

The Guanine Nucleotide Exchange Factors Sec2 and PRONE: Candidate Synapomorphies for the Opisthokonta and the Archaeplastida

Marek Elias

Department of Botany and Department of Philosophy and History of Science, Faculty of Sciences, Charles University in Prague, Prague, Czech Republic

Although recent multigene phylogenetic analyses support close relationship of Metazoa and Fungi (the eukaryotic supergroup Opisthokonta) and monophyly of eukaryotes with the primary plastid, that is, Chloroplastida, Rhodophyta, and Glaucophyta (the supergroup Archaeplastida or Plantae), some authors still challenge this scheme. I found that 2 particular types of guanine nucleotide exchange factors (GEFs, i.e., cofactors of GTPases) might provide a new piece of evidence to resolve this controversy. An exhaustive analysis of available sequence data revealed that Sec2-related proteins, known to serve as GEF for exocytic GTPases of the Rab8/Sec4 subfamily, are restricted to opisthokonts, whereas proteins with the PRONE domain, recently described as novel plant-specific GEFs for RHO family GTPases, occur only in Chloroplastida and Rhodophyta. The results thus point to possible evolutionary innovations in the exocytic apparatus of the ancestral opisthokonts and reveal the probably first plastid-independent trait (i.e., a unique mode of RHO GTPase regulation) exclusive for Chloroplastida + Rhodophyta, further supporting monophyly of these 2 groups.

Main Text

Defining the major phylogenetic branches of eukaryotes has proven a challenging task, but recent years have witnessed a remarkable progress owing to improved methods of phylogenetic inference and growing amount of sequence data from diverse eukaryotes. A widely adopted consensus assumes the existence of some 6 major monophyletic supergroups—Opisthokonta, Amoebozoa, Excavata, Archaeplastida (=Plantae), Chromalveolata, and Rhizaria (Simpson and Roger 2004; Adl et al. 2005). Interrelationship among the supergroups and the position of the root of the eukaryotic tree remain controversial, as actually does the monophyly of most supergroups themselves. Opisthokonta, which comprise Metazoa, Fungi, and their unicellular relatives (e.g., choanoflagellates), are recovered even in single-gene phylogenies, and multigene analyses provide an overwhelming support for their monophyly (e.g., Steenkamp et al. 2006; Burki et al. 2007; Rodriguez-Ezpeleta et al. 2007). In addition, several characters unique to opisthokonts are known, including a specific insertion in the elongation factor EF1- α or the presence of a tyrosyl-tRNA synthase derived by horizontal gene transfer from haloarchaea (Huang et al. 2005; Steenkamp et al. 2006). Nevertheless, some authors remain skeptical about opisthokont monophyly, pointing to a number of traits presumably shared by metazoans and green plants to the exclusion of fungi, for example, a specific form of the pre-mRNA capping enzyme or the Rb-E2F pathway for the cell cycle control (recently discussed by Stiller [2007]). However, phylogenetic distribution of neither of these traits putatively specific for metazoa and green plants has been thoroughly investigated, so their significance remains uncertain.

The supergroup Archaeplastida embraces eukaryotes with the so-called primary plastid, that is, green algae and plants (Chloroplastida or Viridiplantae), red algae (Rhodophyta), and the small algal group Glaucophyta (Adl et al. 2005). Monophyly of Archaeplastida has been

slow in gaining acceptance, in part due to lack of support in most phylogenies up to now. Genome and expressed sequence tag (EST) sequencing have finally generated data sets large enough to show the monophyly of Archaeplastida with at least suggestive statistical support (e.g., Hackett et al. 2007; Rodriguez-Ezpeleta et al. 2007). However, alternative results have been obtained in some multigene analyses (Nozaki et al. 2007), and methodological concerns have been raised, for example, that the signal supporting the monophyly of Archaeplastida might be artificial due to convergent “contamination” of their nuclear genomes by genes from plastid endosymbionts gained independently by individual archaeplastid lineages (Stiller 2007). It is obvious that additional evidence is needed to finally prove or refute monophyly of opisthokonts and archaeplastids.

Here I show that such new evidence may come from a comparative analysis of regulatory pathways dependent on Ras-like GTPases, a diverse superfamily implicated in various transport and signaling pathways in the eukaryotic cell (Takai et al. 2001). The so-called guanine nucleotide exchange factors (GEFs), which activate GTPases by promoting the exchange of bound GDP for GTP, are a heterogeneous group of proteins with structural features shared typically only by GEFs specific to a particular subgroup of the Ras superfamily. In the RAB family, GEFs differ for each subfamily (Segev 2001). For example, the unrelated protein complex TRAPP-I and the Vps39 subunit of the protein complex HOPS act as GEFs specific for the Rab1 and Rab7 subfamilies, respectively. A recent survey of eukaryotic genomes indicates a wide conservation of the 2 complexes, suggesting their early origin in the eukaryotic evolution (Koumandou et al. 2007).

The yeast protein Sec2 and its mammalian orthologs, RAB3IP (=Rabin3 or Rabin8) and RAB3IL1 (=GRAB), tell a different story. They exhibit GEF activity toward exocytic RABs of the Rab8/Sec4 subfamily, exemplified by Sec4 in the yeast and Rab8 and Rab3 in mammals (Walch-Solimena et al. 1997; Luo et al. 2001; Hattula et al. 2002). The Rab8/Sec4 subfamily is widely conserved in eukaryotes (e.g., Rutherford and Moore 2002; Eisen et al. 2006; fig. 1). However, my exhaustive searches of available genome and EST sequence data using a sensitive method implemented in HMMER (see Methods and supplementary fig. S1, Supplementary Material online) did not reveal Sec2

Key words: GTPases, eukaryotes, phylogeny, Opisthokonta, Archaeplastida.

E-mail: melias@natur.cuni.cz.

Mol. Biol. Evol. 25(8):1526–1529. 2008

doi:10.1093/molbev/msn113

Advance Access publication May 13, 2008

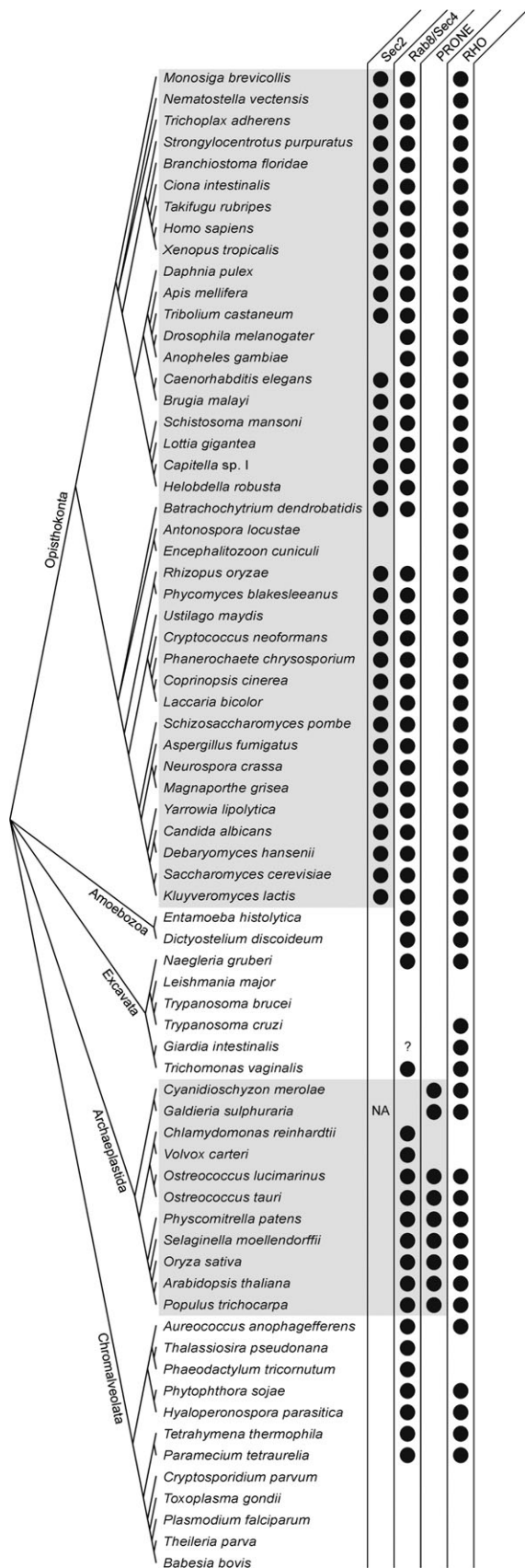


FIG. 1.—Phylogenetic distribution of Sec2 and PRONE GEFs. For comparison, phylogenetic distribution of the Rab8/Sec4 subfamily and

homologs outside opisthokonts (fig. 1). The absence of Sec2 in a few opisthokonts (dipterans and Microsporidia, the latter of which also lack the Rab8/Sec4 subfamily) is apparently secondary because these species are nested within Sec2-containing clades (Adl et al. 2005; Steenkamp et al. 2006). I therefore suggest that the Sec2 GEF is an evolutionary novelty that arose in the lineage leading to the opisthokonts after it had split from other eukaryotes. Interestingly, Field et al. (2007) suggested that 4 other components of the membrane trafficking machinery, GGA, Epsin, Eps15, and Vps27/Hrs, are similarly restricted to the opisthokonts. Compared with the present study, these authors analyzed a more limited sample of genomes and used Blast instead of the more sensitive HMMER, so their conclusions have to be treated as provisional pending additional tests. It is nevertheless becoming apparent that opisthokonts might display a lot of idiosyncrasies in their endomembrane system.

Synapomorphies so far identified for Archaeplastida all relate to the principal defining feature of this supergroup, the primary plastid (e.g., Reyes-Prieto and Bhattacharya 2007a). I was therefore intrigued by finding a RHO-specific GEF shared by representatives of Archaeplastida but not other supergroups. The RHO family is common in most eukaryotic lineages and seems to be involved mainly in signaling to the actin cytoskeleton (Jaffe and Hall 2005; Brembu et al. 2006; Boureux et al. 2007). The 2 types of RHO-specific GEFs characterized in Metazoa and Fungi contain either the DH or the CZH2 domain (Zheng 2001; Meller et al. 2005), and genome surveys revealed a broad distribution of both types across the eukaryotic tree, indicating their ancient origin (Meller et al. 2005; Brembu et al. 2006; Marek Elias, unpublished data).

Angiosperm plants have been recently shown to possess a completely novel RhoGEF type containing a conserved domain dubbed PRONE (Berken et al. 2005; Brembu et al. 2006; Gu et al. 2006). The 2 recently sequenced land plants only distantly related to angiosperms, the lycophyte *Selaginella moellendorffii* and the moss *Physcomitrella patens*, encode a group of 4 and 6 paralogs, respectively, highly similar to angiosperm PRONE proteins (see Supplementary Material online). To identify even distant homologs of the PRONE domain, I probed available genomic and EST sequences with a profile HMM constructed from the embryophyte sequences. Significant matches were found only in green algae from the genus

RHO family of GTPases, known as targets of these 2 types of GEFs, is shown as well. Black dots indicate that at least one representative of a given group of proteins is encoded by the genome of the respective species, whereas an empty space indicate true or (in the case of draft genomes) probable absence. Protein sequences from *Galdieria sulphuraria* are not available for HMMER searches, so presence/absence of Sec2 could not be assessed with certainty (NA). The question mark for *Giardia* Rab8/Sec4 means that assignment of some highly divergent RABs in this species is uncertain. The topology of the schematic tree follows the recent phylogenetic literature (cited in Supplementary Material online). Shaded boxes highlight restriction of Sec2 and PRONE to opisthokonts and archaeplastids, respectively. For details, see Supplementary Material online.

Ostreococcus (*Ostreococcus tauri* and *Ostreococcus lucimarinus*) and in the red alga *Cyanidioschyzon merolae*; Blast further finds a PRONE protein in another red alga, *Galdieria sulphuraria* (not yet available for HMMER searches). The conservation of the PRONE sequences among the distinct taxonomic subgroups (embryophytes, *Ostreococcus*, rhodophytes) is not high, especially in the N-terminal region of the *Ostreococcus* sequences (supplementary fig. S2, Supplementary Material online), but homology of all these sequences is clearly indicated by various homology searches applied (for details, see supplementary methods, Supplementary Material online). No PRONE-related protein could be identified in other species tested (fig. 1). These results are consistent with the fact that the tertiary structure of the PRONE domain is unlike any structure solved so far (Thomas et al. 2007).

Interestingly and in contrast to the chloroplastid PRONE proteins, the rhodophyte homologs both appear to harbor the microtubule interacting and transport (MIT) domain in the region N-terminal to the PRONE domain. The MIT domain is known to serve as an adaptor-mediating interactions with the endosomal ESCRT III protein complex (Tsang et al. 2006), raising the possibility that regulation of the RHO GTPase signaling is coupled to endosomal trafficking in rhodophytes.

The phylogenetic distribution of the PRONE domain is most parsimoniously explained as an evolutionary novelty of a lineage comprising Rhodophyta and Chloroplastida, being primarily absent from other clades. The chlorophytes *Chlamydomonas reinhardtii* and *Volvox carteri* lack both the PRONE domain and RHO GTPases, suggesting a secondary loss of the whole RHO-based module from this lineage. Whether PRONE arose before the radiation of the whole Archaeplastida depends on its status in glaucophytes, a probable basal-most archaeplastid branch (Reyes-Prieto and Bhattacharya 2007b). ESTs available for 2 glaucophytes (*Cyanophora paradoxa* and *Glaucocystis nostochinearum*) do not record a PRONE domain, but complete genomes are required to definitely resolve this issue.

The restriction of the PRONE domain to organisms with the primary plastid might indicate that it arrived with the endosymbiotic cyanobacterium. However, HMMER search of all cyanobacterial protein sequences available as of November 2007 revealed no discernible homologs of the PRONE domain, suggesting that it might be the first plastid-independent character uniting chloroplastids and rhodophytes identified.

In summary, the distribution of the Sec2 and PRONE GEFs is consistent with the “traditional” relationships of Metazoa, Fungi, Chloroplastida, and Rhodophyta, rather than with the iconoclastic view advocated by Stiller (2007). It is of note that a type of the mRNA capping enzyme claimed by Stiller to be a synapomorphy for Metazoa and green plants is actually not only restricted to these 2 taxa but can also be found, at least, in the choanoflagellate *Monosiga brevicollis* (GenPept accession number EDQ85467), which does belong to a lineage closely related to Metazoa, or the excavate *Trichomonas vaginalis* (EAY15722) generally not considered to be specifically related to Metazoa or green plants. Apparently, much more

genomes, especially from currently poorly sampled lineages, are necessary to assess the significance of the various characters, including Sec2 and PRONE, as markers for the major eukaryotic clades.

Supplementary Material

Supplementary methods and figures S1 and S2 are available at *Molecular Biology and Evolution* online (<http://www.mbe.oxfordjournals.org/>).

Acknowledgments

I am indebted to genome sequencing centers for providing sequence data crucial for this study. I thank Fatima Cvrckova for critical reading of the manuscript. This work was supported by research project no. 21620828 of Czech Ministry of Education.

Literature Cited

- Adl SM, Simpson AG, Farmer MA, et al. (25 co-authors). 2005. The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *J Eukaryot Microbiol.* 52:399–451.
- Berken A, Thomas C, Wittinghofer A. 2005. A new family of RhoGEFs activates the Rop molecular switch in plants. *Nature.* 436:1176–1180.
- Boureux A, Vignal E, Faure S, Fort P. 2007. Evolution of the Rho family of ras-like GTPases in eukaryotes. *Mol Biol Evol.* 24:203–216.
- Brembu T, Winge P, Bones AM, Yang Z. 2006. A RHOse by any other name: a comparative analysis of animal and plant Rho GTPases. *Cell Res.* 16:435–445.
- Burki F, Shalchian-Tabrizi K, Minge M, Skjaeveland A, Nikolaev SI, Jakobsen KS, Pawlowski J. 2007. Phylogenomics reshuffles the eukaryotic supergroups. *PLoS ONE.* 2:e790.
- Eisen JA, Coyne RS, Wu M, et al. (50 co-authors). 2006. Macronuclear genome sequence of the ciliate *Tetrahymena thermophila*, a model eukaryote. *PLoS Biol.* 4:e286.
- Field MC, Gabernet-Castello C, Dacks JB. 2007. Reconstructing the evolution of the endocytic system: insights from genomics and molecular cell biology. *Adv Exp Med Biol.* 607:84–96.
- Gu Y, Li S, Lord EM, Yang Z. 2006. Members of a novel class of Arabidopsis Rho guanine nucleotide exchange factors control Rho GTPase-dependent polar growth. *Plant Cell.* 18:366–381.
- Hackett JD, Yoon HS, Li S, Reyes-Prieto A, Rummele SE, Bhattacharya D. 2007. Phylogenomic analysis supports the monophyly of cryptophytes and haptophytes and the association of rhizaria with chromalveolates. *Mol Biol Evol.* 24:1702–1713.
- Hattula K, Furuholm J, Arffman A, Peranen J. 2002. A Rab8-specific GDP/GTP exchange factor is involved in actin remodeling and polarized membrane transport. *Mol Biol Cell.* 13:3268–3280.
- Huang J, Xu Y, Gogarten JP. 2005. The presence of a haloarchaeal type tyrosyl-tRNA synthetase marks the opisthokonts as monophyletic. *Mol Biol Evol.* 22:2142–2146.
- Jaffe AB, Hall A. 2005. Rho GTPases: biochemistry and biology. *Annu Rev Cell Dev Biol.* 21:247–269.
- Koumandou VL, Dacks JB, Coulson RM, Field MC. 2007. Control systems for membrane fusion in the ancestral

- eukaryote; evolution of tethering complexes and SM proteins. *BMC Evol Biol.* 7:29.
- Luo HR, Saiardi A, Nagata E, Ye K, Yu H, Jung TS, Luo X, Jain S, Sawa A, Snyder SH. 2001. GRAB: a physiologic guanine nucleotide exchange factor for Rab3A, which interacts with inositol hexakisphosphate kinase. *Neuron.* 31:439–451.
- Meller N, Merlot S, Guda C. 2005. CZH proteins: a new family of Rho-GEFs. *J Cell Sci.* 118:4937–4946.
- 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.
- Reyes-Prieto A, Bhattacharya D. 2007a. Phylogeny of Calvin cycle enzymes supports Plantae monophyly. *Mol Phylogenet Evol.* 45:384–391.
- Reyes-Prieto A, Bhattacharya D. 2007b. Phylogeny of nuclear encoded plastid targeted proteins supports an early divergence of glaucophytes within Plantae. *Mol Biol Evol.* 24:2358–2361.
- Rodriguez-Ezpeleta N, Brinkmann H, Burger G, Roger AJ, Gray MW, Philippe H, Lang BF. 2007. Toward resolving the eukaryotic tree: the phylogenetic positions of jakobids and cercozoans. *Curr Biol.* 17:1420–1425.
- Rutherford S, Moore I. 2002. The Arabidopsis Rab GTPase family: another enigma variation. *Curr Opin Plant Biol.* 5:518–528.
- Segev N. 2001. Ypt/rab gtpases: regulators of protein trafficking. *Sci STKE.* 2001:RE11.
- Simpson AG, Roger AJ. 2004. The real ‘kingdoms’ of eukaryotes. *Curr Biol.* 14:R693–R696.
- Steenkamp ET, Wright J, Baldauf SL. 2006. The protistan origins of animals and fungi. *Mol Biol Evol.* 23:93–106.
- Stiller JW. 2007. Plastid endosymbiosis, genome evolution and the origin of green plants. *Trends Plant Sci.* 12:391–396.
- Takai Y, Sasaki T, Matozaki T. 2001. Small GTP-binding proteins. *Physiol Rev.* 81:153–208.
- Thomas C, Fricke I, Scrima A, Berken A, Wittinghofer A. 2007. Structural evidence for a common intermediate in small G protein-GEF reactions. *Mol Cell.* 25:141–149.
- Tsang HT, Connell JW, Brown SE, Thompson A, Reid E, Sanderson CM. 2006. A systematic analysis of human CHMP protein interactions: additional MIT domain-containing proteins bind to multiple components of the human ESCRT III complex. *Genomics.* 88:333–346.
- Walch-Solimena C, Collins RN, Novick PJ. 1997. Sec2p mediates nucleotide exchange on Sec4p and is involved in polarized delivery of post-Golgi vesicles. *J Cell Biol.* 137:1495–1509.
- Zheng Y. 2001. Dbl family guanine nucleotide exchange factors. *Trends Biochem Sci.* 26:724–732.

Andrew Roger, Associate Editor

Accepted May 7, 2008