IN FOCUS

Social Interactomes for Enabling Research Communities

Justin Guinney[1], Rodrigo Dienstmann[1], Charles Ferté[2,3,4,5], Stephen Friend[1], and Frank McCormick[6,7,8]

Summary: Data-driven analyses of scientific abstracts with web apps such as "abstract interactomes" provide a new visualization tool for the biomedical research community to interactively navigate a rich assembly of investigators and identify common research topics. Alternative conference formats such as "social interactomes," with structured, albeit informal, discussions among attendees, are able to engage fellows and top investigators, facilitate the exchange of ideas, and encourage data sharing and future collaborations. *Cancer Discov; 4(11); 1265–8. ©2014 AACR.*

INTRODUCTION

Biomedical conferences such as the AACR Annual Meeting serve an important role in the education of scientists and the public alike, through short talks, invited lectures, and poster presentations. These conferences also provide a gathering spot for researchers from all over the world, offering opportunities for direct and personal interactions. Even with the increasing digital connectedness of researchers worldwide, the popularity of scientific conferences points to the important role of "face time" in the scientific enterprise. This raises the important question of whether conferences are achieving their full potential as facilitators of science, and whether there may be alternative formats that are able to improve dialogue and the exchange of ideas among researchers.

In 2013, with the support of AACR, we introduced a unique and experimental session at the Annual Meeting to foster openness and collaboration within a research community. For this pilot project, we targeted researchers working on the *RAS* oncogene and its related pathways, as this is currently a large and active research area. A major goal of the project was to blend data-driven analysis of scientific abstracts with human surveys to develop a snapshot of activity within this community, and to provide a clearer understanding of where the major interests and controversies lie. Another objective was to facilitate group discussions around major research topics among RAS investigators to encourage data sharing and new collaborations. The success of the first RAS session in 2013 led to repeat sessions at the 2014 AACR RAS Oncogenes Conference and the 2014 AACR Annual Meeting. In this article, we present a review of the sessions and examine whether these alternative formats serve a currently unmet need at biomedical conferences.

INTERACTOME

The scientific abstract is the *sine qua non* for encapsulating recent activity of researchers and for the selection of oral and poster presentations by conference organizers. Yet, as a tool for navigating the scientific landscape at conferences, the abstract remains largely underutilized. Our goal, therefore, in building an "interactome" application was to provide conference attendees with a tool to explore their field of interest in a data-driven manner and to identify ideas and people pertinent to their own work. To this end, we applied machine-learning algorithms to the conference abstracts to identify clusters of topics and authors (described below) and embedded the results in an interactive online tool.

We applied the following analytical pipeline to all abstracts submitted to the AACR meetings: (i) We first reduced the set of abstracts to those containing key words related to the RAS community (e.g., KRAS, NRAS, BRAF); (ii) we then applied a lexical "bag-of-words" analysis to the abstracts, including text normalization (i.e., stop-word removal, stemming, punctuation remove, term frequency/inverse document frequency weighting) that resulted in an "n-by-n" abstract similarity matrix; (iii) we applied an iterative clustering algorithm to identify hierarchical groups of related abstracts (1); and (iv) finally we embedded the results into an interactive web tool. Separate interactomes were generated for the 2013 AACR Annual Meeting, the 2014 RAS Oncogenes Conference, and the 2014 AACR Annual Meeting; these applications are accessible from the following url: http://cancer.sagebase.org/interactomes.html. Figure 1 shows a snapshot of the 2014 AACR abstract interactome.

In addition to views of abstract similarity based on a lexical analysis, we also constructed interactomes using geographic location and topic categories. For the geographic clustering, we used the city of the first author to extract GPS coordinates, computed geodesic distances between authors, and clustered abstracts similarly to the lexical analysis above. For the RAS Oncogenes Conference and 2014 AACR Annual Meeting,

[1]Sage Bionetworks, Fred Hutchinson Cancer Research Center, Seattle, Washington. [2]Department of Medical Oncology, Institut Gustave Roussy, Villejuif, France. [3]Department of Biostatistics and Epidemiology, Institut Gustave Roussy, Villejuif, France. [4]Department of Innovative Therapeutics and Early Drug Development, Institut Gustave Roussy, Villejuif, France. [5]INSERM U981, University Paris Sud, Institut Gustave Roussy, Villejuif, France. [6]Leidos Biomedical Research, Inc., Frederick National Laboratory for Cancer Research, Frederick, Maryland. [7]Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, California. [8]Department of Cell and Molecular Pharmacology, University of California, San Francisco, California.

Corresponding Authors: Justin Guinney, Sage Bionetworks, 1100 Fairview Avenue North, Seattle, WA 98109. Phone: 919-357-4762; Fax: 206-667-2062; E-mail: justin.guinney@sagebase.org; and Frank McCormick, UCSF Helen Diller Family Comprehensive Cancer Center, Room 371, 1450 3(rd) Street, P.O. Box 589001, San Francisco, CA 94158-9001. E-mail: mccormic@cc.ucsf.edu

doi: 10.1158/2159-8290.CD-14-0774

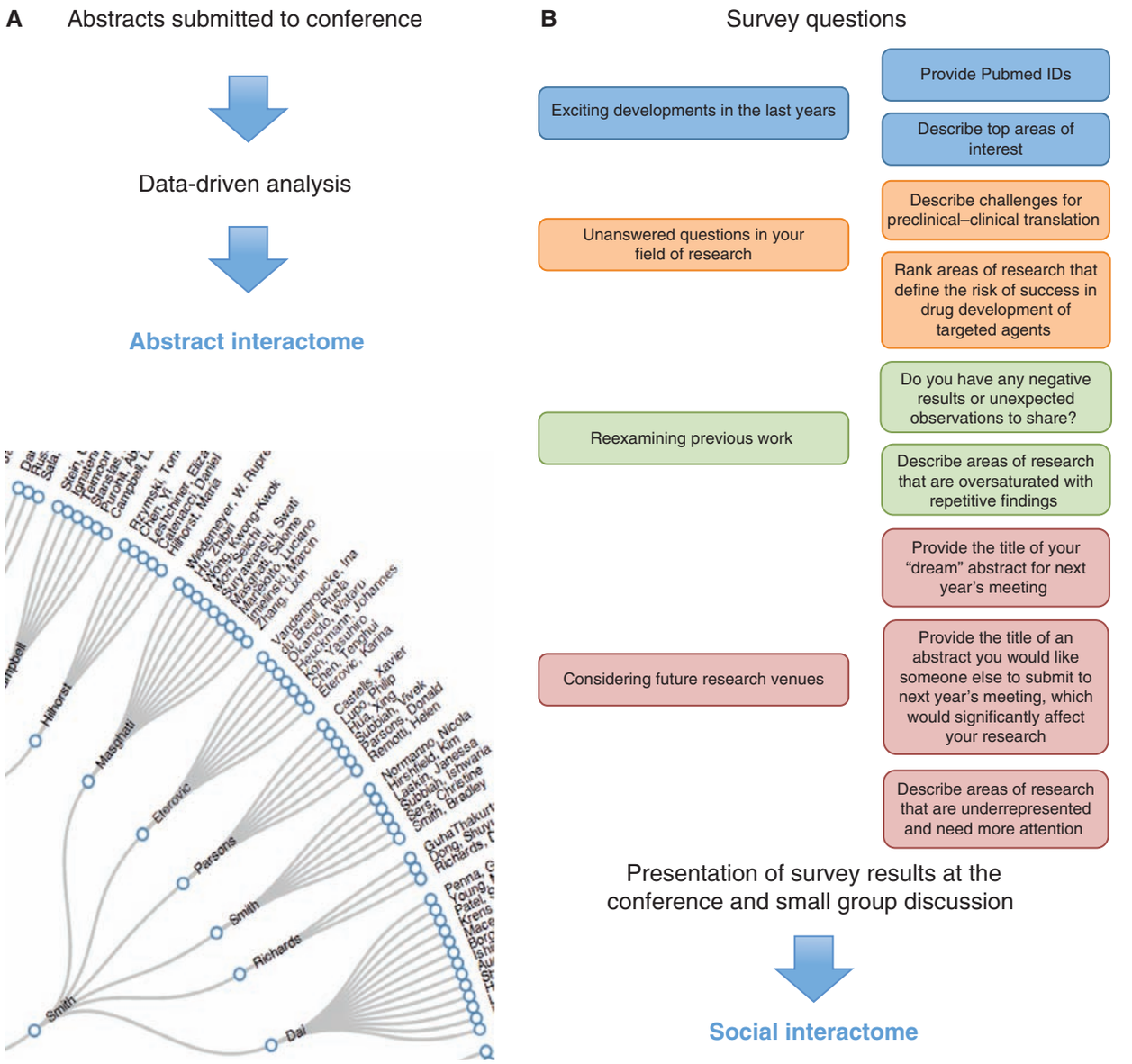

Figure 1. **A,** snapshot of the RAS abstracts interactome from the 2014 AACR Annual Meeting. Each node is associated with a single abstract and labeled with the first author. **B,** outline of survey questions submitted to RAS investigators before the conference. The results guided the discussions in the social interactome session.

we also performed a manual curation of each abstract and categorized them in major fields of research, from biology to therapy and preclinical to clinical translation. This exercise provided another perspective for identifying related research and helped us better understand the current frontiers of investigation. Of note, in all interactomes, the full abstract text, author list, and author institutions are available online, as well as the ability to email the abstract authors directly.

COMMUNITY DEVELOPMENT

By text-mining abstracts, we were able to provide insight into a researcher's interest and the identification of related work. However, it is no surprise that abstracts are an imperfect way of judging similarity of research because (i) some authors submit "boilerplate" or vague abstracts to reveal as little as possible and (ii) word choice and style may influence abstract similarity more than scientific content. For example, authors who submitted multiple abstracts on different topics often clustered together, as they shared common word usage.

In addition, abstracts—like scientific articles—fall short of capturing the complete experience of a researcher, and often omit significant details such as negative results and untested hypotheses. To explore these hidden data, we sent out surveys to attendees covering a range of topics, as described in Fig. 1. We observed high response rates for such open-ended

question surveys, reflecting the enthusiasm of the RAS community and the willingness of researchers to be engaged on their primary research topic.

A key objective of the survey was to identify popular and interesting areas of research within the RAS community. Exciting developments reported by 2013 AACR Annual Meeting attendees involved resistance mechanisms to EGFR/BRAF inhibitors, the role of wild-type RAS isoforms in KRAS-driven tumors, and the intriguing dependency of *KRAS*-mutant models on activation for enhanced activity. In 2014, areas with promising achievements changed considerably as compared with previous years, with major interest in drug design for direct or indirect RAS targeting, preclinical–clinical translation with novel agents and combinations of targeted therapies in RAS/RAF–driven models, the role of oncogenic KRAS in metabolic reprogramming, and its interaction with the tumor microenvironment. Top publications in the field repeatedly mentioned by respondents reported the development of compounds that bind to well-defined surface pockets on oncogenic KRAS in a mutant-specific manner (2, 3) or that inhibit its interaction with plasma membrane/scaffold proteins (4, 5). In line with these findings, in 2013, the most pressing question for RAS scientists concerned the druggability of KRAS, whereas in 2014, investigators were mainly interested in preclinical–clinical translation of KRAS-interacting agents.

We also asked attendees to revisit previous work in the RAS–RAF–MAPK domain and give an impartial opinion on areas of research that they consider oversaturated with repetitive findings or with results that are not reproducible. Many groups reported the failure to validate potential targets described in the literature as synthetic lethals in *KRAS*-mutant models as the major unexpected finding. Apart from high-throughput RNA interference screen studies, publications that attracted less attention from respondents in 2014 as compared with 2013 included those evaluating the preclinical activity of MEK inhibitors in combination with PI3K pathway inhibitors in *KRAS*-mutant models, resistance to selective BRAF inhibitors in melanoma, and genome-sequencing studies assessing the molecular epidemiology of RAS/RAF. On the other hand, recurring topics with the highest interest for future research and that need more investment according to attendees were predominantly preclinical experiments linked to drug design and functional and systems biology approaches for studying MAPK pathway regulation in RAS-driven tumors.

Most importantly, this exercise set the scene for the second part of the RAS interactome sessions, when we organized small group discussions, allowing similarly minded researchers to interact face to face. In these breakout sessions, tables of 10 to 15 people were set up, with placards at each table describing a subtopic within the RAS community, including "models of RAS," "RAS regulators," "metabolism," "microenvironment," "RAS drug targeting," "preclinical–clinical translation," and "clinical validation." Attendees were asked to share their research questions and problems and to find areas where topic discussions and/or future collaborations could help resolve existing research bottlenecks.

To understand the benefit of this breakout session, we sent out a follow-up survey after the conclusion of the 2014 AACR Annual Meeting. We received close to 50 responses, all of them reporting positively on the session and 11 describing new collaborations as a direct result of the breakout discussions. A common request was for greater structure within the discussion sections, with predefined leaders and more specialized topics of discussion. These comments reflected our own underappreciation of the desire for structured, albeit informal, discussions among conference attendees, and will be emphasized at future interactome sessions.

FUTURE OF BIOMEDICAL CONFERENCES

The large attendance, enthusiastic response, and active participation by those at these AACR special sessions demonstrate a desire for alternative forms of interactions at biomedical conferences. Indeed, the need for roundtable dialogues engaging fellow scientists is a clear message of these experiments and should compel AACR and other organizations to expand this type of meeting to other areas. We therefore recommend the following components as important areas of attention for future conference planning.

Technology

Phone apps, web videos, and schedulers have dramatically improved access to session information, as well as the ability to identify presentations that are of interest to attendees. However, to our knowledge, no attempt has been made to predict common areas of interest among conference attendees using abstracts or other information. As shown here, data-driven analyses are feasible but require a sustained commitment and investment on the part of conference organizers. Moreover, as cancer research matures toward a quantitative, big-data science, conferences should reflect this same data-driven mindset.

Expert Dialogues

A point–counterpoint discussion session addressing important yet unresolved questions in a particular research topic can be entertaining and informative, opening a window to debate on controversial issues and network with top scientists in the field. Such forums provide a unique educational opportunity for those who are less familiar with the subject, allow experts to evolve in real time their fundamental concepts on the topic, and guide future research efforts.

Community Overview

The responses from surveys provide a high-level understanding of the major themes and concerns that exist in a particular research area, from the most exciting trends and promising developments to emerging topics that may attract less attention. On-site presentation of the results gives voice to all researchers—principal investigators, postdocs, and PhD students—and exposes attendees to a wide range of subspecialties, giving a broader perspective of challenges and prospects in the field.

Social Interaction

Direct, face-to-face interaction is indispensable for establishing trust and productive collaborations. The most tangible benefit we observed in our experimental sessions was

the exchange of ideas that occurred during the breakout sessions with a focus on specific fields within the broader RAS domain. These breakout discussions allowed researchers to discuss their work and encouraged data sharing and potential collaborations to find solutions to common goals. Moreover, it provided opportunities for young investigators to interact with experts in their field to gain advice and insight. We will be tracking nascent projects that emerged from this session, and hope to invite these scientists to present project outcomes at the 2015 AACR Annual Meeting.

In a world with increasingly sophisticated means of communication across great distances, there still remains a basic need for direct human interaction at scientific conferences. We believe that organizations such as AACR are beginning to recognize their role in supporting research communities and expanding opportunities for interactions. Here, we have presented the results of one such experiment that successfully engaged researchers in a socially constructive and intellectually stimulating manner.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Published online November 3, 2014.

REFERENCES

1. Frey BJ, Dueck D. Clustering by passing messages between data points. Science 2007;315:972–6.
2. Ostrem JM, Peters U, Sos ML, Wells JA, Shokat KM. K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. Nature 2013;503:548–51.
3. Lim SM, Westover KD, Ficarro SB, Harrison RA, Choi HG, Pacold ME, et al. Therapeutic targeting of oncogenic K-Ras by a covalent catalytic site inhibitor. Angew Chem Int Ed Engl 2014;53:199–204.
4. Zimmermann G, Papke B, Ismail S, Vartak N, Chandra A, Hoffmann M, et al. Small molecule inhibition of the KRAS-PDEδ interaction impairs oncogenic KRAS signalling. Nature 2013;497:638–42.
5. van der Hoeven D, Cho KJ, Ma X, Chigurupati S, Parton RG, Hancock JF. Fendiline inhibits K-Ras plasma membrane localization and blocks K-Ras signal transmission. Mol Cell Biol 2013;33:237–51.

Social Interactomes for Enabling Research Communities

Justin Guinney, Rodrigo Dienstmann, Charles Ferté, et al.

Cancer Discovery 2014;4:1265-1268.