

GigaData news: Macaque DOIs published in Nature Biotechnology

Scott Edmunds 

Published October 21, 2011

Citation

Edmunds, S. (n.d.). In *GigaBlog*. GigaBlog. <https://doi.org/10.59350/9cw7c-jdz05>

Keywords

Publishing, Data Citation, Data Publication, DataCite, DOI

Feature Image

Copyright

Copyright © Scott Edmunds 2011. 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.

This week marks another success for the fledgling practice of data citation, with two datasets from our [GigaScience database published](#) in *Nature Biotechnology*. The genomes sequenced by our colleagues at the BGI for the Cynomolgus and Chinese rhesus macaques were initially [released DOIs](#) at our [launch](#) in July, and were amongst the first (at the time) unpublished genomes released in this way. Data citation is an important [concept](#), allowing data producers to obtain an early form of credit for releasing their work, speeding up research by encouraging early data release, and allowing the impact and reuse of data to be tracked.

After the recent [success](#) of our first dataset being [published](#) in the *New England of Medicine* (the genome of the recent [outbreak](#) strain of *E. coli* O104:H4), this is the first time one of our data DOIs has been accepted in a [Nature](#) journal. For data citation to work the assistance of journals is key, and *Nature Biotechnology* has been particularly helpful in promoting the scheme, arguing in an [editorial](#) as far back as 2009 that novel forms of credit for data producers were needed, and suggesting DOIs as an ideal solution for this. The [Datacite](#) consortium was set up in late 2010 to do exactly that, and we would like to thank them and the [British Library](#) for their help in issuing these DOIs.

Macaque species are the most commonly used non-human primate models in medical research, and their genomes will hopefully aid human [disease research](#) and drug discovery. Looking at orthologues of human druggable protein domains in these species is aiding the potential therapeutic exploitation of their 'druggable genome', and has [already lead](#) to BGI producing an exome sequencing platform for the species. On top of their genome assemblies, the DOI [landing pages](#) include links to functionally annotated and coding sequence sets, as well as a [link](#) to a browser and database. After the release of other datasets such as the CHO cell line [genome](#), we are currently collecting another large batch of datasets to be released, so watch [this space](#) for further news and announcements.

References

1. Yan, G. *et al.* Genome sequencing and comparison of two nonhuman primate animal models, the cynomolgus and Chinese rhesus macaques. *Nat Biotech*, (2011). <https://doi.org/10.1038/nbt.1992>
2. Credit where credit is overdue. *Nat Biotech* **27**, 579 (2009). <https://doi.org/10.1038/nbt0709-579>

To cite the two datasets please use the following citations:

3. Yan, G; Zhang, G; Fang, X; Zhang, Y; Li, C; Ling, F; Cooper, DN; Li, O; Li, Y; van Gool, AJ; Du, H; Chen, J; Chen, R; Zhang, P; Huang, Z; Thompson, JR; Meng, Y; Bai, Y; Wang, J; Zhuo, M; Wang, T; Huang, Y; Wei, L; Li, J; Wang, Z; Hu, H; Le, L; Stenson, PD; Li, B; Liu, X; Ball, EV; An, N; Huang, Q; Zhang, Y; Fan, W; Zhang, X; Li, Y; Wang, W; Katze, MG; Su, B; Nielsen, R; Yang, H; Wang, J; Wang, X; Wang, J (2011): Genomic data from the Chinese Rhesus Macaque (*Macaca mulatta lasiota*). *GigaScience*. [doi:10.5524/100002](https://doi.org/10.5524/100002)

<http://dx.doi.org/10.5524/100002>

4. Yan, G; Zhang, G; Fang, X; Zhang, Y; Li, C; Ling, F; Cooper, DN; Li, O; Li, Y; van Gool, AJ; Du, H; Chen, J; Chen, R; Zhang, P; Huang, Z; Thompson, JR; Meng, Y; Bai, Y; Wang, J; Zhuo, M; Wang, T; Huang, Y; Wei, L; Li, J; Wang, Z; Hu, H; Le, L; Stenson, PD; Li, B; Liu, X; Ball, EV; An, N; Huang, Q; Zhang, Y; Fan, W; Zhang, X; Li, Y; Wang, W; Katze, MG; Su, B; Nielsen, R; Yang, H; Wang, J; Wang, X; Wang, J (2011): Genomic data from the Crab Eating Macaque/Cynomolgus Monkey (*Macaca fascicularis*). *GigaScience*. doi:10.5524/100003

<http://dx.doi.org/10.5524/100003>

The post [GigaData news: Macaque DOIs published in Nature Biotechnology](#) appeared first on [GigaBlog](#).