doi: 10.1111/j.1420-9101.2009.01797.x

Functional replacement of a primary metabolic pathway via multiple independent eukaryote-to-eukaryote gene transfers and selective retention

A. M. NEDELCU, A. J. C. BLAKNEY & K. D. LOGUE

Biology Department, University of New Brunswick, Fredericton, NB, Canada

Keywords:

ammonium transporter;
glutamate synthase;
glutamine synthetase;
lateral gene transfer;
Monosiga;
primary metabolic pathway.

Abstract

Although lateral gene transfer (LGT) is now recognized as a major force in the evolution of prokaryotes, the contribution of LGT to the evolution and diversification of eukaryotes is less understood. Notably, transfers of complete pathways are believed to be less likely between eukaryotes, because the successful transfer of a pathway requires the physical clustering of functionally related genes. Here, we report that in one of the closest unicellular relatives of animals, the choanoflagellate, *Monosiga*, three genes whose products work together in the glutamate synthase cycle are of algal origin. The concerted retention of these three independently acquired genes is best explained as the consequence of a series of adaptive replacement events. More generally, this study argues that (i) eukaryote-to-eukaryote transfers of entire metabolic pathways are possible, (ii) adaptive functional replacements of primary pathways can occur, and (iii) functional replacements involving eukaryotic genes are likely to have also contributed to the evolution of eukaryotes. Lastly, these data underscore the potential contribution of algal genes to the evolution of nonphotosynthetic lineages.

Introduction

Lateral gene transfer (LGT) is currently recognized as a major force in the evolution of prokaryotes (e.g. Boucher *et al.*, 2003). However, with the exception of the massive transfers associated with the establishment of mitochondria and plastids, the contribution of LGT (especially eukaryote-to-eukaryote transfer) to the evolution and diversification of eukaryotic lineages is less understood (Andersson, 2005; Keeling & Palmer, 2008). Laterally acquired genes can be added to the recipient's gene complement (gene additions) or simply replace existing endogenous counterparts (functional replacements). The recruitment of novel genes is thought to allow the recipient to adapt to specialized or new ecological niches (such as anaerobic or sugar-rich environments, soil, etc.) as well as new life-styles (parasitism) (e.g. Andersson *et al.*, 2003; Opperdoes & Michels, 2007; Keeling & Palmer, 2008). On the other hand, as functional replacements can be either adaptive or selectively neutral, their impact on the recipient's adaptive or long-term evolutionary potential is less clear (e.g. Huang & Gogarten, 2008; Keeling & Palmer, 2008).

Although most LGTs involve single genes that can function independently or, if replacing a homologue, become part of a mosaic/chimeric pathway [e.g. the carotenoid biosynthetic pathway in chromist algae; (Frommolt *et al.*, 2008)], complete pathways have also been transferred between prokaryotes as well as from prokaryotes to eukaryotes [e.g. the transfer of a pathway involved in vitamin B_6 biosynthesis, from a prokaryotic source to a nematode; (Craig *et al.*, 2008)]. However, most of the recruited bacterial operons involve novel genes associated with nonessential pathways that have the potential to provide the recipient with new adaptive capabilities; the replacement of complete primary metabolic pathways is known to occur only under special environmental conditions [e.g. the acquisition of the arginine biosynthesis operon in Xantomonadales; (Lima & Menck, 2008)]. Notably, transfers of complete

Correspondence: Aurora M. Nedelcu, Department of Biology, University of New Brunswick, PO Box 4400, Fredericton, NB E3B 5A3, Canada. Tel.: +506 458 7463; fax: +506 453 3583; e-mail: anedelcu@unb.ca

© 2009 THE AUTHORS. *J. EVOL. BIOL.* **22** (2009) 1882–1894
JOURNAL COMPILATION © 2009 EUROPEAN SOCIETY FOR EVOLUTIONARY BIOLOGY

pathways (either additions or functional replacements) are believed to be less likely between eukaryotes, because the successful transfer of a complete pathway requires the physical clustering of functionally related genes (such as in operons), and – in contrast to prokaryotes, such clustering is rather infrequent among eukaryotes (Lawrence & Roth, 1996).

To address the impact of eukaryote-to-eukaryote LGT on the long-term evolution and diversification of eukaryotic lineages, we are investigating the potential contribution of algal genes to the evolution of choanoflagellates, a strictly nonphotosynthetic group that comprises both unicellular and colonial forms (Carr *et al.*, 2008). Sequence data for two unicellular choanoflagellate species, *Monosiga brevicollis* and *Monosiga ovata*, are available (http://genome.jgi-psf.org/Monbr1/; http://amoebidia.bcm.umontreal.ca/pepdb/). The two *Monosiga* species inhabit distinct habitats (marine vs. freshwater), and each belongs to one of the three major clades currently recognized within Choanoflagellata (Carr *et al.*, 2008). Phylogenetic analyses using various types of sequences (single genes or concatenated) as well as a number of molecular traits (i.e. metazoan-specific proteins) support the notion that Choanoflagellata is the sister taxon to animals (Metazoa), and that choanoflagellates, animals and fungi form a well-supported monophyletic group, the Opisthokonta, to the exclusion of all photosynthetic lineages (e.g. Cavalier-Smith & Chao, 2003; Nozaki *et al.*, 2007; Abedin & King, 2008; Carr *et al.*, 2008; King *et al.*, 2008; Ruiz-Trillo *et al.*, 2008).

Recently, we reported the presence of several stress-related genes of algal origin (including two ascorbate peroxidase and two metacaspase genes) in *Monosiga* (Nedelcu *et al.*, 2008). Because at least three of these genes represent additions to the choanoflagellate gene complement and appear to have been acquired early in the evolution of choanoflagellates (i.e. before the divergence of the lineages leading to the extant *M. brevicollis* and *M. ovata*), we suggested that they could have contributed to this group's ability to adapt to new environments (e.g. freshwater) and/or new life-styles (e.g. sessile, colonial). Here, we are addressing the potential contribution of a different class of genes, namely, genes involved in a primary metabolic pathway, and a different evolutionary outcome – i.e. functional gene replacement.

In addition to their ability to capture light, photosynthetic organisms differ from animals in their capacity to acquire and utilize nitrogen. Nitrogen is an indispensable element that is incorporated in many important structural and functional molecules, including amino acids and nucleotides. While algae and plants can synthesize organic nitrogen from inorganic sources (such as nitrite, nitrate and ammonium), animals are dependent on organic nitrogen produced by autotrophs (Inokuchi *et al.*, 2002; Katagiri & Nakamura, 2003). Ammonium plays a central role in the metabolism of nitrogen, in both autotrophs and heterotrophs. The main route of ammonium assimilation (either directly up-taken from the environment, synthesized from nitrates, or metabolically-produced) is the so-called glutamine synthetase/glutamate synthase cycle (Lea *et al.*, 1990) Fig. 1). In the first reaction, catalyzed by glutamine synthetase (GS), glutamine is synthesized via the ATP-dependent condensation of ammonium with glutamate. The second reaction, catalyzed by glutamate synthase (GLTS or GOGAT), ensures the regeneration of glutamate via the transfer of the amido group from glutamine to 2-oxoglutarate, to yield two molecules of glutamate. The products of this central cycle (aka the GS/GOGAT or the glutamate synthase cycle) serve as the starting points for the synthesis of all other nitrogen compounds, via a series of aminotransferase reactions (Fig. 1).

The GS/GOGAT cycle is found across all three domains of life, and thus is believed to be a very old pathway (e.g. Raymond, 2005). However, despite its central role in nitrogen metabolism, including amino acid biosynthesis and the neutralization of excess catabolic ammonium, vertebrates appear to have lost the glutamate synthase component. Indeed, while *GLTS* sequences are present in invertebrates and early chordates such as cephalochordates, no *GLTS* homologues have been found among vertebrates. The loss of GLTS in vertebrates is likely related to the fact that animals, in contrast to plants, can

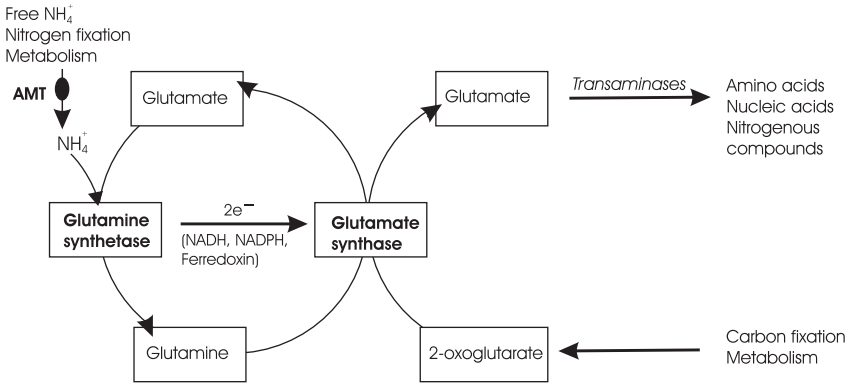


Fig. 1 Schematic representation of the GS/GOGAT pathway (adapted from Raymond, 2005); in bold are the three components discussed in the text.

regenerate the glutamate needed for the first reaction of the GS/GOGAT cycle from dietary nitrogen-containing precursors (Katagiri & Nakamura, 2003).

Interestingly, although nitrogen is a major nutrient for all organisms, its availability is inadequate in many environments (Raymond, 2005). Furthermore, although an important plant nutrient and the preferred nitrogen source (partly because of the lower energetic cost to metabolize ammonium relative to nitrate), ammonium is often limiting for optimal plant growth; in fact, in most soils and in most coastal waters nitrate is more abundant than ammonium (see Inokuchi *et al.*, 2002; Gonzalez-Ballester *et al.*, 2004 for discussion and references). The large nitrogen requirement of plants resulted in their evolving unique strategies to acquire, capture and/or release ammonium, including a number of high-affinity ammonium transporters that belong to the AMT/Rh family (von Wiren *et al.*, 2000; Ludewig *et al.*, 2007). Notably, while present in invertebrates, AMT transporters are missing in vertebrates, which posses a distantly related family of transporters – the so-called Rhesus (Rh) glycoproteins (Huang & Peng, 2005). As in the GLTS case discussed above, the loss of AMTs in vertebrates might be because of their specific metabolic regimes; this is consistent with the loss of AMTs in species that live in nitrogen-rich environments, such as the unicellular parasites, *Plasmodium falciparum* and *Trypanosoma brucei* (Ludewig *et al.*, 2007). Alternatively, the loss may have occurred because the extremely toxic ammonium derived from amino acid catabolism is salvaged and reused by reversing the glutamate dehydrogenase reaction (Huang & Peng, 2005).

We have searched the available genome and EST *Monosiga* databases for sequences encoding components of the GS/GOGAT cycle and ammonium transporters. Here, we report the finding of *GS*, *GLTS* and *AMT* genes of algal origin in *Monosiga*. The independent acquisition of these sequences whose products work together in a central metabolic pathway, the GS/GOGAT cycle, is best explained as the consequence of a series of adaptive replacement events. To our knowledge, this is the first example of functional replacement of a complete primary metabolic pathway involving eukaryote-to-eukaryote gene transfer. Overall, these findings argue that (i) entire pathways can be acquired in the absence of physical clustering of the corresponding genes, (ii) adaptive functional replacement of primary metabolic pathways can occur via multiple independent gene acquisitions and selective retention events, and (iii) functional replacements involving eukaryotic genes are also likely to have contributed to the evolution and diversification of eukaryotes.

Methods

The *M. brevicollis* genome (http://genome.jgi-psf.org/Monbr1/) and the *M. ovata* and *M. brevicollis* EST (http://amoebidia.bcm.umontreal.ca/pepdb/) databases were searched for nitrogen metabolism-related sequences. Homologues from phylogenetically diverse lineages (both prokaryotes and eukaryotes) were retrieved from Uniprot (http://www.uniprot.org/), Interpro (http://www.ebi.ac.uk/interpro/), the Joint Genome Institute (JGI; http://www.jgi.doe.gov), TbEST (http://amoebidia.bcm.umontreal.ca/pepdb/), and several other genome databases (e.g. http://merolae.biol.s.u-tokyo.ac.jp/; http://genomics.msu.edu/galdieria/), using text and BLAST (tblastn and blastp) searches (Altschul *et al.*, 1990). All sequences were checked for the presence of functional domains using SMART, InterProScan and Pfam (http://smart.embl-heidelberg.de/; http://www.sanger.ac.uk/Software/Pfam/; http://www.ebi.ac.uk/InterProScan/), and aligned with Muscle (http://www.drive5.com/muscle/) (Edgar, 2004). Phylogenetic analyses (gaps and unalignable regions excluded) were performed using MrBayes v3.0B4 (mixed amino acid model; 3 500 000 generations; 100 sample frequency; 5000 burnin) and PhyML (http://atgc.lirmm.fr/phyml/; 200 replicates; four-category gamma distribution; proportion of variable sites estimated from the data; best-fit amino acid model indicated by ProtTest) (Huelsenbeck & Ronquist, 2001; Abascal *et al.*, 2005; Guindon *et al.*, 2005). SignalP (http://www.cbs.dtu.dk/services/SignalP/; Emanuelsson *et al.*, 2000) was used to predict signal peptides. Functional enzymatic parameters were retrieved from BRENDA [http://www.brenda-enzymes.info/index.php4; (Schomburg *et al.*, 2004)].

Results

Glutamate synthase

Three main classes of evolutionarily related GLTSs are currently recognized (e.g. Temple *et al.*, 1998; Raymond, 2005; Suzuki & Knaff, 2005; Vanoni *et al.*, 2005). Most eubacteria possess a NADPH-dependent GLTS (EC 1.4.1.13) consisting of two distinct subunits, alpha and beta. On the other hand, eukaryotes (including plants, fungi and invertebrates) express a NADH-dependent GLTS (EC 1.4.1.14) comprising a single long polypeptide derived from the fusion of the bacterial alpha and beta subunits (Andersson & Roger, 2002). However, cyanobacteria and plants (in their chloroplasts) use an additional ferredoxin-dependent GLTS (EC 1.4.7.1), represented by a single polypeptide chain similar in size and sequence to the alpha subunit of the eubacterial NADPH.

Our searches in the *Monosiga* genome and EST databases identified a *GLTS* gene in the *M. brevicollis* genome. The structure of this *GLTS* gene, covering both the alpha and beta subunits, indicates that the encoded protein is a NADH-dependent GLTS. Phylogenetic analyses (both Bayesian and maximum likelihood) including NADH-, NADPH-, and Ferredoxin-dependent GLTSs

from all major lineages confirmed the inclusion of the *M. brevicollis* predicted protein among NADH-dependent GLTSs, but failed to cluster this GLTS with its animal and fungal homologues. Instead, the *M. brevicollis* GLTS branched consistently, and with strong support, with the green algal/land plant group (Fig. 2a). Consistent with a green algal affiliation, the location of several insertions is shared between *M. brevicollis* and green algal/plant GLTS sequences, to the exclusion of animal and fungal homologues (Fig. 2b) (note that there are no insertions that are uniquely shared by *M. brevicollis* and animals and/or fungi; see Fig. S1 for a full alignment). Altogether, these findings argue strongly that the *M. brevicollis GLTS* gene has been acquired laterally, from an algal donor. In the absence of genomic information for *M. ovata*, the timing of this event – i.e. before or after the divergence of the two *Monosiga* lineages, cannot be inferred at this time.

Glutamine synthetase

Glutamine synthetases (E.C. 6.3.1.2) are coded by three distinct gene families, GSI/GSII/GSIII, and the number and type of GS isoenzymes vary greatly among lineages (e.g. Robertson *et al.*, 2001; Raymond, 2005; Robertson & Tartar, 2006). For instance, among eukaryotes, Opisthokonta (fungi and animals) and Plantae (glaucophytes, red and green algae, and land plants) possess GS enzymes of the type II, while lineages within Chromalveolata (i.e. diatoms, oomycetes, haptophytes) have GSIII and/or GSII enzymes.

Our searches in the available *Monosiga* databases identified a *GS* gene in the *M. brevicollis* genome and a partial *GS* sequence in the *M. ovata* EST database; both predicted GS enzymes are of the type II and cluster together (see Fig. S2). However, phylogenetic analyses failed to group the predicted *Monosiga* GS proteins with homologues from their closest relatives, the animals (Fig. 3a and Fig. S2). Instead, the two *Monosiga* GSII sequences consistently grouped within the clade of photosynthetic (and previously photosynthetic) GSII homologues; specifically, at the base of (or within) the chromalveolate clade (in Bayesian analyses; Fig. 3a) or – with less support, at the base of the Plantae/Chromalveolata clade, close to the glaucophyte GSII sequence (in some maximum likelihood analyses; data not shown). Overall, although with variable support, in all analyses, regardless of method, phylogenetic distribution of the taxa included (e.g. with or without prokaryotic sequences) and number of taxa or sites, *Monosiga* GSII sequences grouped with counterparts from photosynthetic lineages, and away from fungal and animal homologues.

In this context, it should be mentioned that most photosynthetic eukaryotes have both cytosol and plastid-targeted GS isoenzymes. However, while in vascular plants, both cytosolic and chloroplast enzymes are of the GSII type [and arose via a recent duplication event in the plant lineage; (Coruzzi *et al.*, 1989)], in diatoms, the cytosolic and plastid-targeted GSs are members of the GSIII and GSII families respectively, with the latter believed to be the result of an endosymbiotic gene transfer from the nuclear genome of the red algal symbiont that gave rise to the diatom plastid (Robertson *et al.*, 2001; Robertson & Tartar, 2006). Interestingly, although GSIII sequences are also known from other chromalveolates (i.e. haptophytes) as well as amoebozoans, the secondarily nonphotosynthetic relatives of diatoms, the oomycetes, possess only GSII sequences; the absence of the ancestral GSIII gene in oomycetes is believed to be the consequence of a functional replacement by the endosymbiont-derived (nuclear-encoded) GSII gene (Robertson & Tartar, 2006). Notably, a GSII sequence was also found in the nonphotosynthetic dinoflagellate, *Oxyrrhis marina*, and its affiliation with a diatom plastid-targeted GSII was interpreted as evidence for the presence of a functional secondary plastid in the evolutionary past of this presently nonphotosynthetic dinoflagellate lineage (Slamovits & Keeling, 2008).

The phylogenetic relationships depicted in Fig. 3a are consistent with those reported by Robertson & Tartar (2006) in (i) failing to cluster the red algal and chromalveolate sequences [likely because of limited taxon sampling and/or the lack of red-algal plastid GSII sequences; (Robertson & Tartar, 2006)] and (ii) placing the chloroplast-targeted green algal GSII outside the clade containing all eukaryotic GSII sequences [possibly reflecting a LGT event from a bacterial source; (Robertson *et al.*, 1999)]. However, in contrast to Robertson & Tartar (2006), some of our analyses recovered the monophyly of chromalveolates (Fig. 3a). Interestingly, in the haptophyte, *Emiliania huxleyi,* we identified two distinct GSII sequences: one that groups with diatom plastid-targeted sequences (this grouping is also supported by the presence of a putative signal peptide, characteristic of sequences targeted to the secondary plastids), and one that affiliates with oomycete cytosolic sequences (Fig. 3a). As *E. huxleyi* possesses cytosolic GS sequences of type III (e.g. Maurin & Le Gal, 1997), the putative cytosolic GSII sequence in *E. huxleyi* represents an addition to the haptophyte gene complement, which suggests a rather complex evolutionary history for GS in the chromalveolate lineage.

Because of this unexpected complexity and the recently recognized issue of multiple LGTs (from both endosymbiotic and food sources) and consecutive endosymbiotic replacement events in lineages with secondary plastids (e.g. Frommolt *et al.*, 2008), which are likely to affect phylogenetic inferences, we also performed analyses restricted to lineages possessing primary plastids. In all analyses (i.e. both Bayesian and maximum likelihood; with or without prokaryotic sequences; with additional taxa), *Monosiga* GSII sequences failed to branch within the Opisthokonta; instead, they consistently branched within the Plantae group (Fig. 3b and Fig. S3).

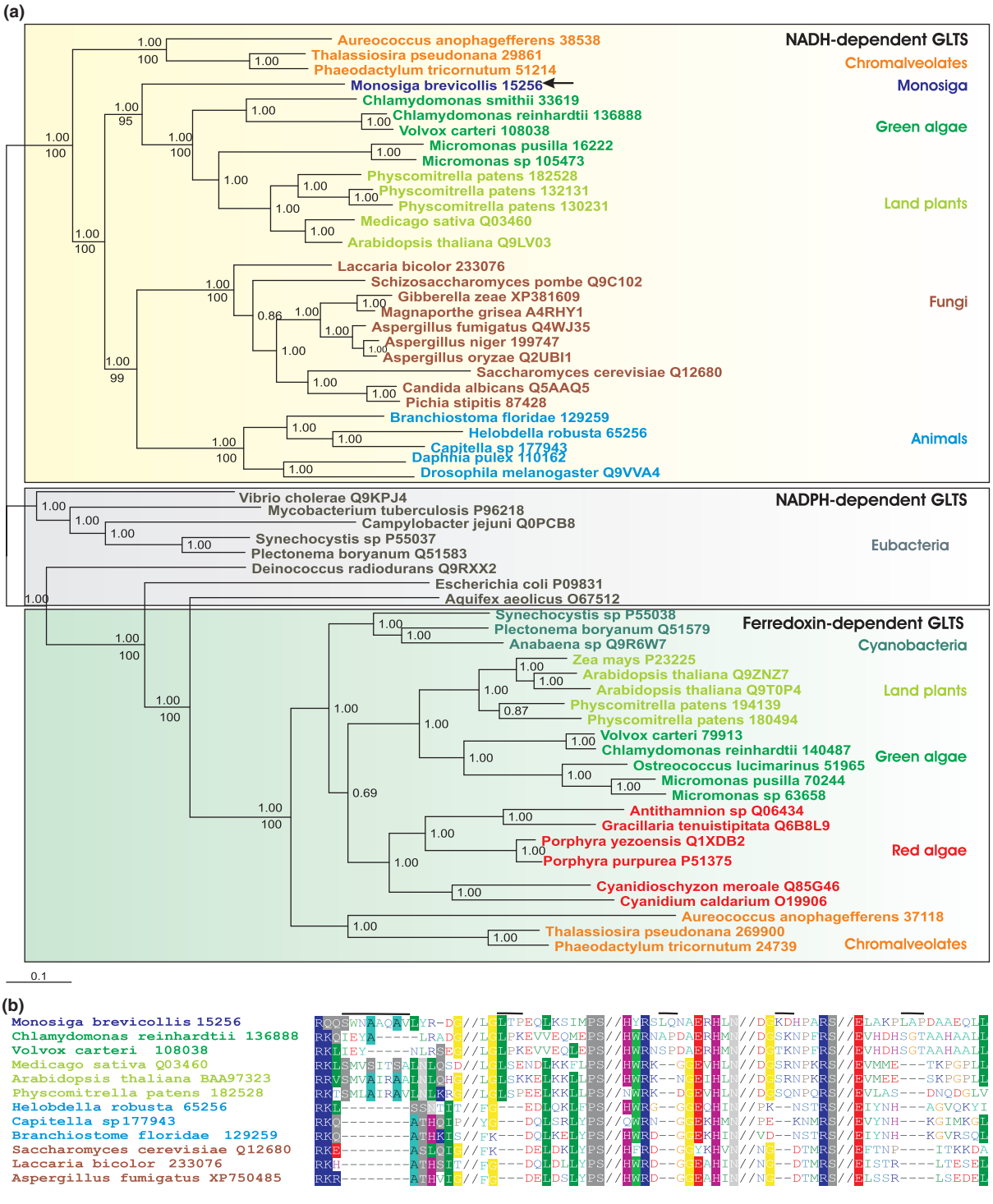


Fig. 2 Glutamate synthases. (a) Bayesian analysis (61 taxa/1208 amino acid sites; numbers at nodes are posterior probabilities) of selected glutamate synthases (corresponding to the alpha subunit) from all three GLTS classes; the apicomplexan GLTS sequences were excluded from the analysis because of their extreme amino acid bias (Andersson & Roger, 2002). Maximum likelihood analyses suggest similar relationships; bootstrap values (200 replicates) for key nodes are indicated below the posterior probability values. Species names are followed by Uniprot IDs – if composed of both letters and numbers, or JGI IDs – if consist of only numbers. (b) Partial alignment showing the location (indicated by horizontal bars) and sequence of several insertions shared by *M. brevicollis* and green algal/plant NADH-dependent GLTSs, to the exclusion of animal and fungal homologues; sequences are colour-coded as in panel a.

Furthermore, the exclusion of the chromalveolate sequences increased the support for the *Monosiga* GSII sequences grouping with homologues from photosynthetic lineages (Fig. 3b).

Overall, the current data indicate that the *Monosiga GSII* genes are of algal origin [note that *GSII* sequences have not been reported in cyanobacteria (Robertson & Tartar, 2006)], but the exact nature of the algal donor cannot be inferred at this time. The presence of algal-related *GSII* genes in both *Monosiga* species, which belong to two early diverged clades within Choanoflagellata (Carr *et al.*, 2008), indicates that the acquisition event took place early in the evolution of this group. This conclusion is consistent with our previous finding of two ascorbate peroxidase and two metacaspase genes that have also been acquired from an algal donor early in the evolutionary history of choanoflagellates (Nedelcu *et al.*, 2008).

Ammonium transporters

Ammonium transporters of the AMT/Rh family have been described in archaea, bacteria, fungi, algae, plants, and invertebrates – but are missing in vertebrates; on the other hand, Rh glycoproteins – abundant in vertebrates, are also found in invertebrates, slime moulds, algae, and some bacteria, but are missing in vascular plants (Huang & Peng, 2005). Nevertheless, the two types of proteins appear to have co-existed for a long time, as both AMT and Rh sequences are found in many lineages, including slime moulds, green algae, oomycetes, and invertebrates (Huang & Peng, 2005).

Our searches in the *M. brevicollis* genome database retrieved up to five *AMT* genes (though only three have reliable gene models) and one gene encoding an Rh-like protein; in addition, we also found an AMT sequence in the *M. ovata* EST database. Interestingly, phylogenetic analyses indicate that the predicted AMT transporters from the two *Monosiga* species belong to two distant clades (Fig. 4a). The three *M. brevicollis* predicted AMTs and some animal homologues cluster together within a large clade of AMT sequences from photosynthetic lineages, including members of the *Arabidopsis* AMT1 family of high-affinity transporters (von Wiren *et al.*, 2000), called here the AMT1 clade. On the other hand, the *M. ovata* AMT sequence branches with homologues from the amoebozoan, *Dictyostelium discoideum*, and the excavate, *Trypanosoma cruzi*, in a distant clade that also contains fungal, oomycetes, green algal and plant sequences, including the well-characterized *Arabidopsis* AMT2 transporter (i.e. the AMT2 clade) (Fig. 4a).

The separation of plant AMTs in two clades, one of which branches with fungal and bacterial homologues has been previously reported (Gonzalez-Ballester *et al.*, 2004; Ludewig *et al.*, 2007). Surprisingly, although the animal AMT sequences are also split into two groups, neither of them clusters with fungal homologues, which form a distant branch in the AMT2 clade (Fig. 4a). The split of animal AMT sequences in two groups as well as the close relationship between one of the animal AMT groups and algal/plant homologues (Fig. 4a) is consistent with previous analyses of AMT and Rh sequences (Huang & Peng, 2005). An affiliation between the AMT sequences from the *M. brevicollis*/animal group and those from photosynthetic lineages (including lineages with secondary plastids, such as the diatoms) is supported by an insertion that is shared by sequences in the AMT1 clade, to the exclusion of all other AMT sequences (Fig. 4b).

Notably, the only lineages that appear to posses the two distinct types of AMT transporters are the green plants (i.e. green algae and land plants), *Monosiga*, and the amoeobozoan, *D. discoideum* (Fig. 4a). The presence of both AMT1- and AMT2-like sequences in these lineages can, in principle, reflect a duplication event that took place before the divergence of the lineages leading to green plants, *Monosiga* and Amoebozoa. However, as Choanoflagellata and Amoebozoa – on the one hand, and Viridiplantae (i.e. green plants) – on the other hand, are representative of the two main eukaryotic lineages, the Bikonts and the Unikonts [believed to have diverged very early in eukaryote evolution; (Stechmann & Cavalier-Smith, 2003)], such a scenario will require independent differential losses in many lineages. These include the loss of AMT1-like sequences in excavates, oomycetes and fungi, and the loss of AMT2-like sequences in diatoms and animals. Furthermore, if this were the case, the AMT1-like sequences from *D. discoideum* and *M. brevicollis* should cluster together – to the exclusion of Plantae and chromalveolate homologues; instead, *D. discoideum* AMT1-like sequences branch away from *M. brevicollis* AMT1 sequences, and affiliate with a specific group of green algal AMT1-like homologues (Fig. 4a).

In this context, LGT events appear to be a more likely explanation for the observed AMT distribution and affiliations. Specifically, the algal-related AMT1 sequences from *D. discoideum* can be interpreted as AMT additions (from a *Chlamydomonas reinhardtii*-like alga) to its existing AMT2 complement. Notably, both *D. discoideum* and *C. reinhardtii* are soil-dwelling species, and LGT events are thought to be facilitated by the donor and recipient inhabiting the same habitat (Andersson, 2005; Andersson *et al.*, 2006). Furthermore, *D. discoideum* is known to have acquired several bacterial genes, many of which are related to living in the soil (Eichinger *et al.*, 2005). Likewise, the presence of AMT1-like sequences in diatoms, while their nonphotosynthetic relatives, the oomycetes, only posses AMT2-like sequences (Fig. 4a), can be interpreted as the result of an endosymbiotic transfer event from the photosynthetic red algal endosymbiont followed by the replacement of the resident AMT2 sequence (note that we have identified AMT1-like sequences in the red algae, *Cyanidioschyzon merolae* (http://merolae.biol.s.u-tokyo.ac.jp/) and *Galdie-*

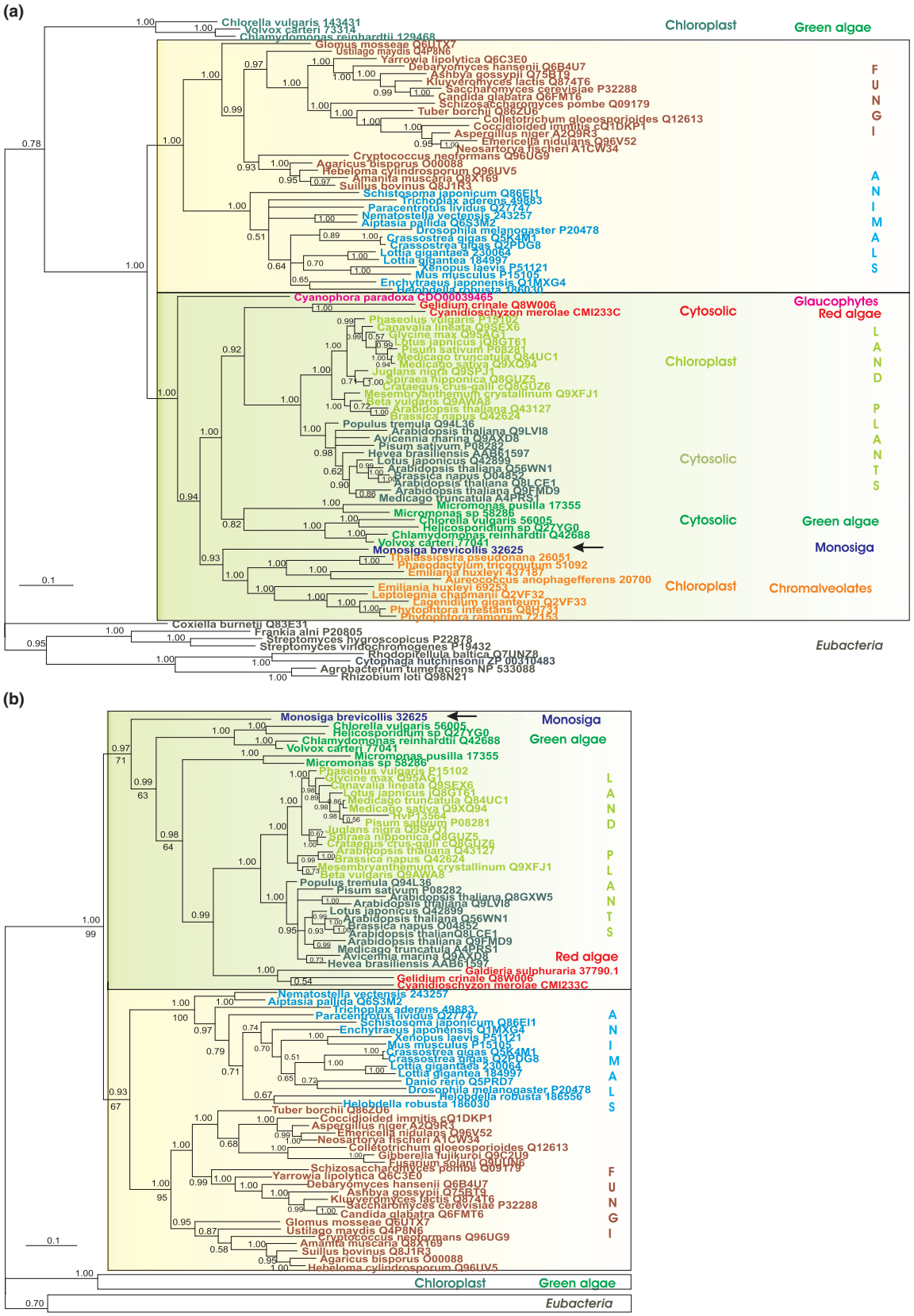

(a)
Chlorella vulgaris 143431
Volvox carteri 73314
Chlamydomonas reinhardtii 129468
Chloroplast
Green algae
Glomus mosseae Q6UTX7
Ustilago maydis Q4P8N6
Yarrowia lipolytica Q6C3E0
Debaryomyces hansenii Q6B4U7
Ashbya gossypii Q75BT9
Kluyveromyces lactis Q874T6
Saccharomyces cerevisiae P32288
Candida glabatra Q6FMT6
Schizosaccharomyces pombe Q09179
Tuber borchii Q86ZU6
Colletotrichum gloeosporioides Q12613
Coccidioides immitis cQ1DKP1
Aspergillus niger A2Q9R3
Emericella nidulans Q96V52
Neosartorya fischeri A1CW34
Cryptococcus neoformans Q96UG9
Agaricus bisporus O00088
Hebeloma cylindrosporum Q96UV5
Amanita muscaria Q8X169
Suillus bovinus Q8J1R3
FUNGI
Schistosoma japonicum Q86EI1
Trichoplax aderens 49883
Paracentrotus lividus Q27747
Nematostella vectensis 243257
Aiptasia pallida Q6S3M2
Drosophila melanogaster P20478
Crassostrea gigas Q5K4M1
Crassostrea gigas Q2PDG8
Lottia gigantaea 230064
Lottia gigantea 184997
Xenopus laevis P51121
Mus musculus P15105
Enchytraeus japonensis Q1MXG4
Helobdella robusta 186030
ANIMALS
Cyanophora paradoxa CDO00039465
Gelidium crinale Q8W006
Cyanidioschyzon merolae CMI233C
Cytosolic
Glaucophytes
Red algae
Phaseolus vulgaris P15102
Canavalia lineata Q9SEX6
Glycine max Q95AG1
Lotus japonicus Q8GT61
Pisum sativum P08281
Medicago truncatula Q84UC1
Medicago sativa Q9XQ94
Juglans nigra Q9SPJ1
Spiraea hipponica Q8GUZ5
Crataegus crus-galli cQ8GUZ6
Mesembryanthemum crystallinum Q9XFJ1
Beta vulgaris Q9AWA8
Arabidopsis thaliana Q43127
Brassica napus Q42624
Populus tremula Q94L36
Arabidopsis thaliana Q9LVI8
Avicennia marina Q9AXD8
Pisum sativum P08282
Hevea brasiliensis AAB61597
Chloroplast
LANDPLANTS
Lotus japonicus Q42899
Arabidopsis thaliana Q56WN1
Brassica napus Q04852
Arabidopsis thaliana Q8LCE1
Arabidopsis thaliana Q9FMD9
Medicago truncatula A4PRS1
Micromonas pusilla 17355
Micromonas sp 58286
Chlorella vulgaris 56005
Helicosporidium sp Q27YG0
Chlamydomonas reinhardtii Q42688
Volvox carteri 77041
Monosiga brevicollis 32625
Cytosolic
Green algae
Monosiga
Thalassiosira pseudonana 26051
Phaeodactylum tricornutum 51092
Emiliania huxleyi 437187
Aureococcus anophagefferens 20700
Emiliania huxleyi 69253
Leptolegnia chapmanii Q2VF32
Lagenidium giganteum Q2VF33
Phytophtora infestans Q8H731
Phytophtora ramorum 72153
Chloroplast
Chromalveolates
0.1
Coxiella burnetii Q83E31
Frankia alni P20805
Streptomyces hygroscopicus P22878
Streptomyces viridochromogenes P19432
Rhodopirellula baltica Q7UNZ8
Cytophaga hutchinsonii ZP 00310483
Agrobacterium tumefaciens NP 533088
Rhizobium loti Q98N21
Eubacteria
(b)
Monosiga brevicollis 32625
Chlorella vulgaris 56005
Helicosporidium sp Q27YG0
Chlamydomonas reinhardtii Q42688
Volvox carteri 77041
Micromonas pusilla 17355
Micromonas sp 58286
Monosiga
Green algae
Phaseolus vulgaris P15102
Glycine max Q95AG1
Canavalia lineata Q9SEX6
Lotus japonicus Q8GT61
Medicago truncatula Q84UC1
Medicago sativa Q9XQ94
HvP13364
Pisum sativum P08281
Juglans nigra Q9SPJ1
Spiraea hipponica Q8GUZ5
Crataegus crus-galli cQ8GUZ6
Arabidopsis thaliana Q43127
Brassica napus Q42624
Mesembryanthemum crystallinum Q9XFJ1
Beta vulgaris Q9AWA8
Populus tremula Q94L36
Pisum sativum P08282
Arabidopsis thaliana Q8GXW5
Arabidopsis thaliana Q9LVI8
Lotus japonicus Q42899
Arabidopsis thaliana Q56WN1
Brassica napus Q04852
Arabidopsis thaliana Q8LCE1
Arabidopsis thaliana Q9FMD9
Medicago truncatula A4PRS1
Avicennia marina Q9AXD8
Hevea brasiliensis AAB61597
LANDPLANTS
Red algae
Galdieria sulphuraria 37790.1
Gelidium crinale Q8W006
Cyanidioschyzon merolae CMI233C
Nematostella vectensis 243257
Aiptasia pallida Q6S3M2
Trichoplax aderens 49883
Paracentrotus lividus Q27747
Schistosoma japonicum Q86EI1
Enchytraeus japonensis Q1MXG4
Xenopus laevis P51121
Mus musculus P15105
Crassostrea gigas Q5K4M1
Crassostrea gigas Q2PDG8
Lottia gigantaea 230064
Lottia gigantea 184997
Danio rerio Q5PRD7
Drosophila melanogaster P20478
Helobdella robusta 186556
Helobdella robusta 186030
ANIMALS
Tuber borchii Q86ZU6
Coccidioides immitis cQ1DKP1
Aspergillus niger A2Q9R3
Emericella nidulans Q96V52
Neosartorya fischeri A1CW34
Colletotrichum gloeosporioides Q12613
Gibberella fujikuroi Q9C2U9
Fusarium solani Q9UTN6
Schizosaccharomyces pombe Q09179
Yarrowia lipolytica Q6C3E0
Debaryomyces hansenii Q6B4U7
Ashbya gossypii Q75BT9
Kluyveromyces lactis Q874T6
Saccharomyces cerevisiae P32288
Candida glabatra Q6FMT6
Glomus mosseae Q6UTX7
Ustilago maydis Q4P8N6
Cryptococcus neoformans Q96UG9
Amanita muscaria Q8X169
Suillus bovinus Q8J1R3
Agaricus bisporus O00088
Hebeloma cylindrosporum Q96UV5
FUNGI
0.1
Chloroplast
Green algae
Eubacteria

Fig. 3 Type II glutamine synthetases. (a) Bayesian analysis (89 taxa/288 amino acid sites; numbers at nodes are posterior probabilities) of selected type II glutamine synthetases from all major lineages for which complete sequences are available (the highly diverged *Trypanosoma*, *Leishmania* and ciliate GSII sequences were excluded from the analysis); eubacterial GSII sequences were used to root the tree (Robertson & Tartar, 2006). The inclusion of the incomplete *M. ovata* GSII EST sequence does not affect the observed relationships, but decreases the number of sites, and thus the support values of some nodes (see Fig. S2). Species names are followed by Uniprot IDs – if composed of both letters and numbers (except for the *Cyanophora paradoxa* sequence – which is a TBestDB ID), or JGI IDs – if consist of only numbers. (b) Bayesian analysis (86 taxa/288 amino acid sites) of selected type II glutamine synthetases excluding lineages that acquired their GSII sequences via endosymbiont gene transfer (i.e. diatoms, oomycetes, haptophytes). The inclusion of the incomplete *M. ovata* GSII EST sequence does not affect the observed relationships (see Fig. S3). Maximum likelihood analyses suggest similar relationships; bootstrap values (200 replicates) for key nodes are indicated below the posterior probability values.

ria sulphuraria (http://genomics.msu.edu/gal dieria/), but have excluded them from the phylogenetic analysis in Fig. 4a because of their diverged sequences).

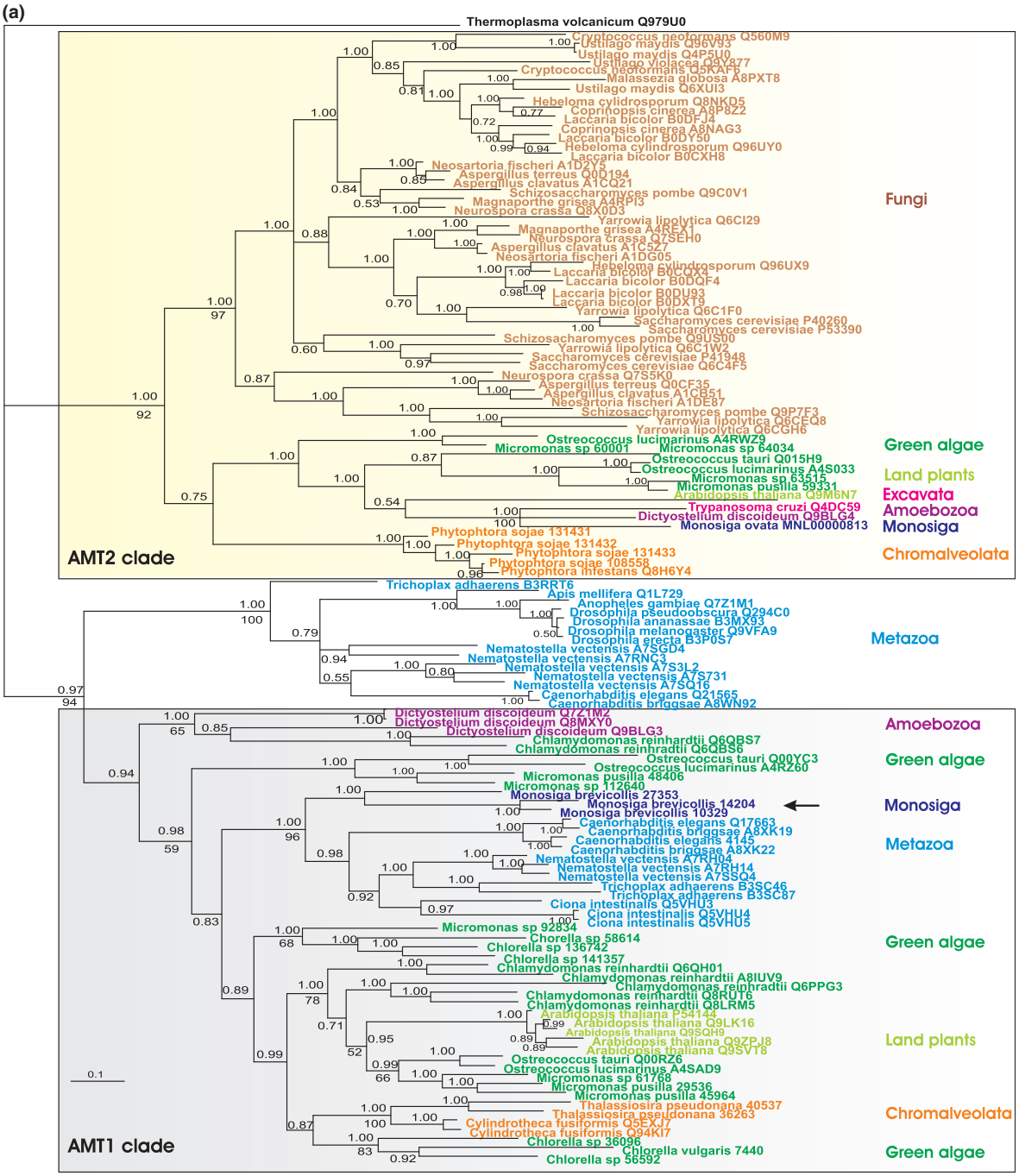
A similar scenario can also be envisioned for *M. brevicollis*: the clustering of the *M. brevicollis* AMT1-like sequences with algal/plant homologues can indicate an algal origin for the former. As AMT2-like sequences are present in Amoebozoa, fungi, and *M. ovata*, it is likely that the last common ancestor of the lineages leading to *M. brevicollis* and *M. ovata* possessed an AMT2 sequence. However, we were not able to identify an AMT2 homologue in the available genomic sequence of *M. brevicollis*, which allows for the possibility that the laterally acquired AMT1 sequence replaced the resident AMT2 in the lineage leading to *M. brevicollis*. Notably, AMT2 sequences are also missing in the *C. reinhardtii* genome, and it was suggested that in this lineage the role of AMT2 transporters was undertaken by some of the AMT1 proteins (Gonzalez-Ballester *et al.*, 2004). On the other hand, although genomic information from *M. ovata* is not available, the fact that AMT1-like sequences are present in both *M. brevicollis* and metazoans indicates that the acquisition event took place before the divergence of the Choanoflagellata and Metazoa, and thus, before the divergence of the two *Monosiga* lineages. Consequently, both AMT1 and AMT2 sequences would have co-existed in the primitive *Monosiga*, and the AMT1 would have replaced AMT2 in the lineage leading to *M. brevicollis*.

Discussion

When discussing LGTs and their evolutionary significance, several issues are especially relevant: the type of gene (novel or homologous to an endogenous gene); its contribution to the host gene complement (gene addition or functional replacement); number of genes transferred (single gene or multiple genes); the type of product encoded (secondary or primary metabolite); functional integration (single protein or part of a pathway); origin (prokaryotic or eukaryotic); effect on the fitness of the host (neutral or adaptive). Many laterally acquired genes are single genes coding for novel proteins that confer an adaptive advantage to the recipient (e.g. Andersson, 2005; Keeling & Palmer, 2008). However, entire pathways can also be acquired, if the functionally related genes are physically clustered on the genome, such as in prokaryotic operons (e.g. Craig *et al.*, 2008). Although it was initially believed that core genes are not likely to be transferred [because the recipient already possesses functional genes whose products are well integrated in complex metabolic pathways; (Lawrence, 1999; Pal *et al.*, 2005; Merkl, 2006)], genes related to primary metabolic pathways are now known to also be acquired laterally (Omelchenko *et al.*, 2003; Lima *et al.*, 2008, 2009). However, functional replacements of entire primary metabolic pathways are thought to occur only under special circumstances (see below).

Generally, several scenarios can be envisioned to account for functional replacement events. For instance: (i) the foreign and native genes initially co-exist, and because of the resulting functional redundancy, one of the two genes is lost; (ii) the two genes initially co-exist, and if not fully equivalent, the acquired gene can be selectively retained; and (iii) the native gene has been lost in response to a prior change in ecological niche or life-style, but when circumstances change again, a foreign homologue is re-acquired. In the first scenario, the fixation of the foreign gene implies stochastic events, which can be facilitated by a 'ratchet mechanism' (Doolittle, 1998). On the other hand, the second scenario entails a selective benefit associated with the replacement, and while such events are theoretically possible, they are more difficult to document (Huang & Gogarten, 2008). Lastly, the third scenario was recently invoked to explain the replacement of the arginine biosynthesis pathway in a group of bacteria (the Xantomonadales), and the re-acquisition of the vitamin B pathway in the parasitic nematode, *Heterodera glycines*; in both cases, it was proposed that the resident pathway was lost following the switch to a parasitic life-style, and was later re-acquired laterally, when the conditions became restrictive again (Craig *et al.*, 2008; Lima & Menck, 2008). Overall, while the events in the former scenario may be considered selectively neutral (Gogarten & Townsend, 2005), the latter two scenarios call for a selective advantage associated with the retention of the acquired gene, and thus, are adaptive.

The data reported here argue that the *M. brevicollis* genes coding for GS, GLTS and one family of ammonium transporters are of algal origin. Furthermore, these



© 2009 THE AUTHORS. *J. EVOL. BIOL.* **22** (2009) 1882–1894
JOURNAL COMPILATION © 2009 EUROPEAN SOCIETY FOR EVOLUTIONARY BIOLOGY

Fig. 4 Ammonium transporters. (a) Bayesian analysis (125 taxa/219 amino acid sites; numbers at nodes are posterior probabilities) of selected AMT sequences; *Thermoplasma volcanicum* AMT sequence was used to root the tree (Huang & Peng, 2005). Maximum likelihood analyses suggest similar relationships; bootstrap values (200 replicates) for key nodes are indicated below the posterior probability values. Species names are followed by Uniprot IDs – if composed of both letters and numbers, or JGI IDs – if consist of only numbers. (b) Partial alignment showing the location and sequence of an insertion shared by members of the AMT1 clade, to the exclusion of all other AMT sequences; sequences are colour-coded as in panel a (the *M. ovata* accession refers to its TBestDB ID).

acquisitions appear to have involved the functional replacement of the resident homologues. As the three genes are not known to be physically clustered in any system, it is most likely that they were acquired independently. Although the independent acquisition and physical integration in the recipient genome of these three genes can be understood as a consequence of stochastic events, it is less probable that chance events were also responsible for the concerted retention of all three genes, at the expense of the native homologues; rather, it is more likely that these three foreign sequences whose products work together in a primary metabolic pathway have been retained because they provided a selective advantage over the endogenous genes. The fact that these genes appear to have been acquired from the same type of donor (i.e. a photosynthetic alga) further supports this suggestion.

Thus, according to the theoretical scenarios described above, the adaptive replacement of the three genes in *M. brevicollis* could be envisioned following scenario 2 (i.e. the selective retention of the acquired homologue) or 3 (i.e. the prior loss of the native gene followed by the re-acquisition of a foreign homologue under new adaptive pressures). The latter scenario requires two major changes (to account for both the loss and the re-acquisition events) in the ecology and/or life-style of the lineage leading to the extant *M. brevicollis*. While this possibility cannot be entirely excluded, there is no indication that such changes have occurred in the evolutionary history of *Monosiga* (all extant *Monosiga* are aquatic, phagotrophic, free-living species). Consequently, the second scenario, which requires that the foreign homologues provided adaptive advantages over the native copies, is more likely.

What could such advantages be? The acquisition of a homologous sequence is usually seen as creating functional redundancy; however, proteins encoded by members of the same family can have distinct biophysico-chemical and/or kinetic properties, even within the same species. For instance, five cytosolic GS isozymes are present in *Arabidopsis thaliana*, and their affinity (i.e. K_M) for ammonium varies from 10 μM (for the high-affinity GSs) to 2450 μM (for the low-affinity GSs). Thus, it is conceivable that the native choanoflagellate and the acquired algal GS and GLTS sequences differed in their enzymatic properties in such a way that that the acquired algal sequences provided a benefit to its recipient in terms of substrate affinity or specific activity. It should be noted that potential interdomain LGT events (between Eubacteria and Archaea) involving GS and GLTS sequences have also been reported, and are believed to possibly represent cases of adaptive orthologous replacements (see Raymond, 2005 for a discussion).

In contrast to Rh glycoproteins, which are thought to act as passive transporters of NH_3, AMTs can specifically and actively (i.e. against a gradient) transport NH_4^+ or cotransport NH_3/H^+ (Ludewig, 2006). AMT/Rh transporters in bacteria, plants and animals are known to differ in their ammonium transport capabilities, and it has been suggested that the functional differences between them are likely to reflect evolutionary adaptations to different ammonium gradients and nitrogen requirements (Ludewig, 2006). In addition, AMTs can also differ vastly in their K_M, even in the same organism: from 0.5 to 170 μM – in plants, and from 7 to 30 μM – in *Chlamydomonas* (Gonzalez-Ballester *et al.*, 2004). Furthermore, the differences between bacterial and plant AMT/Rh transporters are believed to possibly be significant in a competitive soil and provide an evolutionary adaptation to the large nitrogen requirements of plants (Ludewig, 2006).

Interestingly, a rather large number (up to five) of AMT1-like transporters are present in *M. brevicollis*; the existence of a large number of transporters with different but complementary affinities and activities for the substrate in unicellular organisms such as *C. reinhardtii* (which has eight AMT1 transporters, the largest known number of AMT1 in a species) is thought to reflect an adaptive strategy to allow an efficient uptake under changing environmental conditions (Gonzalez-Ballester *et al.*, 2004). As ammonium is rather low in marine environments, it is conceivable that the acquisition of high-affinity ammonium-transporters in the ancestor of *Monosiga* could have provided a selective advantage in this environment, and thus could have been selectively retained. Ammonium is also a constant by-product of amino acid catabolism and de-amination reactions, and at high concentration becomes toxic. Unicellular organisms that lack specialized means to accumulate toxic metabolic products (such as the large vacuole in plant cells) need to have very efficient mechanisms to excrete ammonium at high and toxic concentrations, and also re-uptake the passively lost ammonium, when its extracellular concentration becomes low (Gonzalez-Ballester *et al.*, 2004). In this context, it should be noted that in contrast to many unicellular species, adult *Monosiga* are sessile, and thus, in the absence of motility (which is also

true for land plants), these needs might be greatly intensified.

Although at this time we could only speculate on the adaptive benefit(s) (past or present) of these replacements, it is noteworthy that in plants, in addition to their role in nitrogen assimilation, the enzymes involved in nitrogen metabolism are also thought to play an important role in tolerance against water deficiency and possibly salt stress conditions [e.g. (Ramanjulu & Sudhakar, 1997; Debouba *et al.*, 2006)]. For instance, differences in drought tolerance between two mulberry genotypes were correlated, at least in part, with the ability to maintain greater levels of amino acid pools coupled with a more pronounced re-assimilation of toxic ammonia (Ramanjulu & Sudhakar, 1997). Notably, drought and salinity stress also induce oxidative stress, and in *Monosiga* several stress-related genes, including two ascorbate peroxidases, involved in coping with oxidative stress, are of algal origin as well (Nedelcu *et al.*, 2008). Remarkably, in nonphotosynthetic dynoflagellates, GSII and ascorbate peroxidase genes are also among the sequences thought to have been acquired from their algal endosymbiont and retained after the loss of photosynthesis (Sanchez-Puerta *et al.*, 2007; Slamovits & Keeling, 2008), suggesting that these laterally acquired algal sequences can be generally adaptive in nonphotosynthetic lineages.

Overall, the data presented here indicate that three *Monosiga* genes coding for proteins centrally involved in the metabolism of nitrogen are of algal origin and have been acquired early in the evolution of the choanoflagellates. The concerted retention of these independently acquired genes whose products work together in the GS/GOGAT pathway suggests that the functional replacement of the endogenous homologues was adaptive. More generally, this study argues that (i) eukaryote-to-eukaryote transfers of entire metabolic pathways are possible in the absence of the physical clustering of the corresponding genes, (ii) adaptive functional replacements of primary metabolic pathways can occur via multiple independent gene transfers and the selective retention of the acquired sequences, and (iii) functional replacements involving eukaryotic genes could have contributed to the evolution and diversification of eukaryotes. Furthermore, by adding to our previous finding of stress-related genes of algal origin in *Monosiga*, this report underscores the potential contribution of algal genes to the evolution of nonphotosynthetic lineages.

Acknowledgments

This research was supported by a Discovery Grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada to A.M.N. Many of the sequences analyzed in this study were produced by the US Department of Energy Joint Genome Institute (http://www.jgi.doe.gov/) and the Protist EST Program (http://amoebidia.bcm.umontreal.ca/pepdb/) and are provided for use in this publication only.

References

Abascal, F., Zardoya, R. & Posada, D. 2005. ProtTest, selection of best-fit models of protein evolution. *Bioinformatics* **21**: 2104–2105.

Abedin, M. & King, N. 2008. The premetazoan ancestry of cadherins. *Science* **319**: 946–948.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**: 403–410.

Andersson, J.O. 2005. Lateral gene transfer in eukaryotes. *Cell. Mol. Life Sci.* **62**: 1182–1197.

Andersson, J.O. & Roger, A.J. 2002. Evolutionary analyses of the small subunit of glutamate synthase: gene order conservation, gene fusions, and prokaryote-to-eukaryote lateral gene transfers. *Euk. Cell* **1**: 304–310.

Andersson, J.O., Sjogren, A.M., Davis, L.A.M., Embley, T.M. & Roger, A.J. 2003. Phylogenetic analyses of diplomonad genes reveal frequent lateral gene transfers affecting eukaryotes. *Curr. Biol.* **13**: 94–104.

Andersson, J.O., Hirt, R., Foster, P. & Roger, A. 2006. Evolution of four gene families with patchy phylogenetic distributions: influx of genes into protist genomes. *BMC Evol. Biol.* **6**: 27.

Boucher, Y., Douady, C.J., Papke, R.T., Walsh, D.A., Boudreau, M.E.R., Nesbo, C.L., Case, R.J. & Doolittle, W.F. 2003. Lateral gene transfer and the origins of prokaryotic groups. *Ann. Rev. Genet.* **37**: 283–328.

Carr, M., Leadbeater, B.S.C., Hassan, R., Nelson, M. & Baldauf, S.L. 2008. Molecular phylogeny of choanoflagellates, the sister group to Metazoa. *Proc. Natl Acad. Sci. USA* **105**: 16641–16646.

Cavalier-Smith, T. & Chao, E.E.Y. 2003. Phylogeny of choanozoa, apusozoa, and other protozoa and early eukaryote megaevolution. *J. Mol. Evol.* **56**: 540–563.

Coruzzi, G.M., Edwards, J.W., Tingey, S.V., Tsai, F.-Y. & Walker, E.L. 1989. Glutamine synthetase: molecular evolution of an eclectic multi-gene family. In: *The Molecular Basis of Plant Development* (R. Goldberg & R. Alan, eds), pp. 223–232. Liss Inc., New York.

Craig, J.P., Bekal, S., Hudson, M., Domier, L., Niblack, T. & Lambert, K.N. 2008. Analysis of a horizontally transferred pathway involved in vitamin B-6 biosynthesis from the soybean cyst nematode *Heterodera glycines*. *Mol. Biol. Evol.* **25**: 2085–2098.

Debouba, M., Gouia, H., Valadier, M.H., Ghorbel, M.H. & Suzuki, A. 2006. Salinity-induced tissue-specific diurnal changes in nitrogen assimilatory enzymes in tomato seedlings grown under high or low nitrate medium. *Plant Physiol. Biochem.* **44**: 409–419.

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.

Edgar, R.C. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**: 1–19.

Eichinger, L., Pachebat, J.A., Glockner, G., Rajandream, M.A., Sucgang, R., Berriman, M., Song, J., Olsen, R., Szafranski, K., Xu, Q., Tunggal, B., Kummerfeld, S., Madera, M., Konfortov, B.A., Rivero, F., Bankier, A.T., Lehmann, R., Hamlin, N., Davies, R., Gaudet, P., Fey, P., Pilcher, K., Chen, G., Saunders,

D., Sodergren, E., Davis, P., Kerhornou, A., Nie, X., Hall, N., Anjard, C., Hemphill, L., Bason, N., Farbrother, P., Desany, B., Just, E., Morio, T., Rost, R., Churcher, C., Cooper, J., Haydock, S., van Driessche, N., Cronin, A., Goodhead, I., Muzny, D., Mourier, T., Pain, A., Lu, M., Harper, D., Lindsay, R., Hauser, H., James, K., Quiles, M., Babu, M.M., Saito, T., Buchrieser, C., Wardroper, A., Felder, M., Thangavelu, M., Johnson, D., Knights, A., Loulseged, H., Mungall, K., Oliver, K., Price, C., Quail, M.A., Urushihara, H., Hernandez, J., Rabbinowitsch, E., Steffen, D., Sanders, M., Ma, J., Kohara, Y., Sharp, S., Simmonds, M., Spiegler, S., Tivey, A., Sugano, S., White, B., Walker, D., Woodward, J., Winckler, T., Tanaka, Y., Shaulsky, G., Schleicher, M., Weinstock, G., Rosenthal, A., Cox, E.C., Chisholm, R.L., Gibbs, R., Loomis, W.F., Platzer, M., Kay, R.R., Williams, J., Dear, P.H., Noegel, A.A., Barrell, B. & Kuspa, A. 2005. The genome of the social amoeba *Dictyostelium discoideum*. *Nature* **435**: 43–57.

Emanuelsson, O., Nielsen, H., Brunak, S. & von Heijne, G. 2000. Predicting subcellular localization of proteins based in their N-terminal amino acid sequence. *J. Mol. Biol.* **300**: 1005–1016.

Frommolt, R., Werner, S., Paulsen, H., Goss, R., Wilhelm, C., Zauner, S., Maier, U.G., Grossman, A.R., Bhattacharya, D. & Lohr, M. 2008. Ancient recruitment by chromists of green algal genes encoding enzymes for carotenoid biosynthesis. *Mol. Biol. Evol.* **25**: 2653–2667.

Gogarten, J.P. & Townsend, J.P. 2005. Horizontal gene transfer, genome innovation and evolution. *Nat. Rev. Microbiol.* **3**: 679–687.

Gonzalez-Ballester, D., Camargo, A. & Fernandez, E. 2004. Ammonium transporter genes in *Chlamydomonas*: the nitrate-specific regulatory gene Nit2 is involved in Amt1;1 expression. *Plant Mol. Biol.* **56**: 863–878.

Guindon, S., Lethiec, F., Duroux, P. & Gascuel, O. 2005. PHYML Online – a web server for fast maximum likelihood-based phylogenetic inference. *Nucl. Acids Res.* **33**: W557–W559.

Huang, J.L. & Gogarten, J.P. 2008. Concerted gene recruitment in early plant evolution. *Genome Biol.* **9**: R109.

Huang, C.H. & Peng, J.B. 2005. Evolutionary conservation and diversification of Rh family genes and proteins. *Proc. Natl Acad. Sci. USA* **102**: 15512–15517.

Huelsenbeck, J.P. & Ronquist, F. 2001. MRBAYES, Bayesian inference of phylogenetic trees. *Bioinformatics* **17**: 754–755.

Inokuchi, R., Kuma, K., Miyata, T. & Okada, M. 2002. Nitrogen-assimilating enzymes in land plants and algae: phylogenic and physiological perspectives. *Physiol. Plantarum.* **116**: 1–11.

Katagiri, M. & Nakamura, M. 2003. Reappraisal of the 20th-century version of amino acid metabolism. *Biochem. Biophys. Res. Comm.* **312**: 205–208.

Keeling, P.J. & Palmer, J.D. 2008. Horizontal gene transfer in eukaryotic evolution. *Nat. Rev. Genet.* **9**: 605–618.

King, N., Westbrook, M.J., Young, S.L., Kuo, A., Abedin, M., Chapman, J., Fairclough, S., Hellsten, U., Isogai, Y., Letunic, I., Marr, M., Pincus, D., Putnam, N., Rokas, A., Wright, K.J., Zuzow, R., Dirks, W., Good, M., Goodstein, D., Lemons, D., Li, W.Q., Lyons, J.B., Morris, A., Nichols, S., Richter, D.J., Salamov, A., Bork, P., Lim, W.A., Manning, G., Miller, W.T., McGinnis, W., Shapiro, H., Tjian, R., Grigoriev, I.V. & Rokhsar, D. 2008. The genome of the Choanoflagellate *Monosiga brevicollis* and the origin of metazoans. *Nature* **451**: 783–788.

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

Lawrence, J.G. & Roth, J.R. 1996. Selfish operons: horizontal transfer may drive the evolution of gene clusters. *Genetics* **143**: 1843–1860.

Lea, P., Robinson, S. & Stewart, G. 1990. The enzymology and metabolism of glutamine, glutamate and asparagine. In: *The Biochemistry of Plants* (B. Miflin & P. Lea, eds.), pp. 121–159. Academic Press, San Diego, CA.

Lima, W.C. & Menck, C.F.M. 2008. Replacement of the arginine biosynthesis operon in Xanthomonadales by lateral gene transfer. *J. Mol. Evol.* **66**: 266–275.

Lima, W.C., Paquola, A.C.M., Varani, A.M., Van Sluys, M.A. & Menck, C.F.M. 2008. Laterally transferred genomic islands in Xanthomonadales related to pathogenicity and primary metabolism. *FEMS Microbiol. Lett.* **281**: 87–97.

Lima, W.C., Varani, A.M. & Menck, C.F.M. 2009. NAD biosynthesis evolution in bacteria: lateral gene transfer of Kynurenine pathway in xanthomonadales and flavobacteriales. *Mol. Biol. Evol.* **26**: 399–406.

Ludewig, U. 2006. Ion transport versus gas conduction: function of AMT/Rh-type proteins. *Transf. Clin. Biol.* **13**: 111–116.

Ludewig, U., Neuhduser, B. & Dynowski, M. 2007. Molecular mechanisms of ammonium transport and accumulation in plants. *FEBS Lett.* **581**: 2301–2308.

Maurin, C. & Le Gal, Y. 1997. Isoforms of glutamine synthetase in the marine coccolithophorid *Emiliania huxleyi* Prymnesiophyceae. *Comp. Biochem. Physiol. B-Biochem. Mol. Biol.* **118**: 903–912.

Merkl, R. 2006. A comparative categorization of protein function encoded in bacterial or archeal genomic islands. *J. Mol. Evol.* **62**: 1–14.

Nedelcu, A.M., Miles, I.H., Fagir, A.M. & Karol, K. 2008. Adaptive eukaryote-to-eukaryote lateral gene transfer: stress-related genes of algal origin in the closest unicellular relatives of animals. *J. Evol. Biol.* **21**: 1852–1860.

Nozaki, H., Iseki, M., Hasegawa, M., Misawa, K., Nakada, T., Sasaki, N. & Watanabe, M. 2007. Phylogeny of primary photosynthetic eukaryotes as deduced from slowly evolving nuclear genes. *Mol. Biol. Evol.* **24**: 1592–1595.

Omelchenko, M.V., Makarova, K.S., Wolf, Y.I., Rogozin, I.B. & Koonin, E.V. 2003. Evolution of mosaic operons by horizontal gene transfer and gene displacement in situ. *Genome Biol.* **4**: R55.

Opperdoes, F.R. & Michels, P.A.M. 2007. Horizontal gene transfer in trypanosomatids. *Trends Parasit.* **23**: 470–476.

Pal, C., Papp, B. & Lercher, M.J. 2005. Adaptive evolution of bacterial metabolic networks by horizontal gene transfer. *Nature Genet.* **37**: 1372–1375.

Ramanjulu, S. & Sudhakar, C. 1997. Drought tolerance is partly related to amino acid accumulation and ammonia assimilation: a comparative study in two mulberry genotypes differing in drought sensitivity. *J. Plant Physiol.* **150**: 345–350.

Raymond, J. 2005. The evolution of biological carbon and nitrogen cycling – a genomic perspective. *Mol. Geomicrobiol.* **59**: 211–231.

Robertson, D.L. & Tartar, A. 2006. Evolution of glutamine synthetase in heterokonts: evidence for endosymbiotic gene transfer and the early evolution of photosynthesis. *Mol. Biol. Evol.* **23**: 1048–1055.

Robertson, D.L., Smith, G.J. & Alberte, R.S. 1999. Characterization of a cDNA encoding glutamine synthetase from the marine diatom *Skeletonema costatum* Bacillariophyceae. *J. Phycol.* **35**: 786–797.

Robertson, D.L., Smith, G.J. & Alberte, R.S. 2001. Glutamine synthetase in marine algae: new surprises from an old enzyme. *J. Phycol.* **37**: 793–795.

Ruiz-Trillo, I., Roger, A.J., Burger, G., Gray, M.W. & Lang, B.F. 2008. A phylogenomic investigation into the origin of metazoa. *Mol. Biol. Evol.* **25**: 664–672.

Sanchez-Puerta, M.V., Lippmeier, J.C., Apt, K.E. & Delwiche, C.F. 2007. Plastid genes in a non-photosynthetic dinoflagellate. *Protist* **158**: 105–117.

Schomburg, I., Chang, A., Ebeling, C., Gremse, M., Heldt, C., Huhn, G. & Schomburg, D. 2004. BRENDA, the enzyme database: updates and major new developments. *Nucl. Acids Res.* **32**: D431–D433.

Slamovits, C.H. & Keeling, P.J. 2008. Plastid-derived genes in the nonphotosynthetic alveolate *Oxyrrhis marina*. *Mol. Biol. Evol.* **25**: 1297–1306.

Stechmann, A. & Cavalier-Smith, T. 2003. The root of the eukaryote tree pinpointed. *Curr. Biol.* **13**: R665–R666.

Suzuki, A. & Knaff, D.B. 2005. Glutamate synthase: structural, mechanistic and regulatory properties, and role in the amino acid metabolism. *Photosynth. Res.* **83**: 191–217.

Temple, S.J., Vance, C.P. & Gantt, J.S. 1998. Glutamate synthase and nitrogen assimilation. *Trends Plant Sci.* **3**: 51–56.

Vanoni, M.A., Dossena, L., van den Heuvel, R.H.H. & Curti, B. 2005. Structure–function studies on the complex iron–sulfur flavoprotein glutamate synthase: the key enzyme of ammonia assimilation. *Photosynth. Res.* **83**: 219–238.

von Wiren, N., Gazzarrini, S., Gojon, A. & Frommer, W.B. 2000. The molecular physiology of ammonium uptake and retrieval. *Curr. Opinion Plant Biol.* **3**: 254–261.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Figure S1 Full alignment of predicted NADH-dependent glutamate synthases.

Figure S2 Bayesian analysis of selected type II glutamine synthetases, including an incomplete *Monosiga ovata* GSII EST sequence.

Figure S3 Bayesian analysis of selected type II glutamine synthetases excluding lineages that acquired their GSII sequences via endosymbiont gene transfer, and including an incomplete *Monosiga ovata* GSII EST sequence.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

Received 19 April 2009; revised 29 May 2009; accepted 5 June 2009