Molecular Phylogenetics and Evolution 101 (2016) 133–141

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Complex evolution of two types of cardiolipin synthase in the eukaryotic lineage stramenopiles

Fumiya Noguchi [a,b], Goro Tanifuji [c], Matthew W. Brown [d,e], Katsunori Fujikura [b], Kiyotaka Takishita [a,b,*]

[a] Graduate School of Marine Science and Technology, Tokyo University of Marine Science and Technology, Minato-ku, Tokyo, Japan
[b] Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Yokosuka, Kanagawa, Japan
[c] Graduate School of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Ibaraki, Japan
[d] Department of Biological Sciences, Mississippi State University, MS, USA
[e] Institute for Genomics, Biocomputing & Biotechnology, Mississippi State University, MS, USA

ARTICLE INFO

Article history:
Received 22 December 2015
Revised 29 March 2016
Accepted 6 May 2016
Available online 9 May 2016

Keywords:
Cardiolipin synthase
Differential gene loss
Eukaryotes
Phylogenomics
Stramenopiles

ABSTRACT

The phospholipid cardiolipin is indispensable for eukaryotes to activate mitochondria, and it was previously reported that two phylogenetically distinct types of enzyme synthesizing cardiolipin, one with two phospholipase D domains (CLS_pld) and the other with a CDP-alcohol phosphatidyltransferase domain (CLS_cap), are patchily and complementarily distributed at higher taxonomic (e.g., supergroup) levels of eukaryotes. Stramenopiles, one of the major eukaryotic clades, were considered to exclusively possess CLS_cap. However, through our present surveys with genome or transcriptome data from a broad range of stramenopile taxa, species with both CLS_cap and CLS_pld and species with only CLS_pld or CLS_cap were discovered among this group. Because these homologues of CLS_cap and CLS_pld retrieved from stramenopiles were likely inherited from the last eukaryotic common ancestor, it is reasonable to assume that a common ancestor of all stramenopiles harbored both CLS_cap and CLS_pld. Furthermore, based on the robust organismal phylogeny of stramenopiles unveiled with large-scale phylogenetic analyses, the earliest diverging lineage of stramenopiles (including bicosoecids, placidids, etc.) was found to comprise species with both CLS_cap and CLS_pld along with species with only either CLS_cap or CLS_pld. These findings suggest that a common ancestor of the most basal stramenopile lineage retained these two vertically inherited enzymes and that differential losses of either CLS_cap or CLS_pld occurred in this lineage. On the other hand, in the other stramenopile lineage composed of Ochrophyta, Pseudofungi, and Labyrinthulomycetes (to the exclusion of the most basal lineage), only CLS_cap was found, and therefore a common ancestor of these three groups likely lost CLS_pld. Based on our findings, the evolution of CLS_cap/CLS_pld in stramenopiles appears to be more complex than previously thought.

1. Introduction

Cardiopilin (CL) is an anionic dimeric phospholipid with two phosphate residues and four types of fatty acyl chain. It is widely held that CL in mitochondria plays a pivotal role in stabilizing and/or regulating various mitochondrial proteins such as the respiratory chain complexes/supercomplexes (e.g., Fry and Green, 1981; Eble et al., 1990; Robinson et al., 1990). Furthermore, mitochondrial CL is involved in organizing structure characteristic of mitochondria such as cristae, as well as in aging and apotosis (e.g., Schug and Gottlieb, 2009; Sakamoto et al., 2012). It is also known that bacterial CL is essential to activate proteins associated with energy metabolism and to properly localize membrane proteins (e.g., Yankovskaya et al., 2003; Gold et al., 2010; Arias-Cartin et al., 2011). Based on these findings, it was originally proposed that CL is exclusively located in mitochondrial (inner) and bacterial plasma membranes (White and Frerman, 1967; Zinser et al., 1991), but recently this type of phospholipid has also been found in the plasma membranes of a very limited number of archaeal species and in the peroxisomal membranes of eukaryotes (Lattanzio et al., 2002; Koga and Morii, 2007; Wriessnegger et al., 2007; Lobasso et al., 2008).

CL is known to be biosynthesized by either of two phylogenetically distinct enzymes: CL synthase (CLS) with two phospholipase D domains, CLS_pld, which synthesizes CL from two molecules of phosphatidylglycerols (PGs) (e.g., Nishijima et al., 1988) or CLS with one CDP-alcohol phosphatidyltransferase domain, CLS_cap, which produces this lipid using a PG and a cytidine diphosphate

* Corresponding author at: Japan Agency for Marine-Earth Science and Technology, 2-15 Natsushima-cho, Yokosuka, Kanagawa 237-0061, Japan.
E-mail address: takishitak@jamstec.go.jp (K. Takishita).

http://dx.doi.org/10.1016/j.ympev.2016.05.011
1055-7903/

diacylglycerol (CDP-DAG) as substrates (e.g., Chang et al., 1998; Tuller et al., 1998). It has been believed that CLS_pld and CLS_cap function in bacteria and eukaryotes (mitochondria), respectively (e.g., Schlame, 2008). In contrast to the bacterial-type CL, mitochondrial "immature" CL synthesized by CLS is further remodeled (reacylated), resulting in mature CL generally possessing the same fatty acids at sn-1, 2 sites in one molecule. This eukaryotic CL maturation pathway consists of two steps: in the first step, immature CL is deacylated into monolysocardiolipin (MLCL) with either CL-specific phospholipase (CLD) or calcium-independent phospholipase A2 (iPLA2) beta/gamma (Beranek et al., 2009; Malhotra et al., 2009; Zachman et al., 2010). After the deacylation in the first step, MLCL is then reacylated to generate mature CL with either CoA-independent tafazzin (TAZ) or acylCoA:lysocardiolipin acyltransferase 1 (ALCAT1) (Cao et al., 2004; Gu et al., 2004).

Recently, exceptions to the above-mentioned "simple" hypothesis regarding CLS phylogenetic distribution, in which CLS_pld and CLS_cap are exclusively found in bacteria and eukaryotes, respectively, have been found: actinobacteria and proteobacteria were found to contain CLS_cap-like proteins (Sandoval-Calderón et al., 2009; Tian et al., 2012). In addition, Tian et al. (2012) reported that the eukaryotic supergroups Amoebozoa, Excavata, and Alveolata, a subgroup of the supergroup "SAR" (Adl et al., 2012), have only CLS_pld (without phylogenetic affiliation to any particular bacterial homologues), while the supergroups Opisthokonta (including animals and fungi) and Archaeplastida (including land plants) along with another SAR subgroup stramenopiles possess only CLS_cap (closely related to alpha-proteobacterial homologues). It should be noted that Tian et al. (2012) analyzed a broad range of eukaryotes including non-model organisms other than animals, yeasts, and plants, resulting in the detection of CLS_pld in eukaryotes. It is also remarkable that both CLS_cap and CLS_pld are patchily and complementarily distributed at higher taxonomic levels of eukaryotes. Based on these findings, Tian et al. (2012) proposed the following evolutionary scenarios: a primitive eukaryotic cell prior to acquisition of mitochondria inherited CLS_pld from its ancient bacterial ancestor, and then mitochondria originating from an endosymbiotic event of an alpha-proteobacterium took CLS_cap to the last eukaryotic common ancestor (LECA), meaning that the LECA harbored both CLS_cap and CLS_pld. Subsequently, either CLS_cap or CLS_pld was differentially lost in various eukaryotic lineages.

We previously analyzed the metabolic characteristics of mitochondrion-related organelles (degenerated mitochondria) in the free-living anaerobic stramenopile *Cantina marsupialis* based on RNA-seq data (Noguchi et al., 2015). Through those analyses, we found that this stramenopile species has the gene encoding CLS_pld rather than CLS_cap, which stramenopiles exclusively possess according to Tian et al. (2012). Stramenopiles are a huge monophyletic assemblage of eukaryotes comprising Ochrophyta (composed of a wide variety of photosynthetic lineages including diatoms and brown algae) and two heterotrophic groups, Bigyra (including bicosoecids, labyrinthulomycetes, placidids, and *Blastocystis*) and Pseudofungi (composed of oomycetes, hyphochytriomycetes, and *Developayella*) (Cavalier-Smith and Chao, 2006). Nevertheless, Tian et al. (2012) surveyed the presence or absence of CLS_cap/CLS_pld in a very limited number of phylogenetically narrow representatives of stramenopiles (five species of oomycetes, two species of diatoms, one species of brown alga, and *Blastocystis*). Therefore, it remained uncertain whether many other stramenopile lineages contain CLS_cap or CLS_pld and whether only *Cantina* exceptionally possesses CLS_pld among stramenopiles.

In the present study, using genome or transcriptome data, we examined the presence or absence of the two types of CL biosynthesis enzyme, CLS_cap and CLS_pld, together with the CL maturation (remodeling) enzymes, CLD, iPLA2, TAZ, and ALCAT1, in stramenopile species more comprehensively in terms of phylogeny than in the case of Tian et al. (2012). Consequently, several species of stramenopiles other than *Cantina* were also found to have CLS_pld. We further traced the evolution of CLS_cap/CLS_pld within the radiation of stramenopiles based on the organismal phylogeny of this large eukaryotic group as resolved by large-scale phylogenetic analyses.

2. Materials and methods

2.1. Cultures

The strains *Cafeteria roenbergensis* NIES1012, *Wobblia lunata* NIES1015, and *Developayella elegans* NIES1388 were purchased from the Microbial Culture Collection of the National Institute for Environmental Studies (NIES, 16-2 Onogawa, Tsukuba, Ibaraki 305-8506, Japan). These cells were grown according to the instructions from the NIES.

2.2. Analyses of RNA-seq data from cultured strains

Cells of *C. roenbergensis*, *W. lunata*, and *D. elegans* were harvested by centrifugation at 4170 g for 60 min at 4 °C. Total RNA was isolated from the harvested cells using TRIzol reagent (Life Technologies, USA). Construction of cDNA libraries and paired-end sequencing with Illumina HiSeq 2000 (100 bp per read) were performed by Hokkaido System Science Co., Ltd. Raw sequencing reads were deposited in the Sequence Read Archive under the accession numbers DRX044597–DRX044599. 154–158 million raw sequencing read data were filtered using TRIMMOMATIC software version 0.30 (Lohse et al., 2012) to remove adapter sequences and low-quality bases. Filtered sequences were then assembled into 23,271–49,186 transcript contigs using the TRINITY package (release 2013-02-25) (Grabherr et al., 2011).

2.3. Sequence data collection

In addition to RNA-seq datasets from *C. roenbergensis*, *W. lunata*, and *D. elegans* generated in this study and from *Cantina marsupialis* generated in Noguchi et al. (2015), publicly available genome or transcriptome datasets from 80 taxa of eukaryotes (including 55 taxa of stramenopiles) were downloaded to find orthologous sequences for phylogenomic analyses and/or to identify the genes encoding enzymes for cardiolipin (CL) biosynthesis and maturation. Taxon names for which sequence datasets were collected are listed in Table S1.

2.4. Phylogenomic analyses

Genome/transcriptome data as mentioned above were used as inputs for an in-house pipeline, described below, for the creation of single protein datasets and, subsequently, the phylogenomic data matrix. The organismal data were individually screened for orthologues using either blastp or tblastn, depending on the data type, with the reference orthologue sequences used as queries in BLASTMONKEY from the Barrel-o-Monkeys toolkit (http://roger-lab.biochem.dal.ca). If the sequence dataset was nucleotides, then the tblastn hits were translated to amino acid residues. Blastp was then used to screen these putative orthologues against the OrthoMCL database, and the output for each gene from each organism was compared against a dictionary of orthologous OrthoMCL IDs. Those putative orthologues that did not match orthologous IDs were designated as paralogues and removed. The remaining orthologues from each organism were combined and aligned using