parts. It starts with a brief introduction including a plea for the importance of face-to-face interactions at conferences and a description of options researchers have today to find the appropriate computer software to support their research projects. The next part describes the software development process for bioinformatics software as seen by the authors. They claim that there are basically two approaches that I would call developer-centric and research-centric. The first one seems to assume that developers develop a more general tool, but have difficulties to connect to potential users, while the other one apparently results in a program that suffers from a lack of general usability due to a too narrow focus on a specific biological problem. I'm not quite sure if this is based on the NCBI experience of one of the co-authors and how tools like Bioconductor would fit in here. There is certainly a problem for small scale software projects like those developed for one particular research project. That could be clarified with specific examples possibly from participants in the matchmaking session.

The next part of the paper describes and discusses the session at the conference, emphasizing the focus on face-to-face communication, setup, implementation, feedback of presenters, and future plans. I thought of this as the essential part of the paper. Finally, as a third part the authors included a table with the description of the presented software and contact information.

I believe that this experimental session is a very interesting and important approach, and the authors make very valuable points about the setup and implementation of the session and the outcomes especially for younger researcher. From the success the authors had I hope there will be more sessions like that at future conferences. Of course, the conclusions need to be preliminary based on only one session, however the authors make strong points that many results can be generalized. While the introductory and the second part feel like a unit and are the only ones referred to in the conclusion, I feel like the second part and the table are not really integrated enough. They deal with important aspects of the topic, if the topic is not a mere description of the whereabouts of the session and the conclusions drawn by the authors. If the purpose is purely informative, it could be largely reduced, but if it is part of the argument – as I assume, see also my comments above -, it should be included in the discussion and conclusion, and that would strengthen the message. I also agree with the comments of the other reviewer.

In conclusion, the authors discuss a very interesting and promising approach to improve communication and personal connections especially for younger researcher in the bioinformatics community.

Is the topic of the opinion article discussed accurately in the context of the current literature?

Yes

Are all factual statements correct and adequately supported by citations?

Yes

Are arguments sufficiently supported by evidence from the published literature?

Yes

Are the conclusions drawn balanced and justified on the basis of the presented arguments?

Yes

Competing Interests: No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Referee Report 26 February 2018

doi:10.5256/f1000research.14887.r30750



Robert M. Blumenthal

Department of Medical Microbiology and Immunology, Program in Bioinformatics, University of Toledo College of Medicine and Life Sciences, Toledo, OH, USA

This manuscript summarizes experience and justification for a rapid developer-user meeting format, which was first implemented at the 2017 GLBIO-ISCB meeting. It is a useful summary and may stimulate others to try similar approaches. My comments are entirely on ways to clarify the writing, because the content is fine as is.

- P3 Para2: The heavy use of “e.g.” is distracting and unnecessary – suggest just leaving it out.
- P3 Para4: Top line, “fewer” should be “smaller”; 3rd line delete “an”; 4th line delete “often” (since you use the word “average”). Next column (same para), add a comma after “total usage”; near bottom of para replace “are” with “is” before “intuitive”.
- P3 Para5: 3rd line delete “e.g.”; 7th line, replace “since” with “for one or more of the following reasons:” and delete both occurrences of “and/or”; 12 lines from bottom replace “Second” with “In the second broad approach”; and 3 lines below that remove “an”.
- P3 Para7: replace “analysis” with “analytic”.
- P3 and throughout: Is it F1000Research style to capitalize “Bioinformatics” with every use?

- P4 Para2: top line add “the” before “International”; 3rd line remove “has”; 7th line add “and” before “since”.
- P4 Para3: remove “e.g.”
- P4 Para7: remove “have” before “also allowed the session”.
- P5 Para2: unclear what is meant by “lightweight” – please clarify.

Is the topic of the opinion article discussed accurately in the context of the current literature?

Yes

Are all factual statements correct and adequately supported by citations?

Yes

Are arguments sufficiently supported by evidence from the published literature?

Yes

Are the conclusions drawn balanced and justified on the basis of the presented arguments?

Yes

Competing Interests: No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Discuss this Article

Version 1

Reader Comment 21 Feb 2018

Basel Abu Jamous,

We have published a preprint manuscript for the clust tool at bioRxiv:

<https://www.biorxiv.org/content/early/2018/02/13/221309>

Basel Abu-Jamous and Steven Kelly (2018) Clust: automatic extraction of optimal co-expressed gene clusters from gene expression data. bioRxiv 221309; doi: 10.1101/221309

Please cite this preprint if you were to use the method.

Thanks

Basel Abu-Jamous

Competing Interests: No competing interests were disclosed.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research