Molecular Phylogenetics and Evolution 53 (2009) 565–570

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Expanded phylogenies of canonical and non-canonical types of methionine adenosyltransferase reveal a complex history of these gene families in eukaryotes

Ryoma Kamikawa [a,*], Gabino F. Sanchez-Perez [b,1], Yoshihiko Sako [c], Andrew J. Roger [b], Yuji Inagaki [a]

[a] Center of Computational Sciences and Institute of Biological Sciences, University of Tsukuba, Tsukuba, Ibaraki 305-8577, Japan
[b] Centre for Comparative Genomics and Evolutionary Bioinformatics, Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax, Nova Scotia, Canada
[c] Laboratory of Marine Microbiology, Graduate School of Agriculture, Kyoto University, Kyoto 606-8502, Japan

ARTICLE INFO

Article history:
Received 8 April 2009
Revised 26 June 2009
Accepted 30 June 2009
Available online 3 July 2009

Keywords:
Methionine adenosyltransferase
Lateral gene transfer
Diatoms
Aureococcus
Haptophytes
Cryptomonads

ABSTRACT

Most eukaryotes possess the highly-conserved enzyme methionine adenosyltransferase (MAT) that produces S-adenosyl-L-methionine, a molecule essential to a variety of cellular processes. However, a recent study revealed that genomes of a very few eukaryote lineages encode a highly divergent type of MAT (called MATX), instead of the canonical MAT enzyme. Since MATX-containing eukaryotes are phylogenetically interspersed with MAT-containing organisms, it is likely that the MATX gene was spread into the MAT-containing groups via multiple eukaryote-to-eukaryote lateral gene transfer events. Here, we further investigate the evolutionary history of these gene families by vastly increasing the sampling of species containing MAT (22 new taxa) and MATX (8 new taxa). Our expanded analyses reveal the first example of lateral transfer of a MAT gene between the pelagophycean alga *Aureococcus anophagefferens* and a cryptomonad. The increased MATX sampling also provided new insights into the evolution of MATX. Specifically, our MATX phylogeny robustly grouped the haptophyte homologues with the *Aureococcus* homologue to the exclusion of the diatom homologues, suggesting a transfer of the MATX gene between haptophytes and pelagophytes. Various scenarios of MAT and MATX gene family evolution in diatoms are re-evaluated in light of the new data.

1. Introduction

Methionine adenosyltransferase (MAT) is an enzyme that synthesizes S-adenosyl-L-methionine (SAM), a molecule that serves as the major methyl-group donor to DNA, RNA, phospholipids, and proteins in cells (Cantoni, 1975; Chiang et al., 1996). MAT is a highly-conserved protein that is encoded in the genomes of most eukaryotes, eubacteria, and archaebacteria that has been well studied at the primary, secondary, and tertiary structural levels due to its indispensable function for cell viability (Takusagawa et al., 1996; Gonzalez et al., 2000). Whereas MAT homologues in eukaryotes and eubacteria are highly similar to each other, archaebacterial MAT enzymes are only distantly related to the former two groups (Graham et al., 2000).

Until recently, it was assumed that MAT genes evolved by vertical descent from the last common ancestor of all cells (Sanchez-Perez et al., 2004). However, large-scale expressed sequence tag (EST) analyses from a broad spectrum of eukaryotes challenged this view of MAT evolution. Prymnesiophycean haptophytes, dinoflagellates, euglenids, and the diatom *Thalassiosira pseudonana* appeared to possess a divergent MAT-related protein called MATX, and no transcript encoding the canonical MAT enzyme was found in the EST data of any of these MATX-bearing lineages except *Thalassiosira* (Sanchez-Perez et al., 2008). In phylogenetic trees, all MATX homologues formed a strongly-supported clade to the exclusion of canonical eukaryote MAT enzymes. A recent study demonstrated that the MATX gene from the dinoflagellate *Crypthecodinium cohnii* complemented a mutation in the *Saccharomyces cerevisiae* MAT gene (Ho et al., 2007), indicating that MATX proteins indeed must have MAT enzymatic activity. Although MAT and MATX genes may carry out the same primary function, their evolutionary tempos and modes are clearly distinct (Sanchez-Perez et al., 2008).

Whereas the vast majority of eukaryotes possess a MAT gene, MATX-containing eukaryotes are patchily distributed in the eukaryote tree (Sanchez-Perez et al., 2008). To rationalize this patchy distribution, Sanchez-Perez et al. (2008) proposed a scenario that makes three claims: (i) the MAT gene is ancestral to all extant eukaryotes, and a MATX gene evolved later in one particular eukaryote lineage, (ii) MATX genes were then spread into distantly-related eukaryotic groups/species via lateral gene transfer

* Corresponding author. Fax: +81 29 853 6406.
E-mail address: ryoma@ccs.tsukuba.ac.jp (R. Kamikawa).
[1] Present address: Theoretical Biology and Bioinformatics, Department of Biology, Science Faculty, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands.

1055-7903/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2009.06.016

events perhaps in the context of secondary plastid symbiosis, and (iii) in the recipient cells, the exogenous MATX took over the original MAT function and the latter enzyme was lost (except in *Thalassiosira* which has both MAT and MATX genes). Amongst other claims, Sanchez-Perez et al. (2008) proposed specifically that the MATX gene was transferred between a haptophyte and *Thalassiosira* based on phylogenetic analyses of MATX that grouped the *Thalassiosira* and haptophyte homologues to the exclusion of other sequences.

In this study, we greatly expand the phylogenetic representation of MAT- and MATX-containing lineages; we report the sequences of 22 novel MAT genes and seven novel MATX genes from various eukaryotes. Consistent with Sanchez-Perez et al. (2008), the overall MAT phylogeny we recover suggests that vertical inheritance is the predominant mode of eukaryotic MAT gene evolution. Nevertheless, we report a probable example of lateral transfer of MAT gene between a pelagophycean stramenopile alga *Aureococcus anophagefferens* and a cryptomonad. Our MATX phylogeny that includes newly identified homologues from diatoms and *Aureococcus* indicates that the haptophyte homologues shared the most recent evolutionary ancestry with a MATX from *Aureococcus*, but not with the *Thalassiosira* homologue (or other diatom homologues) as postulated in Sanchez-Perez et al. (2008). Finally, we propose a model that explains the MAT/MATX gene distribution in diatoms.

2. Materials and methods

2.1. Cryptomonad MAT genes

DNA samples of *Cryptomonas paramecium* and *Rhodomonas* sp. were kindly provided by John M. Archibald (Dalhousie University, Halifax, Canada). In order to amplify the cryptomonad MAT gene, we used a forward primer directed at the consensus amino acid (aa) sequence GHPDK, MATA3 (5′-GAGTCIGTSAAYGARGGHCA YCCIGAC), and a reverse primer within the eukaryotic consensus aa sequence TYGGWGAH inside a conserved block, MATB3 (5′-CC RTGIGCNCCCAICCNCCRTAIGT). PCR cycling conditions consisted of 40 cycles of 95 °C for 30 s, 55 °C for 60 s and 72 °C for 90 s. The new sequences from *C. paramecium* and *Rhodomonas* sp. were deposited in GenBank under the Accession Nos. FJ429359 and FJ429360, respectively.

2.2. Diatom MAT and MATX genes

Achnanthes kuwaitensis (NIES1349), *Asterionella glacialis* (NIES417), *Cylindrotheca closterium* (NIES1045), *Ditylum brightwellii* (NIES350), *Skeletonema costatum* (NIES17), *Chaetoceros didymus* (NIES586), and *Thalassionema nitzschioides* (NIES534) were purchased from the National Institute for Environmental Study (http://mcc.nies.go.jp/). *Detonula confervacea* (CCMP353) was purchased from Provasoli-Guillard National Center for Culture of Marine Phytoplankton (http://ccmp.bigelow.org/). Total RNA extraction was performed by using the RNeasy Plant Mini Kit (QIAGEN) following the manufacture's instructions. cDNA synthesis was performed as described in Kamikawa et al. (2008, 2009). The 3′- and 5′-ends of mRNAs were amplified by using the 3′ Full RACE Core kit with the poly(dT) primer (TaKaRa) and 5′ RACE system for Rapid Amplification of cDNA Ends version 2.0 (Invitrogen), respectively.

DNA fragments (approximately 500 bp in length) encoding the N-terminal region commonly conserved between MAT and MATX genes were amplified by using the following degenerate primers: one forward primer, MATForward1 (5′-GTNAAYGARGGNCAYCCN-GAYAA); three reverse primers, MATReverse1 (5′-CATNATNCCYT GRTCNCCNGC), MATReverse2 (5′-ATYTTNCGNCCNGTYAANCC), and MATReverse3 (5′-ATYTTNCGNCCNGTYAANCC). MATReverse2 and MATReverse3 were designed to anneal to the same region in both MAT and MATX genes. Firstly, PCR was performed with the first set of primers MATForward1 and MATReverse2/MATReverse3 using the cDNA sample as the template. Subsequently, by using the first PCR mix as the template, the second PCR was performed with the set of primers, MATForward1 and MATReverse1. PCR cycling conditions consisted of 30 cycles of 95 °C for 30 s, 40 °C for 30 s, and 72 °C for 30 s. The experiments described above amplified MAT gene fragments from *Asterionella*, *Anchnanthes*, *Cylindrotheca*, and *Ditylum*, and MATX gene fragments from *Cylindrotheca*, *Chaetoceros*, and *Skeletonema*.

Partial fragments of *Detonula* MAT gene were amplified from the cDNA sample by using a set of degenerate primers, MATDF (5′-GGMT ACGAYGAYCCHGCCAAGG) and MATDR (5′-CYTGNGTCTTDCCR TCVGGACG). From *Asterionella*, *Detonula*, *Ditylum*, and *Thalassionema* cDNA samples, partial MATX gene fragments were amplified by using a set of degenerate primers, MATXDF (5′-GCC ACSAAGGACAA-CATGG) and MATXDR (5′-CRGCAATRTCGGGAGA YTGCTTG). PCR cycling conditions for the PCR using MATDF and MATDR primers and those using MATXDF and MATXDR primers consisted of 30 cycles of 95 °C for 30 s, 55 °C for 30 s and 72 °C for 45 s. PCR products were cloned into pGEMT-Easy vector (Promega). For each product, 9–15 clones were randomly sequenced on both strands.

We determined the 3′- and 5′-end sequences of diatom MAT and MATX cDNAs by RACE experiments using exact-match primers based on the partial sequences amplified by sets of the degenerate primers (see above). These experiments were performed by using the 3′ Full RACE Core kit (TaKaRa Bio Inc.), and the 5′ RACE system for Rapid Amplification of cDNA Ends version 2.0 (Invitrogen) using the manufacturers' instructions. DNA fragments obtained in the RACE analyses were cloned and sequenced as described above. The new diatom sequences that were experimentally determined in this study were deposited in GenBank/EMBL/DDBJ (Accession Nos. AB461922–AB461932).

2.3. Phylogenetic analysis

We manually aligned the MAT and MATX aa sequences determined and/or identified in this study to the alignments described in Sanchez-Perez et al. (2008). Redundant sequences and ambiguously aligned positions from the modified MAT and MATX alignments were excluded before the phylogenetic analyses described below. A final MAT data set including 62 sequences with 350 aa positions, and a final MATX data set including 17 sequences with 422 aa positions, were subjected to maximum-likelihood (ML) and Bayesian analyses described below.

ML phylogenetic analyses were performed by using RAxML 7.0.0 (Stamatakis, 2006) under the WAG model (Whelan and Goldman, 2001) and a correction for among-site rate variation (ASRV). The initial tree search was started from 10 distinct parsimony starting trees and topologies were evaluated by ML with the WAG model with CAT-approximation of ASRV (the WAG + CAT model). The optimal tree among the 10 final trees was then selected the WAG model with ASRV modelled by a discrete gamma distribution (the WAG + Γ model). Bootstrap analyses (100 replicates) were conducted as described above except the final trees were evaluated with the WAG + Γ model.

We also conducted Bayesian analyses on the MAT and MATX data sets using MrBayes v.3.1.1 (Ronquist and Huelsenbeck, 2003) with the WAG + Γ model. One cold and three heated Markov chain Monte Carlo chains with default chain temperatures were run for 1×10^6 generations, sampling log-likelihood values and trees at 100-generation intervals. The first 3×10^4 generations (i.e., 300 trees) and the 1×10^4 generations (i.e., 100 trees) were

discarded as "burn-in" in the MAT and MATX analyses, respectively. Bayesian posterior probabilities (BPP) and branch-lengths were calculated from the remaining trees.

We additionally explored 10 alternative hypotheses for the relationship between the *Aureococcus* MAT2 and the *R. salina* MAT homologues by the approximately unbiased (AU) test (Shimodaira, 2002). The ML topology was modified by enforcing the *Aureococcus* MAT2 homologue to group with other stramenopile homologues, or the *R. salina* homologue to group with the homologues from *C. paramecium* and *Rhodomonas* sp. The AU test was done using Consel v.0.1 (Shimodaira and Hasegawa, 2001).

3. Results and discussion

3.1. MAT phylogeny

We surveyed MAT genes in phylogenetically divergent eukaryotic lineages (Table 1). Seven novel MAT homologues were experimentally isolated, and 15 homologues were identified in public databases from EST or genome projects. In the nuclear-encoded MAT phylogeny presented in Sanchez-Perez et al. (2008), no sign of lateral gene transfer was detected. This observation extends to most of the newly determined/identified MAT homologues in this study, with the notable exception of a homologues in the pelagophycean stramenopile alga *A. anophagefferens* and the cryptomonad *R. salina* (see below for the details).

In the *Aureococcus* genome sequence data (http://www.jgi.doe.gov), we identified two distinct MAT genes. In the phylogeny of MAT (Fig. 1), the first *Aureococcus* homologue (henceforth designated as "MAT1") was nested in a clade of homologues from other stramenopiles, suggesting that this gene has been vertically inherited through stramenopile evolution. However, the second *Aureococcus* homologue, "MAT2", robustly grouped with the homologue from the cryptomonad *R. salina* (Fig. 1). An exclusive relationship between the *Aureococcus* MAT2 homologue and the *R. salina* homologue is incongruent with a fairly distant organismal relationship between stramenopiles (including pelagophycean algae) and cryptomonads. Current data suggests that stramenopiles are most likely close relatives of alveolates (Arisue et al., 2002), whereas cryptomonads are expected to group with haptophytes (e.g., Patron et al., 2007). Using AU tests, we tested alternative topologies constrained to be consistent with a vertical inheritance pattern for the *Aureococcus* MAT2 homologue or that of the *R. salina* homologue; all alternatives were rejected with p-values $< 1 \times 10^{-3}$ (Supplementary Table S1). This incongruity between the organismal and MAT phylogenies (Fig. 1) suggests that a lateral gene transfer of MAT resulted in the robust clade of *Aureococcus* MAT2 and *R. salina* MAT homologues (henceforth referred to as the "*Aureococcus–R. salina*" clade).

Unfortunately, the phylogeny of MAT is insufficiently resolved to clarify the source and directions of transfers that gave rise to the *Aureococcus–R. salina* MAT clade. It seems unlikely that either stramenopiles or cryptomonads are the original source of the exogenous MAT gene in either of these species, since the clade formed by these taxa showed neither an affinity to other cryptomonad homologues (those from *C. paramecium* and *Rhodomonas* sp.), nor the stramenopile grouping (Fig. 1). Two other scenarios involving lateral gene transfer are possible: (i) either *Aureococcus* or *R. salina* acquired a MAT gene from an unknown eukaryote, and the second gene transfer subsequently took place between these two lineages, or (ii) *Aureococcus* and *R. salina* separately received a MAT gene from members of a third lineage of eukaryotes. Curiously, the *Aureococcus–R. salina* clade was nested within the red algal MAT homologues (Fig. 1), although the bootstrap support for this relationship is not strong. It is possible that a red alga (or red algae) is (are) the ultimate donor(s) of the ancestral gene for *Aureococcus* MAT2 and *R. salina* MAT genes. To

Table 1
New MAT and MATX sequences identified in this study.

	Organisms	Genes	Database or Accession Nos.
Amoebozoans	*Entamoeba histolytica*	MAT	XM_647762, 001913755, 001913609
	Hartmanniella vermiformis	MAT	TBestDB[c]
Cercozoans	*Cercomonas langicauda*	MAT	PEPDB
Cryptomonads	*Cryptomonas paramecium*	MAT[a]	FJ429359
	Rhodomonas sp.	MAT[a]	FJ429360
	Rhodomonas salina	MAT[a]	PEPDB
Excavates	*Andalucia godoyi*	MAT	PEPDB
	Cryptobia salmositica	MAT	AY603961
	Histiona aroides	MAT	TBestDB[c]
	Jakoba libera	MAT	TBestDB[c]
	Malawimonas californiana	MAT	TBestDB[c]
	Tritrichomonas foetus	MAT	CX155231
Rhodophyte	*Cyanidioschyzon merolae*	MAT	AP006496
Stramenopiles			
Pennate diatoms (Pennales)	*Achnanthes kuwaitensis*	MAT[a]	AB461923
	Asterionella glacialis	MAT[a]	AB461922
		MATX[a]	AB461927
	Cylindrotheca closterium	MAT[a]	AB461924
		MATX[a]	AB461928
	Fragilariopsis cylindrus	MAT	EL737868, CF245585
	Thalassionema nitzschioides	MATX[a]	AB461932
Centric diatoms (Centrales)	*Detonula confervacea*	MAT[a]	AB461926
		MATX[a]	AB461930
	Ditylum brightwellii	MAT[a]	AB461925
		MATX[a]	AB461929
	Skeletonema costatum	MATX[a]	AB461931
Phaeophyceae	*Laminaria digitata*	MAT	AW400747, CN467610
Pelagophyte	*Aureococcus anophagefferens*	MAT[b]	DOE Joint Genome Institute[d]
		MATX	

[a] PCR- or RT PCR-based survey.
[b] Two distinctive genes were identified.
[c] http://tbestdb.bcm.umontreal.ca/searches/login.php.
[d] http://www.jgi.doe.gov.

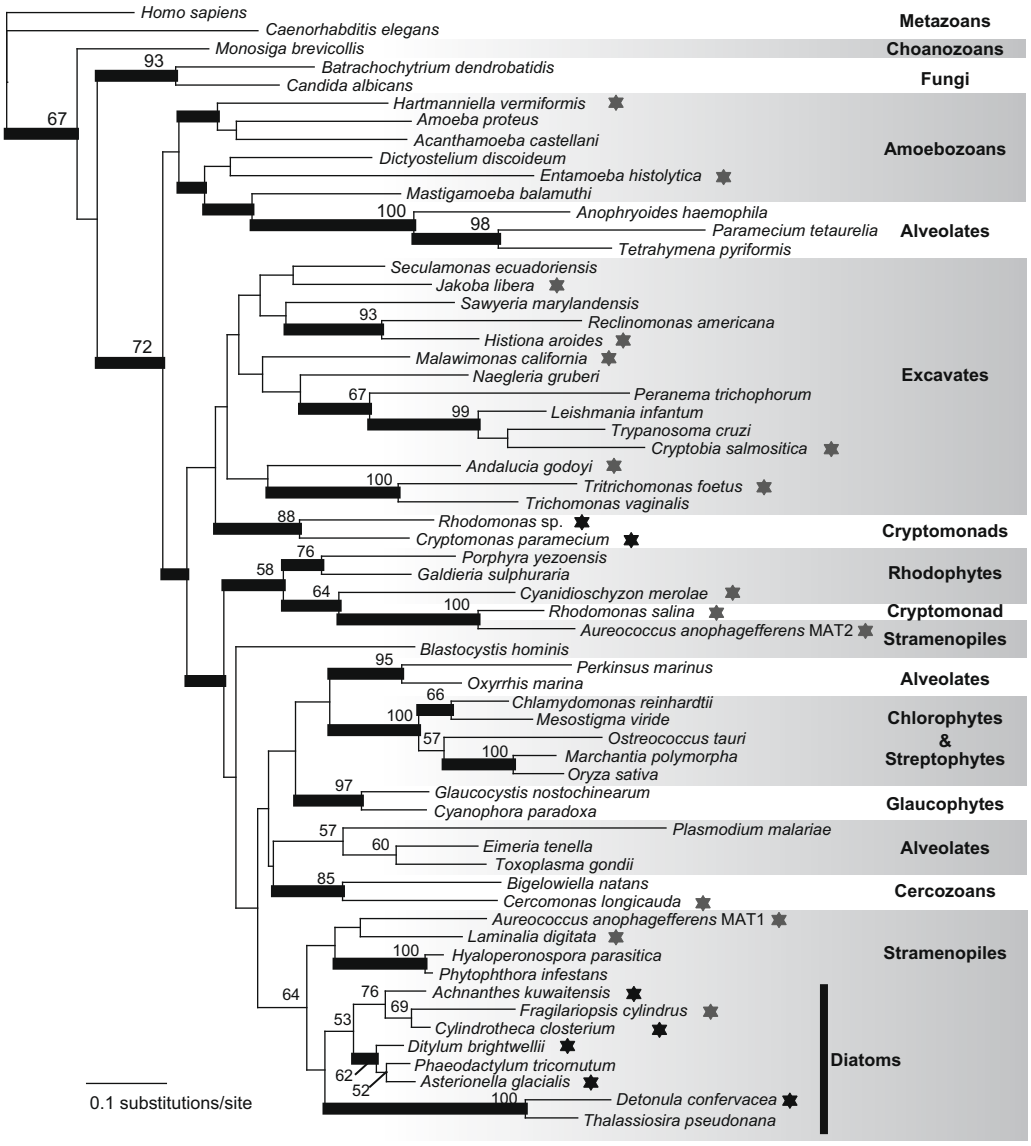

Fig. 1. Phylogenetic relationships among eukaryotic MAT homologues. A maximum-likelihood tree was inferred from the MAT data set (62 sequences; 350 amino acid positions) under the WAG + Γ model. Only bootstrap values for branches greater than 50% are shown. The branches supported by Bayesian posterior probabilities equal to or more than 0.95 are highlighted by thick bars. Black and grey stars represent newly identified sequences by PCR-based and database survey, respectively.

further clarify the origins of these genes, more extensive sampling of MAT genes within the pelagophycae, cryptomonads and red algae will be necessary.

3.2. MATX phylogeny

Prior to this study, MATX homologues were identified in a variety of distantly-related protistan lineages, including euglenids, dinoflagellates, the single diatom species *T. pseudonana*, and haptophytes. Sanchez-Perez et al. (2008) proposed that this distribution is expected if MATX genes were spread into these disparate lineages via lateral transfer events. Importantly, the complete genome of another diatom species *Phaeodactylum tricornutum* was shown to encode only a MAT gene, and not a MATX homologue (Sanchez-Perez et al., 2008). Based on this MAT/MATX gene distribution in these two diatom species and the close relationship observed between MATX homologues from *Thalassiosira* and haptophytes, it was suggested that a MATX gene had been transferred from a haptophyte to *Thalassiosira* (Sanchez-Perez et al., 2008).

We experimentally surveyed MAT and MATX genes in seven diatoms, and MATX transcripts were detected in six of the seven species, *A. glacialis*, *C. closterium*, *D. brightwellii*, *D. confervacea*, *S. costatum*, and *T. nitzschioides* (Table 1). In addition, our database survey identified a MATX gene in *Aureococcus* genome (http://genome.jgi-psf.org/). We estimated a MATX phylogeny including the seven newly determined/identified homologues (Fig. 2A), and found that the homologues from diatoms, haptophytes, and *Aureococcus* all grouped together, to the exclusion of MATX sequences from euglenids and dinoflagellates (Fig. 2A). This grouping was robust (BP = 96%, BPP > 0.95; Fig. 2A), and is supported by a characteristic insertion in the alignment shaded in Fig. 2B. In the diatom–haptophyte–*Aureococcus* clade, the haptophyte and *Aureococcus* homologues showed a specific affinity with moderate bootstrap support (BP = 79%). Note that the overall MATX tree topology and support values were not significantly changed in an analysis that used MAT homologues as the outgroup (data not shown). In light of the specific affinity between the haptophyte and *Aureococcus* homologues, the putative haptophyte-to-*Thalassiosira* gene transfer postulated in Sanchez-Perez et al. (2008) needs to be

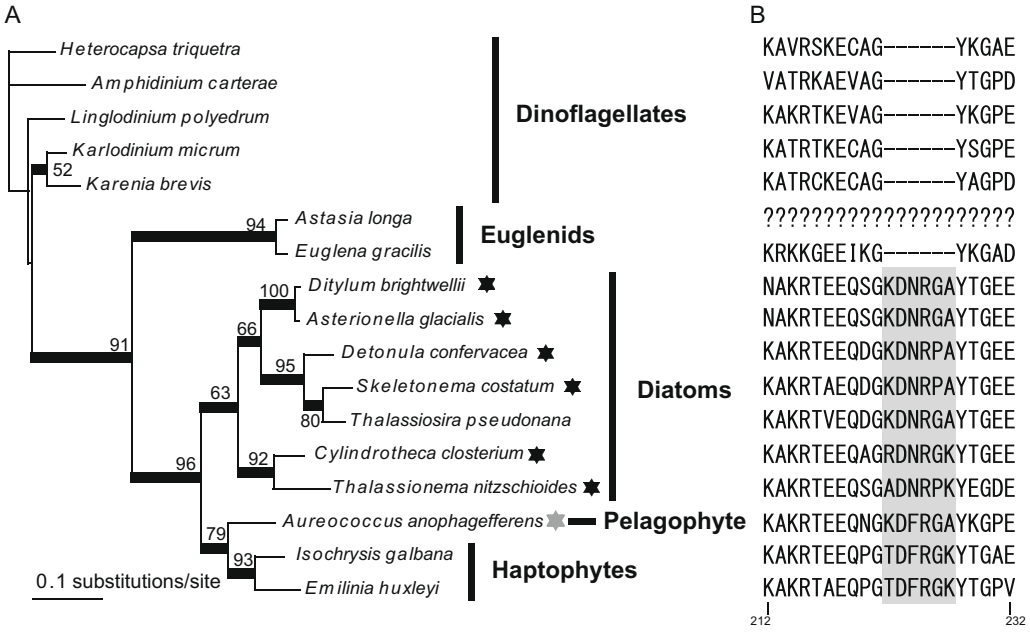

Fig. 2. Phylogenetic relationships among eukaryotic MATX homologues. (A) The maximum-likelihood tree. The MATX data set (17 sequences; 422 amino acid positions) was analyzed under the WAG + Γ model. Only bootstrap greater than 50% are shown. The branches supported by Bayesian posterior probabilities equal to or more than 0.95 are highlighted by thick bars. Details for stars are described in Fig. 1. (B) An amino acid insertion exclusively shared between the diatom, *Aureococcus*, and haptophyte homologues. Numbers below the amino acid alignment correspond to the residue numbers of the *Thalassiosira pseudonana* MATX homologue.

reconsidered. The *Aureococcus*–haptophyte clade can be explained by: (i) a scenario invoking the initial MATX gene transfer from an unknown eukaryote to either *Aureococcus* or a common ancestor of prymnesiophycean haptophytes (e.g., *Emiliania huxleyi*, *Isochrysis galbana*, and *Prymnesium*) followed by a second transfer between the two lineages, or (ii) *Aureococcus* and the ancestral prymnesiophycean haptophyte cell separately acquired a MAT gene from an unknown eukaryote (or two distinct, but closely related eukaryotes). Unfortunately, deeper insights into the evolutionary history of the MATX gene in diatoms, haptophytes, and *Aureococcus* is precluded by the lack of phylogenetic resolution in MATX phylogeny (Fig. 2A). Again, future surveys of MAT/MATX genes from phylogenetically diverse eukaryotes may help to solidify groupings in the trees of these proteins.

3.3. MAT/MATX gene evolution in diatoms

Our MAT/MATX survey revealed that MATX genes are distributed in two large subgroups in diatoms, Pennales and Centrales (Fig. 2A) — *Cylindrotheca*, *Asterionella*, and *Thalassionema* belong to Pennales, while *Ditylum*, *Detonula*, *Skeletonema*, and *Thalassiosira* to Centrales. The MATX distribution suggests that the ancestral diatom cells possessed MATX homologues with the unique insertion (Fig. 2B). If this is correct, then *Phaeodactylum*, whose genome encodes only a MAT homologue, must have secondarily lost its MATX copy. Similarly, since no MATX transcripts have so far been detected in *Fragilariopsis* or *Achnanthes*, these species may also have secondarily lost MATX genes. However, we cannot exclude the possibility that MATX genes are encoded in their genomes but could not be amplified from the cDNAs of these organisms.

In addition to a MATX gene, the ancestral diatom may have possessed a MAT gene, since MAT genes were found in the members of Pennales (e.g., *Asterionella*) and those of Centrales (e.g., *Thalassiosira*) (Fig. 1). Therefore, another intriguing aspect of MAT/MATX gene evolution within diatoms is the possible secondary losses of the MAT gene. *Thalassionema* and *Skeletonema*, where only MATX transcripts have been detected so far, might be the lineages that have secondarily lost MAT genes. To confirm the absence of MAT genes in these two species it will be necessary to fully characterize their genomes.

Prior to this study, *Thalassiosira* was the only known organism to possess genes encoding both MAT and MATX in its nuclear genome (Sanchez-Perez et al., 2008). Our MAT/MATX gene survey extends this "dual-MAT" status to an additional four diatom species (*Cylindrotheca*, *Ditylum*, *Asterionella*, and *Detonula*) and the pelagophycean alga *Aureococcus*. In *Ditylum*, *Asterionella*, and *Aureococcus*, neither the MAT nor the MATX homologues have apparent plastid or mitochondrial targeting peptides at their N-termini (see Supplementary Fig. S1), and thus both of these enzymes may function in the cytosol of these organisms. We postulate that these dual-MAT cells express MAT or MATX in different environmental conditions or in different life cycle stages, as observed for the multiple MAT paralogs studied in yeast and land plants (Thomas and Surdin-Kerjan, 1991; Breusegem et al., 1994; Lee et al., 1997).

The *Thalassiosira* MAT homologue bears an N-terminal extension that is absent from other MAT homologues, except in *Detonula* where a partial N-terminal extension sequence was determined (see Supplementary Fig. S1). The N-terminal extension may be a characteristic to the MAT homologues in *Thalassiosira* and closely related diatoms (e.g., *Detonula*) and was suggested to represent a potential plastid targeting signal (Sanchez-Perez et al., 2008). If the MAT homologues were truly targeted into plastids in *Thalassiosira*, then two enzymes with the same enzymatic activity — MAT and MATX — would function in different sub-cellular compartments (i.e., plastids and cytosol, respectively), avoiding a competition for substrates and the possible negative effects of hetero-oligomeric forms of the enzyme. To pursue this possibility, two uncertainties need to be clarified. Firstly, the N-terminal extension of *Detonula* MAT homologue needs to be completely sequenced as our 5′ RACE experiment failed to amplify the complete 5′-end (see Supplementary Fig. S1). Secondly, the exact localizations of the MAT and MATX homologues in *Thalassiosira* (and *Detonula*) need to be experimentally confirmed.

Acknowledgments

R.K. was a research fellow supported by the Japan Society for the Promotion of Sciences (JSPS) for Young Scientists (Nos. 1803336 and 210528). G.F.S.P. is supported by a Ministerio de Educacion y Ciencia (MEC)/Fulbright postdoctoral fellowship and the Canadian Institutes for Health Research Grant MOP-62809 awarded to A.J.R. Y.I. is supported by a grant from the JSPS (No. 201242).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2009.06.016.

References

Arisue, N., Hashimoto, T., Yoshikawa, H., Nakamura, Y., Nakamura, G., Nakamura, F., Yano, T.A., Hasegawa, M., 2002. Phylogenetic position of *Blastocystis hominis* and of stramenopiles inferred from multiple molecular sequence data. J. Eukaryot. Microbiol. 49, 42–53.

Breusegem, F.V., Dekeyser, R., Gielen, J., Montagu, M.V., Caplan, A., 1994. Characterization of a S-adenosylmethionine synthetase gene in rice. Plant Physiol. 105, 1463–1464.

Cantoni, G.L., 1975. Biological methylation: selected aspects. Annu. Rev. Biochem. 44, 435–451.

Chiang, P.K., Gordon, R.K., Tal, J., Zeng, G.C., Doctor, B.P., Pardhasaradhi, K., McCann, P.P., 1996. S-adenosylmethionine and methylation. FASEB J. 10, 471–480.

Gonzalez, B., Pajares, M.A., Hermoso, J.A., Alvarez, L., Garrido, F., Sufrin, J.R., Sanz-Aparicio, J., 2000. The crystal structure of tetrameric methionine adenosyltransferase from rat liver reveals the methionine-binding site. J. Mol. Biol. 300, 363–375.

Graham, D.E., Bock, C.L., Schalk-Hihi, C., Lu, Z.J., Markham, G.D., 2000. Identification of a highly diverged class of S-adenosylmethionine synthetases in the archaea. J. Biol. Chem. 275, 4055–4059.

Ho, P., Kong, K.F., Chan, Y.H., Tsang, J.S., Wong, J.T., 2007. An unusual S-adenosylmethionine synthetase gene from dinoflagellate is methylated. BMC Mol. Biol. 8, 87.

Kamikawa, R., Inagaki, Y., Sako, Y., 2008. Direct phylogenetic evidence for lateral transfer of elongation factor-like gene. Proc. Natl. Acad. Sci. USA 105, 6965–6969.

Kamikawa, R., Nishimura, H., Sako, Y., 2009. Analysis of the mitochondrial genome, transcripts, and electron transport activity in *Alexandrium catenella* (Gonyaulacales, Dinophyceae). Phycol. Res. 57, 1–11.

Lee, J.H., Chae, H.S., Lee, J.H., Hwang, B., Hahn, K.W., Kang, B.G., Kim, W.T., 1997. Structure and expression of two cDNAs encoding S-adenosyl-L-methionine synthetase of rice (*Oryza sativa* L.). Biochim. Biophys. Acta 1354, 13–18.

Patron, N.J., Inagaki, Y., Keeling, P.J., 2007. Multiple gene phylogenies support the monophyly of cryptomonad and haptophyte host lineages. Curr. Biol. 17, 887–891.

Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.

Sanchez-Perez, G.F., Bautista, J.M., Pajares, M.A., 2004. Methionine adenosyltransferase as a useful molecular systematics tool revealed by phylogenetic and structural analyses. J. Mol. Biol. 335, 693–706.

Sanchez-Perez, G.F., Hampl, V., Simpson, A.G.B., Roger, A.J., 2008. A new divergent type of eukaryotic methionine adenosyltransferase is present in multiple distantly related secondary algal lineages. J. Eukaryot. Microbiol. 55, 374–381.

Shimodaira, H., Hasegawa, M., 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics 17, 246–1247.

Shimodaira, H., 2002. An approximately unbiased test of phylogenetic tree selection. Syst. Biol. 51, 492–508.

Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688–2690.

Takusagawa, F., Kamitori, S., Misaki, S., Markham, G.D., 1996. Crystal structure of S-adenosylmethionine synthetase. J. Biol. Chem. 271, 136–147.

Thomas, D., Surdin-Kerjan, Y., 1991. The synthesis of the two S-adenosylmethionine synthetases is differently regulated in *Saccharomyces cerevisiae*. Mol. Gen. Genet. 226, 224–232.

Whelan, S., Goldman, N., 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol. Biol. Evol. 18, 691–699.