Microbiology Comment provides a platform for readers of **Microbiology** to communicate their personal observations and opinions in a more informal way than through the submission of papers.

Most of us feel, from time to time, that other authors have not acknowledged the work of our own or other groups or have omitted to interpret important aspects of their own data. Perhaps we have observations that, although not sufficient to merit a full paper, add a further dimension to one published by others. In other instances we may have a useful piece of methodology that we would like to share.

The Editors hope that readers will take full advantage of this section and use it to raise matters that hitherto have been confined to a limited audience.

Christopher M. Thomas, Editor-in-chief

Maize Yellow Stripe1, an iron-phytosiderophore uptake transporter, is a member of the oligopeptide transporter (OPT) family

As genome sequencing technologies become increasingly efficient, the need for rapid accurate annotation procedures becomes critical. At present, genome annotation is usually the rate-limiting step in the publication of genomic data. Correct annotation, however, depends upon the veracity of functional assignments for sequenced genes in the current databases. Erroneous assignments can be, and have been, propagated into hundreds of errors resulting from the use of semi-automated genome sequencing procedures that involve minimal human intervention. The correction of such proliferated errors depends on the meticulous examination of assignments using laborious procedures for which no *in silico* substitute for the human brain has yet been invented. When annotation errors are introduced into a database, it is therefore im-

perative that members of the scientific community who initially discover these errors relay their findings to both the authors and the database curators so as to 'nip the problem in the bud'.

In an exciting recent paper, Curie *et al.* (3) identified a transporter in the maize root membrane responsible for the uptake of Fe^{3+} -phytosiderophores. Plants extrude low molecular weight secondary amino acids, called mugineic acids that chelate Fe^{3+} , thereby rendering this cation soluble. The resultant chelation complex is taken up from the soil, providing the iron essential for normal metabolic function. Chelation is the mechanism used by the world's principal grain crops for iron acquisition. Iron deficiency leads to chlorosis, poor yield and decreased nutritional quality (3).

In their study, Curie *et al.* showed that the *yellow stripe1* (*ys1*) mutant in corn is deficient in Fe^{3+} -phytosiderophore uptake, thus leading to the conclusion that YS1 is the transporter. This suggestion was substantiated by showing that [1] *ys1* mRNA levels were increased in roots under iron-deficient conditions and [2] expression of *ys1* in an iron-uptake-deficient strain of *Saccharomyces cerevisiae* restored growth specifically in the presence of Fe^{3+} -phytosiderophores (3).

Curie *et al.* noted that 'the amino acid sequence predicted from the *ys1* cDNA does not show strong similarity to any protein of known function' and that it 'shows weak but significant similarity to a hypothetical yeast protein, YGL114 (36 % similarity), belonging to the major facilitator superfamily' (MFS; reviewed in reference 9). We have examined this last claim and found that YGL114 is not a member of the MFS (TC no. 2.A.1; 10, 11), but instead is a member of the oligopeptide transporter (OPT) family (TC no. 2.A.67; 7, 8). Moreover, YS1 does not show significant sequence similarity to any member of the MFS although it exhibits significant sequence similarity to several functionally characterized yeast porters of the OPT family which transport tri-, tetra-, penta- and hexapeptides (1, 4, 5, 7, 8).

The currently recognized members of the OPT family, retrieved from the current (7/2001) NCBI databases, are presented in Table 1, and the phylogenetic tree of the OPT family is shown in Fig. 1. The tree shows that all of

the prokaryotic proteins cluster loosely together (cluster 1). However, the eukaryotic proteins form four clusters: two well defined plant clusters (clusters 2 and 5), one well defined fungal cluster (cluster 3) consisting of only two proteins, and one very diffuse fungal cluster (cluster 4) consisting of seven proteins from four species. All of the functionally characterized peptide transporters (indicated in bold print in Fig. 1) are in cluster 4, while the maize YS1 protein is in cluster 2. The eight YS1-like (YSL) *Arabidopsis thaliana* proteins noted by Curie *et al.* (3) are in cluster 2. A ninth partial YSL sequence (YSL9) was also found. No member of the yeast cluster 3 and plant cluster 5 is functionally characterized. Although no function is known for a prokaryotic homologue, the sequence-divergent EspB protein of *Myxococcus xanthus* is an early sporulation protein that influences the temporal developmental pattern (2). We postulate that it is a transporter that brings into the cytoplasm a molecule, probably a peptide or amino acid derivative, that influences the activity of EspA (2).

The divergent substrate specificities of Opt1p and YS1 suggest that each cluster of sequence-divergent proteins may be specific for a different class of compounds. However, since phytosiderophores and peptides are amino acid derivatives, and since all functionally characterized members of the family

► GUIDELINES

Communications should be in the form of letters and should be brief and to the point. A single small Table or Figure may be included, as may a limited number of references (cited in the text by numbers, and listed in alphabetical order at the end of the letter). A short title (fewer than 50 characters) should be provided.

Approval for the publication rests with the Editor-in-Chief, who reserves the right to edit letters and/or to make a brief reply. Other interested persons may also be invited to reply. The Editors of **Microbiology** do not necessarily agree with the views expressed in **Microbiology Comment**.

Contributions should be addressed to the Editor-in-Chief via the Editorial Office.

Table 1. Proteins of the oligopeptide transporter (OPT) family

Protein abbreviation	Database description	Organism	Size (aa)	GenBank index no.
Bacteria (cluster 1)				
Orf Cje	Probable integral membrane protein Cj0204	<i>Campylobacter jejuni</i>	665	11281587
Orf Ccr	Oligopeptide transporter, OPT family	<i>Caulobacter crescentus</i>	666	13424225
Orf Hin	Hypothetical protein HI0561	<i>Haemophilus influenzae</i> Rd	661*	2506742
Orf Lsa	Orf5	<i>Lactobacillus sakei</i>	645	1370208
Orf Mtu	Probable membrane protein Rv2395	<i>Mycobacterium tuberculosis</i>	667	7478347
EspB Mxa	Putative membrane protein	<i>Myxococcus xanthus</i>	592	9313036
Orf Nme	Conserved hypothetical protein NMB0060	<i>Neisseria meningitidis</i>	672	11281588
Orf Pal	Hypothetical protein (fragment)	<i>Prevotella albensis</i>	407	5295906
Orf Pae	Conserved hypothetical protein PA3934	<i>Pseudomonas aeruginosa</i>	678	11348115
Orf Xfa	Oligopeptide transporter XF2261	<i>Xylella fastidiosa</i>	653	11281590
Archaea (cluster 1)				
Orf Hsp	Vng6268c	<i>Halobacterium</i> sp. NRC-1	655	10584311
Orf Pho	Hypothetical protein PH0361	<i>Pyrococcus horikoshii</i>	626	7518601
Fungi (cluster 4)				
Opt1p Cal	Opt1p	<i>Candida albicans</i>	783	2367386
Opt1p Sce	Opt1p	<i>Saccharomyces cerevisiae</i>	799	6322249
Opt2p Sce	Peptide transporter, Opt2p	<i>Saccharomyces cerevisiae</i>	877	6325452
Mtd1 Sco	Peptide transporter MTD1	<i>Schizophyllum commune</i>	777	6716399
Isp4 Spo	Sexual differentiation process protein isp4	<i>Schizosaccharomyces pombe</i>	785	9955814
Orf1 Spo	Hypothetical protein SPAC29B 12.10c	<i>Schizosaccharomyces pombe</i>	851	7491035
Orf2 Spo	Major facilitator protein homologue	<i>Schizosaccharomyces pombe</i>	791	7492160
Fungi (cluster 3)				
Orf Sce	Ygl114wp	<i>Saccharomyces cerevisiae</i>	725	6321324
Orf Ncr	Conserved hypothetical protein	<i>Neurospora crassa</i>	738	9368956
Plants (cluster 2)				
Ys1 Zma	Iron-phytosiderophore transporter protein yellow stripe1	<i>Zea mays</i>	682	10770865
Ysl1 Ath	Putative protein	<i>Arabidopsis thaliana</i>	665	14423556
Ysl2 Ath	Contains similarity to integral membrane protein	<i>Arabidopsis thaliana</i>	652	10177899
Ysl3 Ath	EspB-like protein	<i>Arabidopsis thaliana</i>	669	9759194
Ysl4 Ath	Gb AAF23830.1 ~ gene_id: MEE6.7	<i>Arabidopsis thaliana</i>	664	9759146
Ysl5 Ath	Gb AAD49762.1 ~ gene_id: MKP6.21	<i>Arabidopsis thaliana</i>	714	9294153
Ysl6 Ath	Gb AAF23830.1 ~ gene_id: MOJ10.9	<i>Arabidopsis thaliana</i>	676	9279625
Ysl7 Ath	F1E22.10	<i>Arabidopsis thaliana</i>	688	6686396
Ysl8 Ath	At1g48370/F11A17_27	<i>Arabidopsis thaliana</i>	724	14517530
Ysl9 Ath	Gb AAD49762.1 ~ gene_id: MFC19.12 (fragment)	<i>Arabidopsis thaliana</i>	216	9758735
Ysl1 Osa	Contains EST AU108928(E2625)	<i>Oryza sativa</i>	708	11034704
Ysl2 Osa	Similar to <i>Arabidopsis thaliana</i> BAC F1E22 (fragment)	<i>Oryza sativa</i>	181	9558505
Plants (cluster 5)				
Orf I Ath	Sexual differentiation process protein ISP4-like	<i>Arabidopsis thaliana</i>	755	9758213
Orf2 Ath	Isp4 protein homologue T15N24.40	<i>Arabidopsis thaliana</i>	753	7487930
Orf3 Ath	Similar to <i>S. pombe</i> ISP4 (Gb D83992)	<i>Arabidopsis thaliana</i>	722	2160185
Orf4 Ath	Isp4-like protein	<i>Arabidopsis thaliana</i>	729	9759417
Orf5 Ath	Isp4 protein homologue T29A15.220	<i>Arabidopsis thaliana</i>	736	7487931
Orf6 Ath	Isp4 protein	<i>Arabidopsis thaliana</i>	741	9759190
Orf7 Ath	Isp4 protein	<i>Arabidopsis thaliana</i>	733	9759191
Orf8 Ath	Isp4 protein homologue T12H20.7	<i>Arabidopsis thaliana</i>	766	7487929
Orf9 Ath	Isp4-like protein	<i>Arabidopsis thaliana</i>	691	7485100
Orf1 Osa	ESTs D25093(R3139), AU082160(R3139)	<i>Oryza sativa</i>	766	6983869
Orf2 Osa	EST D25093(R3139)	<i>Oryza sativa</i>	786	6983868
Orf3 Osa	Similar to <i>Arabidopsis thaliana</i> BAC T12H20	<i>Oryza sativa</i>	769	6983880

* The original sequence of HI0561 was 633 aa due to an apparent nonsense mutation near the N terminus of the protein, probably due to a sequencing error. The nonsense codon should encode a fully conserved arginyl residue, and if so the protein is extended by 28 residues at the N terminus.

Fig. 1. Phylogenetic tree of the oligopeptide transporter (OPT) family. Protein abbreviations are as indicated in Table 1. The CLUSTAL X program (6) was used to generate the tree. Functionally characterized proteins are presented in bold print. See text for explanation of clustering patterns. The two archaeal proteins are indicated by the dotted line. The *Xylella fastidiosa* protein which unexpectedly clusters with the archaeal proteins did not exhibit the codon usage frequencies of *X. fastidiosa* proteins and therefore may have been acquired by horizontal gene transfer from another organism.

function with inwardly directed polarity, all members of the OPT family may function in the uptake of amino acid-containing compounds and their derivatives (11). It is interesting to note that each of the five clusters shown in Fig. 1 includes proteins with a characteristic size range ($1 < 2 < 3 < 5 < 4$) (see Table 1). Further, while no prokaryote has more than one sequenced OPT family homologue, and many sequenced bacterial genomes (such as those of *Escherichia coli* and *Bacillus subtilis*) do not encode such a protein, both *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* have three paralogs, and *A. thaliana* has nine each in clusters 2 and 5.

In summary, the recent work of Curie *et al.* (3) and our *in silico* analyses have led to the realization that the OPT family is far more complex, diverse and interesting than was

previously recognized. Since no functionally characterized OPT family member had been shown to transport anything but oligopeptides, genome annotators have designated all members of this family as peptide transporters. We wish to alert the scientific community to the fact that genome annotations based on the previous incomplete functional data may well be inaccurate.

Our analyses identify distant OPT family homologues and put the work of Curie *et al.* into a firm evolutionary framework. The phylogenetic approach has shown that OPT family members exhibit clustering according to organism type, but that within a cluster, horizontal gene transfer has probably occurred. The occurrence of whole clusters of previously unrecognized homologues with non-functionally characterized members, including clusters 1 (bacterial), 3 (fungal) and 5

(plant) is particularly noteworthy. We rationalize the relationship between YS1 and peptide transporters based on the amino acid nature of the YS1 substrate siderophores. This article therefore serves as a guide to the functional elucidation of phylogenetically divergent members of this new superfamily. Reverse genetic approaches should prove revealing.

M.-R. Yen^{1,2}, Y.-H. Tseng² and M. H. Saier, Jr¹

¹Division of Biology, University of California at San Diego, La Jolla, CA 92093-0116, USA

²Institute of Molecular Biology, National Chung Hsing University, Taichung, 402, Taiwan, Republic of China

Author for correspondence: M. H. Saier, Jr.
Tel: +1 858 534 4084. Fax: +1 858 534 7108.
e-mail: msaier@ucsd.edu

1. Bourbouloux, A., Shahi, P., Chakladar, A., Delrot, S. & Bachhawat, A. K. (2000). Hgtlp, a high affinity glutathione transporter from the yeast *Saccharomyces cerevisiae*. *J Biol Chem* **275**, 13259–13265.
2. Cho, K. & Zusman, D. R. (1999). Sporulation timing in *Myxococcus xanthus* is controlled by the *espAB* locus. *Mol Microbiol* **34**, 714–725.
3. Curie, C., Panaviene, Z., Loulergue, C., Dellaporta, S. L., Briat, J.-F. & Walker, E. L. (2001). Maize *yellow stripe1* encodes a membrane protein directly involved in Fe(III) uptake. *Nature* **409**, 346–349.
4. Hauser, M., Donhardt, A. M., Barnes, D., Naider, F. & Becker, J. M. (2000). Enkephalins are transported by a novel eukaryotic peptide uptake system. *J Biol Chem* **275**, 3037–3041.
5. Hauser, M., Narita, V., Donhardt, A. M., Naider, F. & Becker, J. M. (2001). Multiplicity and regulation of genes encoding peptide transporters in *Saccharomyces cerevisiae*. *Mol Membr Biol* **18**, 105–112.
6. Jeanmougin, F., Thompson, J. D., Gouy, M., Higgins, D. G. & Gibson, T. J. (1998). Multiple sequence alignment with CLUSTAL X. *Trends Biochem Sci* **23**, 403–405.
7. Lubkowitz, M. A., Hauser, L., Breslav, M., Naider, F. & Becker, J. M. (1997). An oligopeptide transport gene from *Candida albicans*. *Microbiology* **143**, 387–396.
8. Lubkowitz, M. A., Barnes, D., Breslav, M., Burchfield, A., Naider, F. & Becker, J. M. (1998). *Schizosaccharomyces pombe* *isp4* encodes a transporter representing a novel family of oligopeptide transporters. *Mol Microbiol* **28**, 729–741.
9. Pao, S. S., Paulsen, I. T. & Saier, M. H., Jr (1998). The major facilitator superfamily. *Microbiol Mol Biol Rev* **62**, 1–32.
10. Saier, M. H., Jr., Beatty, J. T., Goffeau, A. & 11 other authors (1999). The major facilitator superfamily. *J Mol Microbiol Biotechnol* **1**, 257–279.
11. Saier, M. H., Jr. (2000). A functional/phylogenetic classification system for transmembrane solute transporters. *Microbiol Mol Biol Rev* **64**, 354–411.