

Dynamic DOI's: A Better Way to Share Research

Earl B. Beutler

Earl B. Beutler¹

¹LabArchives, LLC, USA, ebeutler@labarchives.com

BACKGROUND

Over the past several years, there has been a growing movement to share research data with the public driven by issues with the irreproducibility of research results [1-3] as well as the sense that investigation conducted with public funds should, indeed, be shared with the public. Many journals and leading funding bodies, including the NHMRC as well as the U.S. NIH, and NSF have either mandated or encouraged the public sharing of research data. It is clear to most observers that this will soon be a requirement for Australian researchers.

At present, there are a wide variety of repositories in which researchers can make their data public. Some repositories are very specific to a particular field of study and type of data (e.g. GenBank for genomic sequences), while others are very broadly based and have the ability to contain virtually any type of data.

Most repositories, however, simply contain individual files or collections of files with some metadata. They are lacking in context, and there is no indication of when the work was actually conducted, nor any kind of history of these data. For example, the protocol used in the experiment in which the data were gathered is a critical piece of information for those trying to analyze or replicate the published results.

Furthermore, investigation often continues beyond the posting of data to a repository, and should the author wish to make ongoing progress available, there is no mechanism to do so.

A DIFFERENT WAY TO SHARE RESEARCH DATA

With the advent of Electronic Laboratory Notebooks (ELN's), many investigators are now maintaining a digital notebook containing contextual data. Many of these products contain "versioning", where all modified information is maintained, in some case in perpetuity.

LabArchives is one such leading ELN, which provides the ability to share data "in situ". Rather than uploading files to a repository, the ELN itself can serve as the repository itself. Not only does this remove the step of locating and uploading files to a separate service, but it eliminates errors in file selection, includes date and time stamps, and allows viewers (at the option of the author) to see newer and/or older versions of shared data. It can include links to protocols – including to specific versions of protocols that may have been altered post-experiment.

EXAMPLES

LabArchives' "Dynamic DOIs" have been available since 2012, although they have not yet been widely utilized by the research community. There are many reasons for that, including reticence by many researchers to share data, lack of enforcement of funding agency mandates, and restriction of public sharing by many institutions.

The following figures illustrate some of the publicly available data sets which, in many cases, goes far beyond the traditional collection of files that is found many repositories.

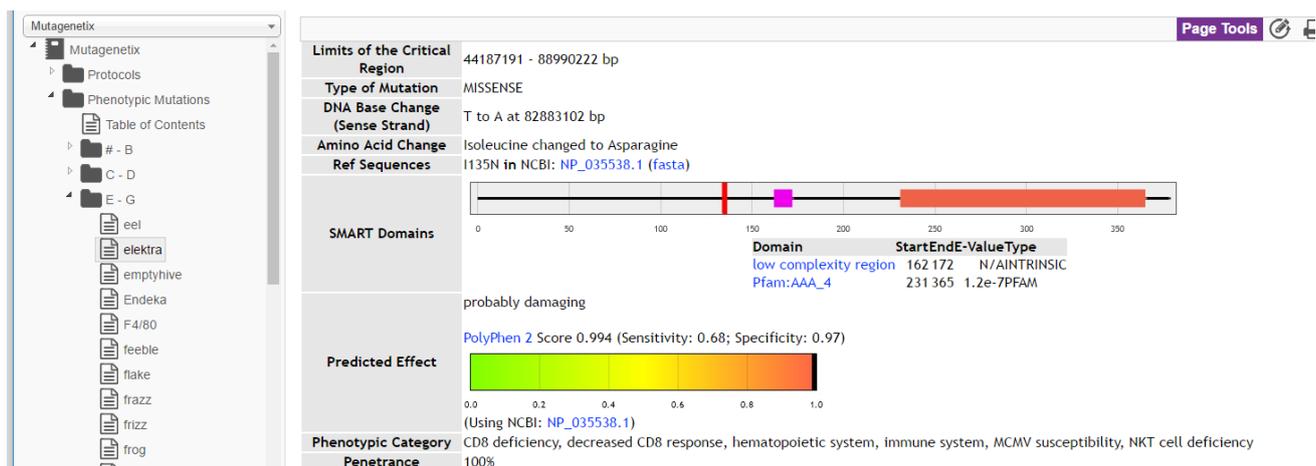


Figure 1: A Page from the Mutagenetix Notebook⁴

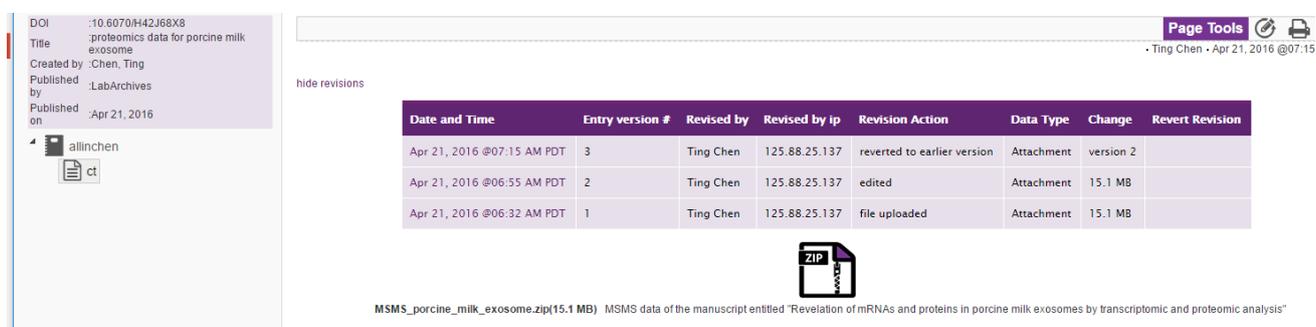


Figure 2: A Page from supplementary data showing older versions of data⁵

REFERENCES

1. Begley, CG, and Ellis, LM, Drug development: Raise standards for preclinical cancer research, *Nature* 2012. 483:531-533
2. Freedman, LP, Cockburn, IM, and Simcoe, TS, The economics of reproducibility in preclinical Research, *PLOS Biology* 2015. 13(6):
3. Loscalzo, J, Irreproducible experimental results: Causes, (mis)interpretations, and consequences, *Circulation* 2012. 125:1211-1214
4. Beutler, BA: Mutagenetix, 2012. DOI: 10.6070/H45Q4T14
5. Chen, T., Xi, Q, Sun, J-J, Ye, R-S, Cheng, X, Sun, R-P, Wang, S-B, Shu, G, Wang, L-N, Zhu, X-T, Jiang, Q-Y, and Zhang, Y-L: Revelations of mRNAs and proteins in porcine milk exosomes by transcriptomic and proteomic analysis, *BMC Vet Res* 2017:13:101. Data DOI 10.6070/H42J68X8