

Reproducible Research Resources for Research(ing) Parasites

Scott Edmunds 

Published June 3, 2016

Citation

Edmunds, S. (n.d.). In *GigaBlog*. GigaBlog. <https://doi.org/10.59350/vhpgg-ec668>

Keywords

Open Access, Publishing, Technology, Genomics, Metagenomics
Feature Image

Copyright

Copyright © Scott Edmunds 2016. Distributed under the terms of the [Creative Commons Attribution 4.0 International License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

****GigaScience has Tapeworms and Scabies! And Reproducible Research.**

IncreaseReproducibility**While there has been [recent controversy \(and hashtags in response\)](#) from some of the more conservative sections of the medical community calling those who use or build on previous data "research parasites", as data publishers we strongly disagree with this. And also feel it is unfair to drag parasites into this when they can teach us a thing or two about good research practice. Parasitology remains a complex field given the often extreme differences between parasites, which all fall under the umbrella definition of an organism that lives in or on another organism (host) and derives nutrients at the host's expense. Published today in *GigaScience* are articles on two parasitic organisms, [scabies](#) and on the [tapeworm *Schistocephalus solidus*](#). Not only are both papers in parasitology, but the way in which these studies are presented showcase a new collaboration with [protocols.io](#) that provides a unique means for reporting the Methods that serves to improve reproducibility. Here the authors take advantage of their open access repository of scientific methods and a collaborative protocol-centered platform, and we for the first time have integrated this into our submission, review and publication process. We now also have a [groups page](#) on the portal where our methods can be stored.

Currently, the most common way of presenting methods in articles is in extremely brief paragraphs as supplemental downloadable PDF files. The result is often incomplete or non-discoverable methodology, which is key for scientists to properly build on scientific discovery. The parasitology articles published today are the first two studies to showcase the seamless integration into the manuscript submission and publication process of clear, detailed, and complete methodology descriptions. The [protocols.io](#) platform enables researchers to submit their methods in a standard format, with no space limitations, that can be directly linked to any article simply through a citable DOIs. These can also be searched online, and best yet, can be versioned allowing for adaptations for future work. Not only does this allow the research community easy access to detailed methods, it also means authors don't have to continually rewrite methods for every paper that uses them. Being [big promoters of data citation](#), in a similar manner this incentivizes good practice and method sharing as users can simply cite and credit the 'recipe' in protocols.io. See [their video](#) for more on how this works.

<https://www.youtube.com/watch?v=r9eBZgTzRNA>

Itching to solve the reproducible research problem

After the research parasites debate it seems fitting that the complexity of making scientific reporting reproducible is demonstrated in papers that capture the complexity of parasitic organisms, and, in these cases, parasites that require many different complicated experimental steps and unusual computational pipelines to study them.

In [the first study](#), researchers from the National Health and Medical Research Council in Australia studied the genome of the human scabies parasite collected from remote disadvantaged and indigenous communities in Northern Australia, where up to 25% of adults and 50% of children acquire scabies infections each year. Scabies infections are linked to bacterial skin infections and rheumatic fever. As a consequence of this children with scabies do

less well, and this is a contributing factor to indigenous Australians having significantly reduced life expectancy and amongst the highest rates of rheumatic heart disease in the world.

Until now studying this species has been challenging. Being fractions of a millimeter in size, the researchers needed to collect, per sample, about 1000 mites to obtain enough DNA for next generation sequencing. In addition to the complications of collecting and pooling the mites, their tiny size also meant they had to deal with contamination from the mite's gut contents. All of these variables can create difficulty in clearly describing how conclusions are derived and how the research can be built on. The lead author Anthony Papenfuss, discussing the challenges of communicating this work has been, stated: "Writing clear and accurate descriptions of the wet lab and bioinformatics methods is a challenge at the best of times. It is especially hard when the design is complex and iterative exploratory analysis using multiple tools is required. It requires great care and time consuming refinement of the text. I think documenting the methods using [protocols.io](#) will make this much easier.". As with all our papers, supporting genomics data is [available from GigaDB](#) and the SRA.

Slide1In our [second paper](#), researchers from Quebec studied the molecular biology of the parasitic tapeworm *Schistocephalus solidus*. Despite *S. solidus* serving as an emblematic study system in parasitology with [two centuries of research](#); however, it has an extremely complicated life-cycle with multiple developmental states and host species (parasitizing crustaceans, fish and birds). As a consequence, while there is much known about its morphology and physiology, identifying which genes are used at each stage of infection, has been comparatively lacking.

The work here includes recreating the different host conditions and collecting living worms from the different life cycles to collect RNA and produce a transcriptome gene catalogue. First author François-Olivier Hébert stated "Describing such a long process of field sampling, experimental infections in the lab using multiple hosts and, of course, the complementary bioinformatic analyses was one of the greatest challenges in this paper". With the new integrated data and method publishing pipeline aiding this, the authors added "We were able to achieve that by making all of our homemade scripts, programs and datasets freely available to the public through *GigaScience*, GigaDB and [protocols.io](#). They represent essential complementary platforms that allowed us to respect our vision of a reproducible science". The extensive supporting data is again [available from our GigaDB repository](#).

We at *GigaScience* are excited to announce this collaboration with concrete examples to show, and with [protocols.io](#) having been highlighted in our instructions for authors, as well as integrated into our data submission pipelines these are the first of what will be an increasingly common part of our our published papers. Keep checking the [GigaScience groups page on protocols.io](#) to see as and when these get published.

Update 3/6/16: BMC have been late to publish the *Schistocephalus solidus* publication, but a provisional version can be seen here: <https://goo.gl/VtpxSp>. Apologies, and are chasing them to fix this as soon as possible.

Update 7/6/16: The *Schistocephalus solidus* paper came out the following day and is working fine now. [protocols.io](#) have also published a [blog post](#) on the integration.

References

Mofiz E. et al., Genomic resources and draft reference assemblies of the human and porcine scabies mites, *Sarcoptes scabiei* var. *hominis* and var. *suis*. *GigaScience*. **5**:23. 2016. DOI:[10.1186/s13742-016-0129-2](https://doi.org/10.1186/s13742-016-0129-2)

Mofiz, E; Holt, D; Seemann, T; Currie, B, J; Fischer, K; Papenfuss, A, T (2016): Draft genome assembly using parasitic mite population NGS DNA sample from mites extracted from host wound environment. Protocols.io. <http://dx.doi.org/10.17504/protocols.io.exwbfpe>

Hebert FO. et al., Reference transcriptome for the parasite *Schistocephalus solidus*: insights into the molecular evolution of parasitism. *GigaScience*. **5**:24. 2016. DOI:[10.1186/s13742-016-0128-3](https://doi.org/10.1186/s13742-016-0128-3)

Herbert, F.O.; Grambauer, S.; Barber, I.; Landry, C.R., Aubin-Horth, N. (2016): Protocols for "Reference transcriptome sequence resource for the study of the Cestode *Schistocephalus solidus*, a threespine stickleback parasite.". Protocols.io. <http://dx.doi.org/10.17504/protocols.io.ew9bfh6>

The post [Reproducible Research Resources for Research\(ing\) Parasites](#) appeared first on [GigaBlog](#).