INSTITUT PASTEUR

Research in Microbiology 162 (2011) 77–91

www.elsevier.com/locate/resmic

A new fusion hypothesis for the origin of Eukarya: better than previous ones, but probably also wrong

Patrick Forterre [a,b,*]

[a] Institut Pasteur, 25 rue du Docteur Roux, 75015 Paris, France
[b] Institut de Génétique et Microbiologie, Univ. Paris-Sud - CNRS UMR 8621, 91405 Orsay Cedex, France

Abstract

I discuss here the possibility that Eukarya originated from the engulfment of a thaumarchaeon by a PCV (Planctomycetes, Verrucomicrobia, Chlamydiae) bacterium, followed by invasions of NCLDV and retroviruses. The thaumarchaeon provided both informational and operational proteins (actins, ESCRT proteins), including some essential proteins absent in other archaeal phyla (Topo IB), whereas the PVC bacterium provided phospholipids, tubulin and the membrane coat proteins required for the formation of the nucleus. Viral invasions introduced many proteins that are specific to modern Eukarya and produced an arms race that favoured the evolution of Eukarya toward increasing complexity. This scenario is the best possible fusion hypothesis that can be presently proposed. However, it still requires several ***ad hoc*** assumptions to explain the origin of the nucleus and the distribution pattern of archaeal and bacterial traits in modern Eukarya. Furthermore, it still fails to explain convincingly the origin of eukaryal viruses and the existence of three distinct lineages of ribosomes. I conclude that Eukarya and their viruses more probably evolved from a specific lineage, according to the three domains scenario originally proposed by Carl Woese.

Keywords: Archaea; Thaumarchaea; Planctomycetes; NCLDV; Retrovirus; Origin of eukaryotes

1. Introduction

The prokaryote/eukaryote dichotomy, set up in 1962 by Stanier and Van Niel (Stanier and Van Niel, 1962), based on cell biology, has been the dominant paradigm in biology for the last 50 years. It remains today accepted by most biologists despite the fundamental work of Carl Woese and colleagues who demonstrated in 1977, based on molecular biology, the existence of three (and not two) major evolutionary cellular lineages, Archaea, Bacteria and Eukarya (Woese and Fox, 1977; Woese et al., 1990). The procaryote paradigm indeed satisfies the traditional idea that evolution always progresses from simple organisms to complex organisms (see a critique of this idea in Forterre and Philippe, 1999). It is inherent in the terms prokaryote/eukaryote themselves, since « *pro* » karyotes implies before « *eu* » karyotes (true nucleus) (Forterre, 1995; Pace, 2009). I suggested in 1992 to replace the terms prokaryote and eukaryote by neutral ones (evolutionary speaking) such as akaryote and synkaryotes (without and with nucleus) to designate the two types of known cellular organization (Forterre, 2002), but this proposal was completely ignored. Another reason explaining the resistance of the prokaryote paradigm is that we (human beings) like dichotomies, so the prokaryote/eukaryote dichotomy is regularly revived, like the phoenix, even if it is contradicted by facts.

In recent years, the prokaryote/eukaryote dichotomy has been on the agenda again under a new appearance, i.e. "fusion" hypotheses in which Eukarya originated from the association of an archaeon and a bacterium (see discussions in Poole and Penny, 2007; Gribaldo et al., 2010; Cavalier-Smith, 2010; Koonin, 2010a,b, Forterre, 2010b). These hypotheses are based on the observation that Eukarya and Archaea share similar informational mechanisms (translation, transcription, replication) very distinct (sometimes non homologous) from

---

* Institut Pasteur, 25 rue du Docteur Roux, 75015 Paris, France. Tel.: +33 169157489.
E-mail address: forterre@pasteur.fr.

0923-2508/$ - see front matter © 2010 Institut Pasteur. Published by Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.resmic.2010.10.005

those of Bacteria, whereas Eukarya and Bacteria share a similar type of membrane phospholipid, very distinct from those of Archaea (see below). Bacteria and Eukarya also exclusively share many proteins (often involved in so-called operational mechanisms) whose origin is disputed. Some authors suppose that these proteins have been all introduced in Eukarya either by the alpha-proteobacterium at the origin of mitochondria or by other bacteria used as preys by ancient Eukarya. In contrast, proponents of "fusion" hypotheses argue that these proteins testify for the existence of a specific bacterium at the origin of Eukarya, whose proteome was supplemented by the informational proteins of an archaeon.

The fusions scenarios that have been proposed by different authors involve various types of associations, symbioses, syntrophy, etc. often justified by metabolic considerations to explain why the two partners first of all associated. I will put all these types of combinations here under the broad term "fusion" for the sake of simplicity. Proponents of fusion scenarios for the origin of Eukarya usually still use the old name archaebacteria, instead of Archaea, thus perpetuating the confusion between Bacteria and prokaryotes. In that scheme, Bacteria (as a general term encompassing eubacteria and archaebacteria) regain the primitive position that they have occupied in the tree of life from the XIX century (the Haeckel's monera). In fusion scenarios, the three primary domains of Woese and colleagues are thus replaced by two primary "prokaryotic" domains, and "eukaryotes" therefore emerged once more after prokaryotes, as a derived secondary domain (see discussion in Gribaldo et al., 2010).

Proponents of fusion scenarios on the origin of Eukarya often play of the confusion between Eukarya in general (including those which lived before the acquisition of mitochondria) and modern Eukarya. The latter indeed originated after Bacteria, since their last common ancestor already had *bona fide* mitochondria. However, in the three domains scenario, the primordial eukaryal lineage (urkaryote *sensu* Woese and Fox, 1977) does not correspond to modern Eukarya, but to a lineage of ancient Eukarya that led to LECA (the Last Eukaryal Common Ancestor) (Fig. 1A). This lineage might be as ancient as those that led to Bacteria and Archaea. In this scenario, FECA (the First Eukaryal Common Ancestor) might have even predated the divergence of the various phyla of modern archaea and bacteria (Fig. 1A). In contrast, in fusion scenarios, FECA only appeared after the diversification of the archaeal and bacterial domains (Fig. 1B). Indeed, with few exceptions, proponents of fusion scenarios for the origin of Eukarya have always proposed that the two partners that associate were close relatives of modern Bacteria and Archaea.

Historically, various archaeal and bacterial partners have been proposed by different authors to sell convincingly their particular fusion hypotheses. Lynn Margulis, suggested in the eighties that the archaeal partner was a relative of *Thermoplasma acidophilum* (Margulis and Bermudes, 1985), a member of the archaeal phylum Euryarchaeota, because *T. acidophilum* is a wall-less microorganism in which "eukaryotic-like histones" were described (biochemically) in the seventies (Stein and Searcy, 1978). In her hypothesis, the bacterial partner was a Spirochete because these bacteria harbour microtubule-like structures reminiscent of the eukaryal cytoskeleton (Margulis et al., 1979). Later on, Rivera and Lake suggested an archaeon of the phylum Crenarchaeota for the archaeal partner, because Crenarchaeota (called Eocytes by Lake) were sister groups of Eukarya in their phylogenetic analyses (Rivera and Lake, 1992, 2004). More recently, in the "*ring of life*" scenario, they identified the bacterial partner as a relative of Proteobacteria and/or Cyanobacteria (Rivera and Lake, 2004). In the so-called, "*hydrogen hypothesis*", Martin and Muller suggested an unusual version of fusion hypotheses in which the bacterial partner corresponds to the alpha-proteobacterium at the origin of mitochondria (Fig. 1B, arrow b). In their scenario, the "archaebacterium" is an unspecified autotrophic hydrogen-dependent anaerobe (Martin and Müller, 1998). One of the more detailed and original fusion hypothesis is the "*syntrophic hypothesis*" proposed by Moreira and Lopez-

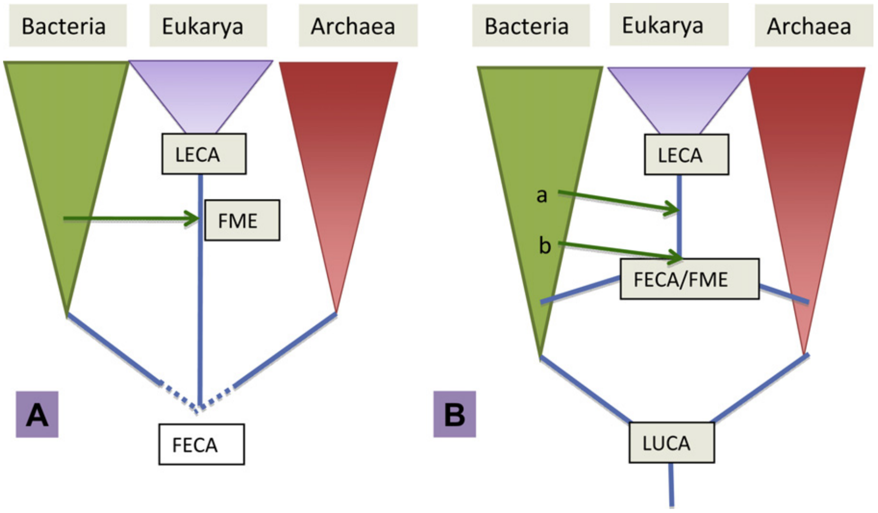

Fig. 1. Two alternative scenarios for the origin and evolution of modern domains. A: Scenario based on the three domains concept (dotted lines symbolize an unrooted tree), B: Fusion scenario based on the prokaryote/eukaryote dichotomy. The green arrows represent the origin of the mitochondrion from an alpha-proteobacterium. FECA: First Eukaryotic (specific) Common Ancestor, LECA: Last Eukaryotic Common Ancestor, FME: First Mitochondrial Eukaryote.