COMMENTARY

A tale of two tardigrades

Thomas A. Richardsa,b,1 and Adam Moniera,c

Horizontal or lateral gene transfer (HGT) involves the transmission of genetic material between separate genomes. It is a major driver of evolutionary innovation in archaea and bacteria (1–4), but the role of HGT in eukaryotes, especially multicellular organisms, remains controversial. The genome sequence of Hypsibius dujardini, a member of the tardigrades, a group of animals that can survive a range of extreme environments, was reported to have undergone extensive HGT, amounting to 17.5% (5) of the genes identified. Even though Boothby et al. performed numerous checks to support their claim, reanalysis using new approaches to identify genome contamination combined with additional sequencing suggests only ~0.4% of the H. dujardini gene repertoire can confidently be identified as derived by HGT (6). These are fascinating animals and no doubt there will be further genome studies on these organisms before the last word on this topic emerges.

There are clear examples of HGT in animal lineages (e.g., refs. 7–10), yet HGT is thought to be infrequent in multicellular organisms (11–13). This is because multicellular organisms often separate their germ-line cells from their somatic cells, so the reproductive genome is in part isolated from contact with foreign sources of DNA. There are exceptions to this rule, such as when intracellular bacteria are present in the reproductive cells of an animal host (14, 15). Recent work has also demonstrated that bona fide HGT, separate from gene transfer arising from the progenitors of endosymbiotic organelles, appears to be rare in 55 genomes spanning the known diversity of the eukaryotes (16). Although more analysis is very much needed on a diversity of different eukaryotes, this finding is consistent with other work suggesting that eukaryotic groups that have shared environments and ecological interactions for over 400 million y (17) have undergone very little gene transfer (18).

Boothby et al. (5) published the first genome of a tardigrade animal, with the primary conclusion that ~one-sixth of the gene content is of HGT ancestry. Using a comparative similarity index approach (19), Boothby et al. (5) identified that 6,663 of the 39,532 predicted genes had sequence similarity profiles, suggesting they are not vertically inherited. Because of the scarcity of HGTs in animals, the extent of transfer identified in the tardigrade genome was quite a surprise. Boothby et al. set out to test this observation, performing several checks to validate the transfers identified. First, they performed additional phylogenetic analyses of a subsample of 107 largely randomly selected candidate transfers (a test HGT dataset), finding phylogenetic support for HGT in 101 cases. To exclude assembly artifacts, they then confirmed their genome assembly using long-read sequencing technologies and used PCR to confirm the genomic scaffold of 104 of the 107-test HGT dataset. Of these PCR amplifications, 59 targeted a section of a genome contig that bridged a gap between a putative HGT gene and a native (vertically derived) gene. Of these “bridging” PCRs, 58 seemed to validate the synteny recovered in their genome assembly, suggesting these candidate HGTs were integrated into the native tardigrade genome. Finally, 57% of all of the putatively transferred genes reported also contained a predicted spliceosomal intron. Spliceosomal introns are a hallmark of eukaryotic encoded genes, supporting Boothby et al.’s initial claim that the prokaryotic genes have been integrated into the tardigrade genome.

Shortly after this publication, a manuscript emerged reporting a second H. dujardini genome sequence (6). Flow cytometry analysis demonstrated a genome size of 110 Mb (6), over 100 Mb smaller than the previously published genome assembly. Preliminary assembly of this second genome amounted to 186 Mb. Using GC content vs. read coverage plots (20), Koutsovoulos et al. (6) identified extensive traces of contamination in both genome assemblies. This tool was then used to clean the second genome assembly, resulting in a draft genome of 135 Mb and encoding 23,021 putative proteins. Koutsovoulos et al. then retested the provenance of the “test HGT dataset” selected by Boothby et al. (5), identifying 51 as informative for HGT analysis. Further analysis also demonstrated that the majority of the HGTs, inferred from the similarity index approach, were likely derived from contamination (6). Finally, analysis of the cleaned genome identified 94 strong HGT candidates, 0.4% of the total repertoire of genes identified: 49 derived from other eukaryotes and 55 from

1To whom correspondence should be addressed. Email: t.a.richards@exeter.ac.uk.

Department of Biosciences, University of Exeter, Exeter EX4 4QD, United Kingdom; (1)Integrated Microbial Biodiversity Program, Canadian Institute for Advanced Research, Toronto, ON, Canada MSG 1Z8; and (2)Québec Océan, Université Laval, Québec, QC, Canada G1V 0A6

Author contributions: T.A.R. and A.M. wrote the paper.

The authors declare no conflict of interest.

See companion article on page 5053.
Although rotifers and the red alga encode functional proteins in eukaryotes seem to be rare (Fig. 1), bacteria. This level of transfer is comparable to that seen in other eukaryotic genomes (5, 6, 13, 19, 21, 28–32). The bar chart illustrates contradictory results (red and gray bars) for HGT into the tardigrade *Hypsibius dujardini* genome. An asterisk indicates an organism that can occupy an extreme environment. Images courtesy of PhylaPic CC-BY 3.0 and (Trichomonas) T. Michael Keesey; (Caenorhabditis) Gareth Monger; (Rotifer) Diego Fontaneto, Elisabeth A. Herniou, Chiara Boschetti, Manuela Caprioli, Giulio Melone, Claudia Ricci, Timothy G. Barraclough, and vectorized by T. Michael Keesey; and (Tardigrade) Michelle Site.

So where are we now? Genes of identifiable HGT ancestry that encode functional proteins in eukaryotes seem to be rare (Fig. 1), although rotifers and the red alga *Galdieria sulphuraria*, which can both survive in “extreme” environments, may still represent exceptions to this rule (19, 21). This trend suggests that barriers to HGT integration do exist in eukaryotes and therefore transfer in eukaryotes is of a different form and frequency than that observed in prokaryotes (16). However, there are clear cases where HGTs have added important functional attributes to the recipient lineage (e.g., refs. 7–10, 22, 23). Consequently, identification of transferred genes is an important challenge that can inform our understanding of how biological systems evolve and function.

With the use of next-generation technologies, genome sequencing is rapidly becoming the default approach of a range of research fields. As sequencing becomes accessible, it is clear that reliable genome assembly is not so straightforward. The conflation between the challenge of genome assembly and the desire to seek HGTs of importance to the evolution of biological function is generating friction. As this field progresses we should look to develop a set of guidelines for identifying HGT. Here we list a set of tentative guidelines for informing HGT analysis.

First, contamination is unavoidable in de novo genome sequencing. As such, de novo genome sequences should be viewed as metagenomes. We must therefore implement cutting edge bioinformatics based on database search similarity patterns/GC/word use/coverage comparisons to identify and remove contamination (6, 24).

Second, during the preparation of DNA samples for genome sequencing, DNA should be tested for contamination using environmental small subunit tDNA-based protocols (e.g., ref. 25), preferably targeting multiple gene-sequence markers (i.e., using both 18S and 16S small subunit PCR primers). This will provide preliminary data for assessing contamination in genome assemblies.

Third, where possible, multiple genome assemblies from different representatives of the same species and from related groups should be sequenced to validate HGTs and polarize the point of gene acquisition.

Fourth, all putative HGTs should be validated using phylogenetic analysis. Phylogenetic trees that demonstrate topological support (bootstrap and alternative topology analysis) for a putative recipient lineage branching with and within a donor group can be used to provide support for HGT (see figure 1 in ref. 26).

Fifth, synteny analysis should be used to confirm that candidate HGTs are nested on native sections of chromosomal DNA (i.e., adjacent to vertically derived genes).

Sixth, transcription of putative HGTs should be validated to confirm intron architecture and test if the gene is transcriptionally functional in the recipient genome.

These guidelines can be modified and added to as more datasets emerge and more cases of HGT are investigated. However, we predict that adherence to these rules will confirm that HGT into eukaryotes does occur but at a lower frequency compared with prokaryotes. Furthermore, we also predict that such criteria will help the field to focus on the HGTs that represent important gains-of-function in the recipient lineages. For us, identifying cases where transfer has amended or added function in the recipient lineage, however rare, represents one of the most exciting prospects of this field of research. Researchers are therefore correct to look for HGT in organisms that are capable of “extraordinary” biological functions, for example colonizing extreme or variant environments in contrast to related taxa (27). However, caution is needed in these endeavors.

---

1 Ochman H, Lawrence JG, Groisman EA (2000) Lateral gene transfer and the nature of bacterial innovation. Nature 405(6784):299–304.
2 Doolittle WF (1999) Lateral genomics. Trends Cell Biol 9(12):M5–M8.
3 Boucher Y, et al. (2003) Lateral gene transfer and the origins of prokaryotic groups. Annu Rev Genet 37:283–328.
4 Nelson-Sathi S, et al. (2012) Acquisition of 1,000 eubacterial genes physiologically transformed a methanogen at the origin of Haloarchaea. Proc Natl Acad Sci USA 109(50):20533–20542.
5 Boothby TC, et al. (2015) Evidence for extensive horizontal gene transfer from the draft genome of a tardigrade. Proc Natl Acad Sci USA 112(52):15976–15981.
6 Koutsovolos G, et al. (2016) No evidence for extensive horizontal gene transfer in the genome of the tardigrade *Hypsibius dujardini*. Proc Natl Acad Sci USA 113:5053–5058.
7 Acuña R, et al. (2012) Adaptive horizontal transfer of a bacterial gene to an invasive insect pest of coffee. Proc Natl Acad Sci USA 109(11):4197–4202.

Richards and Monier
8 Moran NA, Jarvik T (2010) Lateral transfer of genes from fungi underlies carotenoid production in aphids. Science 328(5978):624–627.
9 Wybouw N, et al. (2014) A gene horizontally transferred from bacteria protects arthropods from host plant cyanide poisoning. eLife 3:e02365.
10 Scholl EH, Thome JL, McCarter JP, Bird DM (2003) Horizontally transferred genes in plant-parasitic nematodes: A high-throughput genomic approach. Genome Biol 4(6):R39.
11 Andersson JO, Doolittle WF, Nesbø CL (2001) Genomics. Are there bugs in our genome? Science 292(5523):1848–1850.
12 Stanhope MJ, et al. (2001) Phylogenetic analyses do not support horizontal gene transfers from bacteria to vertebrates. Nature 411(6840):940–944.
13 Crisp A, Boschetti C, Perry M, Tumacliffe A, Micklem G (2015) Expression of multiple horizontally acquired genes is a hallmark of both vertebrate and invertebrate genomes. Genome Biol 16(1):50.
14 Dunning Hotopp JC, et al. (2007) Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. Science 317(5845):1753–1756.
15 Kondo N, Nikoh N, Ijichi N, Shimada M, Fukatsu T (2002) Genome fragment of Wolbachia endosymbiont transferred to X chromosome of host insect. Proc Natl Acad Sci USA 99(22):14280–14285.
16 Ku C, et al. (2015) Endosymbiotic origin and differential loss of eukaryotic genes. Nature 524(7566):427–432.
17 Selosse MA, Le Tacon F (1998) The land flora: A phototroph-fungus partnership? Trends Ecol Evol 13(1):15–20.
18 Richards TA, et al. (2009) Phylogenomic analysis demonstrates a pattern of rare and ancient horizontal gene transfer between plants and fungi. Plant Cell 21(7):1897–1911.
19 Boschetti C, et al. (2012) Biochemical diversification through foreign gene expression in bdelloid rotifers. PLoS Genet 8(11):e1003035.
20 Kumar S, Jones M, Koutsovoulos G, Clarke M, Blaxter M (2013) Blobology: Exploring raw genome data for contaminants, symbionts and parasites using taxon-annotated GC-coverage plots. Front Genet 4:237.
21 Schönknecht G, et al. (2013) Gene transfer from bacteria and archaea facilitated evolution of an extremophilic eukaryote. Science 339(6124):1207–1210.
22 Gojkovic Z, et al. (2004) Horizontal gene transfer promoted evolution of the ability to propagate under anaerobic conditions in yeasts. Mol Genet Genomics 271(4):387–393.
23 Friesen TL, et al. (2006) Emergence of a new disease as a result of interspecific virulence gene transfer. Nat Genet 38(8):953–956.
24 Delmont TO, Eren AM (2016) Identifying contamination with advanced visualization and analysis practices: metagenomic approaches for eukaryotic genome assemblies. PeerJ 4:e1839.
25 Sogin ML, et al. (2006) Microbial diversity in the deep sea and the underexplored “rare biosphere”. Proc Natl Acad Sci USA 103(32):12115–12120.
26 Soanes D, Richards TA (2014) Horizontal gene transfer in eukaryotic plant pathogens. Annu Rev Phytopathol 52(1):583–614.
27 Schönknecht G, Weber APM, Lercher MJ (2014) Horizontal gene acquisitions by eukaryotes as drivers of adaptive evolution. BioEssays 36(1):9–20.
28 Carlton JM, et al. (2007) Draft genome sequence of the sexually transmitted pathogen Trichomonas vaginalis. Science 315(5809):207–212.
29 Dujon B, et al. (2004) Genome evolution in yeasts. Nature 430(6995):35–44.
30 Richards TA, Leonard G, Soanes DM, Talbot NJ (2011) Gene transfer into the fungi. Fungal Biol Rev 25(2):98–110.
31 Loftus B, et al. (2005) The genome of the protozoan parasite Entamoeba histolytica. Nature 437(7028):865–868.
32 Eichinger L, et al. (2005) The genome of the amoeboflagellate Dictyostelium discoideum. Nature 435(7038):43–57.
33 Hirt RP, Alsmark C, Embley TM (2013) Lateral gene transfers and the origins of the eukaryote proteome: A view from microbial parasites. Curr Opin Microbiol 23:155–162.