Letter to the Editor

Lateral Gene Transfer and Metabolic Adaptation in the Human Parasite *Trichomonas vaginalis*

Audrey P. de Koning, Fiona S. L. Brinkman,† Steven J. M. Jones,‡ and Patrick J. Keeling§*

*Department of Zoology, University of British Columbia, Vancouver, British Columbia, Canada; †Department of Microbiology and Immunology, University of British Columbia, Vancouver, British Columbia, Canada; ‡Genome Sequence Centre, British Columbia Cancer Research Centre, Vancouver, British Columbia, Canada; and §Canadian Institute for Advanced Research, Department of Botany, University of British Columbia, Vancouver, British Columbia, Canada

The evolution of a parasitic lifestyle necessitates adaptation to a specialized niche. Examples of common adaptive traits include host interaction systems, metabolic pathways that allow the acquisition of nutrients from the host, and mechanisms to evade host defenses. Such traits could originate by a process of gradual change, but there are mechanisms that would allow potential parasites to adapt very quickly, and one of these mechanisms is lateral gene transfer. Lateral transfer is the process by which genetic information is passed from one genome to an unrelated genome, where it is stably integrated and maintained. There is growing evidence from whole-genome analyses that this process is a very important mechanism in genome evolution, particularly among prokaryotes (Lawrence 1999). Lateral transfer may be especially important in the evolution of a parasitic lifestyle, as infection-related factors could be transmitted, and many of these factors would presumably confer an immediate selective advantage. In bacteria, there are many examples of adaptations that have been acquired through lateral gene transfer between pathogenic species; for instance, pathogenicity islands thought to be derived from a common ancestor are found in the genomes of many distantly related bacterial pathogens (Groisman and Ochman 1996).

While lateral transfer between bacteria is well studied and occurs by known mechanisms (Lawrence 1999), the process of transfer between bacteria and eukaryotes is not as well characterized. Almost all of the transdomain gene transfers that have been well documented involve genes derived from the mitochondrion or the chloroplast, organelles of endosymbiotic origin (Martin and Herrmann 1998). It has been argued that substantial rates of bacteria-to-eukaryote transfer should also take place through phagocytosis of bacteria by eukaryotes (Doolittle 1998). Such transfers are unlikely to take place in ''higher'' eukaryotes like animals or plants, because features such as germ-soma separation and the absence of active phagocytosis reduce the likelihood of passing on any acquired genes. However, many microbial eukaryotes are voracious predators of bacteria, and this creates the opportunity for the frequent uptake and occasional maintenance of bacterial genetic information. Moreover, most parasitic protozoa share their niche with a great many bacterial parasites, and the acquisition of genes from bacterial neighbors could conceivably contribute to the adaptation to parasitism in these protozoa.

We sought evidence of transkingdom lateral transfer using an automated BLAST search to compare protein sequences from parasitic bacteria and protozoa with the nonredundant SwissProt-Trembl database in the SRS system. The *N*-acetylneuraminate lyase gene (EC 4.1.3.3) of the parabasalid protozoan *Trichomonas vaginalis* was identified as one whose primary protein sequence was much more similar to neuraminate lyases of bacterial origin than to any known protein from another eukaryote. The neuraminate lyase gene from *T. vaginalis* is easily alignable and colinear with several bacterial lyase genes, with the single exception of the N-terminus of the *T. vaginalis* gene, which encodes a 24-aa leader that is absent in its bacterial counterparts. Excluding this leader sequence, the *T. vaginalis* gene has 80% sequence identity at the amino acid level to the neuraminate lyase gene in the bacteria *Haemophilus influenzae*. This high degree of similarity has previously been noted (Meysick, Dimock, and Gerber 1996; Traving, Roggentin, and Schauer 1997; Traving and Schauer 1998), but to date, no explanation of this unusual similarity for two such phylogenetically disparate organisms has been forthcoming.

To investigate the phylogenetic relationship between *N*-acetylneuraminate lyase genes, an amino acid sequence alignment was constructed for the gene family that includes *N*-acetylneuraminate lyase, dihydrodipicolinate synthase (EC 4.2.1.52), D-dehydro-4-deoxyoglucarate dehydratase (EC 4.2.1.41), and hydratase-aldolase genes, whose products are structurally and functionally related (Lawrence et al. 1997). Ninety-five sequences were selected from the GenPept and the Unfinished Microbial Genomes databases at NCBI, based on a criterion of lowest expect scores in a BLAST search using the *T. vaginalis* sequence as the query. Sequences were aligned with CLUSTAL X, the alignment was edited manually, and ambiguous positions were removed. These proteins, although structurally and functionally similar, are nevertheless relatively divergent at the primary sequence level, and consequently only 191 unambiguous positions could be identified. Trees were inferred from these positions to confirm the general relationship of the *T. vaginalis* sequence to bacterial *N*-acetylneuraminate lyases (fig. 1). While some relationships were inconsistent between different phylogenetic methods, the relationship between *T. vaginalis* and bacterial *N*-acetylneuraminate lyase genes emerged consis-

Key words: lateral transfer, protozoa, molecular phylogeny, parasitism, *N*-acetyl neuraminate lyase, *Trichomonas vaginalis.*

Address for correspondence and reprints: Audrey P. de Koning, Department of Zoology, University of British Columbia, 6270 University Boulevard, Vancouver, British Columbia, Canada V6T 1Z4. E-mail: dekoning@zoology.ubc.ca.

Mol. Biol. Evol. 17(11):1769–1773. 2000

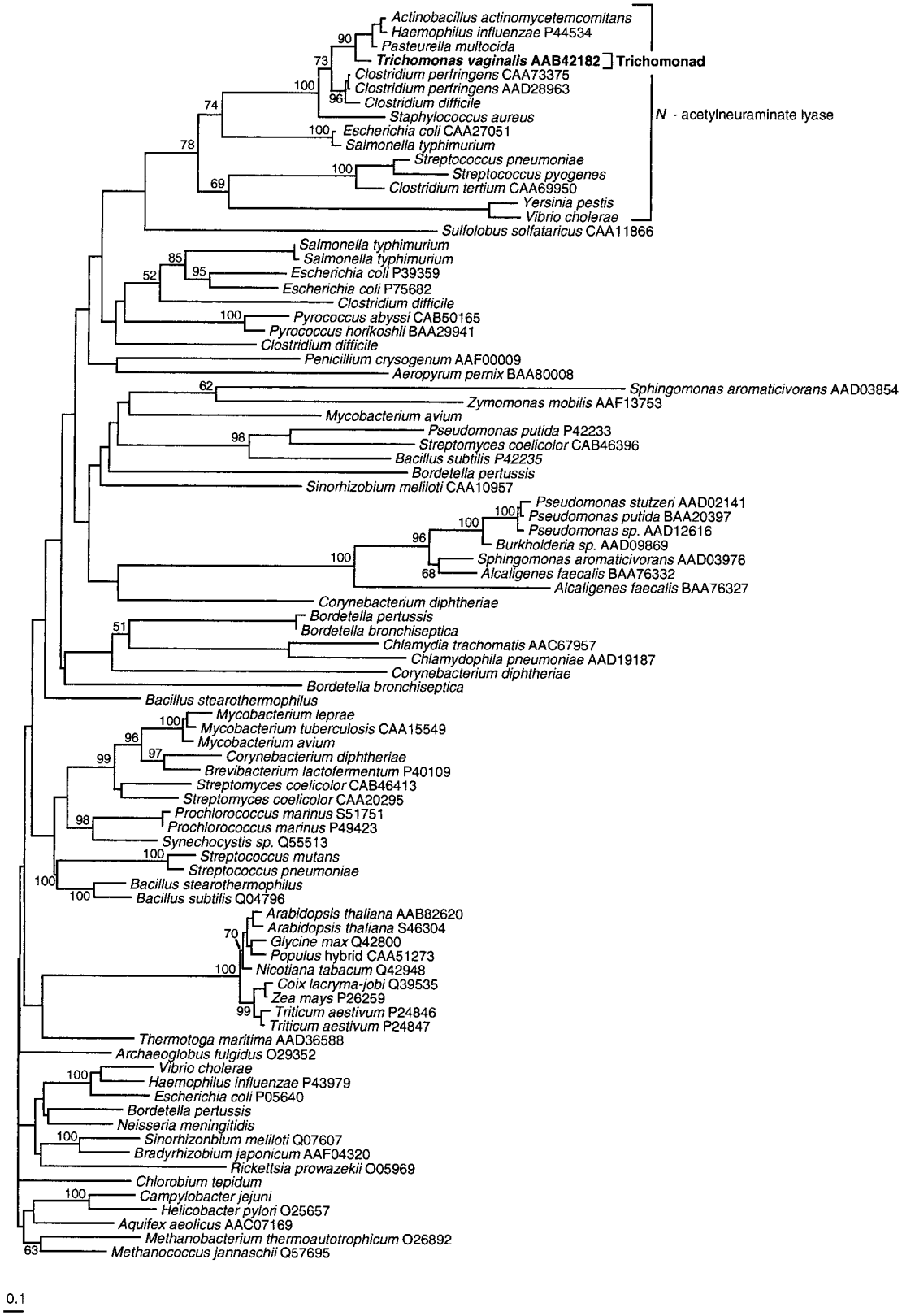

FIG. 1.—Phylogeny of *N*-acetylneuraminate lyase and other members of this gene family. Bootstrap values over 50% are shown for major nodes only. Those with accession numbers are from GenPept, while others were deduced from DNA sequences available from incomplete microbial genome sequencing projects. The phylogeny is a BioNJ tree based on maximum-likelihood distances calculated with PUZZLE, version 4.0.1, using the JTT substitution matrix and amino acid frequencies estimated from the data. Bootstraps were calculated in the same way using the shell script puzzleboot (by M. Holder and A. Roger). For clarity, only the *N*-acetylneuraminate lyase clade has been labeled from this gene family.

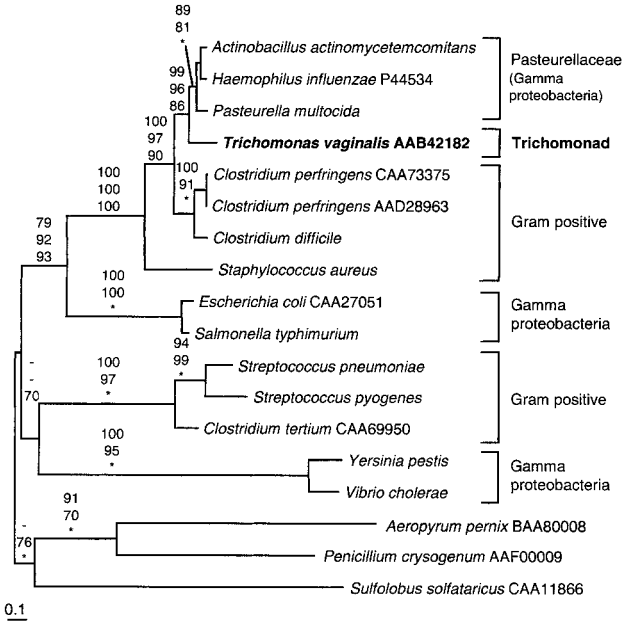

FIG. 2.—Phylogeny of *N*-acetylneuraminate lyases. Labels are as in figure 1, except that numbers at nodes correspond to distance bootstraps (top), quartet puzzling steps (center), and relative estimated log likelihood bootstraps (bottom). Numbers less than 50 are indicated as dashes, and nodes that were constrained in protein maximum likelihood are shown with asterisks. The phylogeny was constructed as in figure 1, except that the distances and bootstrap distances were corrected for site-to-site rate variation using a gamma distribution with eight rate categories plus invariant sites. Quartet puzzling was carried out with the same settings and 10,000 steps. Protein maximum-likelihood trees were the result of an exhaustive search using ProtML, version 2.2, of a partially constrained tree using the JTT substitution matrix. Constraints were based on strongly supported nodes from distance and puzzling trees.

tently and with strong bootstrap support from all methods. Ten alternative topologies were also compared with Kishino-Hasegawa tests, and every tree in which the *T. vaginalis* sequence was outside the *N*-acetylneuraminate lyase clade was rejected at the 1% confidence level.

To further refine the phylogenetic relationships between *T. vaginalis* and the bacterial sequences, a more comprehensive analysis was carried out on a subset of 18 sequences which included the *N*-acetylneuraminate lyase clade and three closely related sequences used as an outgroup. Consideration of only 18 sequences allowed for more extensive analyses, and more importantly, for a great many more sites to be unambiguously aligned (248 positions). All methods used in this analysis consistently showed a strongly supported relationship between the sequences of *T. vaginalis* and members of the Pasteurellaceae family of proteobacteria (fig. 2). Analyses were also carried out using partial amino acid sequences deduced from vertebrate cDNA that are related to this group of enzymes (GenBank accession numbers AA162738 and W79930), but these vertebrate sequences were never observed to be related to the *T. vaginalis*–Pasteurellaceae clade, and they lack insertions common to *T. vaginalis* and pasteurellaceaen enzymes (not shown).

The nesting of the *T. vaginalis* sequence within bacteria, and the convincing specific relationship to Pasteurellaceae strongly implies that the *N*-acetylneuraminate lyase gene was transferred from an ancestor of the Pasteurellaceae to an ancestor of *T. vaginalis.* The phylogeny also suggests that there may have been many lateral transfer events within bacteria, since the relationships among bacterial sequences are not consistent with the known phylogeny of these organisms (e.g., neither the gamma-proteobacteria nor the gram-positive bacteria form distinctive clades). Moreover, *N*-acetylneuraminate lyase is not even present in the majority of completely sequenced bacterial genomes, which lends additional support to the conclusion that this gene was transmitted between bacterial lineages. Because of the evident lateral transfer among bacteria, it is possible that the *T. vaginalis* gene was not transferred from an ancestor of the Pasteurellaceae, but that the ancestors of both the Pasteurellaceae and *T. vaginalis* acquired this gene from some other donor. However, given the high degree of sequence similarity between *T. vaginalis* and Pasteurellaceaen enzymes, this is unlikely, as it would require two independent transfers in a very short period.

Trichomonas vaginalis and other trichomonads frequently internalize bacteria by phagocytosis (Rendon-Maldonado et al. 1998), so over evolutionary time it is inevitable that trichomonad genomes would have been exposed to a great deal of foreign DNA, a factor which has been hypothesized to contribute to bacteria-to-protist lateral transfer events (Doolittle 1998). Indeed, other genes in *Trichomonas* also appear to have a bacterial ancestry, e.g., GAPDH and PFO (Henze et al. 1995; Hrdy and Müller 1995). However, the extremely high degree of similarity between Pasteurellaceaen and *T. vaginalis* sequences is unprecedented for a transfer event between such distantly related organisms, which suggests that the transfer must have occurred very recently in evolution, and makes identification of the participants uncommonly clear.

The organisms involved in this transfer are important because the nature of the participants and the gene involved raises the possibility that this transfer may have affected the adaptation of *Trichomonas* to parasitism. *N*-acetylneuraminate lyase is an enzyme involved in the metabolism of sialic acids, which are a family of monosaccharides that are generally bound to glycolipids, glycoproteins, or oligosaccharides and are found on cell surfaces, some intracellular membranes, and as components of serum and mucus of higher animals. *N*-acetylneuraminate lyase is the final enzyme in the sialic acid degradative pathway and cleaves free sialic acid into acetylmannosamine and pyruvate (Traving and Schauer 1998). This cleavage activity serves a number of purposes in the bacteria in which it is found, but it is thought to primarily provide pyruvate for carbon and energy (Vimr and Troy 1985). Thus, sialic acid degradation allows microbes to scavenge sialic acids from the host for nutrition, which would be especially useful for vertebrate endoparasites, since sialic acid is abundant in higher animals and scarce elsewhere (Schauer et al. 1995). In certain parasitic bacteria, *N*-acetylneuraminate

lyase also plays a role in the synthesis of a sialic acid polymer component of capsules (Ferrero et al. 1996) that help these bacteria evade host defenses (Troy 1995). While acetylneuraminate lyase has not directly been shown to be crucial to infection, the pathways in which it is involved are implicated either in resisting host immune defenses or in enabling the parasite to garner nutrients from its host, both of which are related to parasitism.

The lateral transfer of this enzyme from an ancestor of bacterial parasites of vertebrate epithelia (the Pasteurellaceae) to an ancestor of *T. vaginalis* is intriguing because *Trichomonas* is also an epithelial parasite, specifically of the human urogenital system. Since the pathways in which *N*-acetylneuraminate lyase functions may contribute to the success of epithelial parasites, the transfer of this protein may have facilitated, or at least contributed to, the parasitic lifestyle of *T. vaginalis.* There is some evidence for metabolic pathways involving sialic acid in trichomonads. Sialic acids are found on the surface of *T. vaginalis* (Dias Filho et al. 1992) and a related trichomonad, *Tritrichomonas foetus* (Benchimol et al. 1981), although it is unknown if an acetylneuraminate lyase is involved in their synthesis. Moreover, the *Trichomonas N*-acetylneuraminate lyase has been shown to produce the cleavage product acetylmannosamine in vitro (Meysick, Dimock, and Gerber 1996), and secreted sialidase (another enzyme in the degradative pathway which liberates bound sialic acids from cell surfaces) activity has been detected in *T. vaginalis* (Padilla Vaca and Anaya Velazquez 1997). Interestingly, the *N*-acetylneuraminate lyase of *Trichomonas* also appears to be secreted, since it was first identified by antibodies raised against *T. vaginalis* extracellular factors, and the inferred protein possesses a 24-residue N-terminal extension whose removal increases enzyme activity in vitro (Meysick, Dimock, and Gerber 1996). In addition, the N-terminal end is weakly predicted by SignalP to be a signal peptide that is cleaved at amino acid 16. All other acetylneuraminate lyases studied to date are cytosolic, so the *Trichomonas* enzyme may have a distinctive function, or it may be targeted to a lytic vacuole (a possibility not supported by the identification of this enzyme by a screen for secreted proteins; Meysick, Dimock, and Gerber 1996).

The role of this pathway in trichomonads is unknown, and sialic acid metabolism is involved in a variety of cellular activities (Schauer et al. 1995; Traving and Schauer 1998), so it is not certain whether the trichomonad *N*-acetylneuraminate lyases have the same function as those of the Pasteurellaceae. The recent transfer of this enzyme from one epithelial parasite to another suggests that it may well have the same role, but confirmation of this awaits functional characterization in *T. vaginalis.* The origin of other proteins involved in sialic acid metabolism has also not been investigated in trichomonads, but it would be worthwhile to determine if the entire pathway was acquired through lateral transfer (in certain bacteria, sialidases are known to have been involved in lateral transfers; Roggentin et al. 1993) or if the bacterial acetylneuraminate lyase was integrated into an existing pathway. In addition, it would be interesting to investigate whether the neuraminate lyase gene is also present in other trichomonads. Such a comparison would provide more information about when this gene transfer occurred and could provide a correlation between parasitic versus free-living lifestyles and the presence of neuraminate lyase.

The lateral transfer of *N*-acetylneuraminate lyase from an ancestor of the Pasteurellaceae to a progenitor of *T. vaginalis* is a strongly supported case of a recent lateral transfer from a bacterium to a eukaryote. Moreover, this event may represent an important factor in the evolution of parasitism. Lateral transfer could allow a previously harmless organism to rapidly colonize a new environment by acquiring highly specific biochemical functionality by leaps and bounds, rather than by gradual adaptation.

Acknowledgments

This work was supported by a grant from the Peter Wall Institute for advanced studies. A.P.d.K. is supported by a studentship from the Natural Sciences and Engineering Research Council of Canada. P.J.K. is a scholar of the Canadian Institute for Advanced Research. We thank the Peter Wall Pathogenomics group for helpful discussion and support of this work, and we thank the following sequencing centers, which kindly made preliminary data available. Preliminary sequence data for *Staphylococcus aureus, Streptococcus pneumoniae, Vibrio cholerae, Mycobacterium avium,* and *Chlorobium tepidum* were obtained from the Institute for Genomic Research website at http://www.tigr.org. Preliminary sequence data for *Actinobacillus actinomycetemcomitans, Bacillus stearothermophilus, Streptococcus pyogenes,* and *Streptococcus mutans* were obtained from the University of Oklahoma's Advanced Center for Genome Technology website at http://www.genome.ou.edu. Preliminary sequence data for *Bordetella bronchiseptica, Bordetella pertussis, Clostridium difficile, Campylobacter jejuni, Corynebacterium diphtheriae, Mycobacterium leprae, Neisseria meningitidis, Salmonella typhi,* and *Yersinia pestis* were obtained from the Sanger Centre at http://www.sanger.ac.uk/Projects/Microbes. Preliminary sequence data for *Pasteurella multocida* was obtained from the University of Minnesota *P. multocida* genome sequencing website, http:www.cbc.umn.edu/ResearchProjects/AGAC/Pm/index.html. We also wish to thank the Genome Sequencing Center, Washington University, St. Louis, for communication of DNA sequence *Salmonella typhimurium* data prior to publication.

LITERATURE CITED

BENCHIMOL, M., M. E. PEREIRA, C. A. ELIAS, and W. DE SOUZA. 1981. Cell surface carbohydrates in *Tritrichomonas foetus.* J. Protozool. **28**:337–341.

DIAS FILHO, B. P., A. F. ANDRADE, W. DE SOUZA, M. J. ESTEVES, and J. ANGLUSTER. 1992. Cell surface saccharide differences in drug-susceptible and drug-resistant strains of *Trichomonas vaginalis.* Microbios **71**:55–64.

DOOLITTLE, W. F. 1998. You are what you eat: a gene transfer ratchet could account for bacterial genes in eukaryotic nuclear genomes. Trends Genet. **14**:307–311.

FERRERO, M. A., A. REGLERO, M. FERNANDEZ-LOPEZ, R. ORDAS, and L. B. RODRIGUEZ-APARICIO. 1996. *N*-Acetyl-D-neuraminic acid lyase generates the sialic acid for colominic acid biosynthesis in *Escherichia coli* K1. Biochem. J. **317**: 157–165.

GROISMAN, E. A., and H. OCHMAN. 1996. Pathogenicity islands: bacterial evolution in quantum leaps. Cell **87**:791–794.

HENZE, K., A. BADR, M. WETTERN, R. CERFF, and W. MARTIN. 1995. A nuclear gene of eubacterial origin in *Euglena gracilis* reflects cryptic endosymbioses during protist evolution. Proc. Natl. Acad. Sci. USA **92**:9122–9126.

HRDY, I., and M. MÜLLER. 1995. Primary structure and eubacterial relationships of the pyruvate : ferredoxin oxidoreductase of the amitochondriate eukaryote Trichomonas vaginalis. J. Mol. Evol. **41**:388–396.

LAWRENCE, J. G. 1999. Gene transfer, speciation, and the evolution of bacterial genomes. Curr. Opin. Microbiol. **2**:519–523.

LAWRENCE, M. C., J. A. BARBOSA, B. J. SMITH, N. E. HALL, P. A. PILLING, H. C. OOI, and S. M. MARCUCCIO. 1997. Structure and mechanism of a sub-family of enzymes related to N-acetylneuraminate lyase. J. Mol. Biol. **266**:381–399.

MARTIN, W., and R. G. HERRMANN. 1998. Gene transfer from organelles to the nucleus: how much, what happens, and why? Plant Physiol. **118**:19–17.

MEYSICK, K. C., K. DIMOCK, and G. E. GERBER. 1996. Molecular characterization and expression of a N-acetylneuraminate lyase gene from *Trichomonas vaginalis.* Mol. Biochem. Parasitol. **76**:289–292.

PADILLA VACA, F., and F. ANAYA VELAZQUEZ. 1997. Biochemical properties of a neuraminidase of *Trichomonas vaginalis.* J. Parasitol. **83**:1001–1006.

RENDON-MALDONADO, J. G., M. ESPINOSA-CANTELLANO, A. GONZALEZ-ROBLES, and A. MARTINEZ-PALOMO. 1998. *Trichomonas vaginalis*: in vitro phagocytosis of lactobacilli, vaginal epithelial cells, leukocytes, and erythrocytes. Exp. Parasitol. **89**:241–250.

ROGGENTIN, P., R. SCHAUER, L. L. HOYER, and E. R. VIMR. 1993. The sialidase superfamily and its spread by horizontal gene transfer. Mol. Microbiol. **9**:915–921.

SCHAUER, R., S. KELM, G. REUTER, P. ROGGENTIN, and L. SHAW. 1995. Biochemistry and role of sialic acids. Pp. 7–67 *in* A. ROSENBERG, ed. Biology of the sialic acids. Plenum Press, New York.

TRAVING, C., P. ROGGENTIN, and R. SCHAUER. 1997. Cloning, sequencing and expression of the acylneuraminate lyase gene from *Clostridium perfringens* A99. Glycoconj. J. **14**: 821–830.

TRAVING, C., and R. SCHAUER. 1998. Structure, function and metabolism of sialic acids. Cell Mol. Life Sci. **54**:1330–1349.

TROY, F. A. 1995. Sialobiology and the polysialic acid glycotope: occurrence, structure, function, synthesis, and glycopathology. Pp. 95–144 *in* A. ROSENBERG, ed. Biology of the sialic acids. Plenum Press, New York.

VIMR, E. R., and F. A. TROY. 1985. Identification of an inducible catabolic system for sialic acids (nan) in *Escherichia coli.* J. Bacteriol. **164**:845–853.

B. FRANZ LANG, reviewing editor

Accepted July 11, 2000