Journal of Natural Sciences
December 2015, Vol. 3, No. 2, pp. 01-16
ISSN 2334-2943 (Print) 2334-2951 (Online)

Published by American Research Institute for Policy Development
DOI: 10.15640/jns.v3n2a1
URL: http://dx.doi.org/10.15640/jns.v3n2a1

# Approach to a Comparative Study of the Metabolism of Porphyrins and Chlorophylls

**Pérez-Urria[1], Elena & Avalos, Adolfo[1]**

**Abstract**

Metabolic pathways are series of successive biochemical reactions catalyzed by enzymes which together constitute a process of extraordinary complexity, the cellular metabolism. Like any other biological phenomenon, metabolism is the result of evolution. Different techniques are used to compare characteristics of living entities seeking to analyze the similarity between them. The features or characteristics that define species are of different types: molecular, metabolic, cellular, environmental, behavioral, etc. Depending on the similarity, biological species and any entity of the biological hierarchy, are sorted and classified. This paper focuses on studying a characteristic of organisms that is metabolism of porphyrins and chlorophylls using cladistic procedures, unlike others, allow crawl the changes in this feature. The result of a cladistic analysis is an evolutionary hypothesis, a hypothesis about the evolution of metabolism of porphyrins and chlorophylls. The work raises two objectives: firstly, hypotheses about the evolution of metabolism of porphyrins and chlorophylls; otherwise, hypotheses on the evolution of enzymes involved in cellular metabolism this part, opening the possibility that some of them serve as phylogenetic marker.

**Keywords:** porphyrins and chlorophylls metabolism, cladistics, biological evolution

## 1. Introduction: A look back in time

Metabolic pathways are a series of successive biochemical reactions catalyzed by enzymes which together constitute a process of extraordinary complexity, cellular metabolism. Like any other biological phenomenon, metabolism is the result of evolution.

[1] Department of Plant Biology I (Plant Physiology), Faculty of Biology, Complutense University of Madrid (Spain)

As the history of the species and their family relationships are derived from comparative studies, the history of metabolism can only be reconstructed through comparative analysis of its elements (Cunchillos and Lecointre, 2002, 2003, 2005).

Using the cladistic analysis to build an evolutionary hypothesis of the emergence and evolution of porphyrin and chlorophyll metabolism involves considering photosynthesis from the evolutionary point of view. On the origin of photosynthesis there are basically two hypotheses (Olson, 1999, 2001): one suggests that it originated in the prebiotic atmosphere of the Earth and comes to life; another, based on recent molecular phylogenetic analysis suggests that photosynthesis arose after chemolithotrophs organisms appear. However, we know little about how they were incorporated into the photosynthetic process its multiple biochemical components.

The scientific community agrees that photosynthesis is a process that originates in bacteria because, firstly, there is no archaea with Mg-tetrapyrroles based photosystems and, secondly, photosynthetic eukaryota acquired this ability from cyanobacteria through endosymbiosis. It is possible to achieve a better understanding of the evolution of photosynthesis comparatively studying genes or gene products of photosynthesis in the groups of bacteria that synthesize photosystems: purple bacteria (proteobacteria, with photosystem II), green non-sulfur bacteria (photosystem II) , green sulfur bacteria (photosystem I), heliobacterias (photosystem I) and cyanobacteria (photosystems I and II). The Earth is about 4,500 million years and the beginning of life is in the Archean period, about 3800-4000 million years ago. Stromatolites (the oldest Archaean fossils) containing evidence of biological carbon fixation have an approximate age of 3,800 million years (Schidiowski, 1988) while the oldest cyanobacteria-like cell microfossils have an age of 3,500 million years (Schopf, 1993; Schopf and Packer 1987), ie, appearing 300 million years after the first record of life.

However, Earth was dominated at first by cyanobacteria which were 2,700 million years ago, coinciding with the first appearance of oxygen on earth (Des Marais, 2000). During approximately 1 million years, cyanobacteria increased the oxygen level to almost a quarter of the current level. The success of these organisms could be due not only to the power advantage of photosynthesis but due also to inhibition by oxygen (as toxicant) of potential competitors.

The emergence of the first eukaryotes occurred about 1,800 million years ago, a fact that required 600-800 million years for cyanobacteria were incorporated by endosymbiosis in eukaryotes and evolved into chloroplasts. Thereafter the algae increased the oxygen level until present. Finally, his descendants land plants appeared 500 million years ago.

There is general consensus on the hypothesis suggesting the evolution of photosynthetic pigments from chemoautotrophs: they lived in an environment of chemical imbalance that led to the ability to use the pigments to harness light as an additional source of energy. Established photosynthesis, this might evolve to allow cells use sunlight as the only energy source (Nisbert and Sleep, 2001). On this tour it must be some intermediate stage that could be starring by purple bacteria: the anoxygenic photosynthesis bacteria evolve from this group who had phototaxis by infrared light (Nisbert et al., 1995).

This hypothesis is based on the close relationship between the emission spectrum of geothermal light and absorption spectrum of bacteriochlorophylls b, relationship that serves the authors to raise the possibility that photosynthesis arose in organisms presenting bacteriochlorophylls a or b. They lived in the vicinity of oceanic hydrothermal vents where they could detect a faint infrared radiation.

Accordingly, these phototaxis bacteria with the ability to detect infrared light have lived in an optimum environment which would have been an evolutionary advantage in terms of competition for resources or nutrients. Subsequently adaptation of this primitive photosystem would have allowed organisms to start using the far-red sunlight as it moved into shallower waters. Over time, chlorophylls appear to use light of higher energy (visible) and water photolysis. Photosynthesis in bacteria and plants occurs in two phases, photochemistry, and biochemistry. Photochemical phase involves the participation of photosynthetic pigments including chlorophylls and bacteriochlorophylls of bacteria and plants which are derived from the porphyrin metabolism.

## 2. Map of metabolism of porphyrins and chlorophylls

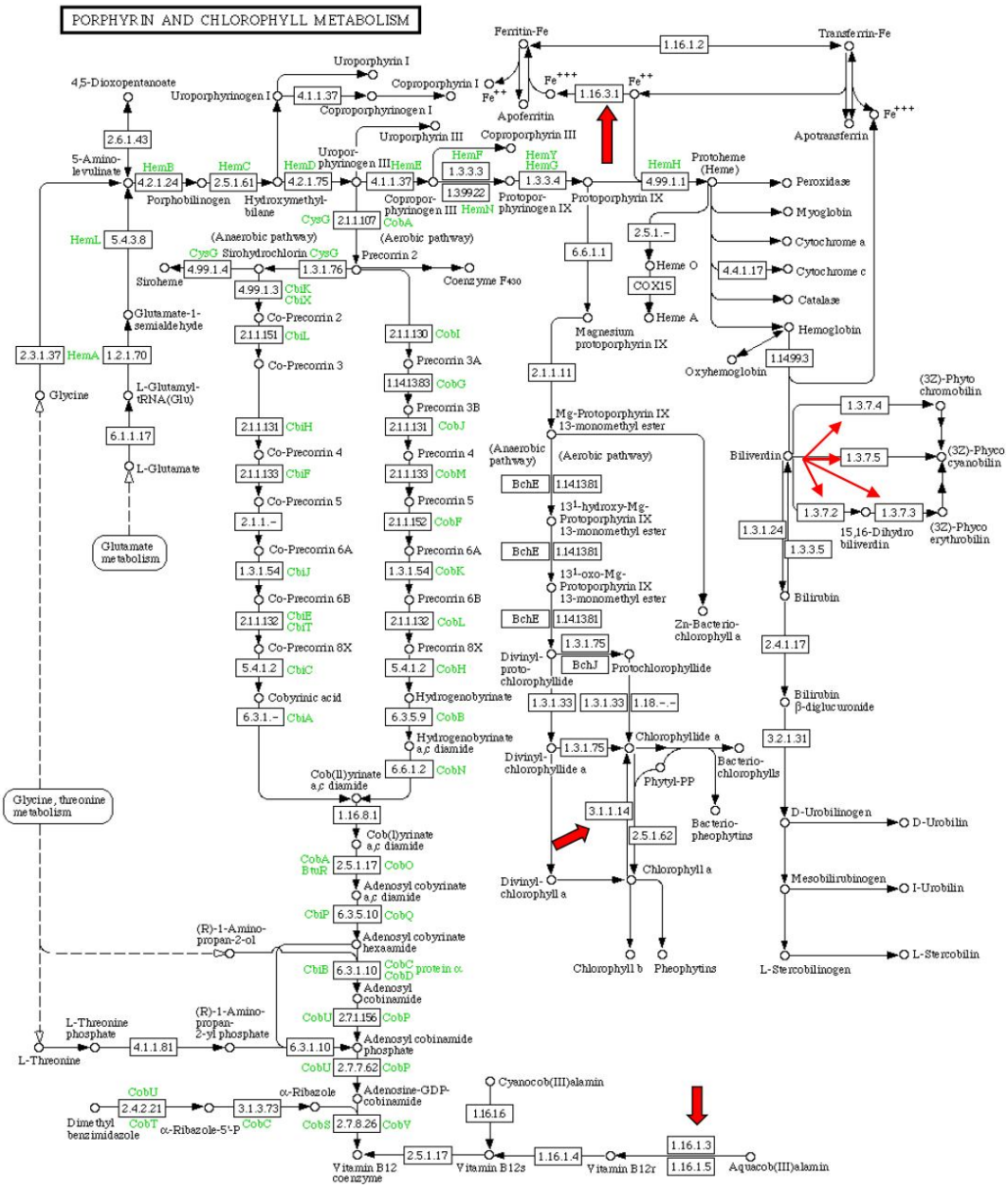
Porphyrins are tetrapyrroles binds covalently to a metal: iron (Fe) to form cytochromes, peroxidase, catalase, myoglobin and hemoglobin; copper (Cu) or nickel (Ni) to form molecules for electron transport in methanogenic bacteria; magnesium (Mg) to form chlorophylls and bacteriochlorophylls.

In the porphyrin biosynthesis pathway (Fig. 1) there are two stages:

1) Synthesis of 5-aminolevulinate (ALA): In all photosynthetic eukaryota and prokaryota, except the α-proteobacteria, ALA is synthesized from glutamate which binds glutamyl-tRNA by glutamyl- tRNA synthase catalyzed reaction (GluRS). Moreover, all non-photosynthetic eukaryota (animals, fungi and apicomplexa) and α-proteobacteria form ALA by condensing glycine and succinyl-CoA in a reaction catalyzed by ALA synthase.

2) Synthesis of protoporphyrin IX from 5-aminolevulinate (ALA), a chain of six reactions, the same in all organisms

**Figure 1: Porphyrins and chlrophylls biosynthesis pathway (Modified from GenomeNet) (red arrows indicate enzymes that undergo a single change in the analysis)**



In the context of evolution it is important to highlight the conceptual difference between "evolution of photosynthesis" and "evolution of photosynthetic organisms": the evolution of photosynthesis involves a limited number of genes and / or gene products while the evolution of photosynthetic organisms involves the complete genome. The phylogeny of the 16S subunit rRNA (Woese, 1987) establishes three domains: Archaea (archaebacteria), Bacteria (eubacteria) and Eucarya (eukaryotes). Although this classification was discussed by other authors (Cavalier-Smith, 1992), the fact is that other phylogenetic studies using new genomic and sequence data provide consistent results with all three domains. However, when phylogeny based on vertical inheritance genes (16S rRNA) and photosynthesis gene phylogenies are compared, inconsistencies appear showing that the former do not necessarily reflect the second: evolution of organisms *versus* evolution of genes.

## 3. Cladistic analysis to approach the evolution of porphyrins metabolism

For the comparative study of various types of entities (molecules, metabolism, organs, organisms, populations, behaviors, geographic distribution, ecosystems) different techniques are used all aimed at analyzing the similarity between the strains under study. Between different procedures, cladistic analysis provides a high degree of objectivity to study but all techniques provide useful and necessary information in the evolutionary context.

The porphyrins and chlorophyll metabolism is a complex process involving numerous chemical reactions catalyzed by enzymes. Its structure (components and functions) in the diversity of organisms is a product of evolution and can only be reconstructed by comparing the components. Tetrapyrroles serve as electron carriers in all domains of life. Heme carrying proteins were postulated to have been present in the last common ancestor of Bacteria and Archaea.

The application of the comparative method requires first determine the set of elements to compare. These elements are referred to as cladistic terminology "operational taxonomic units". They can refer to molecules or species, taxa in general. In this particular case, the set of taxa to be compared is the metabolism of porphyrins and chlorophylls in a set of organisms that constitute the study group (Table 1). Taxon, that is, each organism metabolism study group, is defined by characters that are enzymes (Table 2) involved in this part of cellular metabolism (Fg.1).

Enzymes can be present or absent in a particular taxon (the bchl gene is absent in archaebacteria and present in bacteria) which means that the same character may have different "versions" which are called "character state".

**Table 1: Study Group, metabolism of porphyrins and chlorophylls in 63 species**

| ACRONYM | NAME OF THE SPECIES | DOMAIN | GROUP |
|---|---|---|---|
| **afu** | *Archaeoglobus fulgidus* | Archaea | euryarchaea |
| **ago** | *Ashbya gossypii* | Eukarya | fungi |
| **ana** | *Nostoc* | Bacteria | cyanobacteria |
| **ape** | *Aeropyrum pernix* | Archaea | crenarchaea |
| **ath** | *Arabidopsis thaliana* | Eukarya | plant |
| **atu** | *Agrobacterium fabrum* | Bacteria | α-proteobacteria |
| **bja** | *Bradyrhizobium diazoefficiens* | Bacteria | α-proteobacteria |
| **bsu** | *Bacillus subtilis* | Bacteria | firmicutes |
| **buc** | *Buchnera aphidicola* | Bacteria | γ-proteobacteria |
| **cac** | *Clostridium acetobutylicum* | Bacteria | firmicutes |
| **cal** | *Candida albicans* | Eukarya | fungi |
| **cau** | *Chloroflexus aurantiacus* | Bacteria | chloroflexi |
| **cch** | *Chlorobium chlorochromatii* | Bacteria | chlorobi |
| **cel** | *Caenorhabditis elegans* | Eukarya | metazoa |
| **cho** | *Cryptosporidium hominis* | Eukarya | alveolata |
| **cme** | *Cyanidioschyzon merolae* | Eukarya | rodofite |
| **cne** | *Cryptococcus neoformans* | Eukarya | fungi |
| **cte** | *Chlorobium tepidum* | Bacteria | chlorobi |
| **ddi** | *Dictyostelium discoideum* | Eukarya | amebozoa |
| **det** | *Dehalococcoides mccartyi* | Bacteria | chloroflexi |
| **dge** | *Deinococcus geothermalis* | Bacteria | deinococus |
| **dme** | *Drosophila melanogaster* | Eukarya | metazoa |
| **eco** | *Escherichia coli* | Bacteria | γ-proteobacteria |
| **gox** | *Gluconobacter oxydans* | Bacteria | α-proteobacteria |
| **gvi** | *Gloeobacter violaceus* | Bacteria | cyanobacteria |
| **hal** | *Halobacterium sp. NRC-1* | Archaea | euryarchaea |
| **hbu** | *Hyperthermus butylicus* | Archaea | crenarcheae |
| **hin** | *Haemophilus influenzae* | Bacteria | γ-proteobacteria |
| **lic** | *Leptospira interrogans* | Bacteria | spirochetae |
| **lpf** | *Legionella pneumophila* | Bacteria | γ-proteobacteria |
| **mac** | *Methanosarcina acetivorans* | Archaea | euryarchaea |
| **mca** | *Methylococcus capsulatus* | Bacteria | γ-proteobacteria |
| **mga** | *Mycoplasma gallisepticum* | Bacteria | mollicutes |
| **mja** | *Methanocaldococcus jannaschii* | Archaea | euryarchaea |
| **mka** | *Methanopyrus kandleri* | Archaea | euryarchaea |
| **mth** | *Methanothermobacter thermautotrophicus* | Archaea | euryarchaea |
| **mtu** | *Mycobacterium tuberculosis* | Bacteria | actinobacteria |
| **neq** | *Nanoarchaeum equitans* | Archaea | nonarchaea |
| **neu** | *Nitrosomonas europaea* | Bacteria | β-proteobacteria |
| **nwi** | *Nitrobacter winogradskyi* | Bacteria | α-proteobacteria |
| **pab** | *Pyrococcus abyssi* | Archaea | euryarchaea |
| **pac** | *Propionibacterium acnes* | Bacteria | actinobacteria |

| | | | |
|---|---|---|---|
| **pae** | *Pseudomonas aeruginosa* | Bacteria | γ-proteobacteria |
| **pai** | *Pyrobaculum aerophilum* | Archaea | crenarchaea |
| **pfa** | *Plasmodium falciparum* | Eukarya | alveolata |
| **plt** | *Prevotella intermedia* | Bacteria | bacteroidete |
| **ppr** | *Photobacterium profundum* | Bacteria | γ-proteobacteria |
| **rme** | *Cupriavidus metallidurans* | Bacteria | β-proteobacteria |
| **rpa** | *Rhodopseudomonas palustris* | Bacteria | α-proteobacteria |
| **rpr** | *Rickettsia prowazekii* | Bacteria | α-proteobacteria |
| **rru** | *Rhodospirillum rubrum* | Bacteria | α-proteobacteria |
| **sce** | *Saccharomyces cerevisiae* | Eukarya | fungi |
| **sso** | *Sulfolobus solfataricus* | Archaea | crenarchaea |
| **syc** | *Synechococcus elongatus* | Bacteria | cyanobacteria |
| **syn** | *Synechocystis sp.* | Bacteria | cyanobacteria |
| **tac** | *Thermoplasma acidophilum* | Archaea | euryarchaea |
| **tbd** | *Thiobacillus denitrificans* | Bacteria | β-proteobacteria |
| **tcr** | *Trypanosoma cruzi* | Eukarya | euglenozoa |
| **tde** | *Treponema denticola* | Bacteria | spirochetae |
| **tel** | *Thermosynechococcus elongatus* | Bacteria | cyanobacteria |
| **tth** | *Thermus thermophilus* | Bacteria | deinococcus |
| **zmo** | *Zymomonas mobilis* | Bacteria | α-proteobacteria |

**Table 2: Enzymes, characters selected in the study group**

| Data code | ENZYME NAME | EC | GENE NAME |
|---|---|---|---|
| B/ 1 | 5-aminolevulinate synthase | 2.3.1.37 | ALAS, HemA |
| C/ 2 | glutamyl-tRNA synthetase | 6.1.1.17 | gltX, others |
| D/ 3 | glutamyl-tRNA reductase | 1.2.1.70 | HemA |
| E/ 4 | glutamate-1-semialdehyde 2,1-aminomutase | 5.4.3.8 | HemL |
| F/ 5 | porphobilinogen synthase | 4.2.1.24 | HemB |
| G/ 6 | hydroxymethylbilane synthase | 2.5.1.61 | HemC |
| H/ 7 | uroporphyrinogen-III synthase | 4.2.1.75 | HemD |
| I/ 8 | uroporphyrinogen decarboxylase | 4.1.1.37 | HemE |
| J/ 9 | uroporphyrinogen methyltransferase | 2.1.1.107 | HemX, CysG, CobA |
| K/ 10 | precorrin-2 dehydrogenase | 1.3.1.76 | CysG |
| L/ 11 | sirohydrochlorin ferrochelatase | 4.99.1.4 | CysG |
| M/ 12 | sirohydrochlorin cobaltochelatase | 4.99.1.3 | CbiX, CbiK |
| N/ 13 | cobalt-factor II C20-methyltransferase | 2.1.1.151 | CbiL, CbiL, CbiF |
| O/ 14 | precorrin-3 methyltransferase | 2.1.1.131 | CbiH, CbiL,CobJ |
| P/ 15 | precorrin-3 methylase | 2.1.1.133 | CbiF, CobM |
| Q/ 16 | precorrin-6A reductase | 1.3.1.54 | CobL, CobK |
| S/ 17 | precorrin-6 methyltransferase | 2.1.1.132 | CbiE, CbiT, CobL |
| T/ 18 | precorrin isomerase | 5.4.1.2 | CobH, CbiC |
| U/ 19 | cob(II)yrinic acid a,c-diamide reductase | 1.16.8.1 | |
| V/ 20 | cob(I)yrinic acid a,c-diamide | 2.5.1.17 | CobO, CobA, BtuR |

| | | | |
|---|---|---|---|
| | adenosyltransferase | | |
| W/ 21 | adenosylcobyric acid synthase (glutamine-hydrolysing | 6.3.5.10 | CbiP, CobQ |
| X/ 22 | adenosylcobinamide-phosphate synthase | 6.3.1.10 | CobD, CbiB |
| Y/ 23 | Adenosylcobinamide kinase | 2.7.1.156 | CobU, CobP |
| Z/ 24 | adenosylcobinamide-GDP ribazoletransferase | 2.7.8.26 | CobS |
| AA/ 25 | alpha-ribazole phosphatase | 3.1.3.73 | CobC |
| AB/ 26 | cob(I)yrinic acid a,c-diamide adenosyltransferase | 2.5.1.17 | |
| AC/ 27 | aquacobalamin reductase | 1.16.1.3 | |
| AD/ 28 | precorrin-2 C20-methyltransferase | 2.1.1.130 | CbiL, CobI |
| AE/ 29 | precorrin-3B synthase | 1.14.13.83 | CobG |
| AF/ 30 | precorrin-6A synthase (deacetylating) | 2.1.1.152 | CobF |
| AG/ 31 | coproporphyrinogen oxidase | 1.3.3.3 | HemF |
| AH/ 32 | coproporphyrinogen dehydrogenase | 1.3.99.22 | HemN |
| AJ/ 33 | protoporphyrinogen oxidase | 1.3.3.4 | HemG |
| AK/ 34 | protoporphyrin IX magnesium-chelatase | 6.6.1.1 | ChId, BchI |
| AL/ 35 | magnesium protoporphyrin IX methyltransferase | 2.1.1.11 | ChI, BchI |
| AM/ 36 | Mg-protoporphyrin IX monomethyl ester (oxidative) cyclase | 1.14.13.81 | Acs, Pni |
| AN/ 37 | protochlorophyllide reductase | 1.3.1.33 | |
| AO / 38 | chlorophyllase | 3.1.1.14 | |
| AP / 39 | chlorophyll synthase | 2.5.1.62 | |
| AQ / 40 | ferrochelatase | 4.99.1.1 | HemH |
| AR/ 41 | ferroxidase; ceruloplasmin | 1.16.3.1 | |
| AS / 42 | cytochrome c heme-lyase | 4.4.1.17 | |
| AT / 43 | heme oxygenase | 1.14.99.3 | |
| AU / 44 | phytochromobilin:ferredoxin oxidoreductase | 1.3.7.4 | |
| AV / 45 | phycocyanobilin:ferredoxin oxidoreductase | 1.3.7.5 | |
| AW / 46 | 15,16-dihydrobiliverdin:ferredoxin oxidoreductase | 1.3.7.2 | |
| AX / 47 | phycoerythrobilin:ferredoxin oxidoreductase | 1.3.7.3 | |
| AY / 48 | biliverdin reductase | 1.3.1.24 | |
| AZ / 49 | bilirubin oxidase | 1.3.3.5 | |
| BA / 50 | bilirubin UDP-glucuronosyltransferase | 2.4.1.17 | |
| BB / 51 | beta-glucuronidase | 3.2.1.31 | |

The absence or presences of an enzyme in the metabolism of organisms that constitute the study group are discrete characters that are encoded in a binary system, 0 or 1. This coding means that an enzyme can be in two states: 0-absent or 1-present.

Data have been collected from KEGG: Kyoto Encyclopedia of Genes and Genomes (1995-2015 Kanehisa Laboratories)

The set of taxa and characters are reflected in a "matrix data" showing taxa in rows and states of a character in columns (Tale 3). This matrix is the basis for further analysis.

**Table 3: Data Matrix showing the states of the characters (0-absent enzyme, 1-present enzyme)**

| A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2.3.1.37 | 6.1.1.17 | 1.2.1.70 | 5.4.3.8 | 4.2.1.24 | 2.5.1.61 | 4.2.1.75 | 4.1.1.37 | 2.1.1.107 | 1.3.1.76 | 4.99.1.3 | 4.99.1.4 | 2.1.1.151 | 2.1.1.131 | 2.1.1.133 | 1.3.1.54 |
| ath | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| ago | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| cel | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| cal | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| cne | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| cho | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| cme | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| ddi | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| dme | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| pfa | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| sce | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| tcr | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| ape | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| afu | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| hal | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| hbu | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| mth | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| mja | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| mka | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| mac | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| neq | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| pai | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| pab | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| sso | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 |
| tac | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| ana | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 |
| gvi | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 |
| syc | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 |
| syn | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 |
| tel | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 1 |
| atu | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 |
| bja | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 |
| gox | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| nwi | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| rpa | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 |
| rru | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 |
| rpr | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| zmo | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| neu | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| rme | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| fbd | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| buc | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| eco | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| hin | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| lpf | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| mca | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 |
| ppr | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| pae | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 |
| bsu | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| cac | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| mga | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| lpl | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| tth | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 |
| dge | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 |
| det | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| cte | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| cch | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| plt | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| mlu | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 |
| pac | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |
| lic | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 |
| tde | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 1 |
| cau | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 |

**Table 3. cont.**

| R | S | T | U | V | W | X | Y | Z | AA | AB | AC | AD | AE | AF | AG | AH |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2.1.1.132 | 5.4.1.2 | 1.16.8.1 | 2.5.1.17 | 6.3.5.10 | 6.3.1.10 | 2.7.1.156 | 2.7.8.26 | 3.1.3.73 | 2.5.1.17 | 1.16.1.3 | 2.1.1.130 | 1.14.13.83 | 2.1.1.152 | 1.3.3.3. | 1.3.99.22 |
| ath | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |
| ago | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| cel | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| cal | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| cne | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| cho | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| cme | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |
| ddi | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| dme | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| pfa | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| sce | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| tcr | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| ape | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| afu | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| hal | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| hbu | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| mth | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| mja | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| mka | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| mac | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| neq | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| pai | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| pab | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| sso | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| tac | 1 | 1 | 0 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| ana | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 1 |
| gvi | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 1 |
| syc | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 1 |
| syn | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| tel | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 1 |
| atu | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 1 |
| bja | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 1 |
| gox | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |
| nwi | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| rpa | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 1 |
| rru | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 |
| rpr | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |
| zmo | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |
| neu | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| rme | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| tbd | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| buc | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| eco | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 1 |
| hin | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| lpf | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |
| mca | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 1 |
| ppr | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 1 |
| pae | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 1 |
| bsu | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| cac | 1 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| mga | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| lpl | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| tth | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| dge | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 |
| det | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| cte | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| cch | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| plt | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| mtu | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 1 |
| pac | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 |
| lic | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 |
| tde | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| cau | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |

| AI | AJ | AK | AL | AM | AN | AO | AP | AQ | AR | AS | AT | AU | AV | AW | AX | AY | AZ | BA | BB |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1.3.3.4 | 6.6.1.1 | 2.1.1.11 | 1.14.13.81 | 1.3.1.33 | 3.1.1.14 | 2.5.1.62 | 4.99.1.1 | 1.16.3.1 | 4.4.1.17 | 1.14.99.3 | 1.3.7.4 | 1.3.7.5 | 1.3.7.2 | 1.3.7.3 | 1.3.1.24 | 1.3.3.5 | 2.4.1.17 | 3.2.1.31 |
| ath | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| ago | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| cel | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |
| cal | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| cne | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| cho | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| cme | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| ddi | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| dme | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 |
| pfa | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| sce | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 |
| tcr | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| ape | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| afu | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| hal | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| hbu | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| mth | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| mja | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| mka | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| mac | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| neq | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| pai | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| pab | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| sso | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| tac | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| ana | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 |
| gvi | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 |
| syc | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| syn | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 |
| tel | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| atu | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| bja | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| gox | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| nwi | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| rpa | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| rru | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| rpr | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| zmo | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| neu | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| rme | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| tbd | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| buc | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| eco | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| hin | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| lpf | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| mca | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| ppr | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| pae | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| bsu | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| cac | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| mga | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| lpl | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| tth | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| dge | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| det | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| cte | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| cch | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| plt | 0 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| mtu | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| pac | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| lic | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| tde | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| cau | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |

A data matrix as the above in which the reflected characters are metabolic enzymes in two states, present (1) or absent (0), serves for the cladistic analysis of enzymes and accordingly the metabolism of porphyrins and chlorophylls in organisms of the study group.

The cladistic analysis was performed with the MIX program in Phylip (Felsestein, 2001) (http://evolution.genetics.washington.edu/phylip.html) which estimates phylogenies by Wagner parsimony method for discrete character data with two states (0 and 1) (also available an online version of the program http://mobyle.pasteur.fr/cgi-bin/portal.py#jobs::overview). Cladograms and trees are drawn with PHY.FI application (Fredslund, 2006).

The result of this analysis is reflected by a cladogram representing an evolutionary hypothesis discussed below.

## 4. One hypothesis about the evolution of porphyrins and chlorophylls

The cladistic analysis following Wagner Parsimony Method for discrete characters in two states provides 100 equally parsimonious trees (L = 233). Figure 2 shows the majority rule consensus tree and Figure 3 shows the strict consensus tree of the 100 most parsimonious and equally parsimonious trees.

In the majority rule consensus tree two large groups or clades are observed: one formed by Bacteria and Archaea and another group of Bacteria and Eukaryota (Fig. 2). This result is confirmed by the strict consensus tree (Fig. 3). The ancestral metabolism corresponds to *Nanoarchaeum equitans* (neq) leaving immediately after the ancestor (anc, the hypothetical ancestor with all characters set to 0, plesiomorphic, ie, absent).

All enzymes considered undergo at least one change of state. Those experiencing a single change (bold in the Table 2) are apomorphic characters which define well metabolism (red arrows in Fig. 1). All other enzymes experience more than one change, they are homoplasy, indicate convergent evolution, the adaptive nature of metabolism and its components.

In the clade Metabolism **[BACTERIA + ARCHAEA]** the group of cyanobacteria (tel, gvi, syn, ana and syc) is well defined by the enzyme **phycocyanobilin:ferredoxin oxidoreductase (EC 1.3.7.5)** which experience a unique change and is a sinapomorphic character for the cyanobacteria clade. Within this group of cyanobacteria, *Gloeobacter violaceus* (gvi) is defined by two apomorphic characters: enzymes **15,16-dihydrobiliverdin:ferredoxin oxidoreductase (EC 1.3.7.2)** and **phycoerythrobilin:ferredoxin oxidoreductase (EC 1.3.7.3).**

On the basis of this group of cyanobacteria, in a previous node, is the metabolism of *Chloroflexus aurantiacus* (cau) defined by an autoapomorphy, the **bilirubin oxidase** enzyme (EC 1.3.3.5). The clade Metabolism **[BACTERIA + EUKARYA]** includes among eukaryotes *Arabidopsis thaliana* (ath), plant whose metabolism is well defined by two apomorphies, enzymes **chlorophyllase (EC 3.1.1.14)** and **phytochromobilin: ferredoxin oxidoreductase (EC 1.3.7.4)**.

All other groups or clades are not based on apomorphic characters, no homologies to well define species metabolism. Therefore they are analogies and a case of convergent or parallel evolution. The similarity is due to homology and analogy. But nevertheless analogies or homoplasies not serve to discover phylogenetic relationships. Therefore they can not be established assumptions about relationships between sister groups based on the metabolism of porphyrins and chlorophylls.

Resulting cladograms both majority consensus as strict consensus (Figs. 2 and 3) reflects relationships derived from adaptive changes that lead to the expression, or not, of a gene and possible synthesis or catalytic activity of an enzyme as well as the acquisition or loss metabolic capabilities as an adaptive response to the environment. Moreover, lateral gene transfer is a key process in the early stages of cellular evolution that draws a complex network of relationships that mask and / or replace the vertical transfer and evolution model based on ancestry with modification. On the other hand, much of the diversity found in algae is due to secondary and tertiary endosymbiotic events (Keeling, 2010).

**Figure 2. Mayority rule consense tree**

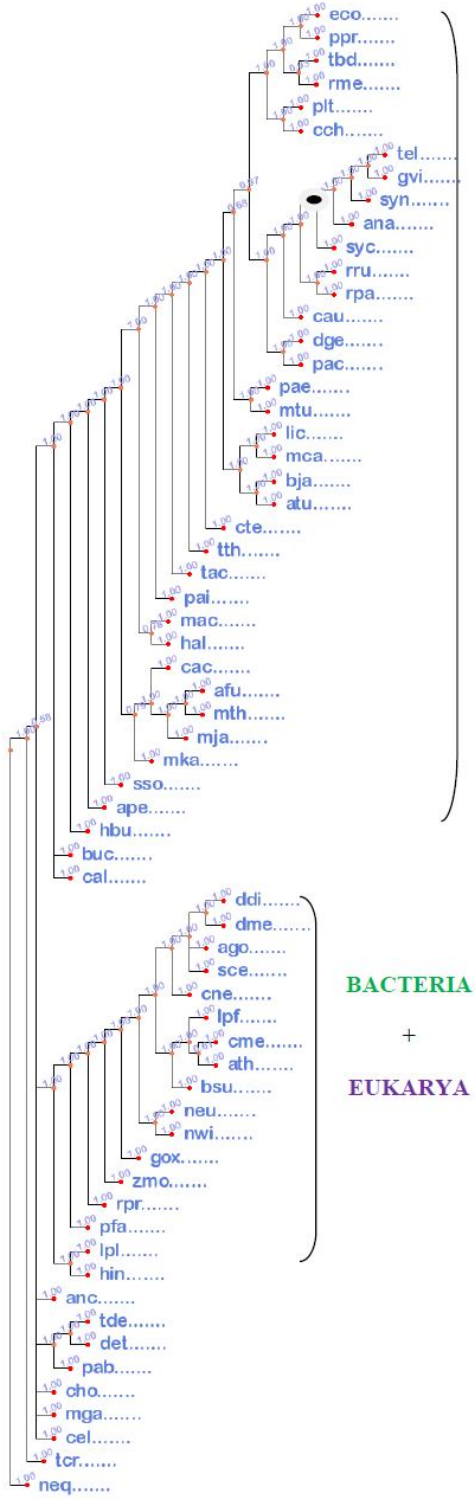
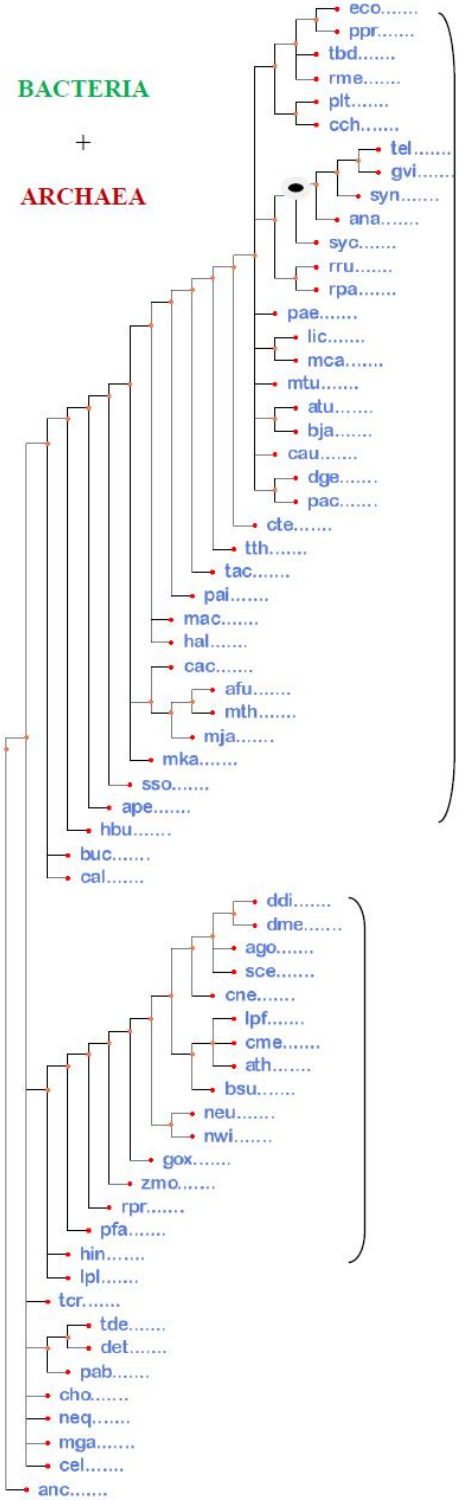



**Figure 3. Strict consense tree**



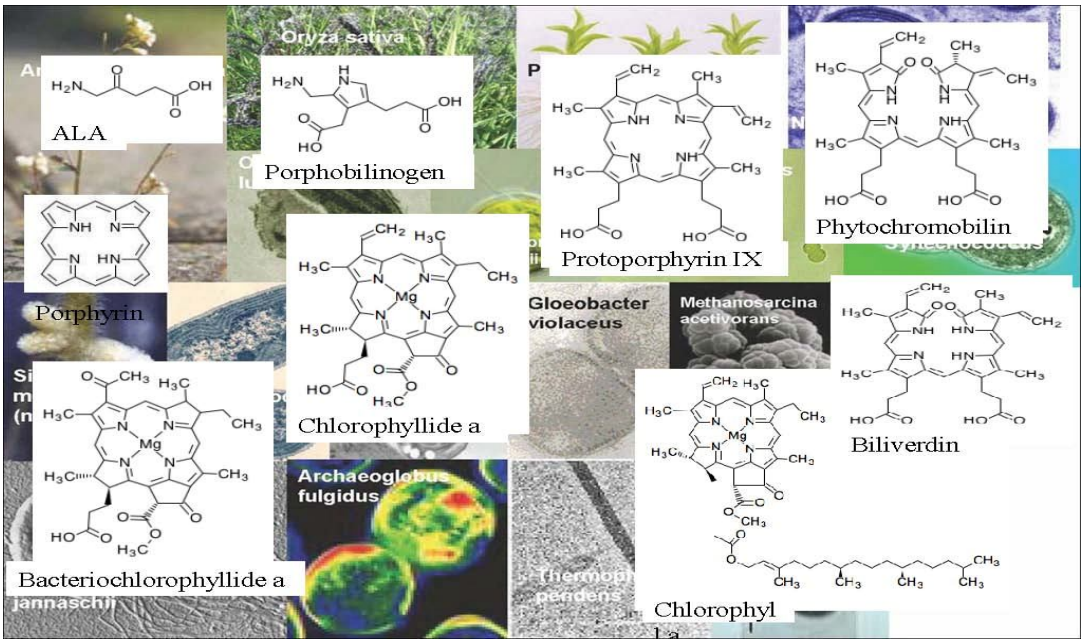
The metabolism compared in the study group makes clear enzymes that change once. Some of these elements as apomorphic homologies could be phylogenetic markers. This is the case of phycocyanobilin:ferredoxin oxidoreductase (EC 1.3.7.5) for cyanobacteria and chlorophyllase (EC 3.1.1.14) for plants.

Green algae and its progeny lost the cyanobacterial phycobilisome light-harvesting system, whereas red algae (cme, *Cyanidioschyzon merolae*) (on the cladogram with *Arabidopsis thaliana*) and glaucophytes retained it. The chloroplasts in algae and plants are derived from cyanobacteria and endosymbiosis gave rise to photosynthetic eukaryotes (Margulis, 1992). The evolution of photosynthesis is a complex process involving different sources and routes of its many components, so that its history can not be described as a simple, linear process. However, it seems certain that the emergence of Mg-tetrapyrrole and apoproteins of the reaction centers are key events that led to the development of the photosynthetic process. The flow of electrons acceptor–is channeled by protein complexes that always contain metallo-organic cofactors. Membrane-bound complexes couple the transfer of electrons across the membrane to the generation of an ion gradient and transmembrane electrical potential. This chemiosmotic mechanism was likely present in the last common ancestor and has been carried forward to the three presently persisting domains of life, Bacteria, Archaea, and Eukarya (Lane et al., 2010).

The Granick (Granick, 1965) and Retrograde (Horowitz, 1945) hypotheses on the establishment of metabolic pathways are complementary. Granick established as hypotheses that the intermediate compounds of the modern biosynthetic pathways were the final products of early pathways and thus the evolution of the pathway can be traced from the beginning to the end. However this contrasts with the retrograde hypothesis which posits that present biosynthetic pathways are set up in the reverse order to their evolutionary history and occurred through gene duplications. But The Granick and retrograde hypotheses are not mutually exclusive because the retrograde hypothesis is a consequence of the depletion of base molecules present in the primordial soup, molecules that follow the Granick hypothesis may be more derived.

There are still many aspects of the evolution of photosynthesis unresolved due in part to the existence of highly diversified components (Fig.4). One way to learn more about this issue is to address the systematic, descriptive and comparative study of genes and gene products of photosynthesis in the diversity of phototrophic organisms. For this purpose contributes undoubtedly the availability of numerous molecular data and the use of phylogenetic analysis tools. For example and based on the results of this study, comparative analysis of enzymes glutamyl-tRNA synthetase (EC 6.1.1.17), porphobilinogen synthase (EC 4.2.1.24) and protoporphyrinogen oxidase (EC 1.3.3.4) arises With all this new knowledge will continue to emerge thus reconstruct the evolutionary history of photosynthesis.

**Figure 4: Some molecules of metabolism of porphyrins and chlorophylls**




## 5. References

Cavalier-Smith, T. (1992). Bacteria and Eucaryotes. Nature, 356: 570.

Cunchillos, C. y Lecointre, G. (2002). Early steps of metabolism evolution inferred by cladistic analysis of amino acid catabolic pathways. C.R. Biologies, 325: 119-129.

Cunchillos, C. y Lecointre, G. (2003). Evolution of Amino Acid Metabolism inferred through Cladistic Analysis. The Journal of Biological Chemistry, 278 (48), nov.: 47960-47970.

Cunchillos, C. y Lecointre, G. (2005). Integrating the Universal Metabolism into a Phylogenetic Analysis. Mol. Biol. Evol., 22 (1): 1-11.

Des Marais, D.J. (2000). When did photosynthesis emerge on Earth? Science, 289: 1703-1705.

Enzyme Nomenclature (1973). Recommendations (1972) of the International Union of Pure and Applied Chemistry and The International Union of Biochemistry. Elsevier Scientific Publishing Company, Amsterdam.

Felsenstein, J. (2001). PHYLIP (phylogeny inference package). version 3.6a3. Department of Genetics. University of Washington, Seattle.

Fredslund, J. (2006). *PHY·FI:* fast and easy online creation and manipulation of phylogeny color figures. BMC Bioinformatics 7:315.

Granick, S. 1965. Evolution of heme and chlorophyll. In Evolving Genes and Proteins, ed. V Bryson, HJ Vogel, pp. 67–88. New York: Academic.

Horowitz, N.J. 1945. On the evolution of biochemical synthesis. Proc. Natl. Acad. Sci. USA 31:153–57

KEGG: Kyoto Encyclopedia of Genes and Genomes 1995-2015 Kanehisa Laboratories. Available on http://www.genome.jp/kegg/

Keeling, P.J. 2010. The endosymbiotic origin, diversification and fate of plastids. *Philos. Trans. R. Soc. Lond. B* 365:729–48.

Lane, N, Allen, J.F, Martin, W. 2010. How did LUCA make a living? Chemiosmosis in the origin of life. BioEssays 32:271–80.

Margulis, L. 1992. Symbiosis in Cell Evolution. San Francisco: Freeman.

Nisbert, E.G. y Sleep, N.H. (2001). The habitat and nature of early life. Nature, 409: 1083-1091.

Olson, J.M. (1999). Early evolution of chlorophyll-based photosystems. Chemtracts, 12: 468-482.

Olson, J.M. (2001). Evolution of Photosynthesis (1970), re-examined thirty years later. Photosynth. Res., 68:95-112.

Schidiowski, M. (1988). A 3.800-million-year isotopic record of life from carbon in sedimentary rocks. Nature, 333: 313-318.

Shopf J.W. (1993). Microfossils of the early archean apex chert: new evidence ofthe antiquity of life. Science, 260: 640-646.

Schopf, J.W. y Packer, B.M. (1987). Early Archean (3.3-billion to 3.5-billion year old) microfossils from Warrawoona Group. Australia. Science, 237: 70-73.

Woese, C.R. (1987). Bacterial evolution. Microbiol. Rev., 51: 221-271.