Published online 27 June 2015

Nucleic Acids Research, 2015, Vol. 43, No. 14 ***6739–6746***
doi: 10.1093/nar/gkv631

The ancestor of modern Holozoa acquired the CCA-adding enzyme from Alphaproteobacteria by horizontal gene transfer

Heike Betat[1], Tobias Mede[2], Sandy Tretbar[1], Lydia Steiner[2], Peter F. Stadler[3,4,5,6,7,8,9], Mario Mörl[1] and Sonja J. Prohaska[2,4,*]

[1] Institute for Biochemistry, University of Leipzig, Brüderstraße 34, D-04103 Leipzig, Germany, [2] Computational EvoDevo Group, Department of Computer Science, University of Leipzig, Härtelstraße 16-18, D-04107 Leipzig, Germany, [3] Bioinformatics Group, Department of Computer Science, University of Leipzig, Härtelstraße 16-18, D-04107 Leipzig, Germany, [4] Interdisciplinary Center for Bioinformatics, University of Leipzig, Härtelstraße 16-18, D-04107 Leipzig, Germany, [5]Max-Planck-Institute for Mathematics in the Sciences, Inselstraße 22, D-04103 Leipzig, Germany, [6]Fraunhofer Institut für Zelltherapie und Immunologie, Perlickstraße 1, D-04103 Leipzig, Germany, [7]Department of Theoretical Chemistry, University of Vienna, Währingerstraße 17, A-1090 Wien, Austria, [8]Center for non-coding RNA in Technology and Health, University of Copenhagen, Grønnegårdsvej 3, DK-1870 Frederiksberg, Denmark and [9]Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA

Received January 30, 2015; Revised May 28, 2015; Accepted June 7, 2015

ABSTRACT

Transfer RNAs (tRNAs) require the absolutely conserved sequence motif CCA at their 3′-ends, representing the site of aminoacylation. In the majority of organisms, this trinucleotide sequence is not encoded in the genome and thus has to be added post-transcriptionally by the CCA-adding enzyme, a specialized nucleotidyltransferase. In eukaryotic genomes this ubiquitous and highly conserved enzyme family is usually represented by a single gene copy. Analysis of published sequence data allows us to pin down the unusual evolution of eukaryotic CCA-adding enzymes. We show that the CCA-adding enzymes of animals originated from a horizontal gene transfer event in the stem lineage of Holozoa, i.e. Metazoa (animals) and their unicellular relatives, the Choanozoa. The tRNA nucleotidyltransferase, acquired from an α-proteobacterium, replaced the ancestral enzyme in Metazoa. However, in Choanoflagellata, the group of Choanozoa that is closest to Metazoa, both the ancestral and the horizontally transferred CCA-adding enzymes have survived. Furthermore, our data refute a mitochondrial origin of the animal tRNA nucleotidyltransferases.

INTRODUCTION

In all three kingdoms of life, tRNAs carry the universally conserved sequence CCA at their 3′ end, representing an indispensable prerequisite for aminoacylation by the cognate aminoacyl-tRNA synthetases (1). Hence, tRNAs with accurate CCA ends are required for protein biosynthesis and, consequently, of vital importance for the cell. According to tRNA databases (2–4), only 60–70% of the bacterial tRNAs carry an encoded CCA-end. In most other organisms, including all eukaryotic and most archaeal species, the CCA triplet is not encoded and has to be added post-transcriptionally (3,5–8). This is also true for tRNA genes in mitochondria and plastids. The enzyme responsible for the synthesis of the CCA sequence is the ATP(CTP):tRNA nucleotidyltransferase or 'CCA-adding enzyme', a highly specialized RNA polymerase that recognizes all tRNA transcripts of a cell. In eukaryotes, this enzyme is also found in mitochondria, where a distinct set of tRNAs is encoded (9,10). In mammals and many other eukaryotes, these mitochondrial tRNAs show substantial deviations from the standard cloverleaf tRNA secondary structure (11). In the case of nematodes, this situation comes to an extreme, as all mitochondrial tRNAs show bizarre structures lacking complete individual arms of the cloverleaf (12–15). Yet, the CCA-adding enzyme recognizes these deviant transcripts and readily synthesizes the required CCA-termini, indicating a broad substrate range for CCA-addition, while the bacterial enzymes seem to be more restricted to standard tRNA structures (16).

The CCA-adding enzyme is also involved in the maintenance and repair of genomically encoded CCA-sequences, as it completes partial CCA-ends by the addition of the missing nucleotides (17–20). Accordingly, even in organisms such as *Escherichia coli* that encode the CCA triplet in their

*To whom correspondence should be addressed. Tel: +49 341 9716703; Fax: +49 341 9716679; Email: sonja@bioinf.uni-leipzig.de

tRNA genes, a knockout of the CCA-adding enzyme leads to an increase of tRNAs with incomplete CCA-ends and growth impairment (8,21).

Based on a common sequence signature, CCA-adding enzymes are classified as members of the polymerase β superfamily (22). Sequence and structural homologies in the catalytic core led to a subdivision into two different and only distantly related enzyme classes (23,24). While archaeal CCA-adding enzymes have been assigned to class I, eubacterial and eukaryotic enzymes belong to class II, characterized by a highly conserved catalytic core region in the N-terminal part (25). Interestingly, both classes add the CCA sequence at high accuracy without the need of a nucleic acid template. Class I enzymes function as ribonucleoproteins, where a conserved arginine residue and several phosphate positions of the sugar-phosphate backbone of the bound tRNA form hydrogen bonds to the incoming nucleotides (26,27). Class II CCA-adding enzymes carry a set of amino acids in the nucleotide-binding pocket that acts as a protein-based template and forms Watson–Crick-like hydrogen bonds with the incoming nucleotides (25). The orientation of the amino acid side chains determines the nucleotide specificity of this binding pocket and allows an interaction with either CTP or ATP (25). This specificity switch is mediated by a flexible loop element, acting as a lever to reorient the templating amino acids (28–30). In some bacteria the addition of the CCA triplet is performed by the collaboration of a pair of closely related tRNA nucleotidyltransferases with partial activities. One enzyme catalyzes the addition of the two C residues, representing a CC-adding enzyme. The second nucleotidyltransferase (A-adding enzyme) completes the CCA-end by the addition of the terminal A residue (29,31–33). Originally found only in a handful of bacterial species (31,32), extensive data base searches revealed that similar collaborating tRNA nucleotidyltransferase pairs with partial activities are more widespread than originally believed (34). Recently, the molecular mechanisms underlying the restricted activities could be clarified: In CC-adding enzymes, the above mentioned flexible loop region, acting as a lever, is missing due to a short deletion in the corresponding gene. Hence, these enzymes cannot switch the nucleotide specificity of the amino acid template, locking the nucleotide-binding pocket in a CTP-only state (29). A-adding enzymes, on the other hand, appear to selectively bind only tRNAs carrying the two C residues at their 3′-end, while transcripts lacking these residues are not recognized (34).

Phylogenetic analyses complemented the growing understanding of the mechanisms of CCA addition and revealed an interesting and unexpected evolution of class II tRNA nucleotidyltransferases. Although it remains an open question whether A-adding enzymes represent the ancestral state or whether they are derived from CCA-adding enzymes, biochemical and phylogenetic data support the hypothesis that CC-adding enzymes originated multiple times through deletion of the flexible loop region (29,34). Besides the conventional CCA-adding enzymes, a subclass termed IIa that is related to A-adding enzymes has been identified (34), suggesting that CCA-adding enzymes emerged at least twice during evolution.

Here, we provide a further piece of the puzzle to resolve the evolution of class II CCA-adding enzymes in eukaryotes. Usually, a single nuclear gene encodes the CCA-adding enzyme that is delivered to the cytosol, the mitochondria and the plastids, acting on the different sets of tRNAs (9,10,35). More than a decade ago, Reichert *et al.* (9) noted a surprisingly high similarity between CCA-adding enzymes of mammals and bacteria. Phylogenetic analysis suggested that the animal CCA-adding enzymes might have been derived from Alphaproteobacteria, which also represent the endosymbiotic progenitor of mitochondria. They proposed that the α-proteobacterial CCA gene may have been transferred to the nucleus during mitochondrial evolution. Here, we show that the animal CCA-adding enzyme is indeed of bacterial origin, but was not derived from the mitochondrial genome. Instead, the gene originated from a later horizontal gene transfer and eventually replaced the ancestral eukaryotic enzyme in the stem-lineage of Holozoa, i.e. the clade consisting of animals and closely related single-cell organisms. Plants and fungi, on the other hand, are grouped in different clades, and still carry the ancestral eukaryotic CCA-adding enzyme, which shows several deviations in conserved elements.

MATERIALS AND METHODS

Phylogeny

We use a coarse-grained phylogeny covering all major clades of eukaryots and Alphaproteobacteria as far as hypothetical CCA sequences are available. The eukaryotic phylogeny is taken from Paps *et al*. (36). For convenience, we adopt the split of eukaryots into Unikonta and Bikonta, even though only the monophyly of unikonts is undisputed. Unikonts comprise Amoebozoa (with no emergent flagella) and Opisthokonta (commonly bearing one flagellum). The latter include Metazoa, Fungi and lesser known unicellular organisms that were formerly assorted to protists. Metazoa together with their close unicellular relatives, the 'Choanozoa', form the taxonomic group of Holozoa (see Figure 2). Recent advances in phylogenetic analysis have led to the conclusion that 'Choanozoa' do not form a monophyletic group. We use the term here to designate 'Holozoa without Metazoa'. The group of opisthokont amoebae, known as Nucleariida (including *Fonticula alba*), does not belong to the Choanozoa. Instead, several studies have placed Nucleariida as the sister group to Fungi (37).

In our analysis, we use 12 choanozoan species: *Monosiga brevicollis* and *Salpingoeca rosetta* (Choanoflagellata), *Ministeria vibrans* and *Capsaspora owczarzaki* (Filasterea), *Amoebidium parasiticum*, *Abeoforma whisleri*, *Ichthyophonus hoferi*, *Sphaerothecum destruens*, *Pirum gemmata*, *Creolimax fragrantissima* and *Spheraforma arctica* (Ichthyosporea) and *Corallochytrium limacisporum* (Corallochytrea). The life style of Choanozoa ranges from free living single-celled aquatic to endosymbiotic or even parasitic forms. Their phylogenetic relationship is unclear. For example, Filasterea either belong to Holozoa or form their sister group (38,39). According to Paps *et al*. (36), Ichthyosporea diverged first, followed by Corallochytrea and Filasterea, putting Filasterea closer to Metazoa. The

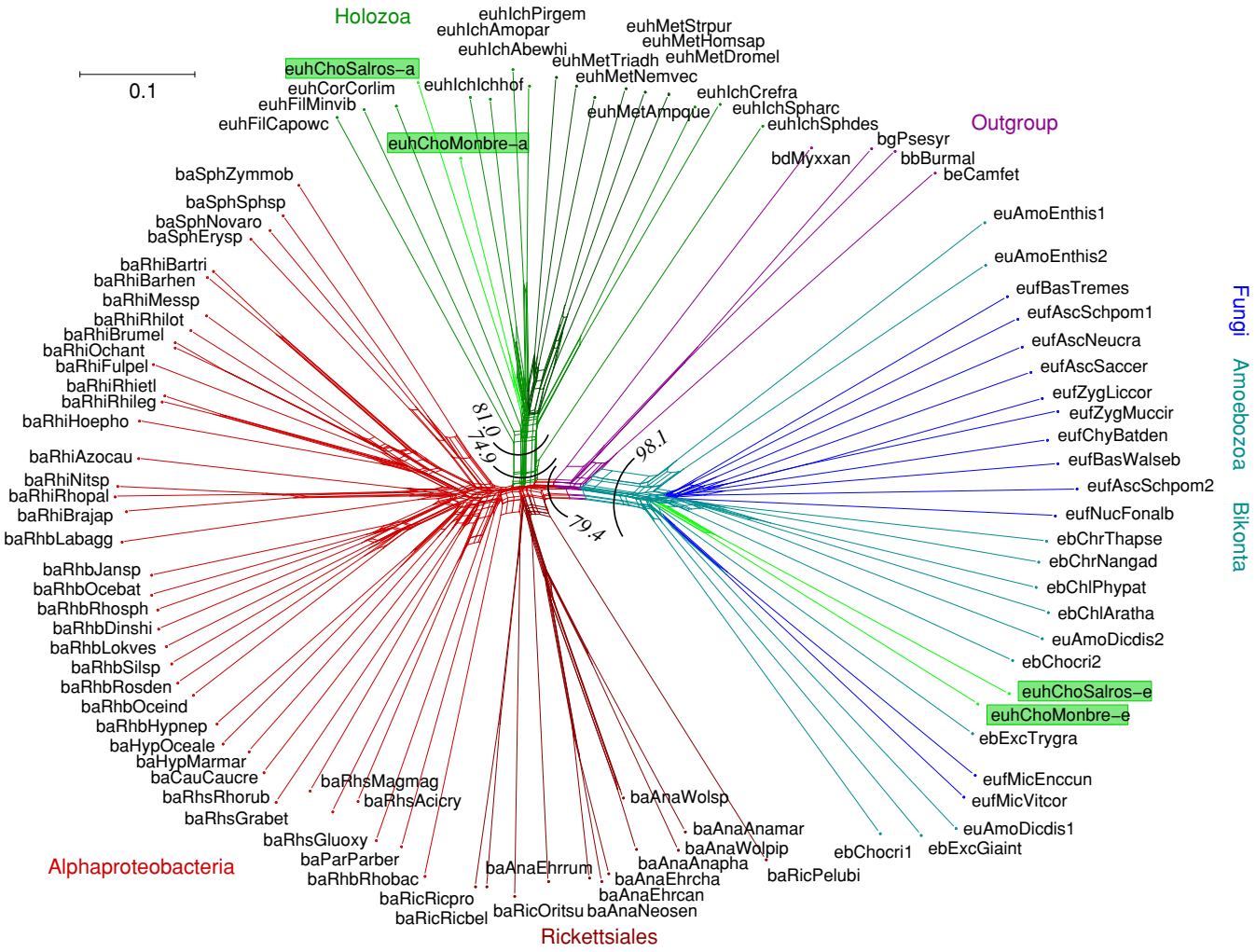

Figure 1. Phylogenetic network of CCA-adding enzymes from Eukarya and Alphaproteobacteria. CCA-adding enzymes from other subgroups of Proteobacteria are used as outgroup. Numbers in italics are bootstrap support values for key splits. aCCAs and eCCAs from choanoflagellates are boxed in green. Sequence names are composed of the taxonomic group prefix (see Supplementary File 6) and the first three letters of the genus and species name (see Supplementary File 1). euh—Holozoa; euf—Fungi; euAmo—Amoebozoa, eb—non-opisthokonts; ba—Alphaproteobacteria; baRic, baAna—Rickettsiales. Monbre—*Monosiga brevicollis*; Salros—*Salpingoeca rosetta*; bb,bg,bd,be—outgroup species from the Proteobacteria subclasses β, γ, δ and ε.

positioning of Choanoflagellata as the sister group of the Metazoa (40) appears to be undisputed.

In this contribution, we consider the branching order within Choanozoa, with the exception of the supported positioning of Choanoflagellata, to be unresolved due to the limited number of genetically and morphologically examined representatives.

Our phylogeny of Alphaproteobacteria is adopted from Williams e*t al*. (41). It covers the same 8 groups, i.e. Rhizobiales, Rhodobacteriaceae, Caulobacteriales, Hyphomonadaceae, Parvularculales, Sphingomonadales, Rhodospirillales and Rickettsiales. Mitochondria very likely emerged at the basis of Rickettsiales before the split of Anaplasmataceae (*Erlichia sp.*, *Anaplasma sp.*, *Wolbachia sp.*) and Rickettsiaceae (*Rickettsia sp.*).

Sequences

The protein sequences of 98 tRNA nucleotidyltransferases are used to compute a gene phylogeny. The sequence set consists of 36 unikont sequences (6 from Metazoa, 14 from 12 Choanozoa, 12 from Fungi and 4 from Amoebozoa), 8 bikont sequences (4 from 3 Archaeplastida, 2 from Chromalveolata and 2 from Excavata), 50 α-proteobacterial sequences and 4 sequences representing the β, γ, δ and ε (sub)classes of the phylum Proteobacteria that serve as an outgroup. The sequences of Choanozoa other than Choanoflagellata are obtained by searching RNA-seq libraries of *A. parasiticum*, *A. whisleri*, *I. hoferi*, *S. destruens*, *P. gemmata*, *C. fragrantissima* and *S. arctica* (Ichthyosporea), *M. vibrans* (Filasterea) and *Corallochytrium limacisporum* kindly provided by I. Ruiz-Trillo and A. de Mendoza Soler (Universitat Pompeu Fabra, Barcelona, Spain). All other sequences are taken from the NCBI database. A list of accession numbers is provided as

Supplementary File 1. The nine Choanozoa sequences derived from RNA-seq are given in Supplementary File 2.

Alignments and trees

The 98 sequences are aligned using `muscle` (42) with default parameter settings. As N- and C-terminal regions are very divergent in length and sequence we manually verify the alignment of the conserved motifs as reported by Neuenfeldt *et al.* (29) (see also Supplementary File 3). The alignment (length 1023 nt) is available as Supplementary File 4. We apply `noisy` (43) to remove phylogenetically uninformative alignment columns, treating gap symbols as missing data which reduces the alignment length to 618 nt. A phylogenetic network is computed using the NeighborNet algorithm (44,45) as implemented in `SplitsTree4` (46) with 1000 bootstrap replicates. To compute a classical maximum likelihood phylogeny we use a condensed version of the original alignment obtained by removing 152 and 46 alignment columns from the start and the end, respectively, because taxon coverage is poor in this region. This results in a final alignment length of 825 nt. With `RAxML` (47) we then compute trees using the GAMMA model of rate heterogeneity, the BLOSUM62 protein substitution model and 100 bootstrap replicates. The `RAxML` tree is available as Supplementary File 5.

RESULTS

To analyze the evolution of class II nucleotidyltransferases in more detail, we collected a set of 98 protein sequences from representatives of a broad spectrum of eukaryotic and bacterial taxa. All of them were classified as CCA-adding enzymes according to the criteria established by Hoffmeier *et al.* (30). Their phylogenetic relationships were analyzed as standard maximum likelihood (`RAxML`) trees and as phylogenetic networks (`NeighborNet`).

Holozoan and α-protobacterial CCA-adding enzymes cluster together

Both, the phylogenetic network and the maximum likelihood tree show a clear and robust separation of metazoan CCA-adding enzymes from those of all other major clades of eukaryotes (i.e. Fungi, Amoebozoa and Bikonta) (Figure 1 and Supplementary File 5).

The six metazoan sequences cluster together (bootstrap support with and without *Drosophila melanogaster*, 38.1 and 47.0%, respectively) and share a common split with the 12 choanozoan sequences (bootstrap support with and without *S. destruens*, 81.0 and 74.9%, respectively). Together, these metazoan and choanozoan CCA-adding enzymes clearly form a monophyletic group (Figure 1) that is firmly placed within the Alphaproteobacteria. Accordingly, we refer to these holozoan CCA-adding enzymes as **aCCA** enzymes.

Sequences of the bikonts and the other two opisthokont clades, i.e. Amoebozoa and Fungi, share the longest and best supported split (bootstrap support 98.1%) close to the root, which is identified by the branching of the outgroups (proteobacteria from the β, γ, δ and ε subgroups (bbBurmal, bgPsesyr, bdMyxxan and beCamfet in Figure 1). We therefore refer to this set of enzymes as the ancestral eukaryotic type of CCA-adding enzymes, **eCCA** enzymes for short.

Two types of CCA-adding enzymes in Choanoflagellata

Surprisingly, we identified two genes for CCA-adding enzymes in the genomes of the choanoflagellates *M. brevicollis* and *S. rosetta*. While the first one clusters together with the Holozoa, the second enzyme unambiguously belongs to the ancestral eukaryotic type of CCA-adding enzyme (eCCA). Thus, the two copies for CCA-adding enzymes in *M. brevicollis* and *S. rosetta* encode both an ancestral eukaryotic eCCA enzyme and an aCCA enzyme.

In contrast to Choanoflagellata, we found no traces of an ancestral eCCA in the RNA-seq data from any of the other 10 choanozoan species.

DISCUSSION

Horizontal transfer of a bacterial CCA gene

The pylogenetic analysis, placing the holozoan CCA sequences within the Alphaproteobacteria, supports the hypothesis of a eubacterial origin of animal tRNA nucleotidyltransferases (9). Furthermore, this indicates that the horizontal gene transfer of the bacterial tRNA nucleotidyltransferase sequence took place at the basis of the holozoan lineage (Figure 2). In a large scale analysis including CCA-like sequences from all bacterial groups (666 sequences total) holozoan sequences clearly cluster with Alphaproteobacteria rather than any other bacterial group (data not shown). However, given the single gene phylogeny, it is not possible to reliably resolve the association with a particular α-proteobacterial subgroup, such as the *Rickettsiales*.

The ancestral Holozoa probably were single-celled, aquatic, free living organisms that either fed on bacteria or hosted a prokaryotic endosymbiont. Even today, prokaryotic endosymbionts are common in protists (48). Therefore, endosymbiosis might have laid the foundation for a transfer of the CCA gene from an ancestral α-proteobacterial to a holozoan host genome.

Surprisingly, this analysis revealed the existence of both, an ancestral eukaryotic (eCCA) and an α-proteobacterial-like (aCCA) enzyme in the Choanoflagellata *M. brevicollis* and *S. rosetta*, a constellation that is probably the result of a horizontal gene transfer at the root of Choanoflagellata.

Evolutionary scenarios

The phylogenetic distribution of CCA-adding enzymes in Opisthokonta can logically be explained by two distinct scenarios (see Figure 2):

Retention Scenario: after a single horizontal gene transfer event both the ancestral eukaryotic eCCA and the α-proteobacterial-like aCCA were present in the holozoan stem lineage. However, only Choanoflagellata, the sister group of Metazoa, retained both types of CCA enzymes. In Metazoa, Filasterea, Corallochytrea and Ichthyosporea the eCCA gene was lost in independent events. As the phylogenetic relationship among Filasterea, Corallochytrea and

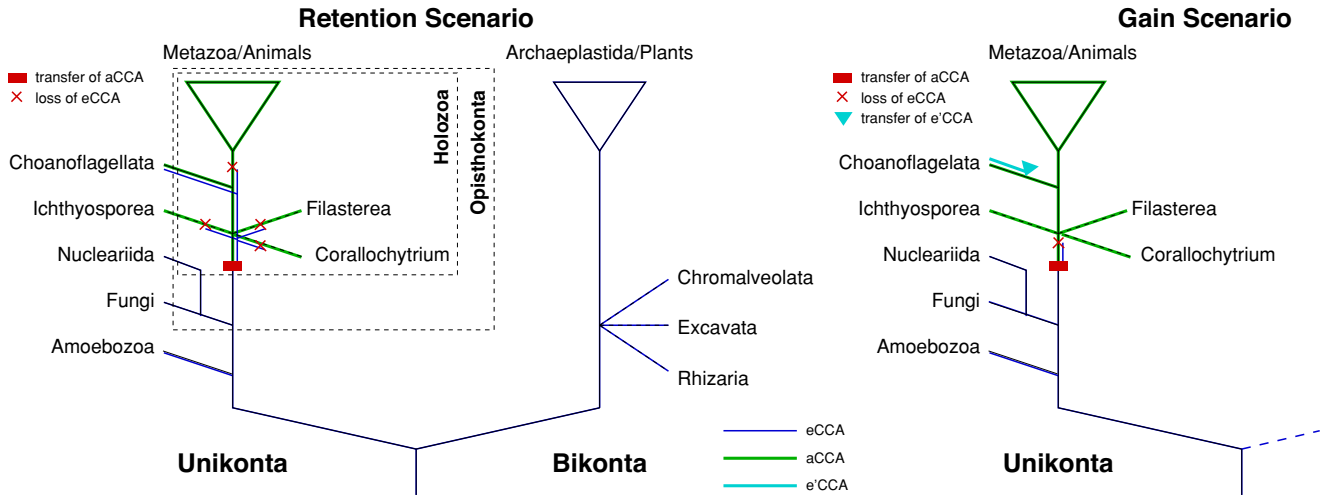

Figure 2. Two evolutionary scenarios explaining the co-occurence of eCCA and aCCA in Choanoflagellata. The solid box indicates the proposed transfer of aCCA from Alphaproteobacteria to the stem lineage of the Holozoa. Subsequent events are marked according to the legend. For a description of the phylogeny see 'Materials and Methods' section.

Ichthysporea is not completely resolved to date, these losses have to be explained parsimoniously by two to four deletion events, one in the metazoa stem lineage and one to three within the Choanozoa. For instance, the minimum of two deletions would require that Filasterea, Corallochytrea and Ichthyosporea lost their eCCA in a single event.

Gain Scenario: the ancestral eCCA was lost shortly after the horizontal transfer of the aCCA gene in the holozoan stem lineage and the Choanoflagellata independently acquired a different eukaryotic CCA gene (e'CCA) from another eukaryotic lineage. Currently, it is not possible to unambiguously determine the origin of a putative acquired eCCA gene as proposed by the *Gain Scenario*. This is because the eCCA of choanoflagellates is not particularly close to the eCCA enzymes of any of the other groups of Eukarya.

Both hypotheses rely only on a coarse-grained and undisputed view of the eukaryotic phylogeny. The unsure positioning of Filasterea, Corallochytrea and Ichthyosporea at the basis of Holozoa only affects the exact number of gain or loss events. However, given the low quality of most choanozoan genomes and the necessarily incomplete genome coverage by transcriptome data, we cannot rule out the possibility that additional functional or pseudogenized eCCA or aCCA genes might be detectable in the future that could change the picture.

In conclusion, we consider the *Retain Scenario* to be the most plausible course of events. It suggest that the imported aCCA gene replaced the ancestral eukaryotic eCCA gene, leading to a number of independent losses of the eCCA gene in the choanozoan groups.

CCA-adding enzyme function and tRNA interaction

Replacement of the ancestral eukaryotic eCCA requires the bacterial-like aCCA to acquire all its essential functions, namely CCA addition and repair at tRNAs in both cytosol and mitochondria. aCCA and eCCA enzymes are closely related proteins and belong to class II of CCA-adding enzymes with a highly conserved N-terminal part. Interestingly, conservation analysis of 18 aCCA enzyme and 26 eCCA enzyme sequences revealed some characteristic differences in the five catalytic core motifs, allowing a discrimination between the aCCA and eCCA enzymes (see Supplementary File 3). A prominent difference is the sequence composition in the flexible loop region that was already used for classification of enzyme families (30). However, none of these signatures allows for a functional differentiation between these two enzyme subtypes, as no impact on the catalytic activity is described for these positions and biochemical characterization of the eCCA enzyme from *Saccharomyces cerevisiae* showed catalytic features and kinetic parameters similar to those of aCCA enzymes (9,49–54). Hence, the question remains as to whether the loss of one of the two functional CCA genes was a random event or whether the retention of the α-proteobacterial CCA-adding enzyme led to an evolutionary advantage for the corresponding organisms. Both yeast and Arabidopsis eCCA enzymes are known to be imported into mitochondria like their aCCA counterparts in Metazoa (35,55,56). Accordingly, these enzymes have also organellar tRNAs as substrates, in addition to the cytosolic transcripts. In contrast to the metazoan situation, the organellar tRNAs of yeast (mitochondria) as well as plants (mitochondria and chloroplasts) do not show comparable structural deviations but fold into the standard cloverleaf structure (57). Hence, it is very likely that the eCCA enzymes are not adapted to bizarre tRNAs as found in mammalian mitochondria.

In Choanoflagellata, only the aCCA sequence of *S. rosetta* carries a putative mitochondrial target sequence in agreement with the mitochondrial localization of the aCCA enzyme in Metazoa. Whether this localization prediction (`Mitoprot` (58), `MultiLoc` (59)) also indicates an adaptation to the mitochondrial tRNA set is unclear, as the mitochondrial tRNAs of Choanoflagellata do not show comparable structural deviations.

In vitro experiments indicate that bacterial CCA enzymes are obviously not able to accept aberrant mitochondrial tRNAs as substrates (16,17) and the question remains as to why the acquired aCCA enzymes in Metazoa tolerate such structural deviations. Like other tRNA-interacting systems (aminoacyl tRNA synthetases, elongation factors, ribosomes), these enzymes co-evolved with the mitochondrial tRNAs. It is possible that the bacterial-type CCA-adding enzymes were better suited for this adaptation than the original eukaryotic enzymes. However, the sequence composition of these enzymes does not show any feature that supports such a scenario. While the eCCA enzymes show a distinct sequence composition in the flexible loop region compared to the aCCA enzymes, it is highly unlikely that this feature has an impact on tRNA recognition, as it is mainly involved in specific nucleotide binding (29,30). As a consequence, it is currently impossible to identify a mechanistic cause why the aCCA enzymes co-evolved with bizarre mitochondrial tRNAs, whereas organisms with eCCA enzymes encode exclusively conventional tRNAs in their mitochondrial and nuclear genomes.

The modern metazoan CCA-adding enzyme has no mitochondrial origin

Regarding the origin of the holozoan aCCA enzymes, their subtree is clearly rooted within the Alphaproteobacteria. The resolution of the single gene phylogeny does not allow a precise localization of the transfer event within the phylogeny of this class, although we observe an affinity to the Rickettsiales, endosymbiontic/endoparasitic Alphaproteobacteria considered to have a common ancestor with the mitochondrial progenitor (60). This coincides with the best estimate for the origin of mitochondria, which were placed within the Alphaproteobacteria (61,62). A later analysis implicated an early point in the evolution of the Rickettsiales, before the divergence of Anaplasmataceae and Rickettsiaceae, as source of the mitochondria (41). Since then, mitochondrial genes have been lost or transferred to the nuclear genome.

The coincidence of these inferred origins of mitochondria and the CCA-adding enzyme led to the speculation that animal CCA-adding enzymes might be a remnant of the mitochondrial endosymbiosis event (9). However, the vast number of bacterial and eukaryotic genomes sequenced during the last decade now allows us to pin down the evolution of class II CCA-adding enzymes with a much higher resolution, rendering the scenario of an endosymbiotic origin of animal CCA-adding enzymes very unlikely. (i) The acquisition of the bacterial CCA-adding enzyme in Holozoa unambiguously occurred after the divergence of the Holozoa and Fungi, while mitochondria became linked to eukaryotes at the very root of the eukaryotic subtree, before the divergence of Unikonta and Bikonta (Figure 2). (ii) No trace of a gene for a CCA-adding enzyme has been found in any of the thousands of sequenced mitogenomes, including that of the basal bikont *Reclinomonas americana*, which most closely resembles the ancestral proto-mitochondrial genome (63). This strongly suggests that the CCA-adding enzyme of the ancestral mitogenome must have been lost or transferred to the nuclear genome very early in the evolution of the Eukarya. This latter scenario, however, is at odds with observation (i). Taken together, a direct mitochondrial origin of the holozoan aCCA gene would require an unlikely large number of independent losses in all other Eukarya lineages.

CONCLUSION

The presence of more than one gene of class II tRNA nucleotidyltransferases has previously been reported for several bacterial species, where the encoded enzymes exhibit only partial activities for CC- and A-addition, respectively (29,31–34). In the Choanoflagellata, such split activities between the two encoded eCCA and aCCA enzymes is highly unlikely, as both enzymes carry a sequence signature in the flexible loop element that does not correspond to that of A-adding enzymes (30). Furthermore, CC-adding enzymes are characterized by a deletion of this loop sequence (29). A corresponding gap is not found in the Choanoflagellate CCA-adding enzymes, excluding an activity restricted to CC-addition. Hence, both proteins probably represent fully active CCA-adding enzymes. It will be interesting to see whether subfunctionalization with respect to localization to different cellular compartments or specificity for subgroups of tRNAs can be observed. While different translation start sites indeed can contribute to the different subcellular localization of CCA-adding enzymes (64), there is no indication so far that such differences change the enzymatic activity. Furthermore, post-translational modifications that might affect the enzymatic behavior are also not described for CCA-adding enzymes. Hence, these surprising enzyme combinations await detailed biochemical characterization to clarify their individual biological function.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENT

We thank I. Ruiz-Trillo and A. de Mendoza Soler from the Universitat Pompeu Fabra, Barcelona, Spain for providing RNA-seq data of 10 choanozoan species.

FUNDING

Deutsche Forschungsgemeinschaft [MI439/14-1 to P.F.S., MO 634/8-1 to M.M., in part].
Conflict of interest statement. None declared.

REFERENCES

1. Sprinzl,M. and Cramer,F. (1979) The C-C-A end of tRNA and its role in protein biosynthesis. *Prog. Nucleic Acid Res. Mol. Biol.*, **22**, 1–69.
2. Chan,P.P. and Lowe,T.M. (2009) GtRNAdb: a database of transfer RNA genes detected in genomic sequence. *Nucleic Acids Res.*, **37**, D93–D97.
3. Jühling,F., Mörl,M., Hartmann,R.K., Sprinzl,M., Stadler,P.F. and Pütz,J. (2009) tRNAdb 2009: compilation of tRNA sequences and tRNA genes. *Nucleic Acids Res.*, **37**, D159–D162.

4. Abe,T., Ikemura,T., Sugahara,J., Kanai,A., Ohara,Y., Uehara,H., Kinouchi,M., Kanaya,S., Yamada,Y., Muto,A. *et al.* (2011) tRNADB-CE 2011: tRNA gene database curated manually by experts. *Nucleic Acids Res.*, **39**, D210–D213.
5. Marck,C. and Grosjean,H. (2002) tRNomics: analysis of tRNA genes from 50 genomes of Eukarya, Archaea, and Bacteria reveals anticodon-sparing strategies and domain-specific features. *RNA*, **8**, 1189–1232.
6. Wegscheid,B. and Hartmann,R.K. (2006) The precursor tRNA 3'-CCA interaction with *Escherichia coli* RNase P RNA is essential for catalysis by RNase P *in vivo*. *RNA*, **12**, 2135–2148.
7. Xiong,Y. and Steitz,T.A. (2006) A story with a good ending: tRNA 3'-end maturation by CCA-adding enzymes. *Curr. Opin. Struct. Biol.*, **16**, 12–17.
8. Zhu,L. and Deutscher,M.P. (1987) tRNA nucleotidyltransferase is not essential for *Escherichia coli* viability. *EMBO J.*, **6**, 2473–2477.
9. Reichert,A.S., Thurlow,D.L. and Mörl,M. (2001) A eubacterial origin for the human tRNA nucleotidyltransferase? *Biol. Chem.*, **382**, 1431–1438.
10. Nagaike,T., Suzuki,T., Tomari,Y., Takemoto-Hori,C., Negayama,F., Watanabe,K. and Ueda,T. (2001) Identification and characterization of mammalian mitochondrial tRNA nucleotidyltransferases. *J. Biol. Chem.*, **276**, 40041–40049.
11. Helm,M., Brulé,H., Friede,D., Giegé,R., Pütz,D. and Florentz,C. (2000) Search for characteristic structural features of mammalian mitochondrial tRNAs. *RNA*, **6**, 1356–1379.
12. Wolstenholme,D.R., Macfarlane,J.L., Okimoto,R., Clary,D.O. and Wahleithner,J.A. (1987) Bizarre tRNAs inferred from DNA sequences of mitochondrial genomes of nematode worms. *Proc. Natl. Acad. Sci. U.S.A.*, **84**, 1324–1328.
13. Okimoto,R., Macfarlane,J., Clary,D. and Wolstenholme,D. (1992) The mitochondrial genomes of two nematodes, Caenorhabditis elegans and Ascaris suum. *Genetics*, **130**, 471–498.
14. Jühling,F., Pütz,J., Florentz,C. and Stadler,P.F. (2012) Armless Mitochondrial tRNAs in Enoplea (Nematoda). *RNA Biol.*, **9**, 1161–1166.
15. Wende,S., Platzer,E.G., Jühling,F., Pütz,J., Florentz,C., Stadler,P.F. and Mörl,M. (2014) Biological evidence for the world's smallest tRNAs. *Biochimie*, **100**, 151–158.
16. Tomari,Y., Suzuki,T. and Ueda,T. (2002) tRNA recognition by CCA-adding enzyme. *Nucleic Acids Symposium Series*, **2**, 77–78.
17. Lizano,E., Scheibe,M., Rammelt,C., Betat,H. and Mörl,M. (2008) A comparative analysis of CCA-adding enzymes from human and *E. coli*: differences in CCA addition and tRNA 3'-end repair. *Biochimie*, **90**, 762–772.
18. Czech,A., Wende,S., Mörl,M., Pan,T. and Ignatova,Z. (2013) Reversible and rapid transfer-RNA deactivation as a mechanism of translational repression in stress. *PLoS Genet.*, **9**, e1003767.
19. Hou,Y.-M.H. (2010) CCA Addition to tRNA: implications for tRNA quality control. *IUBMB Life*, **62**, 251–260.
20. Wilusz,J.E., Whipple,J.M., Phizicky,E. and MandSharp,P.A. (2011) tRNAs marked with CCACCA are targeted for degradation. *Science*, **334**, 817–821.
21. Reuven,N.B., Zhou,Z. and Deutscher,M.P. (1997) Functional overlap of tRNA nucleotidyltransferase, poly (A) polymerase I, and polynucleotide phosphorylase. *J. Biol. Chem.*, **272**, 33255–33259.
22. Aravind,L. and Koonin,E.V. (1999) DNA polymerase beta-like nucleotidyltransferase superfamily: identification of three new families, classification and evolutionary history. *Nucleic Acids Res.*, **27**, 1609–1618.
23. Yue,D., Maizels,N. and Weiner,A.M. (1996) CCA-adding enzymes and poly(A) polymerases are all members of the same nucleotidyltransferase superfamily: characterization of the CCA-adding enzyme from the archaeal hyperthermophile *Sulfolobus shibatae*. *RNA*, **2**, 895–908.
24. Martin,G. and Keller,W. (2007) RNA-specific ribonucleotidyl transferases. *RNA*, **13**, 1834–1849.
25. Li,F., Xiong,Y., Wang,J., Cho,H.D., Tomita,K., Weiner,A.M. and Steitz,T.A. (2002) Crystal structures of the *Bacillus stearothermophilus* CCA-adding enzyme and its complexes with ATP or CTP. *Cell*, **111**, 815–824.
26. Xiong,Y. and Steitz,T.A. (2004) Mechanism of transfer RNA maturation by CCA-adding enzyme without using an oligonucleotide template. *Nature*, **430**, 640–645.
27. Vörtler,S. and Mörl,M. (2010) tRNA-nucleotidyltransferases: highly unusual RNA polymerases with vital functions. *FEBS Lett.*, **584**, 297–302.
28. Just,A., Butter,F., Trenkmann,M., Heitkam,T., Mörl,M. and Betat,H. (2008) A comparative analysis of two conserved motifs in bacterial poly(A) polymerase and CCA-adding enzyme. *Nucleic Acids Res.*, **36**, 5212–5220.
29. Neuenfeldt,A., Just,A., Betat,H. and Mörl,M. (2008) Evolution of tRNA nucleotidyltransferases: a small deletion generated CC-adding enzymes. *Proc. Natl. Acad. Sci. U.S.A.*, **105**, 7953–7958.
30. Hoffmeier,A., Betat,H., Bluschke,A., Günther,R., Junghanns,S., Hofmann,H.-J. and Mörl,M. (2010) Unusual evolution of a catalytic core element in CCA-adding enzymes. *Nucleic Acids Res.*, **38**, 4436–4447.
31. Tomita,K. and Weiner,A.M. (2001) Collaboration between CC- and A-adding enzymes to build and repair the 3'-terminal CCA of tRNA in *Aquifex aeolicus*. *Science*, **294**, 1334–1336.
32. Tomita,K. and Weiner,A.M. (2002) Closely related CC- and A-adding enzymes collaborate to construct and repair the 3′-terminal CCA of tRNA in *Synechocystis sp.* and *Deinococcus radiodurans*. *J. Biol. Chem.*, **277**, 48192–48198.
33. Bralley,P., Cozad,M. and Jones,G.H. (2009) *Geobacter sulfurreducens* contains separate C- and A-adding tRNA nucleotidyltransferases and a poly(A) polymerase. *J. Bacteriol.*, **191**, 109–114.
34. Tretbar,S., Neuenfeldt,A., Betat,H. and Mörl,M. (2011) An inhibitory C-terminal region dictates the specificity of A-adding enzymes. *Proc. Natl. Acad. Sci. U.S.A.*, **108**, 21040–21045.
35. von Braun,S.S., Sabetti,A., Hanic-Joyce,P.J., Gu,J., Schleiff,E. and Joyce,P.B. (2007) Dual targeting of the tRNA nucleotidyltransferase in plants: not just the signal. *J. Exp. Bot.*, **58**, 4083–4093.
36. Paps,J., Medina-Chacón,L.A., Marshall,W., Suga,H. and Ruiz-Trillo,I. (2013) Molecular phylogeny of unikonts: new insights into the position of apusomonads and ancyromonads and the internal relationships of opisthokonts. *Protist*, **164**, 2–12.
37. Liu,Y., Steenkamp,E.T., Brinkmann,H., Forget,L., Philippe,H. and Lang,B.F. (2009) Phylogenomic analyses predict sistergroup relationship of nucleariids and fungi and paraphyly of zygomycetes with significant support. *BMC Evol. Biol.*, **9**, 272.
38. Shalchian-Tabrizi,K., Minge,M.A., Espelund,M., Orr,R., Ruden,T., Jakobsen,K.S. and Cavalier-Smith,T. (2008) Multigene phylogeny of choanozoa and the origin of animals. *PLoS One*, **3**, e2098.
39. del Campo,J. and Ruiz-Trillo,I. (2013) Environmental survey meta-analysis reveals hidden diversity among unicellular opisthokonts. *Mol. Biol. Evol.*, **30**, 802–805.
40. Carr,M., Leadbeater,B.S., Hassan,R., Nelson,M. and Baldauf,S.L. (2008) Molecular phylogeny of choanoflagellates, the sister group to Metazoa. *Proc. Natl. Acad. Sci. U.S.A.*, **105**, 16641–16646.
41. Williams,K.P., Sobral,B.W. and Dickerman,A.W. (2007) A robust species tree for the *Alphaproteobacteria*. *J. Bacteriol.*, **189**, 4578–4586.
42. Edgar,R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.*, **32**, 1792–1797.
43. Dress,A.W., Flamm,C., Fritzsch,G., Grünewald,S., Kruspe,M., Prohaska,S.J. and Stadler,P.F. (2008) Noisy: identification of problematic columns in multiple sequence alignments. *Algorithms Mol. Biol.*, **3**, 7188–7183.
44. Bryant,D., Moulton,V. and Spillner,A. (2004) NeighborNet: an agglomerative method for the construction of planar phylogenetic networks. *Mol. Biol. Evol.*, **21**, 255–265.
45. Levy,D. and Pachter,L. (2011) The neighbor-net algorithm. *Adv. Appl. Math.*, **47**, 240–258.
46. Huson,D.H. and Bryant,D. (2006) Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.*, **23**, 254–267.
47. Stamatakis,A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, **30**, 1312–1313.
48. Wylezich,C., Karpov,S.A., Mylnikov,A.P., Anderson,R. and Jürgens,K. (2012) Ecologically relevant choanoflagellates collected from hypoxic water masses of the Baltic Sea have untypical mitochondrial cristae. *BMC Microbiol.*, **12**, 271.
49. Hegg,L.A. and Thurlow,D.L. (1990) Cytidines in tRNAs that are required intact by ATP/CTP: tRNA nucleotidyltransferases from *Escherichia coli* and *Saccharomyces cerevisiae*. *Nucleic Acids Res.*, **18**, 5975–5979.

50. Hegg,L.A., Kou,M. and Thurlow,D.L. (1990) Recognition of the tRNA-like structure in tobacco mosaic viral RNA by ATP/CTP: tRNA nucleotidyltransferases from *Escherichia coli* and *Saccharomyces cerevisiae*. *J. Biol. Chem.*, **265**, 17441–17445.
51. Chen,J., Kirchner,G., Aebi,M. and Martin,N. (1990) Purification and properties of yeast ATP (CTP): tRNA nucleotidyltransferase from wild type and overproducing cells. *J. Biol. Chem.*, **265**, 16221–16224.
52. Shan,X., Russell,T.A., Paul,S.M., Kushner,D.B. and Joyce,P. (2008) Characterization of a temperature-sensitive mutation that impairs the function of yeast tRNA nucleotidyltransferase. *Yeast*, **25**, 219–233.
53. Leibovitch,M., Bublak,D., Hanic-Joyce,P.J., Tillmann,B., Flinner,N., Amsel,D., Scharf,K.-D., Mirus,O., Joyce,P.B. and Schleiff,E. (2013) The folding capacity of the mature domain of the dual-targeted plant tRNA nucleotidyltransferase influences organelle selection. *Biochem. J.*, **453**, 401–412.
54. Deutscher,M.P. (1982) tRNA nucleotidyltransferase. In: Boyer,PD (ed). *The Enzymes*, 3rd edn. Academic Press, NY, Vol. **XV**, pp. 183–217.
55. Wolfe,C.L., Hopper,A.K. and Martin,N.C. (1996) Mechanisms leading to and the consequences of altering the normal distribution of ATP (CTP): tRNA nucleotidyltransferase in yeast. *J. Biol. Chem.*, **271**, 4679–4686.
56. Zimmer,S.L., Schein,A., Zipor,G., Stern,D.B. and Schuster,G. (2009) Polyadenylation in Arabidopsis and Chlamydomonas organelles: the input of nucleotidyltransferases, poly (A) polymerases and polynucleotide phosphorylase. *Plant J.*, **59**, 88–99.
57. Dirheimer,G., Keith,G., Dumas,P. and Westhof,E. (1995) Primary, secondary, and tertiary structures of tRNAs. In: Söll,D and RajBhandary,U (eds). *RNA: Structure, Biosynthesis and Function*. American Society for Microbiology, Washington, DC, pp. 93–126.
58. Claros,M.G. and Vincens,P. (1996) Computational method to predict mitochondrially imported proteins and their targeting sequences. *Eur. J. Biochem.*, **241**, 779–786.
59. Höglund,A., Dönnes,P., Blum,T., Adolph,H.-W. and Kohlbacher,O. (2006) MultiLoc: prediction of protein subcellular localization using N-terminal targeting sequences, sequence motifs and amino acid composition. *Bioinformatics*, **22**, 1158–1165.
60. Gray,M.W., Burger,G. and Lang,B.F. (2001) The origin and early evolution of mitochondria. *Genome Biol.*, **2**, 1018–1021.
61. Yang,D., Oyaizu,Y., Oyaizu,H., Olsen,G.J. and Woese,C.R. (1985) Mitochondrial origins. *Proc. Natl. Acad. Sci. U.S.A.*, **82**, 4443–4447.
62. Hedges,S.B., Chen,H., Kumar,S., Wang,D. Y.C., Thompson,A.S. and Watanabe,H. (2001) A genomic timescale for the origin of eukaryotes. *BMC Evol. Biol.*, **1**, 4.
63. Lang,B.F., Burger,G., O'Kelly,C.J., Cedergren,R., Golding,G.B., Lemieux,C., Sankoff,D., Turmel,M. and Gray,M.W. (1997) An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. *Nature*, **387**, 493–497.
64. Chen,J.Y., Joyce,P., Wolfe,C., Steffen,M. and Martin,N. (1992) Cytoplasmic and mitochondrial tRNA nucleotidyltransferase activities are derived from the same gene in the yeast *Saccharomyces cerevisiae*. *J. Biol. Chem.*, **267**, 14879–14883.