Published in final edited form as:

Nat Genet. 2017 March 30; 49(4): 483–484. doi:10.1038/ng.3830.

Celebrating parasites

Casey S Greene¹, Lana X Garmire², Jack A Gilbert³, Marylyn D Ritchie⁴, and Lawrence E Hunter⁵

¹Department of Systems Pharmacology and Translational Therapeutics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA

²Cancer Epidemiology Program, University of Hawaii Cancer Center, University of Hawaii, Honolulu, Hawaii, USA

³Department of Surgery, University of Chicago School of Medicine, Chicago, Illinois, USA

⁴Biomedical and Translational Informatics Program, Geisinger Health System, Danville, Pennsylvania, USA

⁵Department of Pharmacology, University of Colorado School of Medicine, Aurora, Colorado, USA

To the Editor

In an editorial published last year¹, Dan Longo and Jeffrey Drazen introduced us to ‘research parasites’. These individuals “had nothing to do with the design and execution of the study but use another group’s data for their own ends, possibly stealing from the research productivity planned by the data gatherers, or even use the data to try to disprove what the original investigators had posited” (ref. 1). The editorial sparked discussion about the role of secondary data analysis in the scientific process, both in official letters to the editor and informal commentary online. In light of the term’s widespread publicity, we chose to use it to honor individuals who practice the craft of data reanalysis for novel ends.

At the Pacific Symposium on Biocomputing (PSB) 2017, we presented the inaugural Research Parasite Awards to researchers selected for their rigorous analysis of publicly accessible data. We specifically sought to honor those whose work extended, replicated or disproved what the original investigators had posited who were not involved in the experimental design or data generation, published independently of the original investigators while appropriately crediting them, and provided their own research products—including source code and intermediate or final results—in a manner that enhanced reproducibility.

We opened a call for nominations and applications in April 2016 and received 41 completed applications. From these, we selected an exemplar of Junior Research Parasitism and a Sustained Parasite. The Junior Parasite Award highlighted work performed as a trainee, while the Sustained Parasite Award required contributions over at least five years of independent research.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

The inaugural Junior Parasite Award recipient was Kun-Hsing Yu of Stanford University for a recent study that employed public omics and imaging data sets. Yu developed an approach to improve prediction of patient survival². He and his coauthors reanalyzed histopathology images from The Cancer Genome Atlas (TCGA) and extracted features using CellProfiler. They then employed different machine learning approaches using packages from the R programming language. Yu and coauthors provided source code and data under an open license, enabling future parasites to build on this work².

The inaugural recipient of the Sustained Parasite Award was Erick Turner from Oregon Health & Science University. Turner's research revealed pervasive selective reporting of clinical studies and showed how this skewed the broader biomedical literature. For example, more than 90% of publications on antidepressant trials reported a positive outcome; however, the US Food and Drug Administration (FDA)-registered studies underlying that body of literature only had positive outcomes 51% of the time³. Trials with negative or ambiguous results were spun as having positive results and published at a lower rate than trials with positive findings. Turner continued to identify reporting biases for multiple drug classes⁴.

These two awardees, in different ways, captured the essence of research parasitism. Their work enhanced understanding of the world and promoted rigorous biomedical research.

The act of rigorous secondary data analysis is critical for maintaining the accuracy and efficiency of scientific discovery. As scientists, we make predictions, perform experiments and generate data to test those predictions. When we ask rigorous questions, we obtain more accurate findings that can prevent harm. For example, Vioxx was evaluated for use in treating pain associated with rheumatoid arthritis⁵. Questions were raised shortly thereafter about its cardiovascular effects⁶. Independent researchers, using data from multiple studies, identified a drug-associated increase in cardiovascular event risk⁷. These research parasites identified important side-effects of this drug, correcting incomplete information on the drug's safety profile.

Journals, although they conduct peer review, do not validate each experimental result or claim. Research parasites fill this gap. In the case of the initial VIGOR study⁵ published in *The New England Journal of Medicine*, Jeffrey Drazen, the editor-in-chief of the journal, stated: "We can't be in the business of policing every bit of data that we put out. We think that that's the role of people who know the field" (ref. 6). Research parasites help to maintain the self-correcting nature of scientific inquiry. Scientists who perform rigorous parasitism put scientific work to the test, and their results may support or challenge what we think we know.

Parasites also improve efficiency: many data sets were originally designed for specific questions, but these data may also answer distinct but related questions. Investigators can refocus data sets via meta-analysis to reveal general patterns that become apparent only with many studies. Data sets can also be individually useful. New researchers can often bring their own creative ideas to existing data, leading to novel breakthroughs and disruptive innovations.

The robust culture of research parasitism is not limited to clinical or biomedical research. A team of eagle-eyed parasites recently identified inconsistencies in a study of food consumption at a buffet restaurant⁸. Others raised issues with the genetic structure of polar bear populations⁹ and social welfare spending¹⁰. This process by which scientists selectively retest assumptions and previous findings forms the foundation of scientific discovery in every discipline.

Under some proposals for data reuse, data would be shared with researchers working in concert with the investigators who initially analyzed the data¹. We expect that this would counteract the recent focus of the US National Institutes of Health (NIH) on rigor, transparency and reproducibility. Any procedure that includes data generators as gatekeepers has the potential to compromise rigor and robustness. As gatekeepers, researchers could withhold data from those with contrary views or a reputation of challenging the status quo. We must expect data sharing to lead to some conclusions being challenged and, ultimately, refuted. If this fails to occur, it indicates a problem with the process and not the correctness of conclusions.

Good research parasites also have an obligation to credit their research hosts. We do not feel that parasites individually have an obligation to benefit those from whom they gather data (that is, in research symbiosis). However, we do feel that parasites should advocate strongly for recognition and support for those who generate and share valuable data without restricting their use.

We enjoyed the work that research parasites shared with us this year and expect research parasites to continue to separate real findings from erroneous conclusions. To celebrate such contributions, we are announcing an open call for applications for the 2018 Research Parasite Awards. Winners receive partial travel support to PSB, a prize and recognition on the award website. Applications are due by 5 p.m. Hawaii Standard Time on 30 September 2017. For full instructions, visit the website at <http://research-parasite.com/>. We have a positive and inclusive view of parasitism and particularly encourage individuals from groups that are under-represented in their discipline to apply.

Acknowledgments

Numerous individuals and organisms contributed time and money toward the 2017 Research Parasite Awards. We would like to thank the PSB organizers, who helped us hone our selection criteria and award committee, and *GigaScience*, the Gordon and Betty Moore Foundation, *Nature Genetics* and *Scientific Data*, which provided financial and travel support for the awards. The authors were supported in part by grants from the Gordon and Betty Moore Foundation (GBMF 4552 to C.S.G.) and the US National Institutes of Health (R01LM012373 and K01ES025434 to L.X.G. and R01LM008111 to L.E.H.).

References

1. Longo DL, Drazen JM. *N Engl J Med*. 2016; 374:276–277. [PubMed: 26789876]
2. Yu KH, et al. *Nat Commun*. 2016; 7:12474. [PubMed: 27527408]
3. Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. *N Engl J Med*. 2008; 358:252–260. [PubMed: 18199864]
4. Roest AM, et al. *JAMA Psychiatry*. 2015; 72:500. [PubMed: 25806940]
5. Bombardier C, et al. *N Engl J Med*. 2000; 343:1520–1528. [PubMed: 11087881]

6. Armstrong D. Wall Street Journal. May 15.2006
7. Mukherjee D, et al. J Am Med Assoc. 2001; 286:954.
8. van der Zee, T., Anaya, J., Brown, NJL. PeerJ Preprints. 2017. <http://dx.doi.org/10.7287/peerj.preprints.2748v1>
9. Malenfant RM, Davis CS, Cullingham CI, Coltman DW. PLoS One. 2016; 11:e0148967. [PubMed: 26974333]
10. Breznau N. Sociol Sci. 2015; 2:440–441.