INTRODUCTION AND REASONING
ON A HYPOTHETICAL AND
VERY SINGULAR ORGANISM

The classification of living organisms presents numerous problems related to the multiple levels on which the criterion for carrying out the classification itself might be based. For instance, the classification based on protein and gene sequences has identified the three domains of life: Archaea, Bacteria and Eukarya (Winker & Woese, 1991) but these do not seem to be monolithic on the basis of the sequences themselves (Lake, 1987). By contrast, the “five kingdoms” scheme and the prokaryote-eukaryote dichotomy, maintained by Margulis & Guerrero (1991) and by Mayr (1998), respectively, are based on criteria that are different from those based on sequences of macromolecules. These criteria may be different in nature such as, for example, the one suggested by Cavalier-Smith (2010) who bases his classification on the topology and the chemistry of membranes and reaches the conclusion that eubacteria were the only ones in a direct relationship with the earliest forms of life and from which archaeabacteria and eukaryotes evolved much later on.

A classification problem also seems to stem from the following hypothetical finding. Let us assume that we have found an organism that still possesses a trait in an evolutionary state of transition, i.e. a paleokaryote (an organism that, for example, uses hairpins instead of complete tRNAs in the synthesis of its proteins), the aim of this paper is to clarify the singular, ancestral and, in some cases, unique characters possessed by Nanoarchaeum equitans. The absence of operons in the genome of N. equitans is considered as an ancestral transitional trait, which analysis equates to hairpins used by the hypothetical paleokaryote and thus leads to the following conclusions: (i) N. equitans might be the representative of a new phylum of Archaea; (ii) it is a living fossil; and (iii) it represents the root of the Archaea domain or of the tree of life (i.e. rooted in the Nanoarchaeota phylum). All these conclusions seem to be likely in light of the analysis here conducted. Whereas, it is not equally clear whether N. equitans may be considered a true paleokaryote and is the representative of a new domain of life, although these two possibilities might be supported by a further analysis of its biology.

Key words: taxonomy, biological classification, domains of life, ancestral traits.

Hypothesising the finding of an organism that still possesses a trait in an evolutionary state of transition, i.e. a paleokaryote (an organism that, for example, uses hairpins instead of complete tRNAs in the synthesis of its proteins), the aim of this paper is to clarify the singular, ancestral and, in some cases, unique characters possessed by Nanoarchaeum equitans. The absence of operons in the genome of N. equitans is considered as an ancestral transitional trait, which analysis equates to hairpins used by the hypothetical paleokaryote and thus leads to the following conclusions: (i) N. equitans might be the representative of a new phylum of Archaea; (ii) it is a living fossil; and (iii) it represents the root of the Archaea domain or of the tree of life (i.e. rooted in the Nanoarchaeota phylum). All these conclusions seem to be likely in light of the analysis here conducted. Whereas, it is not equally clear whether N. equitans may be considered a true paleokaryote and is the representative of a new domain of life, although these two possibilities might be supported by a further analysis of its biology.

Key words: taxonomy, biological classification, domains of life, ancestral traits.
in this hypothetical organism to achieve protein synthesis. Clearly, these RNA hairpin structures on which protein synthesis is carried out in this organism might also be the ancestral (plesiomorphic) forms, i.e. the precursors of the tRNA molecule which in all other organisms were presumably substituted by the tRNA cloverleaf structure. We also recognise their ancestry in the observation that it would have been practically impossible to derive hairpin molecules from all complete tRNA molecules, because this would have required a formidable selective pressure in favour of the hairpin, which seems somewhat difficult to find (Di Giulio 2006a, 2006b, 2009b). In conclusion, I hypothesise that a unicellular organism has been found having a single molecule still in a transitional stage (i.e. the hairpin) which is clearly distinguishable from the molecule’s final form (the complete tRNA molecule) and which can also be clearly seen as a primitive form (the hairpin) compared to the form used in all other organisms (the tRNA molecule), thus allowing us to establish its ancestry in an intuitive and rigorous way.

How can we classify this organism? We can certainly say that it is a paleokaryote, by which we mean that it is an organism presenting ancestral traits still in a transitional stage (Di Giulio 2006a, 2011), i.e. used by the primordial ‘system’ and subsequently evolving into the modern forms found in other organisms. However, this recognition of the state of paleokaryoticity is more formal than substantial as it would seem to say little on how this organism should be classified. More directly, what is its real relationship with the other domains: is it a new domain of life or not? By definition, a paleokaryote possesses at least one plesiomorphic trait still in transition. Hence, this organism ought also to be a new domain of life because that primitive transitional trait is unique, singular, phylogenetically deep and, certainly, highly differentiating and comparable to ‘first level’ differences – as are the differences existing between the membranes of Bacteria and Archaea, the absence/presence of the nucleus between prokaryotes and eukaryotes, and other phylogenetically deep characteristics – and such as to define that organism as a new domain of life, or at least as an organism that, if part of one of the three domains of life, should lie at the root of one of them. In other words, the fact that the paleokaryote possesses a primitive transitional trait is such as to imply that this is part of a new domain of life. This is because that trait would make it unclassifiable in the other domains since this characteristic might not only be unshared by any organisms from the other domains, but it might be so idiosyncratic and singular in nature as to set the hypothetical organism outside the other classification schemes. It should be pointed out, however, that the hypothetical organism with a trait still in a transitional stage might end up sharing a high number of ‘phylogenetically deep’ characters with just one of the three domains and, therefore, this organism represents the root of this domain (or, more precisely, the root of the tree of life) rather than constituting a new domain of life. That is to say, we should consider a quantitative aspect in addition to the qualitative aspect so far taken into consideration.

It should also be specified that a paleokaryote, as such, should manifest other singular and unique traits because it would be unlikely that a paleokaryote would be defined by just one. Therefore, these other singular traits, if present, should better define whether the paleokaryote can be classified in a new domain of life, or not.

Following the later specifications which seem to clarify the relationship between the paleokaryote and the other domains of life, it is worth stressing that a plesiomorphic transitional trait, e.g. hairpins instead of tRNAs, would most probably set this organism outside the classification schemes because such a trait would have no equivalents in the entire biosphere, thus almost certainly placing the paleokaryote at the root of the tree of life and, I believe, in a new domain of life.

**NANOARCHAEM EQUITANS IS A VERY SINGULAR ORGANISM**

On the basis of the small subunit of ribosomal RNA, *Nanoarchaeum equitans* was identified as a new phylum of Archaea (Huber et al., 2002). In the identification of *N. equitans*, difficulties were already identified because the probes normally used to amplify ribosomal RNA turned out to be ineffectual in amplifying the rRNA of *N. equitans* (Huber et al., 2002). Furthermore, its rRNA sequence was unique and singular among the Archaea although it presented some secondary structures typical of Archaea (Huber et al., 2002). When its genome was sequenced (Waters et al., 2003) it became clear that *N. equitans* presented a truly unusual set of characters. Firstly, the large number of split genes: at least eleven proteins are split in *N. equitans*, i.e. a protein such as alanyl-tRNA synthetase, which is normally codified in a single gene is, in *N. equitans*, codified in two completely different
genes (Waters et al., 2003). Coherently, six tRNA genes are split in the sense that the tRNA molecule, which is normally codified in a single gene, is codified in \textit{N. equitans} in two genes codifying only half of the tRNA molecule and located in non-contiguous sites on its genome (Randau et al., 2005). Consistent with these first two characteristics of the genome of \textit{N. equitans} is the observation of the almost total absence of conserved operons on its genome (Waters et al., 2003; Makarova & Koonin, 2005). For instance the super-operon of ribosomal proteins, which is conserved in all the Archaea and Bacteria, is almost totally absent in \textit{N. equitans} and only few fragments are present (Makarova & Koonin, 2005). Therefore, \textit{N. equitans} seems to be the only nearly operon-less prokaryote (Makarova & Koonin, 2005).

\textit{Nanoarchaeum equitans} has not been extensively studied and is therefore not well-characterised at a molecular level, but many observations stress its singularity as follows: (i) the absence of RNase P, the enzyme that universally takes part in the maturation of the tRNA molecule (Randau et al., 2008a; Lai et al., 2010); (ii) one of the two archaenal histones possesses, in \textit{N. equitans}, a unique four-residue insertion which closely resembles the one found in the eukaryotic histones and would therefore seem to be an intermediary towards the H3 histones typical of Eukarya (Friedrich-Jahn et al., 2009); and (iii) the B DNA polymerase of \textit{N. equitans} seems to have very unusual characteristics in that it would seem to utilise deaminated bases as uracyl (which these polymerases are normally unable to use) (Choi et al., 2008).

**THE SINGULAR TRAITS OF \textit{NANOARCHEAUM EQUITANS} ARE ALSO ANCESTRAL CHARACTERS**

\textbf{The tRNA split genes}

The split genes of tRNAs of \textit{Nanoarchaeum equitans} have been shown, through a ‘mathematical’ proof, to be the ancestral form of tRNA genes (Di Giulio, 2009b) and there are numerous other arguments in favour of this hypothesis (Di Giulio, 2006a, b, 2008a, b, c). Contrary to this conclusion is the hypothesis of Randau & Söll (2008), who maintain that the region of the anticodon loop of tRNA genes became the attachment site of an enormous variety of mobile genetic elements and that this resulted in the evolution of tRNA split genes. Therefore, they proposed that the universal presence of the intron in the anticodon loop of tRNA genes is not an ancestral but a derived trait, since this provided tRNA genes with a precious protection mechanism against the integration of viruses and autonomous genetic elements in that the intron removed the integration site from these mobile genetic elements. I have already criticised Randau and Soll’s hypothesis (Di Giulio, 2008c, 2009a, b). What I wish to add here is that in the majority of Archaea (about 90%), and also in \textit{N. equitans} and in a high percentage of Bacteria (about 40%), an immune system (the CRISPR/Cas system) exist; its specific function is to combat and neutralise all kinds of mobile genetic elements (Haurwitz et al., 2010). Therefore, the hypothesis of Randau & Söll (2008) would be questioned since the mechanism on which it is founded (the integration of mobile genetic elements in the anticodon loop of tRNA genes) and which gave rise to the piece genes of tRNA, should not have been a strong selective pressure promoting the evolution of tRNA split genes because the majority of autonomous genetic elements are removed from these organisms by means of the CRISPR/Cas system (Haurwitz et al., 2010). In other words, the very existence of the CRISPR/Cas system would greatly weaken the hypothesis of Randau & Söll (2008) because it would deprive it of the selective pressure (the integration of mobile genetic elements in the anticodon loop of tRNA genes) as these elements would be eliminated primarily by the CRISPR/Cas system and not, as suggested by Randau & Söll (2008), by means of their integration in the anticodon loop of tRNA genes.

\textbf{The split genes of proteins}

The split genes of the proteins in \textit{Nanoarchaeum equitans} have been recognised as the plesiomorphic form of these genes in an analysis in which the point where these genes are split was used to predict the position of introns in the homologous eukaryotic genes (Di Giulio, 2008d). Indeed, in agreement with the exon theory of genes (Gilbert et al., 1997), the introns played a fundamental role in assembling early genes and, therefore, the successful identification of the position of the introns in eukaryotic genes, on the basis of the homologous genes of Archaea and of the split genes of \textit{N. equitans}, would define the ancestrality of the latter genes (Di Giulio, 2008d).

\textbf{The absence of operons in the genome of \textit{Nanoarchaeum equitans}}

The almost total absence of operons in the genome of \textit{N. equitans} (Waters et al., 2003; Makarova & Koonin,
and the split genes of tRNAs and proteins seem to be two sides of the same coin (Di Giulio, 2007, 2008b). Indeed, they might represent the manifestation of the same evolutionary stage in which the ancestral genomes find themselves, since both the split genes and the absence of operons would seem to indicate that the genes and parts of genes were not yet joined in N. equitans, thus testifying to the ancestrality of the genome of N. equitans (Di Giulio, 2007, 2008b). It is more natural to think that ancestral genomes did not have operons because the latter seem to be highly evolved aggregates of genes and thus suitable for responding to the slightest environmental variations. Moreover, the idea that ancestral genomes already possessed operons does not seem sensible because, in the evolutionary transition from RNA genomes to DNA genomes, the operons should not have formed immediately since primarily DNA genomes were evolving. Operons evolved only later on, evidently to better respond to environmental perturbations. Therefore it is more likely that ancestral genomes had scattered genes and not gene aggregates like operons, and thus, according to this reasoning, both the split genes and the absence of operons might be ancestral traits and hence two sides of the same coin (see this point also in Di Giulio, 2008b).

On the other hand, the view that operons might not be ancestral traits is reviewed by Fani et al. (2005).

Forterre et al. (2009) maintain that the absence of operons in N. equitans is a derived trait because Igno
coccus hospitalis, the host of N. equitans, presents a ‘similar’ situation in which 180 gene clusters (typically conserved in Archaea) are disrupted in I. hospitalis (Podar et al., 2008). The fact that the genome of N. equitans has very few or no operons (Makarova & Koonin, 2005) while that of I. hospitalis has disrupted operons (Podar et al., 2008) does not seem, in my view, to be the consequence of a reduction process in their genomes, as maintained by Forterre et al. (2009) but it is instead due to the absence in these two organisms of transposable genetic elements which impeded the movement of genes, and thus the formation of operons in N. equitans, favouring the formation of disrupted operons in I. hospitalis. I must here point out, more generally, that the split genes of tRNAs and proteins have also been observed in other archaea (Di Giulio, 2008d; Fujishima et al., 2009), but no organism has these three characteristics simultaneously except N. equitans, which has indeed been described as a molecular fossil (Di Giulio, 2006b).

### Ribonuclease P

Ribonuclease P (RNase P) is a ribonucleoprotein essential for the maturation of the 5' end of tRNAs. The catalytic component of RNase P is an RNA that is universally conserved and the protein components between Bacteria and Archaea domains are not homologous, while those between Archaea and Eukarya are partly homologous (Hartmann & Hartmann, 2003).

Randau et al. (2008) suggested that the leader region of tRNAs was possessed by the ancestor of N. equitans, i.e. it is an ancestral trait, above all on the basis of the universality of the RNA component of RNase P. However, as mentioned above, the protein components of RNase P are not homologous between the Archaea and Bacteria domains, which would seem to indicate a late evolutionary phase for the evolutionary completion of the structure of RNase P. However, as the catalytically active component of this enzyme is homologous between Archaea and Bacteria, this would seem to favour its function on the leader sequences of tRNAs and therefore infer a presumed ancestrality. It must nevertheless be taken into account that the leader sequences of tRNAs were most probably not present in early genomes because, in the RNA -> DNA transition, it would have been surprising if the ancestral tRNAs already had leader sequences because these seem to play a regulatory function that, in an RNA world, does not seem to have any great value. Therefore, I believe that the tRNAs of N. equitans without leader sequences are the plesiomorphic condition and that, at this evolutionary stage, the RNA component of RNase P played at least a partly different role from its current one or presented a function similar to this but on a different molecule, such as 5S rRNA. And it was only with the evolution of the protein components and with the origin of leader sequences that the function of RNase P evolved into the maturation of the 5’ end of tRNAs. If this is true, both the absence of leader sequences of tRNAs and the absence of RNase P in N. equitans would be plesiomorphic traits, consistently with the ancestrality of tRNA split genes.

### Synthesis

We would have no doubt in recognising an organism that used only hairpin structures instead of tRNAs for its protein synthesis as a paleokaryote, whereas we would have great difficulty in recognising N. equitans as a paleokaryote only on the basis of the absence of operons in its genome. If the absence of operons is
that when the first DNA genomes were formed (without operons), this would imply a still rapidly and progressively evolving situation, i.e. with a tempo and a mode more typical of a progenote than of a genote (Di Giulio, 2011). Hence, an ancestral genome without operons might have belonged to a progenote. If this were true, then a progenote would by definition also have been a paleokaryote because a still evolving genotype-phenotype relationship would undeniably imply that the progenote possessed ancestral transitional traits and, therefore, elements that fall within the definition of paleokaryote. This leads us to the conclusion that the absence of operons might be a trait possessed by a paleokaryote. Therefore, as *N. equitans* has no operons, it shows that it has at least one trait that is characteristic of a paleokaryote.

**CONCLUSIONS**

The possibility that *N. equitans* possesses at least one trait, the absence of operons belonging to a paleokaryote, might, in the prokaryotic world be an index of certain paleokaryoticity. What is more difficult to establish is whether *N. equitans* is a true paleokaryote. Even if the absence of operons is not immediately perceived as a transitional trait, unlike hairpins in protein synthesis, this could, if ancestral, define *N. equitans* as a true paleokaryote, partly in consideration of the presence of singular and unique traits as well as others that are certainly ancestral (see above). Whereas, the weaker conclusions, i.e. that *N. equitans* is a new phylum of Archaea (Huber et al., 2002) or is a living fossil (Di Giulio, 2006b) or represents the root of the Archaea domain or of the tree of life (i.e. rooted in the Nanoarchaeota phylum) (Di Giulio, 2007), all seem highly likely, even if contrary to what has been reported in the literature (Brochier et al., 2005; Marakova & Koonin, 2005; Forterre et al., 2009). In conclusion, this work suggests that there is a possibility that *N. equitans* may be a paleokaryote and, perhaps, also the representative of a new domain of life, or the root of the tree of life, but this will become clear only with further analysis of its biology.

**REFERENCES**


Cavalier-Smith T, 2010. Deep phylogeny, ancestral groups


Di Giulio M, 2008c. Transfer RNA genes in pieces are an ancestral character. EMBO Reports, 9: 820.


Di Giulio M, 2011. The Last Universal Common Ancestor (LUCA) and the ancestors of Archaea and Bacteria were progenotes. Journal of Molecular Evolution, 72: 119-126.


