Nucleic Acids Research, 2003, Vol. 31, No. 1 325–327
DOI: 10.1093/nar/gkg055

The Mitochondrial Protein Import Machinery of Plants (MPIMP) database

Ryan Lister[1], Monika W. Murcha[1,2] and James Whelan[1,*]

[1]Plant Molecular Biology Group, Biochemistry and Molecular Biology, School of Biomedical and Chemical Sciences and [2]Plant Biology, School of Natural and Agricultural Sciences, University of Western Australia, 35 Stirling Highway, Crawley 6009, Western Australia

Received August 23, 2002; Revised and Accepted October 2, 2002

ABSTRACT

The Mitochondrial Protein Import Machinery of Plants database (MPIMP) is an Internet-accessible database containing detailed information on the protein import apparatus of plant mitochondria. The *Arabidopsis* genome was searched for components of the mitochondrial protein import apparatus using components from the well-characterized model system of *Saccharomyces cerevisiae*. Twenty six homologues of 34 components could be found, encompassing the essential components for the general and carrier import pathways. The database is available through the Internet at http://millar3.biochem.uwa.edu.au/~lister/index.html.

INTRODUCTION

The thousand or more proteins thought to be present in the mitochondrion must be imported from the complex mixture of all cytosolically synthesised proteins (1). The mitochondrial import apparatus carries out this act of discrimination, which is responsible for the import and correct intraorganelle sorting of nuclear encoded mitochondrial proteins (2,3). Our knowledge of the components and mechanisms of the plant import apparatus is poor compared to the intensively studied and more amenable systems of *Saccharomyces cerevisiae* and *Neurospora crassa*. Several hundred million years of diversification has taken place between organisms since the single endosymbiotic event leading to the formation of mitochondria took place (4). In *Arabidopsis*, gene families encode the two outer membrane receptors and the inner membrane translocase components, suggesting functional diversification between members. Additionally, structural predictions on several components, particularly both inner membrane translocases, indicates that they differ structurally compared to the well characterised yeast components. Yeast as a model system does not reflect the cellular and developmental complexity of higher plants. Therefore, identification of the components of the protein import apparatus in plants is a necessary first step to understand mitochondrial biogenesis in plants. Identification of the import machinery in plants has enabled the generation of a model for the plant mitochondrial protein import pathways (Fig. 1).

DATABASE DESCRIPTION

The Mitochondrial Protein Import Machinery of Plants (MPIMP) database currently contains detailed information on the protein import components identified from the *Arabidopsis* genome database. Included are putative homologues to import components that have eluded biochemical characterization.

Information is available online at http://millar3.biochem.uwa.edu.au/~lister/index.html, with hyperlinks leading to detailed information, including multiple sequence alignments between yeast and *Arabidopsis* homologues, hydropathy predictions, Stanford Microarray Data, secondary structure predictions and protein motifs. Also, links to ESTs, coding sequences, genomic sequences and protein sequences through TAIR and NCBI are present. Additional sequence information is presented for each component, including size, EST number, and similarity to the yeast homolog. An image of a subset of the data in MPIMP is shown (Fig. 2).

METHODS

Identification of *Arabidopsis* homologues of the yeast mitochondrial protein import apparatus

All bioinformatic programs were used with the default settings unless specified otherwise. Sequence information for the yeast import components was obtained from the *Saccharomyces* Genome Database. The yeast gene and protein sequences were used to search GenBank and TIGR *Arabidopsis* sequence databases for homologues by BLASTN, BLASTP and TBLASTN alignment (5). Identification of *Arabidopsis* ESTs was achieved by searching the TIGR *Arabidopsis* Gene Index with BLASTN. Genomic clones were identified by searching GenBank with BLASTN. Protein sequences were deduced from nucleic acid sequences

*To whom correspondence should be addressed. Tel: +61 893801749; Fax: +61 893801148; Email: seamus@cyllene.uwa.edu.au

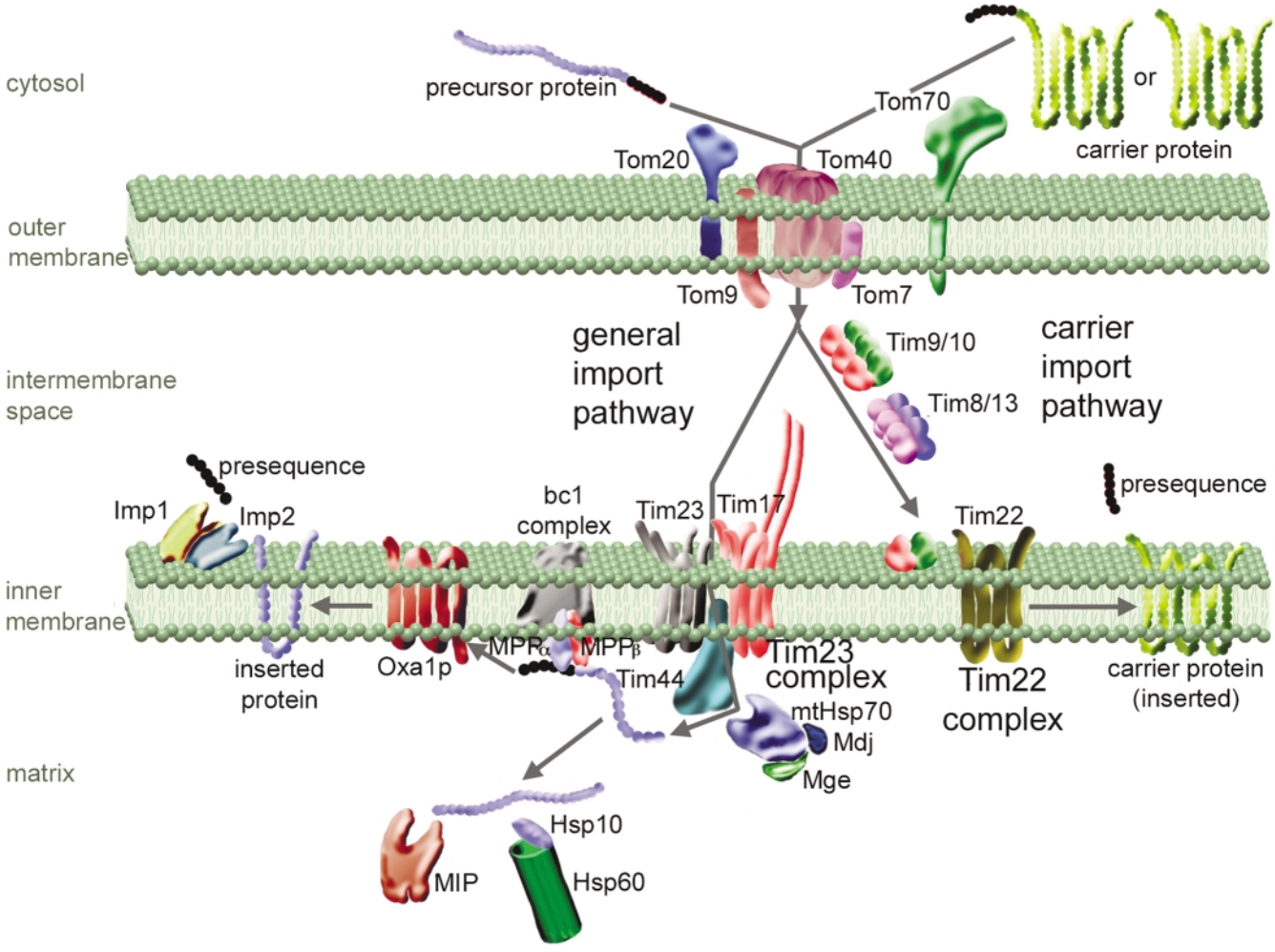


Figure 1. A homology based model of the protein import pathways of plant mitochondria. Precursor proteins in the cytosol can be divided into two distinct classes depending on whether they follow the general import pathway or carrier import pathway. Precursors proteins following the general import pathway usually contain a cleavable presequence and bind the Tom20 receptor on the outer membrane. They are translocated across the outer and inner mitochondrial membranes via Tom40 to Tim17, Tim23 and pulled into mitochondria via the action of mitochondrial HSP70 anchored to Tim44. The presequence is removed by MPP after which the protein may take a number of routes. It may be folded into a functional unit by the action of mitochondrial HSP60, processed further by MIP, or re-directed across the inner membrane via the Oxa1p translocase and may be processed by Imp1 and/or 2. Proteins belonging to the mitochondrial carrier family may or may not contain a cleavable presequence. They bind Tom70 and are translocated to the Tim22 complex by a number of small intermembrane space proteins. The protein is inserted into the inner membrane and the presequence is removed by an unknown peptidase.

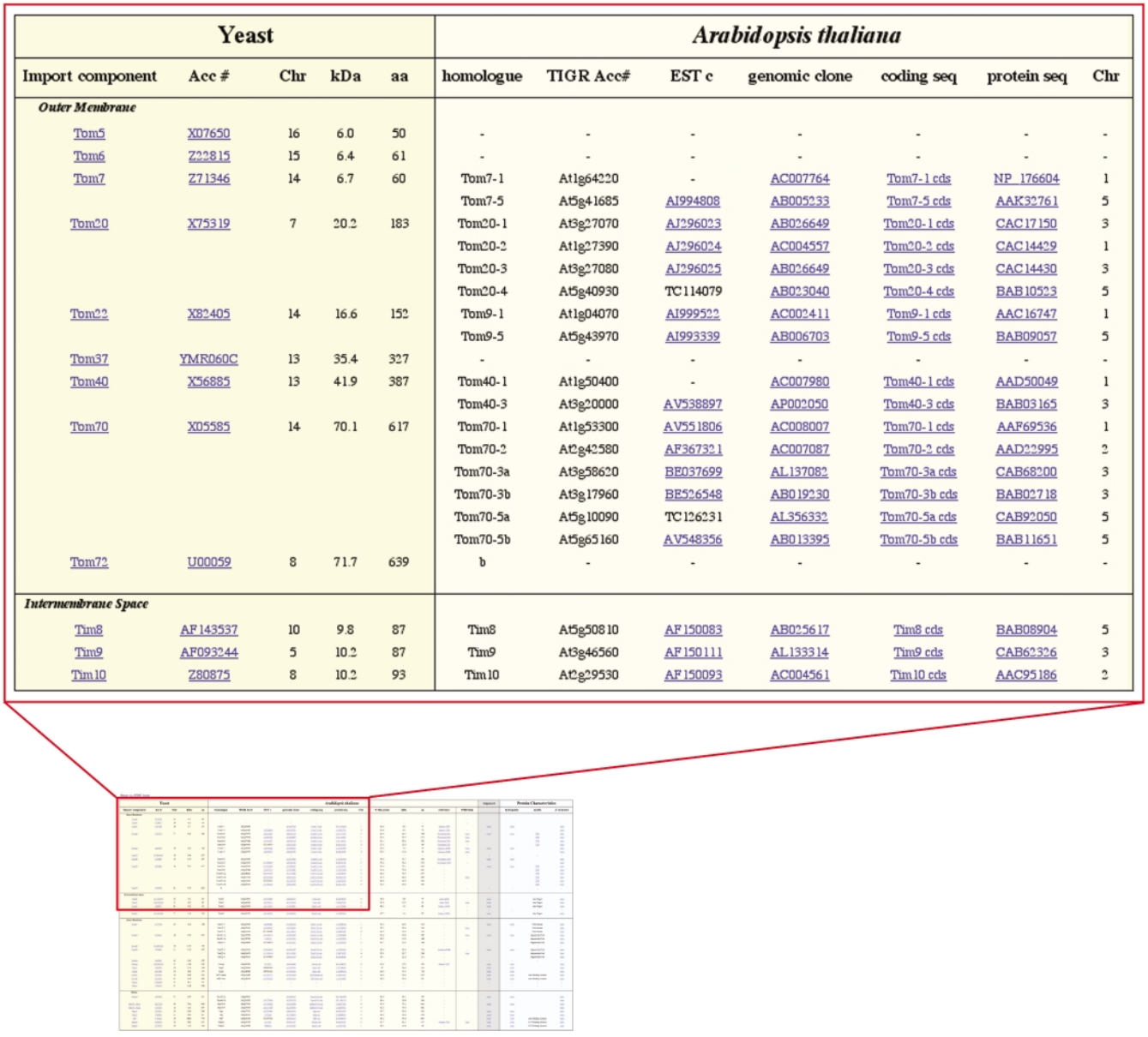
using Translate (University of Wisconsin Genetics Computer Group).

Analysis of *Arabidopsis* mitochondrial import components

Protein alignments were generated using PileUp (University of Wisconsin Genetics Computer Group). Residue shading was done with PrettyBox (University of Wisconsin Genetics Computer Group). Similarity between yeast and *Arabidopsis* proteins was calculated using Gap (University of Wisconsin Genetics Computer Group). Prediction of transmembrane alpha-helices was performed using the DAS transmembrane prediction server (6). Prediction of protein secondary structure was done with Garnier (7). TPR motifs were identified using REP (8).

FUTURE ADDITIONS

In the future, the MPIMP database will be updated with information about the mitochondrial protein import apparatus from other plant species, such as rice. This will allow comparison of the genes present in different species. Additional analysis of the import component genes will be included in the database as it is generated. Links to all research relating to the plant mitochondrial import components will be added as it is published.



Yeast					*Arabidopsis thaliana*						
Import component	Acc #	Chr	kDa	aa	homologue	TIGR Acc#	EST e	genomic clone	coding seq	protein seq	Chr
Outer Membrane											
Tom5	X07650	16	6.0	50	-	-	-	-	-	-	-
Tom6	Z22815	15	6.4	61	-	-	-	-	-	-	-
Tom7	Z71346	14	6.7	60	Tom7-1	At1g64220	-	AC007764	Tom7-1 cds	NP_176604	1
					Tom7-5	At5g41685	AI994808	AB005233	Tom7-5 cds	AAK32761	5
Tom20	X75319	7	20.2	183	Tom20-1	At3g27070	AJ296023	AB026649	Tom20-1 cds	CAC17150	3
					Tom20-2	At1g27390	AJ296024	AC004557	Tom20-2 cds	CAC14429	1
					Tom20-3	At3g27080	AJ296025	AB026649	Tom20-3 cds	CAC14430	3
					Tom20-4	At5g40930	TC114079	AB023040	Tom20-4 cds	BAB10523	5
Tom22	X82405	14	16.6	152	Tom9-1	At1g04070	AI999522	AC002411	Tom9-1 cds	AAC16747	1
					Tom9-5	At5g43970	AI993339	AB006703	Tom9-5 cds	BAB09057	5
Tom37	YMR060C	13	35.4	327	-	-	-	-	-	-	-
Tom40	X56885	13	41.9	387	Tom40-1	At1g50400	-	AC007980	Tom40-1 cds	AAD50049	1
					Tom40-3	At3g20000	AV538897	AP002050	Tom40-3 cds	BAB03165	3
Tom70	X05585	14	70.1	617	Tom70-1	At1g53300	AV551806	AC008007	Tom70-1 cds	AAF69536	1
					Tom70-2	At2g42580	AF367321	AC007087	Tom70-2 cds	AAD22995	2
					Tom70-3a	At3g58620	BE037699	AL137082	Tom70-3a cds	CAB68200	3
					Tom70-3b	At3g17960	BE526548	AB019230	Tom70-3b cds	BAB02718	3
					Tom70-5a	At5g10090	TC126231	AL356332	Tom70-5a cds	CAB92050	5
					Tom70-5b	At5g65160	AV548356	AB013395	Tom70-5b cds	BAB11651	5
Tom72	U00059	8	71.7	639	b	-	-	-	-	-	-
Intermembrane Space											
Tim8	AF143537	10	9.8	87	Tim8	At5g50810	AF150083	AB025617	Tim8 cds	BAB08904	5
Tim9	AF093244	5	10.2	87	Tim9	At3g46560	AF150111	AL133314	Tim9 cds	CAB62326	3
Tim10	Z80875	8	10.2	93	Tim10	At2g29530	AF150093	AC004561	Tim10 cds	AAC95186	2

Figure 2. Image of the MPIMP database, containing functional hypertext links (blue) to information.

ACKNOWLEDGEMENTS

This work was supported by Australian Research Council Grants to J.W. and Australian Postgraduate Award to R.L. and M.W.M.

REFERENCES

1. Emanuelsson,O., Nielsen,H., Brunak,S. and von Heijne,G. (2000) Predicting subcellular localisation of proteins based on their N-terminal amino acid sequence. *J. Mol. Biol.*, **300**, 1005–1016.
2. Herrmann,J.M. and Neupert,W. (2000) Protein Transport into Mitochondria. *Curr. Opin. Microbiol.*, **3**, 210–214.
3. Pfanner,N. and Geissler,A. (2001) Versatility of the mitochondrial protein import machinery. *Nature Rev. Mol. Cell Biol.*, **2**, 339–349.
4. Gray,M.W., Burger,G. and Lang,B.E. (1999) Mitochondrial evolution. *Science*, **283**, 1476–1481.
5. Altschul,S.F., Madden,T.L., Schäffer,A.A., Zhang,J., Zhang,Z., Miller,W., Lipman,D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**, 3389–3402.
6. Cserzo,M., Wallin,E., Simon,I., von Heijne,G., Elofsson,A. (1997) Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the dense alignment surface method. *Protein Eng.*, **10**, 673–676.
7. Pearson,W.R. and Lipman,D.J. (1988) Improved tools for biological sequence analysis. *Proc. Natl Acad. Sci. USA*, **85**, 2444–2448.
8. Andrade,M.A., Ponting,C., Gibson,T. and Bork,P. (2001) Identification of protein repeats and statistical significance of sequence comparisons. *J. Mol. Biol.*, **298**, 521–537.