Correspondence

Probable horizontal transfer of a gene between a protist and a cnidarian

Robert E. Steele[1,2], Steven E. Hampson[3], Nicholas A. Stover[4], Dennis F. Kibler[3] and Hans R. Bode[2,5]

Horizontal gene transfer has been well documented as a significant feature of genome evolution in prokaryotes [1–3]. The frequency and significance of this process in eukaryotic evolution is much less clear. Confirming that a gene has entered a species by a horizontal route can be difficult, and many previously reported cases of horizontal gene transfer from a prokaryote to a eukaryote have subsequently been invalidated [4]. Because of potential contamination, extra care must be taken when putative horizontal gene transfers are detected in sequence datasets from organisms that may harbor a significant number of parasites and commensals. In a small number of metazoan phyla, a significant fraction of the mRNA contains a leader sequence that is attached by *trans*-splicing [5]. These spliced leaders are distinct for each phylum and thus provide a tag that can be used to determine the origin of a *trans*-spliced mRNA. Thus, presence of a spliced leader on a mRNA from a *trans*-splicing species would confirm that the gene was not derived from a non *trans*-splicing contaminating organism.

Hydra, a member of the basal phylum Cnidaria, is one of the few metazoan taxa in which *trans*-spliced leaders are added [5,6]. In a large EST (expressed sequence tag) project with *Hydra magnipapillata* (www.hydrabase.org), we have identified a cDNA homologous to the *flp* genes of the parabasalid protist *Trichomonas vaginalis* (Figure 1A,B). The *flp1* and *flp2* genes of *T. vaginalis* are of unknown function, but their expression is dramatically down-regulated in response to increasing levels of iron [7]. A third *flp* gene (*flp3*) was identified by searching the unfinished *T. vaginalis* genome sequence with the *flp1* and *flp2* sequences. The *flp* genes are not found in any other EST dataset, nor are they found in any of the other sequenced genomes. The *flp* gene is conspicuously absent from the genome datasets of other protists (e.g., *Plasmodium falciparum*, *Giardia lamblia*, the 8x coverage of the *Tetrahymena thermophila* macronuclear genome, and the approximately 12x coverage of the *Entamoeba histolytica* genome). A BLAST comparison of the *Hydra flp* amino acid sequence with the *T. vaginalis* sequences yields E values of 6×10^{-22} and 5×10^{-20} for *flp1* and *flp2*, respectively, supporting the identification of the *Hydra* protein as a *flp* homolog.

The *Hydra flp* cDNA contains a spliced leader identical to spliced leader B (Figure 1A,C), which we previously identified in *Hydra vulgaris* [6]. Among the GenBank entries from a *Hydra magnipapillata* EST project carried out at the National Institute of Genetics in Japan, we identified two *flp* cDNAs that are identical in sequence to the one in Figure 1A, except for the sequence at the 5¢ end. The 5¢ sequence of these cDNAs resembles spliced leader B, but is clearly distinct from it. The spliced leader RNA genes we identified previously in *H. vulgaris* are located within the 5S rDNA repeats [6]. To confirm that this novel 5¢ sequence in *H. magnipapillata* is a spliced leader, we used PCR to amplify 5S rDNA repeats from this species. Sequence analysis of one of these repeats showed that it contained a gene with the new spliced leader followed by the expected GT splice donor dinucleotide (Figure 1C). We have termed this new sequence ‘spliced leader D’. Thus, both types of *flp* cDNAs contain *Hydra* spliced leader sequences, indicating that neither of them originated from a contaminating organism. One could, however, argue that the *flp* genes identified as derived from *Trichomonas* actually come from a metazoan. However, the *Trichomonas flp* genes have been isolated as cDNA clones and genomic clones and probing of *Trichomonas* DNA and RNA blots probed with a *flp* cDNA clone gave strong signals [7]. Furthermore, *Trichomonas* is grown in the laboratory in a defined culture medium [7]. Thus, it is unlikely that the cloned *Trichomonas flp* genes are a contamination from a metazoan.

The anomalous phylogenetic distribution of *flp* suggests that this gene has been horizontally transferred [1]. The preferred approach for testing such a hypothesis is phylogenetic analysis with an appropriate outgroup species [8]. However, the limited occurrence of the *flp* gene makes such an analysis impossible. An alternative explanation for the distribution of *flp* genes is that they were ancestrally present in all eukaryotes and secondarily lost from most taxa. However, this seems unlikely as the large amount of sequence available for diverse taxa did not reveal *flp* genes. Rather, it seems most likely that the *flp* gene was horizontally transferred. The direction of transfer is at present uncertain. We favor the idea of transfer from a parabasalid protist to a cnidarian, but additional information on the distribution of *flp* genes in parabasalid protists and cnidarians will be required to resolve this question. It will be of obvious interest to learn more about the function of the *flp* genes in both *Hydra* and *Trichomonas,* in order to understand the biological significance of this gene transfer. The conserved cysteines and histidines in the *Trichomonas* and *Hydra flp* proteins may mediate metal binding, and could explain the strong response of *flp* gene expression to iron levels in *Trichomonas* [7].

Acknowledgments
The *Hydra* EST project is being carried out in collaboration with the Genome Sequencing Center at Washington University. The

Figure 1. Sequence analysis of a *Hydra flp* gene.
(A) Nucleotide and predicted amino acid sequence of a *Hydra flp* cDNA. The spliced leader sequence is underlined. (B) Alignment of the predicted *Hydra flp* amino acid sequence and the sequences of *flp1*, *flp2*, and *flp3* from *Trichomonas vaginalis*. The alignment was performed using CLUSTAL_X [9]. The *T. vaginalis flp1* and *flp2* sequences were taken from [7]. The *T. vaginalis flp3* sequence was identified by BLAST analysis of the unfinished *T. vaginalis* genome sequence, available from The Institute for Genomic Research. Amino acids identical between the *Hydra* sequence and one or more of the *T. vaginalis* sequences are shaded in gray. (C) Spliced leader B (underlined in the upper sequence) [6] and spliced leader D (underlined in the middle sequence) sequences occur at the 5' ends of *Hydra flp* cDNAs. Bottom sequence, genomic DNA containing spliced leader D (underlined) followed by a GT splice donor dinucleotide (shaded residues). The genomic sequence for the spliced leader D gene was obtained from a PCR product generated using primers for the 5S rRNA genes, which flank spliced leader genes in *Hydra* [6]. The sequence has been deposited in GenBank under accession number AY533704. The *Hydra flp* sequences have been deposited as third party annotation entries in GenBank under accession numbers BK004161 and BK004162.

project is generously supported by grant IBN-0120591 from the National Science Foundation.

References

1. Philippe, H., and Douady, C.J. (2003). Horizontal gene transfer and phylogenetics. Curr. Opin. Microbiol. *6*, 498–505.
2. Eisen, J.A. (2000). Horizontal gene transfer among microbial genomes: new insights from complete genome analysis. Curr. Opin. Genet. Dev. *10*, 606–611.
3. Jain, R., Rivera, M.C., Moore, J.E., and Lake, J.A. (2002). Horizontal gene transfer in microbial genome evolution. Theor. Popul. Biol. *61*, 489–495.
4. Genereux, D.P., and Logsdon, J.M., Jr. (2003). Much ado about bacteria-to-vertebrate lateral gene transfer. Trends Genet. *19*, 191–195.
5. Nilsen, T.W. (2001). Evolutionary origin of SL-addition trans-splicing: still an enigma. Trends Genet. *17*, 678–680.
6. Stover, N.A., and Steele, R.E. (2001). Trans-spliced leader addition to mRNAs in a cnidarian. Proc. Natl. Acad. Sci. USA *98*, 5693–5698.
7. Crouch, M.V., and Alderete, J.F. (2001). *Trichomonas vaginalis* has two fibronectin-like iron-regulated genes. Arch. Med. Res. *32*, 102–107.
8. Stanhope, M.J., Lupas, A., Italia, M.J., Koretke, K.K., Volker, C., and Brown, J.R. (2001). Phylogenetic analyses do not support horizontal gene transfers from bacteria to vertebrates. Nature *411*, 940–944.
9. Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., and Higgins, D.G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. *25*, 4876–4882.

[1]Department of Biological Chemistry, University of California, Irvine, USA. [2]Developmental Biology Center, University of California, Irvine, USA. [3]Department of Computer Science, University of California, Irvine, USA. [4]Department of Ecology and Evolutionary Biology, Princeton University, USA. [5]Department of Developmental and Cell Biology, University of California, Irvine, USA