

the highly variable number of spores/basidium compared with their non-lichenized sister species, are further evidence that a high level of stress on the fungus is associated with a transition to the lichen symbiosis<sup>28</sup>. This may suggest that many attempts to form stable lichen symbioses occur in nature, but only rarely does a specific fungal lineage have all the requirements to survive the costs associated with a transition to this symbiotic state. □

## Methods

### DNA sequencing, taxon and character sampling

Total DNA was isolated, and the SSU and LSU nuclear rDNA were amplified, sequenced and aligned as described in ref. 16. Regions of the alignments where the placements of gaps were ambiguous were removed from the MCMC phylogenetic tree sampling analyses. Taxa were selected to represent all main ascomycete orders<sup>2</sup> known to include lichenized species (13 out of 15 orders) and nearly all main orders of Ascomycota known to include only non-lichenized species (11 out of 31 orders). At least 16 of the unsampled non-lichenized orders almost certainly fall entirely within existing non-lichenized clades (Fig. 2) and their sampling will not affect the results presented here. This is because the reconstruction of ancestral states is not weighted by the number of descendant taxa that have a particular state; rather, the reconstructed state depends on the relative frequencies of the states in the descendants and their phylogenetic distribution. If all of a group of unsampled taxa share a most recent common ancestor and the same state with a species already included in our study, our reconstructions are unchanged. Basidiomycota (*Athelia bombacina* and *Coprinus cinereus*) sequences were included as outgroups. (The voucher/GenBank information is available as Supplementary Information.) We generated a total of 20 SSU and 24 LSU nuclear rDNA sequences in this study. All these sequences were deposited in GenBank under accession numbers AF356653–AF356696. Ten SSU and 20 LSU sequences were from ref. 16, and the remaining 24 SSU and 10 LSU sequences were from GenBank.

### MCMC phylogenetic tree sampling

We used MCMC methods<sup>7</sup> within a Bayesian framework to estimate the posterior probability of phylogenetic trees. The MCMC procedure ensures that trees are sampled in proportion to their probability of occurrence under the model of gene-sequence evolution. We generated 200,000 phylogenetic trees using the MCMC procedure, sampling every tenth one to assure that successive samples were independent<sup>7,29</sup>. We then removed the first 100 trees in the sample to avoid including any trees that might have been sampled before convergence of the Markov chain. We used the general time-reversible model of gene-sequence evolution combined with gamma rate heterogeneity to estimate the likelihood of each tree<sup>30</sup>. Information on the state of each species (lichen-forming/non-lichen-forming) was excluded from the MCMC sampling procedure to ensure that the distribution of tree topologies was not influenced by this trait. A series of runs using the BAMBE<sup>7</sup> 'global' and 'local' options was conducted to ensure that the Markov chain converged to the same region in the universe of trees.

### Reconstruction of gains and losses, and ancestral states

We used a continuous time Markov model of trait evolution, as implemented in the computer program Discrete (available from M.P.), allowing independent gains and losses in each branch of the phylogenetic tree<sup>29</sup>. Parameters specifying rates of gain ( $q_{01}$ ) and loss ( $q_{10}$ ) of lichenization were calculated separately for each tree sampled in the MCMC procedure, following procedures in ref. 9. Because trees are represented in the sample in proportion to their likelihood, investigating the rates over all trees automatically weights our results by the likelihood of a particular tree type. A detailed description of the analyses performed for this study will be published as a book chapter (M.P. and E.L., manuscript in preparation).

Received 21 December 2000; accepted 30 April 2001.

- Hawksworth, D. L. The fungal dimension of biodiversity: magnitude, significance, and conservation. *Mycol. Res.* **95**, 641–655 (1991).
- Hawksworth, D. L., Kirk, P. M., Sutton, B. C. & Pegler, D. N. *Dictionary of the Fungi* (CAB, Wallingford, 1995).
- Barinaga, M. Origins of lichen fungi explored. *Science* **268**, 1437 (1995).
- Gargas, A., DePriest, P. T., Grube, M. & Tehler, A. Multiple origins of lichen symbioses in fungi suggested by SSU rDNA phylogeny. *Science* **268**, 1492–1495 (1995).
- Honegger, R. in *Lichen Biology* (ed. Nash, T. H.) 24–36 (Cambridge Univ. Press, Cambridge, 1996).
- Hafellner, J. in *Handbook of Lichenology* Vol. 3 (ed. Galun, M.) 41–52 (CRC, Boca Raton, 1988).
- Larget, B. & Simon, D. L. Markov chain monte carlo algorithms for the Bayesian analysis of phylogenetic trees. *Mol. Biol. Evol.* **16**, 750–759 (1999).
- Huelsensbeck, J. P., Rannala, B. & Masly, J. P. Accommodating phylogenetic uncertainty in evolutionary studies. *Science* **288**, 2349–2350 (2000).
- Pagel, M. The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies. *Syst. Biol.* **48**, 612–622 (1999).
- Berbee, M. L., Yoshimura, A., Sugiyama, J. & Taylor, J. W. Is *Penicillium* monophyletic? An evaluation of phylogeny in the family Trichocomaceae from 18S, 5.8S, and ITS ribosomal DNA sequence data. *Mycologia* **87**, 210–222 (1995).
- Alexopoulos, C. J., Mims, C. W. & Blackwell, M. *Introductory Mycology* (John Wiley & Sons, New York, 1996).
- Suh, S.-O. & Blackwell, M. Molecular phylogeny of the cleistothecial fungi placed in Cephalothecaceae and Pseudeurotiaceae. *Mycologia* **91**, 836–848 (1999).

- Berbee, M. L. & Taylor, J. W. From 18S ribosomal sequence data to evolution of morphology among the fungi. *Can. J. Bot. (Suppl. 1)* **73**, S677–S683 (1995).
- Spatafora, J. W. Ascomal evolution of filamentous ascomycetes: evidence from molecular data. *Can. J. Bot. (Suppl. 1)* **73**, S811–S815 (1995).
- Berbee, M. L., Carmean, D. A. & Winka, K. Ribosomal DNA and resolution of branching order among the Ascomycota: how many nucleotides are enough? *Mol. Phylog. Evol.* **17**, 337–344 (2000).
- Bhattacharya, D. et al. Widespread occurrence of spliceosomal introns in the rDNA genes of ascomycetes. *Mol. Biol. Evol.* **17**, 1971–1984 (2000).
- Liu, Y. J., Whelen, S. & Hall, B. D. Phylogenetic relationships among ascomycetes: evidence from the RNA polymerase II subunit. *Mol. Biol. Evol.* **16**, 1799–1808 (1999).
- Collins, T. M., Wimberger, P. H. & Naylor, G. J. P. Compositional bias, character-state bias, and character-state reconstruction using parsimony. *Syst. Biol.* **43**, 482–496 (1994).
- Maddison, D. R. Phylogenetic methods for inferring the evolutionary history and process of change in discretely valued characters. *Ann. Rev. Entomol.* **39**, 267–292 (1994).
- Pagel, M. Detecting correlated evolution on phylogenies: a general method for the comparative analysis of discrete characters. *Proc. R. Soc. Lond. B* **255**, 37–45 (1994).
- Pagel, M. Inferring the historical patterns of biological evolution. *Nature* **401**, 877–884 (1999).
- Elix, J. A. in *Lichen Biology* (ed. Nash, T. H.) 154–180 (Cambridge Univ. Press, Cambridge, 1996).
- Rambold, G. & Triebel, D. The inter-lecanoralean associations. *Bibliotheca Lichenologica* **48**, 1–201 (1992).
- Hibbett, D. S., Gilbert, L.-B. & Donoghue, M. J. Evolutionary instability of ectomycorrhizal symbioses in basidiomycetes. *Nature* **407**, 506–508 (2000).
- Lutzoni, F. & Vilgalys, R. *Omphalina* (Basidiomycota, Agaricales) as a model system for the study of coevolution in lichenized fungi. *Crypt. Bot.* **5**, 82–97 (1995).
- Lutzoni, F. Phylogeny of lichen- and non lichen-forming omphalinoid mushrooms and the utility of testing for combinability among multiple data sets. *Syst. Biol.* **46**, 373–406 (1997).
- Lutzoni, F. & Pagel, M. Accelerated evolution as a consequence of transition to mutualism. *Proc. Natl Acad. Sci. USA* **94**, 11422–11427 (1997).
- Kramer, I. & Lutzoni, F. in *Plant Responses to Environmental Stresses: From Phytohormones to Genome Reorganization* (ed. Lerner, H. R.) 591–628 (Marcel Dekker, New York, 1999).
- Wilson, I. & Balding, D. Genealogical inference from microsatellite data. *Genetics* **150**, 499–510 (1998).
- Hillis, D. M., Moritz, C. & Mable, B. K. *Molecular Systematics* 2nd edn (Sinauer, Sunderland, 1996).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

## Acknowledgements

We thank P. Lewis and K. Pryer for insightful suggestions; J. Spatafora, D. Armaleo, C. and W. Culbertson for initiating the first phase of this project; U. Söchting, J. Spatafora, J. Johnson and S. LaGreca for providing unpublished sequences; J. Bêlisle for assistance with Figs 1 and 2; J. Crodian for technical assistance; I. Brodo, D. Gernandt, C. Keller, T. Lumbsch, J. Miadlikowska, J. Platt and U. Söchting for providing lichen material or DNA samples; F. Kauff, T. Bjelland, B. Büdel, P. M. Jørgensen and M. Schultz for the identification of some lichen specimens; and M. Blackwell for comments on the manuscript. This work was supported by a grant from the US National Science Foundation to E.L. and the Pritzker Foundation Fund of The Field Museum. M.P. is supported by the Leverhulme Trust and the Biotechnology and Biological Sciences Research Council of the UK.

Correspondence and requests for materials should be addressed to E.L. (e-mail: flutzoni@fmmh.org).

# Phylogenetic analyses do not support horizontal gene transfers from bacteria to vertebrates

Michael J. Stanhope, Andrei Lupas, Michael J. Italia, Kristin K. Koretke, Craig Volker & James R. Brown

Bioinformatics, GlaxoSmithKline, 1250 South Collegeville Road, UP1345, Collegeville, Pennsylvania 19426, USA

Horizontal gene transfer (HGT) has long been recognized as a principal force in the evolution of genomes<sup>1</sup>. Genome sequences of Archaea and Bacteria have revealed the existence of genes whose similarity to loci in distantly related organisms is explained most parsimoniously by HGT events<sup>2–4</sup>. In most multicellular organisms, such genetic fixation can occur only in the germ line. Therefore, it is notable that the publication of the human genome reports 113 incidents of direct HGT between bacteria

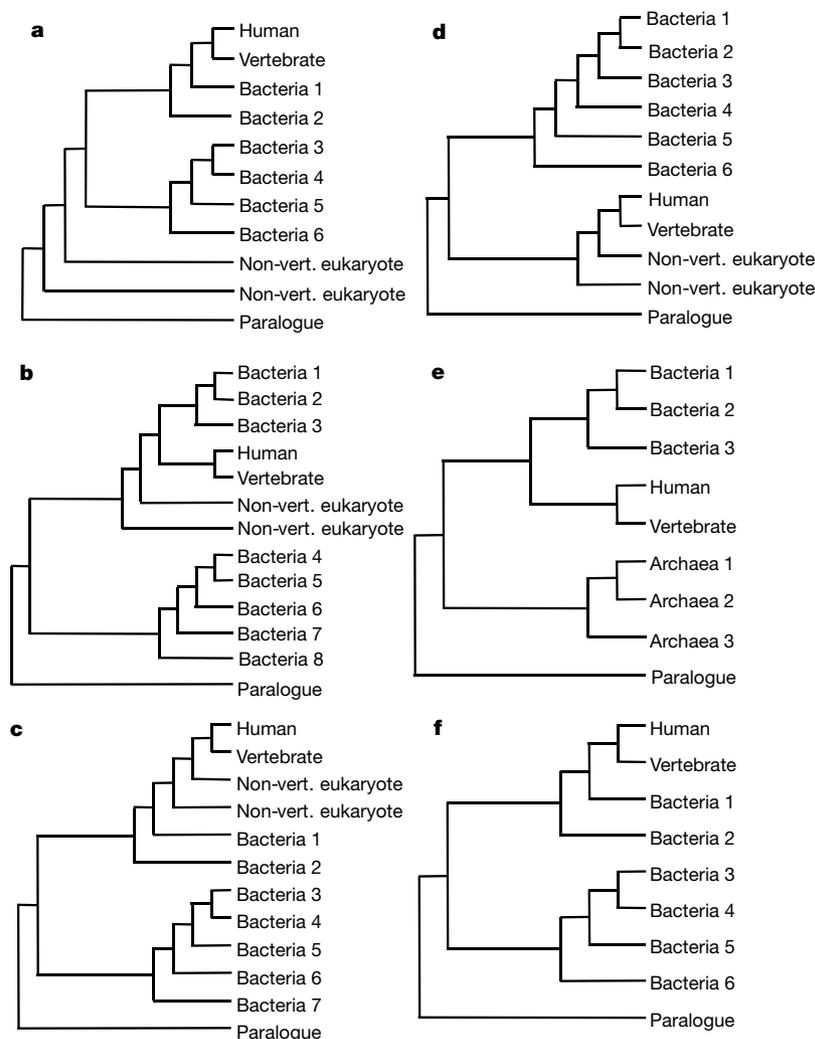
and vertebrates<sup>5</sup>, without any apparent occurrence in evolutionary intermediates, that is, non-vertebrate eukaryotes. Phylogenetic analysis arguably provides the most objective approach for determining the occurrence and directionality of HGT<sup>6,7</sup>. Here we report a phylogenetic analysis of 28 proposed HGT genes, whose presence in the human genome had been confirmed by polymerase chain reaction (PCR)<sup>5</sup>. The results indicate that most putative HGT genes are present in more anciently derived eukaryotes (many such sequences available in non-vertebrate EST databases) and can be explained in terms of descent through common ancestry. They are, therefore, unlikely to be examples of direct HGT from bacteria to vertebrates.

Most of the phylogenetic analyses (see Methods and Fig. 1 for a description of the phylogenetic principles regarding acceptance or rejection of HGT) for the 28 loci analysed here supported the monophyly of eukaryotes, with a non-vertebrate ('non-vertebrate' refers to invertebrate animals as well as fungi, plants and protists) eukaryote at the base of this clade (Table 1 and Fig. 2). Such outcomes were congruent between both parsimony and distance-based methods of phylogenetic analysis. The explanations for why such a conclusion was not reached previously are somewhat varied. We found that orthologues to various proteins, purported to be HGTs in the recent human genome publication<sup>5</sup>, could actually be

found in the expressed sequence tags (ESTs) of non-vertebrates, which are accessible through the "EST others database" on the NCBI BLAST page (<http://www.ncbi.nlm.nih.gov/BLAST>).

A common search result was the collection of a homologue from the slime mould *Dictyostelium discoideum*, which when included in the alignment resulted in eukaryotic monophyly with *D. discoideum* at the base. One such specific example involves the protein known as formiminotransferase cyclodeaminase (accession number AAG01853.1; this and other accession numbers used here are from Table 24 in ref. 5). This seems to be a duplicate enzyme in eukaryotes, with a cyclodeaminase carboxy-terminal domain and a formiminotransferase amino-terminal domain. The only bacterium at present that can clearly be shown to have this arrangement is *Streptococcus pyogenes*. All other taxa either have separate proteins or just the cyclodeaminase. Alignments constructed on this C-terminal half (chosen because more taxa were available for this enzyme) resulted in phylogenetic trees with 100% bootstrap support for the monophyly of eukaryotes, with *D. discoideum* (EST assembly of AU060687.1 and C94072.1) at the base of the clade (Fig. 2a).

Another example of an orthologue that was missed owing to overlooking the EST others database involves monoamine oxidase (AAA59548.1 and AAB27229.1), cited as an example of a gene



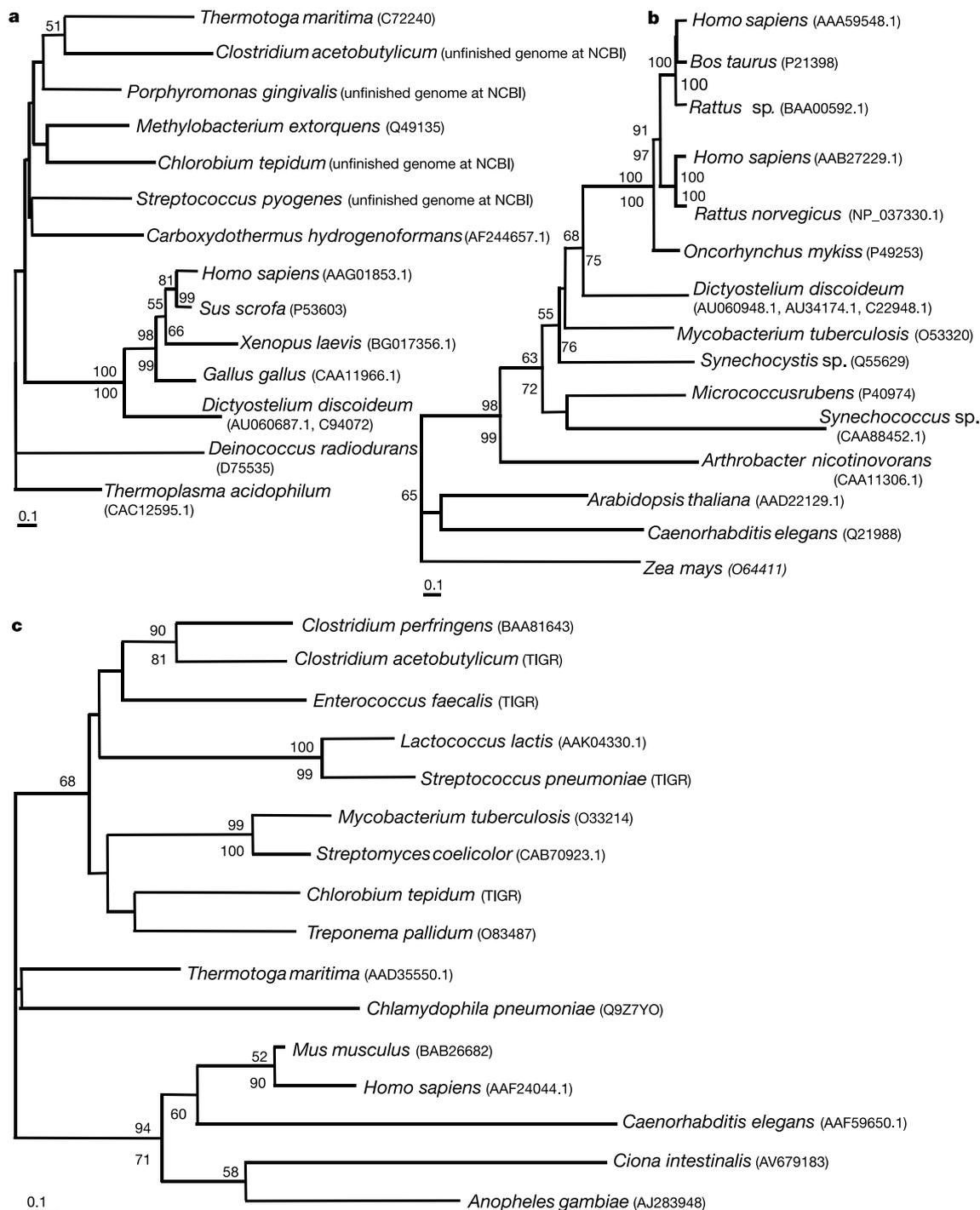
**Figure 1** Hypothetical phylogenetic trees showing various evolutionary models relating to acceptance or rejection of HGT. Paralogue, an anciently derived paralogous sequence; Non-vert. eukaryote, non-vertebrate eukaryote. **a**, Phylogenetic history in support of bacteria-vertebrate HGT. **b**, Phylogenetic history in support of vertebrate-bacteria HGT.

**c-e**, Phylogenetic history rejecting any sort of HGT involving bacteria and vertebrates. **f**, Phylogenetic history which, owing to the absence of a non-vertebrate eukaryote, is considered ambiguous with regard to bacteria-vertebrate HGT.

horizontally transferred into vertebrates, which probably conferred a physiological selective advantage and was thus fixed in vertebrate evolution<sup>5</sup>. This gene may indeed have an important neurological function<sup>8–10</sup>; however, the presence of a *D. discoideum* orthologue at the base of monophyletic eukaryotes (Fig. 2b) argues strongly against the acquisition of this locus by means of HGT.

A different methodological reason for several of the genes in the human genome report<sup>5</sup> being considered as bacteria–vertebrate HGTs, was that phylogenetics was not the analytical approach, and

that the conclusions were instead derived largely from top BLAST hit results. In several instances the top BLAST hit was indeed a bacterial species, whereas further down the list of significant BLAST hits one finds a non-vertebrate eukaryote. When such sequences were properly aligned, the resulting phylogenetic trees often supported the monophyly of eukaryotes with the non-vertebrate eukaryote at the base. For example, searches against the non-redundant (NR) database involving the short peptide AAF24044.1 resulted in relatively few hits, with a top BLAST hit



**Figure 2** Neighbour-joining phylogenetic trees on the basis of PAM Dayhoff sequence-divergence matrixes, showing evolutionary history of various loci from Table 1. Branch lengths are drawn proportional to the amount of sequence change. Numbers at various nodes indicate bootstrap support values for both neighbour-joining (top) and parsimony

(bottom) analyses; only values in excess of 50% are indicated. **a**, AAG01853.1, formiminotransferase cyclodeaminase. **b**, AAA59548.1, monoamine oxidase A and AAB27229.1, monoamine oxidase B. **c**, AAF24044.1, uncharacterized protein.

outside Mammalia corresponding to the bacterial species *Thermotoga maritima*, and with *Caenorhabditis elegans* listed much further down the list. Despite this arrangement of significant BLAST hits, once the sequences were all aligned and phylogenetic analyses performed, *C. elegans* fell in as the sister group to mammals, and eukaryotes were monophyletic. This conclusion is substantiated further through a search of the EST others database, revealing a sequence for *Anopheles gambiae* (mosquito; AJ283948) and *Ciona intestinalis* (ascidian; AV679183), with the resulting topology again supporting eukaryote monophyly, this time inclusive of two mammals and three invertebrates from disparate phyla (Fig. 2c).

If a particular human gene was either a highly divergent member of a gene family widely found in both vertebrates and non-vertebrates, or short in length but highly matched a motif within a larger bacterial protein, then vertebrates and bacteria were often top hits in BLAST searches. An example is the 222-amino-acid surfactin synthetase domain (IGI\_M1\_ctg25107\_24) for which the top BLAST hit is Srf1, a 3,587-amino-acid protein in *Bacillus subtilis*. Subsequent phylogenetic analyses on the conserved regions between bacteria and humans, as well as fungi and Nematoda, resulted in a monophyletic cluster of eukaryotic sequences. Ribosomal protein S6-glutamic acid ligase (IGI\_M1\_ctg12741\_7) exists in humans, mouse, frog and fly (these latter two were detected in EST others database), and all eukaryotic sequences are monophyletic with high bootstrap support.

One of the twenty-eight loci in question was rejected as a possible bacteria-vertebrate HGT (BAA11432.1) because of its very limited bacterial spectrum. This gene did not strictly adhere to our phylogenetic criteria for rejection, or to the reverse transfer; however, it had very few bacterial orthologues (only *Mycobacterium tuberculosis* RV1834 (Q50599) and *Pseudomonas aeruginosa* PA3429

(C83216)), with a broader representation of vertebrates (fish, frog and a few mammals). The two bacteria did not form a clade in the resulting phylogeny. The most parsimonious explanation of this spectrum and branching arrangement is the evolution of an  $\alpha/\beta$  hydrolase in vertebrates, which was then transferred independently into *Mycobacterium* and *Pseudomonas*. In our opinion this limited spectrum is a compelling argument against bacteria-vertebrate HGT. Such an assortment of taxa would necessitate that this locus evolved in bacteria, was subsequently lost in a host of independent lineages (remembering that there are a significant number of complete bacterial genome sequences available), retained in only *Mycobacterium* and *Pseudomonas*, and then one or another of these taxa transferred this gene into a vertebrate lineage germ line. A similar, although more complex case, is presented by a family of paralogous human transporters (CAB81772.1, AAB59448.1, AAA36608.1 and AAC41747.1; see Table 1), which contains two further members (CAC00574.1, AW904970) listed in the Supplementary Table of ref. 5.

In our analyses of these data there are a few examples where the evolutionary history is unclear (listed under ambiguous in Table 1), and thus for which a possible bacteria-vertebrate transfer cannot be rejected. However, no such phylogenetic criteria exist for these same genes to substantiate the claims of bacteria-vertebrate HGT. Most of our analyses and phylogenetic topologies are highly consistent with the view that vertebrates and bacteria share these loci through common ancestry, involving a succession of non-vertebrate eukaryote intermediates. A further point arising from our analysis is that the evolutionary relationships among proteins cannot be concluded solely from the ranking of database hits in homology searches (for example, BLAST reports). This is not a new conceptual point (see refs 7, 12, 13), but one that seems to have been overlooked in this instance. Phylogenetic analysis must be a central component of any

**Table 1 Alternative analysis of proposed vertebrate acquisitions of bacterial genes**

Human protein* (accession number)	Phylogenetic support of vert-bac HGT (Fig. 1b)	Phylogenetic rejection of bac-vert HGT (Fig. 1c, d, e)	Other rationale for rejecting bac-vert HGT	Ambiguous: absence of non-vertebrate eukaryote (Fig. 1f)
AAG01853.1		Yes†		
CAB81772.1			Yes‡	
AAB59448.1			Yes‡	
AAA36608.1			Yes‡	
AAC41747.1			Yes‡	
BAA11432.1			Yes§	
CAB59628.1	Yes			
BAA91273.1		Yes†		
CAA75608.1		Yes		
AAA59548.1		Yes†		
AAB27229.1		Yes†		
AAF12736.1		Yes†		
AAA51565.1				Yes
BAA92632.1			Yes	
BAA34458.1		Yes		
AAF24044.1		Yes†		
BAA91839.1				Yes
BAA92073.1			Yes¶	
BAA92133.1				Yes
BAA91174.1				Yes
AAA60043.1		Yes		
BAA86552.1		Yes†		
IGI_M1_ctg12741_7		Yes†		
IGI_M1_ctg13238_61		Yes†		
IGI_M1_ctg13305_116		Yes†		
IGI_M1_ctg14420_10		Yes†		
IGI_M1_ctg16010_18		Yes		
IGI_M1_ctg25107_24		Yes		

For the 28 loci there was no phylogenetic evidence to support the claim of bacteria-vertebrate HGT (Fig. 1a). Sequence BAA34458.1 did not correspond to a  $\beta$ -lactamase-like hydrolase.

\* Accession numbers correspond to those given in Table 24 of ref. 5. Sequences IGI\_M1\_ctg19153\_147 and IGI\_M1\_ctg16227\_58 from Table 24 were not found in the IPI\_1 database of putative human proteins.

† Searches in which a sequence was recovered from EST others, which proved instrumental in rejecting the bacteria-vertebrate hypothesis.

‡ Large family of paralogous transporters in metazoa; few, phylogenetically diverse bacteria.

§ Few, phylogenetically diverse bacteria.

|| Unique N terminus fused with C-terminal decarboxylase, widely found in Bacteria, Archaea and Eucarya.

¶ Bacterial and human genes are not orthologous.

protein family or genome annotation effort. Importantly, phylogenetic reconstruction is critical to synthesizing, from the growing wealth of sequence data, a more comprehensive view of genome evolution. □

## Methods

### Data collation and analysis

Our analysis involved a set of 28 loci that had been proposed as instances of bacteria–vertebrate HGT, and for which PCR had been used to verify their presence in the human genome<sup>5</sup>. Each protein sequence was searched using the appropriate BLASTP or TBLASTN<sup>14</sup> algorithm against the following databases downloaded on 19 February 2001 from NCBI: non-redundant proteins; EST others nucleotides (GenBank non-mouse and non-human EST entries); unfinished microbial genome nucleotides; and the complete genome sequences of the nematode *C. elegans* and the insect *Drosophila melanogaster*. ClustalW<sup>15</sup> was used to align the resulting set of significant homologues, and the alignments were then refined by eye using the GCG or GeneDoc sequence editors. In several instances BLAST searches resulted in a large number of homologous sequences, which served as input for a preliminary phylogeny designed to assess paralogy (sequence homology due to gene duplication) and orthology (sequence homology due to speciation). This tree then served as a framework on which to exclude the more distantly related paralogues in a follow-up alignment and phylogeny. All potentially ambiguous gaps (that is, those involving complex insertion or deletion events) were removed before phylogenetic analysis. We coded all remaining gaps as missing data.

The amino-acid sequence alignments were analysed using maximum parsimony and neighbour joining, as implemented in the phylogenetic package PHYLIP<sup>16</sup>. Parsimony analyses randomized the input order 20–50 times, depending on the number of sequences in the alignment (that is, greater number of random additions for larger alignments). The distance matrixes that served as input for neighbour-joining analyses were calculated using the point-accepted-mutation (PAM) Dayhoff substitution model. Clade strength was assessed with bootstrap, using 1,000 replications. The alignments are available in Supplementary Information.

### Phylogenetic principles for accepting or rejecting HGT

The branching arrangement of the resulting phylogenies is a critical component of an HGT assessment. In the case of a gene transfer from bacteria to vertebrates, the necessary phylogeny would show eukaryote paraphyly with vertebrates separated from non-vertebrates by a paraphyletic (a group that contains some but not all of the descendants of a common ancestor) assemblage of bacterial lineages, where some of the bacteria were more closely related to vertebrates than other bacteria, and in which most (or all) of available bacterial sequences form a clade with vertebrates (Fig. 1a). Horizontal gene transfer of the opposite direction (vertebrate–bacteria) would be expected to have a topology in which one or a few bacteria were sister group to the vertebrates, joined next by non-vertebrate taxa, and the remaining bacteria (if any) were outside this clade (Fig. 1b). Both such trees would need a root, which could come from Archaea or a paralogous gene. The clearest phylogenetic rejection of proposed bacteria–vertebrate HGT would involve eukaryotic monophyly, with a non-vertebrate eukaryote contained within that clade (Fig. 1c, d). Another example of phylogenetic rejection would include the respective monophylies of Archaea, Eucarya and Bacteria (or conversely, just Eucarya and Bacteria) in a tree rooted at a paralogous sequence, but with only vertebrates represented as eukaryotes (that is, bacterial monophyly negates the bacteria–vertebrate HGT hypothesis; Fig. 1e). Several ambiguous cases could arise, the most likely relating to the absence of non-vertebrate orthologous genes. A paraphyletic group of bacteria forming a clade with vertebrates, joined next by either Archaea or a paralogous gene family, could be a consequence of the poor representation of non-vertebrate genomes in contemporary databases, or could reflect an actual HGT from bacteria to vertebrates (Fig. 1f). The absence of a sequence from either a nematode (*C. elegans*) or fruit fly (*D. melanogaster*)—whose nearly complete genomes are available—in such a topology is not sufficient evidence to conclude bacteria–vertebrate HGT. Until a much broader sampling of non-vertebrate genomes are completed these cases should remain ambiguous.

Received 18 April; accepted 24 May 2001.

- Gray, M. W. The endosymbiont hypothesis revisited. *Int. Rev. Cytol.* **141**, 233–357 (1992).
- Nelson, K. E. *et al.* Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of *Thermotoga maritima*. *Nature* **399**, 323–329 (1999).
- Ruepp, A. *et al.* The genome sequence of the thermoacidophilic scavenger *Thermoplasma acidophilum*. *Nature* **407**, 508–513 (2000).
- Doolittle, W. F. Phylogenetic classification and the universal tree. *Science* **284**, 2124–2128 (1999).
- International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
- Smith, M. W., Feng, D.-F. & Doolittle, R. F. Evolution by acquisition: the case for horizontal gene transfers. *Trends Biochem. Sci.* **17**, 489–493 (1992).
- Logsdon, J. M. Jr & Faguy, D. M. Evolutionary genomics: *Thermotoga* heats up lateral gene transfer. *Curr. Biol.* **9**, R747–R751 (1999).
- Brunner, H. G., Nelen, M., Breakefield, X. O., Ropers, H. H. & van Oost, B. B. A. Abnormal behavior associated with a point mutation in the structural gene for monoamine oxidase A. *Science* **262**, 578–580 (1993).
- Cases, O. *et al.* Aggressive behavior and altered amounts of brain serotonin and norepinephrine in mice lacking MAOA. *Science* **268**, 1763–1766 (1995).
- Deckert, J. *et al.* Excess of high activity monoamine oxidase A gene promoter alleles in female patients with panic disorder. *Hum. Mol. Genet.* **8**, 621–624 (1999).

- Brown, J. R. & Doolittle, W. F. Archaea and the prokaryote–eukaryote transition. *Microbiol. Mol. Biol. Rev.* **61**, 456–502 (1997).
- Kyrpides, N. C. & Olsen, G. J. Archaeal and bacterial hyperthermophiles—horizontal gene exchange or common ancestry? *Trends Genet.* **15**, 298–299 (1999).
- Eisen, J. A. Phylogenomics—improving functional predictions for uncharacterized genes by evolutionary analysis. *PCR Methods Appl.* **8**, 163–167 (1998).
- Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
- Thompson, J. D., Higgins, D. G. & Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680 (1994).
- Felsenstein, J. PHYLIP (Phylogeny Inference Package) version 3.6. (Department of Genetics, University of Washington, Seattle, 2000).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Correspondence and requests for materials should be addressed to J.R.B. (e-mail: [James\\_R\\_Brown@sbphrd.com](mailto:James_R_Brown@sbphrd.com)) or M.J.S. (e-mail: [Michael\\_J\\_Stanhope@sbphrd.com](mailto:Michael_J_Stanhope@sbphrd.com)).

## Divergent sexual selection enhances reproductive isolation in sticklebacks

Janette Wenrick Boughman

Department of Zoology, University of British Columbia, Vancouver BC V6T 1Z4, Canada

Sexual selection may facilitate speciation because it can cause rapid evolutionary diversification of male mating signals and female preferences. Divergence in these traits can then contribute to reproductive isolation<sup>1–3</sup>. The sensory drive hypothesis predicts that three mechanisms underlie divergence in sexually selected traits<sup>4</sup>: (1) habitat-specific transmission of male signals<sup>5–7</sup>; (2) adaptation of female perceptual sensitivity to local ecological conditions<sup>8</sup>; and (3) matching of male signals to female perceptual sensitivity<sup>4,9</sup>. I test these mechanisms in threespine sticklebacks (*Gasterosteus* spp.) that live in different light environments. Here I show that female perceptual sensitivity to red light varies with the extent of redshift in the light environment, and contributes to divergent preferences. Male nuptial colour varies with environment and is tuned to female perceptual sensitivity. The extent of divergence among populations in both male signal colour and female preference for red is correlated with the extent of reproductive isolation in these recently diverged species. These results demonstrate that divergent sexual selection generated by sensory drive contributes to speciation.

Sexual selection can act through sensory drive to favour different mating signals and preferences in different environments. The sensory drive hypothesis predicts three mechanisms that may cause such divergence. First, females are likely to prefer conspicuous signals. Because habitat influences the physics of signal transmission, relative conspicuousness of signals should vary with habitat<sup>4</sup>. With visual signals, conspicuousness is enhanced when a signal differs from the background light. In relatively clear, fresh water, blue and red are high-contrast signal colours<sup>10</sup>. In tea-stained fresh water, red signals are likely to be masked by the redshifted background light, but black should be high contrast<sup>11</sup>. Second, perceptual sensitivity should vary with habitat<sup>8,12</sup> and lead to variation in preference. For example, the spectral quality of ambient light is likely to influence colour perception<sup>12</sup>. Several sources of selection may favour perceptual systems that work effectively in the local environment, because the visual environment can affect prey and predator detection as well as mate detection<sup>12,13</sup>. Third, male signals that match female perception are easier to detect and