

A Functional Screen Identifies Lateral Transfer of β -Glucuronidase (*gus*) from Bacteria to Fungi

Peter Wenzl, Laurie Wong,¹ Kim Kwang-won,² and Richard A. Jefferson

Center for the Application of Molecular Biology to International Agriculture (CAMBIA), Canberra, Australia

Lateral gene transfer (LGT) from prokaryotes to microbial eukaryotes is usually detected by chance through genome-sequencing projects. Here, we explore a different, hypothesis-driven approach. We show that the fitness advantage associated with the transferred gene, typically invoked only in retrospect, can be used to design a functional screen capable of identifying postulated LGT cases. We hypothesized that β -glucuronidase (*gus*) genes may be prone to LGT from bacteria to fungi (thought to lack *gus*) because this would enable fungi to utilize glucuronides in vertebrate urine as a carbon source. Using an enrichment procedure based on a glucose-releasing glucuronide analog (cellobiouronic acid), we isolated two *gus*⁺ ascomycete fungi from soils (*Penicillium canescens* and *Scopulariopsis* sp.). A phylogenetic analysis suggested that their *gus* genes, as well as the *gus* genes identified in genomic sequences of the ascomycetes *Aspergillus nidulans* and *Gibberella zeae*, had been introgressed laterally from high-GC gram⁺ bacteria. Two such bacteria (*Arthrobacter* spp.), isolated together with the *gus*⁺ fungi, appeared to be the descendants of a bacterial donor organism from which *gus* had been transferred to fungi. This scenario was independently supported by similar substrate affinities of the encoded β -glucuronidases, the absence of introns from fungal *gus* genes, and the similarity between the signal peptide-encoding 5' extensions of some fungal *gus* genes and the *Arthrobacter* sequences upstream of *gus*. Differences in the sequences of the fungal 5' extensions suggested at least two separate introgression events after the divergence of the two main Euascomycete classes. We suggest that deposition of glucuronides on soils as a result of the colonization of land by vertebrates may have favored LGT of *gus* from bacteria to fungi in soils.

Introduction

The frequency of lateral gene transfer (LGT) from prokaryotes to eukaryotes is a matter of ongoing discussion. Vertebrates with sequestered germlines do not appear to acquire bacterial genes frequently (Salzberg et al. 2001; Stanhope et al. 2001), but microbial eukaryotes may introgress them more readily (Garcia-Vallvé, Romeu, and Palau 2000; Boucher and Doolittle 2000; Wu, Henze, and Müller 2001; Andersson et al. 2003). A fundamental limitation in detecting such LGT events is that they are typically invoked in retrospect to account for an inconsistency between the phylogeny of a gene in question and the taxonomic relationships among the host organisms. Alternative explanations are almost always possible. A persuasive argument for LGT, therefore, should not be exclusively based on sequence data, but incorporate other aspects of biology and ecology, ideally of both donor and recipient organism (Eisen 2000). Potential LGT cases, however, are usually identified by chance as a byproduct of genome-sequencing projects, which makes it difficult to link sequence data with other biological features (Braun et al. 2000).

Here, we explore a different, hypothesis-driven approach. We suggest that ecological and physiological considerations can be used to identify bacterial genes that may be prone to LGT because they could confer a fitness advantage to microbial eukaryotes that lack them. A functional screen based on this advantage should enable the

isolation of descendants of organisms that have acquired such bacterial genes through LGT.

We hypothesized that β -glucuronidase genes (*gus*; family 2 of glycosyl hydrolases) may be prone to LGT from bacteria to fungi. Vertebrates detoxify a range of xenobiotic and endogenous compounds through conjugation to D-glucuronic acid via a β -O-glycosidic linkage. Bacterial β -glucuronidases are scavenging enzymes that hydrolyze this linkage, thus, providing access to glucuronic acid as a carbon source (Tukey and Strassburg 2000). However, *gus* genes have not been found in fungi.

Fungi can degrade polysaccharides that contain glucuronic or galacturonic acid residues (Kitamoto et al. 2001; Kuroyama et al. 2001). The fungal enzymes that metabolize such polysaccharides, however, may not cleave glucuronides efficiently because of the distinct chemical nature of the nonsugar residues in glucuronides (Sutherland 1995; Kuroyama et al. 2001). Yet, fungi could probably utilize the glucuronic acid residue in glucuronides as a carbon source because they can metabolize galacturonic acid, an epimer of glucuronic acid (Nakagawa et al. 2000; Alves de Cunha et al. 2003). Laterally introgressed bacterial *gus* genes, therefore, could confer an adaptive advantage to fungi by providing access to glucuronic acid as a carbon source.

We used the ability to grow on a glucuronide as the sole carbon source as the selection principle to screen soil extracts for fungi with β -glucuronidase activity. β -Glucuronidases are quite insensitive to the chemical properties of the residue linked to glucuronic acid. We, therefore, selected cellobiouronic acid (Cba; 4-O- $[\beta$ -D-glucurononyl]-D-glucose), in which glucose replaces the nonsugar moiety, as a proxy for vertebrate glucuronides. Using Cba as a substrate avoided accumulation of potentially toxic compounds upon hydrolysis and increased the effectiveness of the screen by releasing two metabolizable sugar residues.

¹ Present address: Cell GeneSys, South San Francisco, California.

² Present address: Department of Biochemistry, Faculty of Biology, Kim Il Sung University, Pyongyang, DPR Korea.

Key words: Lateral gene transfer, *gus*, β -glucuronidase, bacteria, fungi.

E-mail peter@cambia.org.

Mol. Biol. Evol. 22(2):308–316. 2005

doi:10.1093/molbev/msi018

Advance Access publication October 13, 2004

The Cba screen identified several *gus*⁺ fungi. We selected two morphologically distinct isolates to investigate whether their *gus* genes, as well as two more *gus* genes discovered in genome sequences of other fungi, may have been introgressed from bacteria through LGT.

Materials and Methods

Isolation of *gus*⁺ Fungi and Bacteria from Soils

Aqueous extracts of 43 soil samples from the Canberra region (Australia) were plated on M9 minimal medium supplemented with Balch's vitamin and Ho-Le's microelement mixtures (Sambrook, Fritsch, and Maniatis 1989; Gerhardt et al. 1994). Cellobiouronic acid was prepared by acid hydrolysis from the exopolysaccharide of the bacterium *Sphingomonas paucimobilis* (Gellan Gum) and added as the sole carbon source (20 mM) (Jefferson 1998; Hashimoto et al. 1998). The medium also contained 50 mg/l of 5-bromo-4-chloro-3-indolyl- β -D-glucuronide (X-GlcA), which, upon hydrolysis by β -glucuronidases, produces a blue precipitate. The plates were incubated at 28°C.

Blue-staining filamentous fungi were purified from adhering bacteria by six consecutive subcultivations on YPD plates containing a mixture of three antibiotics (ampicillin, streptomycin, and nalidixic acid, all 50 mg/l) (Sambrook, Fritsch, and Maniatis 1989). The fungi were subsequently transferred back to the Cba medium to reconfirm β -glucuronidase activity. No bacteria were visible upon microscopic examination after growing the purified fungal isolates in liquid or on solid YPD medium without the antibiotics.

Blue-staining bacterial colonies growing on the Cba medium were purified by subcultivations until the colony morphology was homogenous.

Taxonomic Identification of *gus*⁺ Fungi and Bacteria Isolated from Soils

Genomic DNA of two selected fungal isolates (RPK and RP38.3) was extracted using the DNAzol kit (Invitrogen, Carlsbad, Calif.). DNA from selected bacteria was extracted according to Ausubel et al. (1988), after an initial 3-hour incubation with 2 mg/ml of lysozyme. Fungi were identified by sequencing part of their small subunit (SSU) ribosomal rRNA genes, amplified with primers N3 + N6 (GenBank accession numbers AY773329 and AY773330), and sequencing their intergenic transcribed spacers (ITS) 1 and 2, amplified with primers ITS1 + ITS4 (GenBank accession numbers AY773331 and AY773332) (White et al. 1990). Bacteria were identified by sequencing part of their SSU rRNA genes, amplified with primers P3 + 1100R (GenBank accession numbers AY773327 and AY773328) (Lane 1991; Muyzer, de Waal, and Uitterlinden 1993). The sequences obtained were used as queries for BlastN searches at GenBank (Altschul et al. 1990).

Sequencing of Fungal and Bacterial *gus* Genes

A pair of degenerate primers flanking a 1.2-kb fragment of *gus* were designed based on the sequences

of bacterial and mammalian *gus* genes: 5'-AAT TAA CCC TCA CTA AAG GGA YTT YTW YAA YTA YGC IGG-3' (*gus*-fwd; T3 handle underlined) and 5'-GTA ATA CGA CTC ACT ATA GGG RAA RTC IGC RAA RAA CCA-3' (*gus*-rev; T7 handle underlined). Genomic DNA (50 ng from fungal isolates; 5 ng from bacterial isolates) was added to 20 μ l of PCR mix containing 1 mM of each primer, 0.25 mM dNTPs, and 0.5 U REDTaq in Sigma PCR buffer (Sigma, St. Louis, Mo.). After an initial denaturation at 94°C (2 min), *gus* fragments were amplified in 35 cycles of 94°C (20 s), 48°C (40 s), and 72°C (2.5 min), followed by a final extension at 72°C for 5 min. PCR products were gel-purified (Ultrafree-DA; Millipore, Bedford, Md.) and sequenced with BigDye Terminator Mix (PerkinElmer ABI, Foster City, Calif.).

Fungal and bacterial sequences were extended in both directions by sequencing a series of partly overlapping fragments. These fragments were generated in several rounds of PCR amplifications using a primer annealing to end of the known sequence in combination with a random primer (GenBank accession numbers AY773333 to AY773337) (Chun et al. 1997). Genes upstream and downstream of *gus* genes were identified through BlastX searches at GenBank and annotated based on their similarity with known or predicted proteins of related species (Altschul et al. 1990).

Identification of Additional *gus* Genes in Genome Sequences of Fungi

Whole-genome-shotgun (wgs) sequences of fungi at GenBank were searched for *gus* genes using the BlastX program and the *Escherichia coli* enzyme as a query sequence (Altschul et al. 1990).

Construction of a Phylogram

A multiple alignment of β -glucuronidases, prepared with ClustalW and the BLOSSOM 26 protein weight matrix (Henikoff and Henikoff 1992; Thompson, Higgins, and Gibson 1994), was refined by hand and is available upon request. It comprised the fungal and *Arthrobacter* enzymes identified in this study (GenBank accession numbers AY773333 to AY773337), a *Salmonella* sequence taken from the patent literature (Jefferson and Mayer 1999), and the β -glucuronidase sequences deposited at GenBank under the following accession numbers: XP_320660 (*Anopheles gambiae*; annotation corrected by adding 59 amino acids at position 551 according to the EST AL93893), NP_493548 (*Caenorhabditis elegans*), O18835 (*Canis familiaris*), O77695 (*Cercopithecus aethiops*), NP_738989 (*Corynebacterium efficiens*; putative sequencing error causing a frame shift at amino acid position 456 corrected based on genomic sequence NC_004369), NP_561063 (*Clostridium perfringens*), NP_725851 (*Drosophila melanogaster*), NP_416134 (*Escherichia coli*), O97524 (*Felis catus*), ZP_00122729 (*Haemophilus somnus*), P08236 (*Homo sapiens*), AAK07836 (*Lactobacillus gasserii*), P12265 (*Mus musculus*), P06760 (*Rattus norvegicus*), NP_837302 (*Shigella flexneri*), AAK29422 (*Staphylococcus* sp.), NP_687713 (*Streptococcus agalactiae*),

NP_344348 (*Sulfolobus solfataricus*), NP_228868 (*Thermotoga maritima*), and NC_002955 (*Streptococcus equi*; contig equi130g4.q1t, positions 60234 to 62021, translated in silico in frame -3).

Approximate phylogenetic relationships, based on the Jones-Taylor-Thornton substitution model and a constant amino acid substitution rate, were inferred using the maximum-likelihood approach implemented in the PROML program of the PHYLIP version 3.6 software package (Felsenstein 1989; Jones, Taylor, and Thornton 1992). The best tree obtained from 10 randomized sequence input orders with global rearrangements was used to estimate the shape parameter α of a Γ distribution modeling the variability of the substitution rates of nonconserved amino acids as follows. Invariable amino acid residues (6.8%) were removed from the alignment and α was computed with the CODEML program of the PAML version 3.12 software package, using an eight-categories approximation of the Γ distribution (Yang 1997). The phylogenetic analysis was subsequently refined by rerunning the PROML program as specified above, but, this time, based on a mixed model comprising eight discrete Γ categories ($\alpha = 1.35$) plus one invariable category (the tree topology was unaffected, but the branch lengths changed). Clade strengths were evaluated by analyzing 250 bootstrap replicates with the PROML program based on a model comprising one invariable plus four Γ categories ($\alpha = 1.35$).

The tree topology was compared with an alternative (hypothetical) topology representing vertical transmission of β -glucuronidases from an ancestral eukaryote to both fungi and Metazoa, using the likelihood ratio test described by Kishino and Hasegawa (1989).

Comparison of Substrate Affinities

Fungal isolates RPK and RP38.3 and bacterial isolate RP10, were grown in liquid Cba medium at 29°C/200 rpm. An *Escherichia coli* strain (CE1) was grown in LB medium containing 1 mM isopropyl-thio- β -D-glucuronide (IPTG) at 37°C/200 rpm. A *Staphylococcus* β -glucuronidase gene, previously cloned into an *E. coli* expression vector (amp^r) under the control of the lac promoter (pTANE95mAI), was introduced into an *E. coli* strain from which the *gus* operon had been deleted (JEMA99.9 [T. Nguyen and J. Mayer, personal communication]). Transformants were grown in LB medium supplemented with 100 mg/l ampicillin and 40 mg/L IPTG at 37°C/200 rpm.

After 1 to 4 days of growth, proteins were extracted from hyphae pulverized with liquid nitrogen (fungi) or from bacteria disintegrated by beating with glass beads. The extraction buffer contained 40 mM piperazine-N,N'-bis(ethanesulfonic acid) (pH 7.0), 2 mM dithiothreitol, 1 mM ethylenediaminetetraacetic acid, 1 mM phenylmethylsulfonyl-fluoride, and 0.1% (v/v) Triton X-100. The supernatants obtained by centrifugation at 17,000 $\times g$ (4°C, 10 min) were passed through Sephadex G25 spin columns equilibrated with extraction buffer. A β -glucuronidase preparation from *Bos taurus* (Sigma, St. Louis, Mo.) was used as an example for mammalian β -glucuronidases.

The β -glucuronidase activity in the enzyme preparations was assayed in quadruplicate in 160 μ l of extraction buffer containing 0.1 mg/ml bovine serum albumin and various concentrations of either 4-methylumbelliferyl- β -D-glucuronide (MUG), para-nitrophenyl- β -D-glucuronide (pNPG), or phenolphthalein- β -D-glucuronide (PHTG) as a substrate. The reactions were incubated at 30°C and stopped by addition of 40 μ l of 2 M Na₂CO₃. Hydrolysis of MUG was measured fluorimetrically (excitation: 360 nm; emission: 465 nm). Hydrolysis of the other substrates was measured colorimetrically (pNPG at 405 nm; PHTG at 560 nm) (Tecan SpectraFluorPlus microplate reader, Grödig, Austria). A Michaelis/Menten-type equation that contained a term describing substrate inhibition (K_i) was fitted to the experimental data using the Marquardt-Levenberg algorithm (Marquardt 1963).

Sequence Analysis of 5'-Regions of Fungal *gus* Genes

N-termini of fungal β -glucuronidases were screened for signal peptides using the neural network program SIGNALP version 1.1 (Nielsen, Brunak, and von Heijne 1999).

The DNA sequence upstream of the initiation codon of *gus* of *Arthrobacter* isolate RP40.7 was manually aligned with the 5' extensions of the *Aspergillus* and *Penicillium gus* genes. A partly degenerate consensus sequence was created from the *Aspergillus* and *Penicillium* sequences to evaluate the statistical significance of this alignment. The consensus sequence plus 300 randomized sequences of identical nucleotide composition and degeneracy were aligned individually against the *Arthrobacter* sequence using the GAP program of the GCG software package (Accelrys, Burlington, Mass.). The scores for matches and mismatches were adjusted to 1 and 0, respectively, the gap weight was varied between 3 and 30, the gap-length weight was varied between 0.3 and 3, and end gaps were penalized. The statistical significance of the fungi versus *Arthrobacter* alignment was estimated based on the mean and standard deviation of the alignment quality scores obtained with the randomized sequences. Similar alignments were performed to compare the *Scopulariopsis* with the *Arthrobacter*, the *Corynebacterium*, and the *Aspergillus/Penicillium* sequences.

Results

Simultaneous Isolation of *gus*⁺ Fungi and Bacteria

We isolated seven *gus*⁺ fungi from a total of 43 soil extracts. Two morphologically distinct isolates that produced diffuse blue halos in the agar medium surrounding their hyphae were selected for further analysis. Both were members of the Euascomycetes (Pezizomycotina) subphylum of ascomycete fungi. One isolate (RPK) was *Penicillium canescens* (Eurotiomycetes class). The other (RP38.3) was tentatively identified as *Scopulariopsis* sp. because it was closely related to *Scopulariopsis brevicaulis* from the Sordariomycetes class (table 1).

We also isolated 42 blue-staining bacteria from 30 soil samples. The vast majority of these isolates (88%)

Table 1
Taxonomic Identification of Fungal and Bacterial Isolates with β -Glucuronidase Activity

Organism	Isolate	ITS1 – 5.8S rRNA Gene – ITS2 – 28S rRNA Gene (Partial)		SSU rRNA Gene (Partial)	
		Closest Match	Identity	Closest Match	Identity
Fungus	RP38.3	— ^a	—	<i>Scopulariopsis brevicaulis</i> (AY083220)	99.8 % (826/828)
Fungus	RPK	<i>Penicillium canescens</i> (AF033493)	100 % (528/528)	<i>Penicillium sacculum</i> ^b (AB027410)	99.9 % (832/833)
				<i>Penicillium herquei</i> (AB086834)	99.9 % (832/833)
				<i>Eladia saccula</i> (AB031391)	99.9 % (832/833)
				<i>Eupenicillium</i> sp. (AY297772)	99.9 % (832/833)
Bacterium	RP10	n.a. ^c	n.a.	<i>Arthrobacter</i> sp. strain 19B (AB017649)	100 % (664/664)
Bacterium	RP40.7	n.a.	n.a.	<i>Arthrobacter aurescens</i> (AF467106, AF388032)	100 % (664/664)
				<i>Arthrobacter nitroguaiacolicus</i> (AJ512504)	100 % (664/664)

^a No continuous match spanning both ITS and the 5.8S rRNA gene.

^b The database did not contain the *Penicillium canescens* sequence.

^c Not applicable.

were *Arthrobacter* species. We selected two isolates from different soil samples (RP10 and RP40.7) to characterize their *gus* genes (table 1).

Fungal *gus* Genes

We amplified and sequenced the β -glucuronidase (*gus*) loci of the fungal isolates RP38.3 (*Scopulariopsis* sp., Sordariomycetes) and RPK (*Penicillium canescens*, Eurotiomycetes). For confirmation, we also amplified and sequenced *gus* from the *P. canescens* strain DSM1215 obtained from the German Collection of Microorganisms and Cell Cultures. In addition, we identified two *gus* genes in genomic shotgun sequences of *Aspergillus nidulans* (Eurotiomycetes; GenBank accession number AACD01000093, positions 285949 to 287784, frame +1) and *Gibberella zeae* (Sordariomycetes; GenBank accession number AACM01000315, positions 77805 to 76006, frame –3).

All fungal sequences contained a continuous open reading frame encoding a β -glucuronidase belonging to family-2 of glycosyl hydrolases (*gus*). They exhibited several features characteristic of eukaryotic genes; for example, the nucleotides at the predicted transcriptional start (adenine surrounded by one pyrimidine nucleotide upstream and four to five downstream), TATA box motifs at –32 bp/–19 bp (RPK) and –40 bp (RP38.3), the latter surrounded by guanidine nucleotides at characteristic positions, a Kozak motif (CCACC) immediately upstream of the initiation codon (RP38.3), and polyA signals at positions 406 bp (RPK; TATATA) and 449 bp (RP38.3; AATAAA) downstream of the termination codons (Kozak 1986; Watson et al. 1987; Knippers et al. 1990). The *gus* genes were surrounded by different genes in the four fungi; some of these genes, such as the putative allantoin permease gene of *Penicillium canescens* and the sugar permease gene of *Gibberella zeae*, contained introns (fig. 1). Sequence similarity among the four fungi in this region was restricted to *gus* (data not shown).

The predicted amino acid sequences of the four fungal β -glucuronidases were most similar to those of the β -glucuronidases from the two *Arthrobacter* species

isolated together with the fungi in this study; the identity values ranged from 52.7% to 56.6% (based on pairwise alignments generated with ClustalW and the BLOSSOM 26 protein weight matrix) (Henikoff and Henikoff 1992; Thompson, Higgins, and Gibson 1994).

Arthrobacter gus Genes

We also amplified and sequenced the β -glucuronidase (*gus*) loci of the *Arthrobacter* isolates RP10 and RP40.7. The two sequences were 77% identical in the region they overlapped. The *gus* genes of the two isolates, which encoded proteins that were 87% identical to each other, were flanked by genes encoding oxidoreductases, presumably members of the same operon as *gus*.

Phylogenetic Relationships Among β -Glucuronidases

A maximum-likelihood phylogenetic analysis of prokaryotic and eukaryotic β -glucuronidases yielded three groups with bootstrap support values between 78% and 79% (see asterisks in figure 1). Group I contained enzymes from low-GC gram⁺ bacteria plus a nested sequence of a gram[–] γ -proteobacterium (*Haemophilus somnus*), which formed a robust subgroup with two *Streptococcus* enzymes (100% bootstrap support). Group II contained metazoan β -glucuronidases from *Caenorhabditis*, insects, and mammals (fig. 1).

Group III was an assortment of both fungal and bacterial β -glucuronidases. It consisted of a basally branching clade of γ -proteobacterial enzymes (100% bootstrap support) and a nested cluster of enzymes from fungi and high-GC gram⁺ bacteria (63% bootstrap support). This mixed cluster comprised a basally branching enzyme from *Corynebacterium efficiens* and a nested group containing the enzymes of the two *Arthrobacter* species isolated in this study, plus all fungal enzymes (81% bootstrap support). The *Arthrobacter* enzymes appeared to be more closely related to the β -glucuronidases of the Eurotiomycetes (*Penicillium canescens* and *Aspergillus nidulans*) than to those of the Sordariomycetes (*Scopulariopsis* sp. and *Gibberella zeae*) (fig. 1).

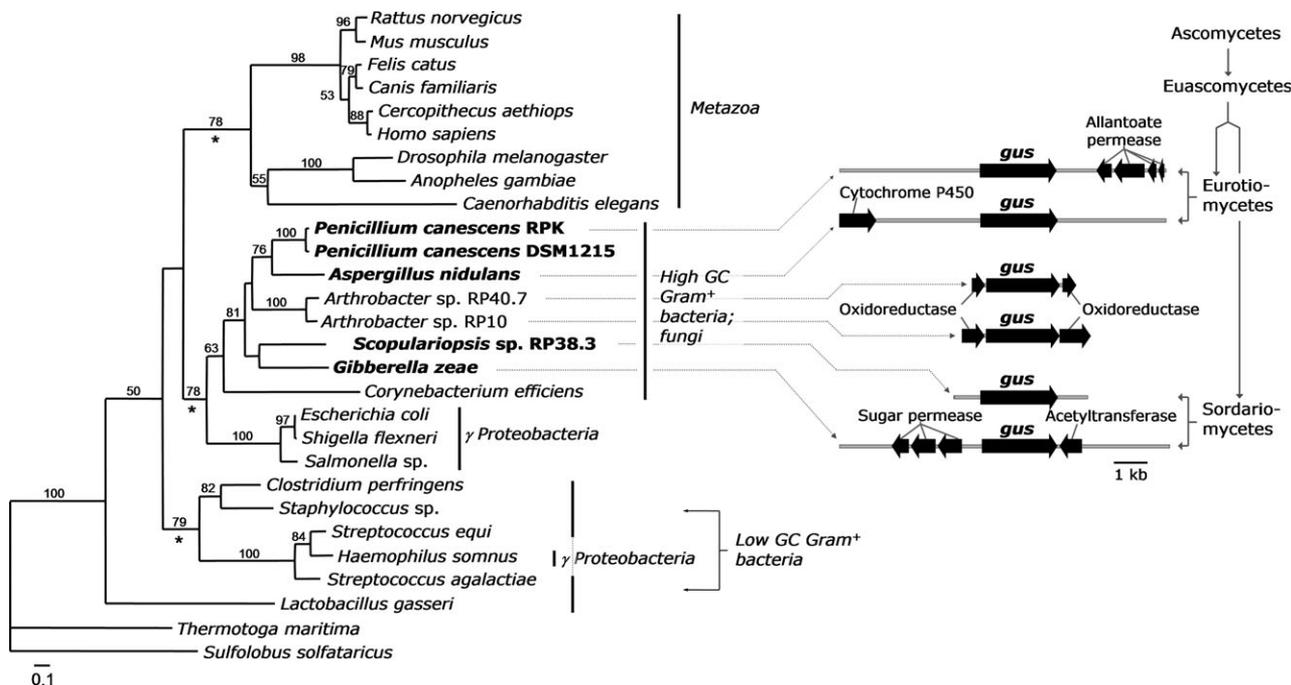


FIG. 1.—Phylogenetic relationships between fungal and other β -glucuronidases based on the protein maximum-likelihood approach described in *Materials and Methods*. Bootstrap support values greater than 50% are indicated at internal branches of the phylogram. Three main clusters with bootstrap support values between 78% and 79% are labeled with asterisks. The β -glucuronidase (*gus*) loci of fungi and *Arthrobacter* species on the right side of the figure are aligned with each other according to the sequences of their *gus* genes.

The topology of the phylogram in figure 1 was significantly superior to that of an alternative (hypothetical) phylogram representing vertical transmission of β -glucuronidases from an ancestral eukaryote to both fungi and Metazoa ($P = 0.0004$). This hypothetical phylogram had the same topology as that in figure 1, with the exception of the four fungal enzymes. The latter had been removed from their cluster with high-GC gram⁺ bacteria and joined (without altering their topology relative to each other) with the metazoan cluster to create a eukaryotic clade.

Biochemical Properties of the Encoded β -Glucuronidases

A comparison of the affinities of β -glucuronidases towards three different glucuronide substrates suggested that the enzymatic properties of the fungal enzymes were most similar to those of the *Arthrobacter* enzyme (fig. 2).

5'-Extensions of Fungal *gus* Genes Encoding N-Terminal Signal Peptides

All fungal β -glucuronidases, with exception of the *Gibberella* enzyme, had N-terminal signal peptides (fig. 3a). The extensions of the two Eurotiomycetes enzymes (*Aspergillus nidulans* and *Penicillium canescens*) were similar in length and were 51% identical at that DNA level (fig. 3a, b). They were quite different, both in length and sequence, to that of the *Scopulariopsis* enzyme (Sordariomycetes). None of the fungal N-terminal extensions resembled that of vertebrates (fig. 3a). The *gus* genes of the two *Arthrobacter* species that had been isolated

together with *gus*⁺ fungi lacked N-terminal signal peptides (fig. 3a).

When the 5' extensions of the two Eurotiomycetes genes were aligned with the region upstream of the *gus* initiation codon of *Arthrobacter* isolate RP40.7, 41% of the *Aspergillus* sequence and 36% of the *Penicillium* sequence were identical to the *Arthrobacter* sequence (fig. 3b). Alignments of randomized versions of an *Aspergillus*/*Penicillium* consensus sequence with the *Arthrobacter* sequence suggested that these identity levels were statistically significant ($P < 0.002$). If the *Arthrobacter* sequence immediately upstream of that shown in figure 3b was used for the alignment, the statistical significance was abolished ($P > 0.6$).

Similar alignments of the corresponding *Scopulariopsis* sequence with either the *Arthrobacter* or the *Corynebacterium* sequence were not statistically significant ($P > 0.7$). The same was true for an alignment of the *Scopulariopsis* with the *Aspergillus*/*Penicillium* consensus sequence ($P > 0.7$).

Discussion

The goal of this study was to investigate whether β -glucuronidase (*gus*) genes had been laterally transferred from bacteria to fungi. A screen based on the presumed ability of *gus*⁺ fungi to grow on a medium containing the glucuronide analog cellobiouronic acid as the sole carbon source identified a group of soil fungi with β -glucuronidase activity. We amplified the *gus* loci of two fungal isolates (RPK [*Penicillium canescens*] and RP38.3 [*Scopulariopsis* sp.], both ascomycetes). The presence of

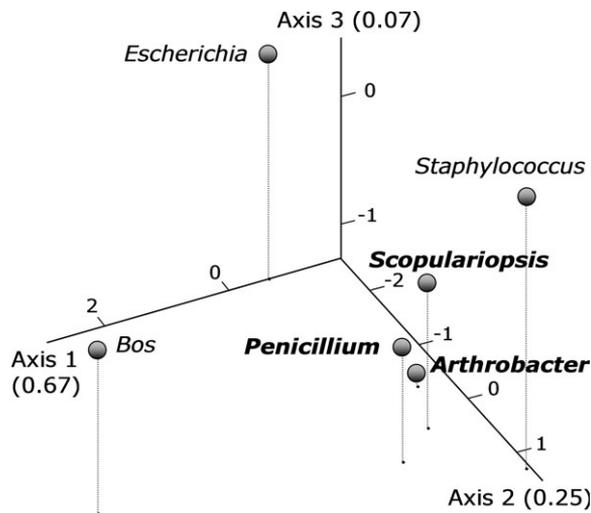


FIG. 2.—Biochemical similarities among selected β -glucuronidases. The similarities in the affinities of β -glucuronidases towards three glucuronide substrates were visualized by extracting the principal components of a matrix containing log-transformed values of K_m and K_i for MUG, pNPG, and PHTG (if substrate inhibition was undetectable, K_i was set to five times the maximum concentration assayed). Individual K_m and K_i values were (in μM): 51 ± 9 and $4,100 \pm 1,300$ (MUG), 210 ± 20 and $7,300 \pm 1,800$ (PHG), and 110 ± 20 and $> 10,000$ (pNPG) for *P. canescens*; 26 ± 2 and $4,200 \pm 500$ (MUG), 190 ± 20 and $3,300 \pm 500$ (PHG), and 51 ± 6 and $5,500 \pm 1,000$ (pNPG) for *Scopulariopsis* sp.; 13 ± 1 and $3,100 \pm 400$ (MUG), 42 ± 6 and $3,700 \pm 700$ (PHG), and 110 ± 30 and $> 10,000$ (pNPG) for *Arthro bacter* sp.; 110 ± 10 and $> 10,000$ (MUG), 35 ± 10 and 180 ± 50 (PHG), and 120 ± 10 and $> 10,000$ (pNPG) for *Escherichia coli*; 56 ± 9 and 660 ± 60 (MUG), 240 ± 20 and $5,800 \pm 1,300$ (PHG), and 45 ± 13 and $1,200 \pm 300$ (pNPG) for *Staphylococcus* sp.; and $1,600 \pm 300$ and $> 100,000$ (MUG), $1,100 \pm 200$ and $6,800 \pm 1,600$ (PHG), and $12,000 \pm 3,000$ and $> 100,000$ (pNPG) for *Bos taurus*.

regulatory sequence motifs required for eukaryotic transcription and translation and the presence of introns in the gene downstream of *gus* in *Penicillium canescens* confirmed the eukaryotic nature of the isolated *gus* genes (fig. 1, and see *Results*). We also identified two more *gus* genes in genomic shotgun sequences of other ascomycetes (*Aspergillus nidulans* and *Gibberella zeae*).

Growth experiments confirmed that the two isolated fungi could grow on glucuronic acid as the sole carbon source (data not shown), consistent with the idea that lateral acquisition of bacterial *gus* genes would have enabled them to utilize the glucuronic acid moiety of vertebrate glucuronides as a carbon source. We, therefore, tested the LGT hypothesis by using various lines of evidence.

Ascomycete Fungi Acquired *gus* from High-GC Gram⁺ Bacteria Through Lateral Gene Transfer

Considering that the last common ancestor of bacteria and fungi lived approximately 2.7 to 3.0 billion years ago, more than 50% amino acid identity between the β -glucuronidases from fungi and bacteria (*Arthro bacter* spp.) immediately suggested interkingdom LGT (Nei, Xu, and Glazko 2001; Hedges et al. 2001).

An analysis of the phylogenetic relationships among β -glucuronidases indeed produced a mixed cluster of enzymes from fungi and *Arthro bacter* species (high-GC

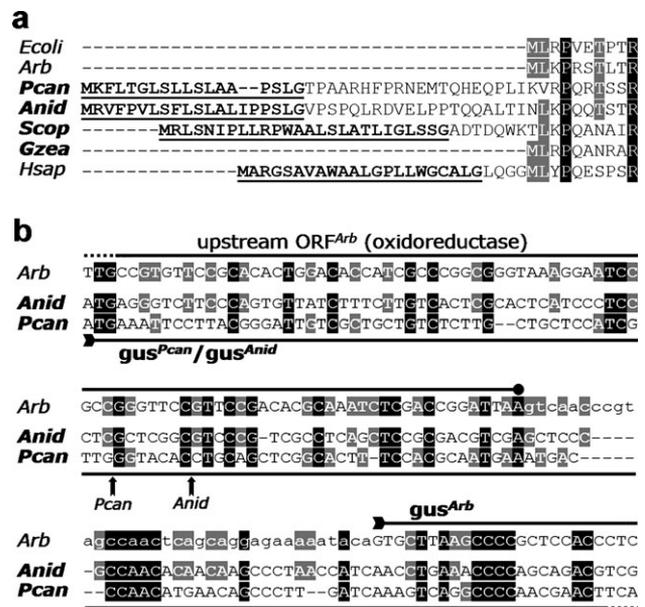


FIG. 3.—N-terminal sequences of β -glucuronidase enzymes and *gus* genes. (a) Multiple alignment of the N-termini of selected β -glucuronidases. Black, and gray shades designate identical, and conserved amino acids, respectively. N-terminal signal peptides, predicted with the SIGNALP version 1.1 program, are underlined (Nielsen, Brunak, and von Heijne 1999). (b) Comparison between the signal peptide-encoding 5' extensions of the *gus* genes from Eurotiomycetes (*Aspergillus nidulans* and *Penicillium canescens* isolate RPK) and the sequence upstream of the *gus* initiation codon of *Arthro bacter* sp. isolate RP40.7 (the sequence of the other *Arthro bacter* isolate was similar but had a small deletion in this region). Nucleotides identical to the *Arthro bacter* sequence are shaded in black (identical in both fungi) or gray (identical in one of the two fungi). Open reading frames are designated by uppercase letters in the DNA sequences, and the solid lines above (*Arthro bacter*) or below (Eurotiomycetes) indicate the alignment. The predicted signal peptide cleavage sites in the Eurotiomycetes sequences are indicated by small arrows below the alignment. Species names are abbreviated as follows: Ecoli, *Escherichia coli*; Arb, *Arthro bacter* sp.; Pcan, *Penicillium canescens*; Anid, *Aspergillus nidulans*; Scop, *Scopulariopsis* sp.; Gzea, *Gibberella zeae*, and Hsap, *Homo sapiens*.

gram⁺ bacteria). This cluster was robustly nested within a lineage leading to *Corynebacterium efficiens*, another high-GC gram⁺ bacterium. The closest relatives of this mixed cluster were β -glucuronidases of gram⁻ γ -proteobacteria. As a group, the enzymes from fungi, high-GC gram⁺, and γ -proteobacteria were well separated from their metazoan counterparts. The last formed a monophyletic cluster consistent with vertical transmission of *gus* genes during metazoan evolution (fig. 1).

This topology is most parsimoniously explained by lateral transfer from high-GC gram⁺ bacteria to fungi. A phylogenetic topology consistent with a primitive eukaryotic ancestor of both fungal and metazoan β -glucuronidases explained protein sequence similarities significantly less well than the phylogenetic tree shown in figure 1 (see *Results*). The absence of introns from all four identified fungal *gus* genes and the similar biochemical properties of the β -glucuronidases from fungi and *Arthro bacter* are also consistent with this LGT scenario (figs. 1 and 2).

The sequences of the signal peptide-encoding 5' extensions of the *gus* genes from the two Eurotiomycetes

(*Aspergillus nidulans* and *Penicillium canescens*) provided an independent piece of evidence that added further weight to the postulated LGT event. These 5' extensions appear to have evolved from the *Arthrobacter* sequence immediately upstream of the *gus* initiation codon, which presumably was transferred to fungi together with *gus*. This *Arthrobacter* sequence comprised 30 codons of an upstream ORF encoding an oxidoreductase plus 36 bp of intergenic sequence (fig. 3b). After lateral transfer to fungi, translation of the introgressed *Arthrobacter gus* gene may have been forced to initiate at an in-frame ATG codon either upstream or downstream of the bacterial initiation codon, because none of the isolated *Arthrobacter gus* genes initiates with ATG (see figure 3b for the RP40.7 isolate). The alignment of the 5' extensions of the two Eurotiomycetes *gus* genes with the *Arthrobacter* sequence suggested that the upstream ORF of *Arthrobacter* may indeed have had a codon similar to ATG at the position that corresponds to the fungal initiation codons (fig. 3b). After lateral introgression into fungi, the peptide encoded by the 5'-part of the upstream sequence from *Arthrobacter* appears to have evolved into a functional signal peptide.

The two *Arthrobacter* species whose β -glucuronidases formed a mixed cluster with fungal enzymes were isolated from the same set of soil samples as the *gus*⁺ fungi. Both donor and recipient organisms of the LGT event(s), therefore, may have lived in the same habitat. *Arthrobacter* species may have been the "logical" donors of bacterial *gus* genes for lateral transfer to fungi because they were the dominant group of cultivatable *gus*⁺ bacteria in our screen (see *Results*), and they typically constitute a numerically important fraction of the bacterial flora of soils (Jones and Keddie 1999).

Unless lateral acquisition of *gus* by fungi took place at an early stage of fungal evolution, the distribution of *gus* among fungi should be restricted to one (or a few) lineage(s) comprising the descendants of the recipient organism(s). Full or partial genome sequences of 16 species from other ascomycete subphyla including the Hemiascomycetes (seven *Saccharomyces* species, two *Candida* species, two *Kluyveromyces* species, *Debaryomyces hansenii*, *Eremothecium gossypii*, *Naumovia castellii*, and *Yarrowia lipolytica*) and the Archiascomyces (*Schizosaccharomyces pombe*) have not revealed any *gus* genes to this date. The same applies to other fungal phyla such as the Basidiomycota (three *Cryptococcus* species, *Coprinopsis cinerea*, *Phanerochaete chrysosporium*, and *Ustilago maydis*) and the Microsporidia (*Encephalitozoon cuniculi*). Within the Euscomycetes, *Neurospora crassa* and *Magnaporthe grisea* (Sordariomycetes) and three *Aspergillus* species as well as *Coccidioides posadasii* (Eurotiomycetes) do not appear to contain *gus*.

Although limited, this data set suggests that *gus* may be restricted to a few members of two Euscomycetes classes (Eurotiomycetes and Sordariomycetes). Such a pattern would be more consistent with a single LGT event in an immediate progenitor of the Euscomycetes, or two separate LGT events in the two Euscomycetes classes, than with vertical transmission from a primitive fungal ancestor followed by widespread differential gene loss. However, *gus* may also have been lost in some descendants of the

Aspergillus/Penicillium progenitor that introgressed it from bacteria, because sequencing efforts have so far failed to identify *gus* in three out of four *Aspergillus* species.

Bacterial *gus* Genes Appear to Have Been Introgressed at Least Twice

To address the question of whether *gus* was introgressed into Euscomycetes more than once, we compared the signal peptide–encoding 5' extensions of the *gus* genes from Eurotiomycetes (*Aspergillus nidulans* and *Penicillium canescens*) and Sordariomycetes (*Scopulariopsis* sp.; *Gibberella zeae* had no 5' extension). We hypothesized that these regions could be different for separate introgression events involving different donor organisms because we established that in the Eurotiomycetes, it had been derived from the genomic region of the bacterial donor that was immediately upstream of *gus* (fig. 3b).

The 5' extension of *gus* from *Scopulariopsis* sp. did not show any sequence similarity to the 5' extensions of the Eurotiomycetes species or the upstream region of *Arthrobacter* species (see *Results*). This finding points to a different bacterial donor organism or to a different mechanism through which the signal peptide–encoding *Scopulariopsis* extension was created. Separate LGT events in the two Euscomycetes classes, therefore, are a more plausible scenario than a single event involving an ancestor of the two classes. It may even be possible that the Sordariomycetes acquired *gus* more than once. This possibility exists because the gene of *Gibberella zeae* has no 5' extension and does not appear to be as closely related to the other Sordariomycetes gene (*Scopulariopsis* sp.) as are the two Eurotiomycetes *gus* genes to each other.

Lateral Introgression of *gus* by Fungi May Have Been a Consequence of the Colonization of Land by Vertebrates

An ancestral Eurotiomycetes species presumably introgressed a bacterial *gus* gene between 670 and 60 MYA; that is, after the divergence of the Eurotiomycetes and Sordariomycetes (670 to 500 MYA) but before divergence of *Penicillium* and *Aspergillus* (60 MYA) (Buades and Moya 1996; Taylor, Hass, and Kerp 1999; Heckman et al. 2001). Ascomycete fungi colonized land together with plants between 700 and 460 MYA (Gray 1985; Heckman et al. 2001; Berbee 2001). At that time, β -glucuronides consisting of nonsugar residues conjugated to glucuronic acid were probably not present in significant amounts in terrestrial ecosystems. Vertebrates, apparently the only animals that secrete glucuronides as detoxification products, colonized land only from approximately 360 MYA onwards (Kumar and Hedges 1998; Tukey and Strassburg 2000). Deposition of vertebrate urine on soils may have created a fitness advantage for soil fungi with laterally acquired bacterial *gus* genes because *gus* would have enabled them to utilize the glucuronic acid moiety of urine glucuronides as a carbon source.

This scenario raises the question of why high-GC gram⁺ soil bacteria (the bacterial *gus* donors) had *gus* genes if there were no appropriate substrates in soils before the arrival of terrestrial vertebrates. Contrary to established phylogenetic relationships among bacterial phyla, the *gus*

genes from high-GC gram⁺ bacteria formed a well-supported cluster with a basally branching group consisting of three Enterobacteriaceae (gram⁻ γ -proteobacteria) (fig. 1 [Daubin, Gouy, and Perrière 2001; Brown et al. 2001]). These species are gastrointestinal parasites or pathogens of vertebrates, and their β -glucuronidases are thought to release glucuronic acid from glucuronides secreted in the bile (Wilson, Hughes, and Jefferson 1992). We speculate that their *gus* genes, presumably introduced into soils through vertebrate feces, could initially have been introgressed into high-GC gram⁺ soil bacteria for the same reason as they were subsequently transferred from high-GC gram⁺ bacteria to fungi. LGT of *gus* between bacteria clearly has occurred during evolution, as evident from the lateral introgression of *gus* of a low-GC gram⁺ bacterium into gram⁻ *Haemophilus somnus*, also detected in this study (fig. 1).

Conclusions

Taken together, our data outline a compelling case of lateral gene transfer from high-GC gram⁺ soil bacteria to ascomycete fungi. Other cases of LGT from high-GC gram⁺ bacteria to fungi include isopenicillin-N-synthetase and chymotrypsin genes (Buades and Moya 1996; Screen and Leger 2000; Rosewich and Kistler 2000). The fact that we identified the β -glucuronidase case as a result of a targeted screen raises the question of whether such interkingdom transfers are more common in soil microbial communities than previously thought, particularly because bacterial conjugation mechanisms are capable of transferring genes to fungi (Heinemann and Sprague 1989). We suggest that more such LGT cases might be identified using comparable functional screens based on the hypothesized fitness advantages of laterally transferred genes.

Supplementary Material

Sequences Submitted to GenBank

- AY773327: *Arthrobacter* sp. isolate RP10 16S ribosomal RNA gene, partial sequence
 AY773328: *Arthrobacter* sp. isolate RP40.7 16S ribosomal RNA gene, partial sequence
 AY773329: *Penicillium canescens* isolate RPK 18S ribosomal RNA gene, partial sequence
 AY773330: *Scopulariopsis* sp. isolate RP38.3 18S ribosomal RNA gene, partial sequence
 AY773331: *Penicillium canescens* isolate RPK internal transcribed spacer 1, 5.8S ribosomal RNA gene, internal transcribed spacer 2; and 28S ribosomal RNA gene (partial)
 AY773332: *Scopulariopsis* sp. isolate RP38.3 internal transcribed spacer 1, 5.8S ribosomal RNA gene, internal transcribed spacer 2; and 28S ribosomal RNA gene (partial)
 AY773333: *Penicillium canescens* isolate RPK β -glucuronidase (*gus*) gene locus
 AY773334: *Penicillium canescens* strain DSM1215 β -glucuronidase (*gus*) gene locus
 AY773335: *Scopulariopsis* sp. isolate RP38.3 β -glucuronidase (*gus*) gene locus

AY773336: *Arthrobacter* sp. isolate RP10 β -glucuronidase (*gus*) gene locus

AY773337: *Arthrobacter* sp. isolate RP40.7 β -glucuronidase (*gus*) gene locus

Acknowledgments

We thank Rohan Puttner for his assistance during the isolation of fungi and bacteria from soils, Subbu Pucha and Jorge Mayer for the taxonomic identification of most soil bacteria, Paul Keese for help with designing degenerate primers, Eric Huttner, Heidi Mitchell, and Brian Weir for their critical comments on the manuscript, and two anonymous reviewers for their valuable comments and suggestions.

Literature Cited

- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**:403–410.
- Alves da Cunha, M. A., A. M. Barbosa, E. C. Giese, and R. F. Dekker. 2003. The effect of carbohydrate carbon sources on the production of constitutive and inducible laccases by *Botryosphaeria* sp. *J. Basic Microbiol.* **43**:385–392.
- Andersson, J. O., Å. M. Sjögren, L. A. M. Davis, T. M. Embley, and A. J. Roger. 2003. Phylogenetic analyses of diplomonad genes reveal frequent lateral gene transfers affecting eukaryotes. *Curr. Biol.* **13**:94–104.
- Ausubel, F. M., R. Rent, R. E. Kingston, D. D. Moore, J. G. Seidman, and K. Struhl. 1988. *Current protocols in molecular biology*. Wiley, New York.
- Berbee, M. L. 2001. The phylogeny of plant and animal pathogens in the Ascomycota. *Physiol. Mol. Plant Pathol.* **59**:165–187.
- Boucher, Y., and W. F. Doolittle. 2000. The role of lateral gene transfer in the evolution of isoprenoid biosynthesis pathways. *Mol. Microbiol.* **37**:703–716.
- Braun, E. L., A. L. Halpern, M. A. Nelson, and D. O. Natvig. 2000. Large-scale comparison of fungal sequence information: mechanisms of innovation in *Neurospora crassa* and gene loss in *Saccharomyces cerevisiae*. *Genome Res.* **10**:416–430.
- Brown, J. R., C. J. Douady, M. J. Italia, W. E. Marshall, and M. J. Stanhope. 2001. Universal trees based on large combined protein sequence data sets. *Nat. Genet.* **28**:281–285.
- Buades, C., and A. Moya. 1996. Phylogenetic analysis of the isopenicillin-N-synthetase horizontal gene transfer. *J. Mol. Evol.* **42**:537–542.
- Chun, K. T., H. J. Edenberg, M. R. Kelley, and M. G. Goebel. 1997. Rapid amplification of uncharacterized transposon-tagged DNA sequences from genomic DNA. *Yeast* **13**:233–240.
- Daubin, V., M. Gouy, and G. Perrière. 2001. Bacterial molecular phylogeny using supertree approach. *Genome Informat.* **12**:155–164.
- Eisen, J. A. 2000. Horizontal gene transfer among microbial genomes: new insights from complete genome analysis. *Curr. Opin. Genet. Develop.* **10**:606–611.
- Felsenstein, J. 1989. PHYLIP: phylogeny inference package. Version 3.2. *Cladistics* **5**:164–166.
- García-Vallvé, S., A. Romeu, and J. Palau. 2000. Horizontal gene transfer of glycosyl hydrolases of the rumen fungi. *Mol. Biol. Evol.* **17**:352–361.

- Gerhardt, P., R. G. E. Murray, W. A. Wood, and N. R. Krieg. 1994. Methods for general and molecular bacteriology. ASM, Washington, DC.
- Gray, J. 1985. The microfossil record of early land plants: advances in understanding of early terrestrialization 1970–1984. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **309**:167–195.
- Hashimoto, W., N. Sato, S. Kimura, and K. Murata. 1998. Polysaccharide lyase: molecular cloning of gellan lyase gene and formation of the lyase from a huge precursor protein in *Bacillus* sp. GL1. *Arch. Biochem. Biophys.* **354**:31–39.
- Heckman, D. S., D. M. Geiser, B. R. Eidell, R. L. Stauffer, N. L. Kardos, and S. B. Hedges. 2001. Molecular evidence for the early colonization of land by fungi and plants. *Science* **293**:1129–1133.
- Hedges, S. B., H. Chen, S. Kumar, D. Y. C. Wang, A. S. Thompson, and H. Watanabe. 2001. A genomic timescale for the origin of eukaryotes. *BMC Evol. Biol.* **1**:4.
- Heinemann, J. A., and G. F. Sprague Jr. 1989. Bacterial conjugative plasmids mobilize DNA transfer between bacteria and yeast. *Nature* **340**:205–209.
- Henikoff, S., and J. G. Henikoff. 1992. Amino acid substitution matrices for protein blocks. *Proc. Natl. Acad. Sci. USA* **89**:10915–10919.
- Jefferson, R. A. 1998. Preparation of cellobiuronic acid from polysaccharide. US patent 6268493.
- Jefferson, R. A., and J. E. Mayer. 1999. Microbial β -glucuronidase genes, gene products and uses thereof. PCT patent application WO 00/55333.
- Jones, D., and R. M. Keddie. 1999. The genus *Arthrobacter*. in M. Dworkin et al., eds. The prokaryotes: an evolving electronic resource for the microbiological community. 2nd edition, release 3.0, May 5 1999. <http://141.150.157.117:8080/prokPUB/index.htm>.
- Jones, D. T., W. R. Taylor, and J. M. Thornton. 1992. The rapid generation of mutation data matrices from protein sequences. *CABIOS* **8**:275–282.
- Kishino, H., and M. Hasegawa. 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. *J. Mol. Evol.* **29**:170–179.
- Kitamoto, N., S. Yoshino-Yasuda, K. Ohmiya, and N. Tsukagoshi. 2001. Sequence analysis and overexpression of a pectin lyase gene (pel1) from *Aspergillus oryzae* KBN616. *Biosci. Biotechnol. Biochem.* **65**:209–12.
- Kumar, S., and B. Hedges. 1998. A molecular timescale for vertebrate evolution. *Nature* **392**:917–920.
- Kuroyama, H., N. Tsutsui, Y. Hashimoto, and Y. Tsumuraya. 2001. Purification and characterization of a β -glucuronidase from *Aspergillus niger*. *Carbohydrate Res.* **333**:27–39.
- Lane, D. J. 1991. 16S/23S rRNA sequencing. Pp. 115–175 in E. Stackebrandt and M. Goodfellow, eds. *Nucleic acid techniques in bacterial systematics*. Wiley, New York.
- Marquardt, D. W. 1963. An algorithm for least squares estimation of parameters. *J. Soc. Industr. Appl. Math.* **11**:431–441.
- Muyzer, G., E. C. de Waal, and A. G. Uitterlinden. 1993. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl. Environ. Microbiol.* **59**:695–700.
- Nakagawa, T., T. Miyaji, H. Yurimoto, Y. Sakai, N. Kato, and N. Tomizuka. 2000. A methylotrophic pathway participates in pectin utilization by *Candida boidinii*. *Appl. Environ. Microbiol.* **66**:4253–4257.
- Nei, M., P. Xu, and G. Glazko. 2001. Estimation of divergence times from multiprotein sequences for a few mammalian species and several distantly related organism. *Proc. Natl. Acad. Sci. USA* **98**:2497–2502.
- Nielsen, H., S. Brunak, and G. von Heijne. 1999. Machine learning approaches to the prediction of signal peptides and other protein sorting signals. *Prot. Eng.* **12**:3–9.
- Rosewich, U. L., and H. C. Kistler. 2000. Role of horizontal gene transfer in the evolution of fungi. *Annu. Rev. Phytopathol.* **38**:325–363.
- Salzberg, S. L., O. White, J. Peterson, and J. A. Eisen. 2001. Microbial genes in the human genome: lateral transfer or gene loss? *Science* **292**:1903–1906.
- Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: A laboratory manual*. Cold Spring Laboratory, Cold Spring Harbor, NY.
- Screen, S. E., and R. J. S. Leger. 2000. Cloning, expression, and substrate specificity of a fungal chymotrypsin. *J. Biol. Chem.* **275**:6689–6694.
- Stanhope, M. J., A. Lupas, M. J. Italia, K. K. Koretke, C. Volker, and J. R. Brown. 2001. Phylogenetic analyses do not support horizontal gene transfer from bacteria to vertebrates. *Nature* **411**:940–944.
- Sutherland, I. W. 1995. Polysaccharide lyases. *FEMS Microbiol. Rev.* **16**:323–347.
- Taylor, T. N., H. Hass, and H. Kerp. 1999. The oldest fossil ascomycetes. *Nature* **399**:648.
- Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**:4673–4680.
- Tukey, R. H., and C. P. Strassburg. 2000. Human UDP-glucuronosyltransferases: metabolism, expression, and disease. *Annu. Rev. Pharmacol. Toxicol.* **40**:581–616.
- White, T. J., T. Bruns, S. Lee, and J. Taylor. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. Pp. 315–322 in M. A. Innis, D. H. Gelfand, J. J. Sninsky, and T. J. White, eds. *PCR protocols: A Guide to methods and applications*. Academic Press, San Diego.
- Wilson, K. J., S. G. Hughes, and R. A. Jefferson. 1992. The *Escherichia coli* gus operon: induction and expression of the gus operon in *E. coli* and the occurrence and use of GUS in other bacteria. In S. R. Gallagher, ed. *GUS protocols: Using the gus gene as a reporter of gene expression*. Academic Press, San Diego.
- Wu, G., K. Henze, and M. Müller. 2001. Evolutionary relationships of the glucokinase from the amitochondriate protist, *Trichomonas vaginalis*. *Gene* **264**:265–271.
- Yang, Z. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. *CABIOS* **13**:555–556.

Laura Katz, Associate Editor

Accepted October 4, 2004