RNA world and protocells

Planets and Astrobiology, Academic Year 2019-2020

Giovanni Vladilo (INAF-OATs)

Origin of molecular replication and metabolism

- Conceptual "chicken-egg" problem
  - In present-day cells, nucleic acids and proteins are responsible for replication and metabolic functions, respectively
  - The formation of each one of these two types of macromolecules requires the previous existence of the other one

    The synthesis of nucleic acids is catalyzed by proteins (enzymes)

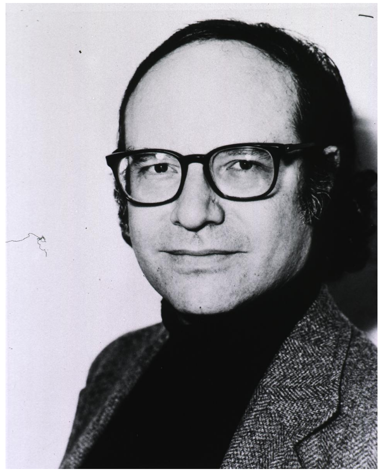
    The synthesis of proteins requires the instructions stored in the nucleic acids
- Who came first?
  - Proteins or nucleic acids ?
  - Replication/genetic or metabolic functions ?
- Different approaches have been adopted to tackle this problem
  - Old approach: "Metabolism first" <u>or</u> "genes first"
  - Present-day approach: <u>search for macromolecules that show both properties</u>

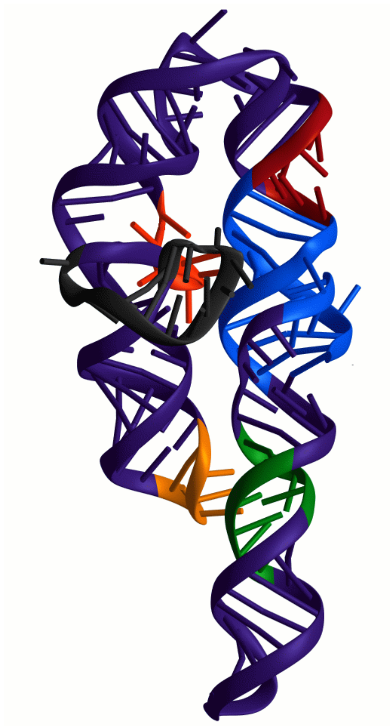
The “RNA world”

- Present-day, main stream theory in studies on life’s origin
- Introduced by Walter Gilbert (1986) after the discovery of ribozymes
  - RNA molecules with catalytic properties
- According to this theory, the genetic system is the first to emerge, but with self-catalytic properties
  - Present-day ribozymes would be a sort of molecular fossiles of an ancient “RNA world”
- Present-day DNA-world would have emerged at a later stage because of its advantages
  - DNA provides greater genetic stability

    The lack of an oxygen atom in the sugar (deoxyribose instead of ribose) makes DNA less reactive than RNA
  - The DNA world has an extremely greater flexibility

    due to the introduction of proteins specialized in a large variety of metabolic functions





Life as a kinetic state of matter

Addy Pross

Example of the kinetic power of self replication

A, B: reactants

X: catalyst

- Comparison between normal and self-catalytic reactions
  - start with 1 molecule of catalyst X
  - assume reaction rate 1μs in both cases
- Time required to build up a mole of products (6 x 10^{23})
  - Normal case: 20 billon years
  - Self-catalytic case: 79 μs
- The kinetic control of chemical reactions could be the key for understanding the origin of life (in chemistry, the term "kinetics" is related to the rate of chemical reactions)
  - see literature by Addy Pross

$$A + B \xrightarrow{X} C$$

$$A + B \xrightarrow{X} X$$

Replication and molecular evolution

Imperfect replication and chemical selection are supposed to be the key ingredients of some form of molecular evolution that has supposedly lead to the molecular machinery that we see today

In a broad sense, *molecular replication and chemical selection is an extension back in time of the concept of Darwinian evolution (reproduction and natural selection)* which, strictly speaking, takes place only after the first living organisms are born

Darwinian evolution, whether of species or molecules, is not teleological

It works *a posteriori*, in the sense that it favours the most suitable variations for a given function that already exists

The capability of replication is probably the key function for the initial selection of biomolecules

Compartments

In order to develop protocells, the early products of the RNA-world must have been enclosed in compartments

Compartmentalization is required to prevent the dispersion of genetic information and to concentrate the action of cooperative biochemical processes in an enclosed space

Membranes delimit a set of structures and reactions that can be transmitted as a specific heritage, paving the road for the onset of Darwinian evolution

Early membranes

In present-day life the compartments are provided by the phospholipid bilayers of the cell membranes

Phospholipds are the result of an evolutionary process, and their synthesis requires enzymatically catalyzed reactions not available for the first protocells

Early membranes could have been constituted by simple fatty acids

Simple fatty acids can be spontaneously generated in prebiotic chemistry, as demonstrated, for example, by their presence in the Murchison meteorite

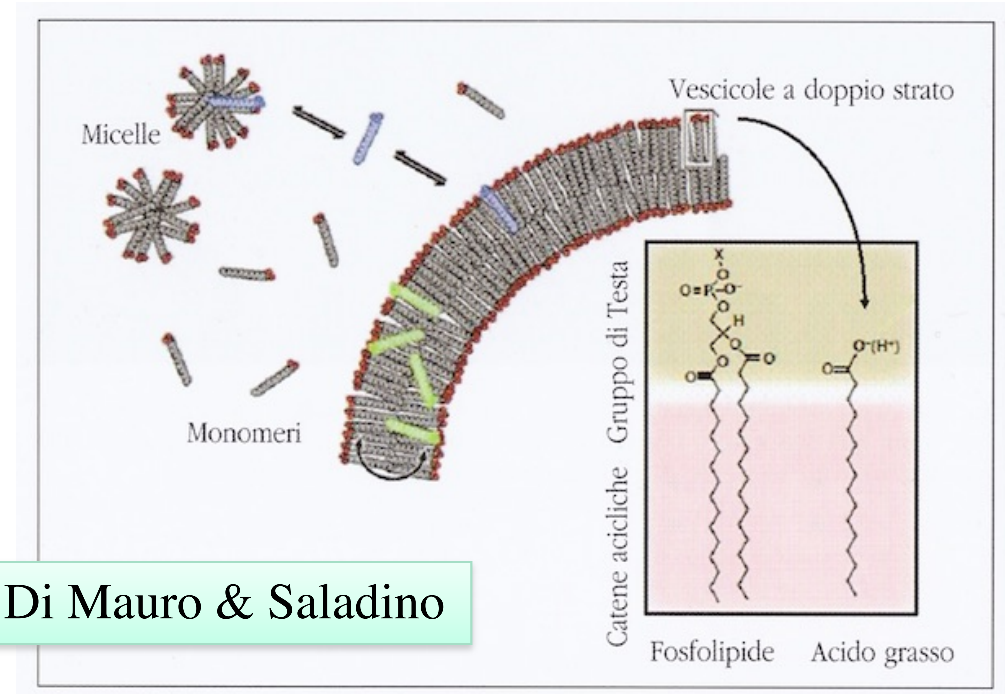


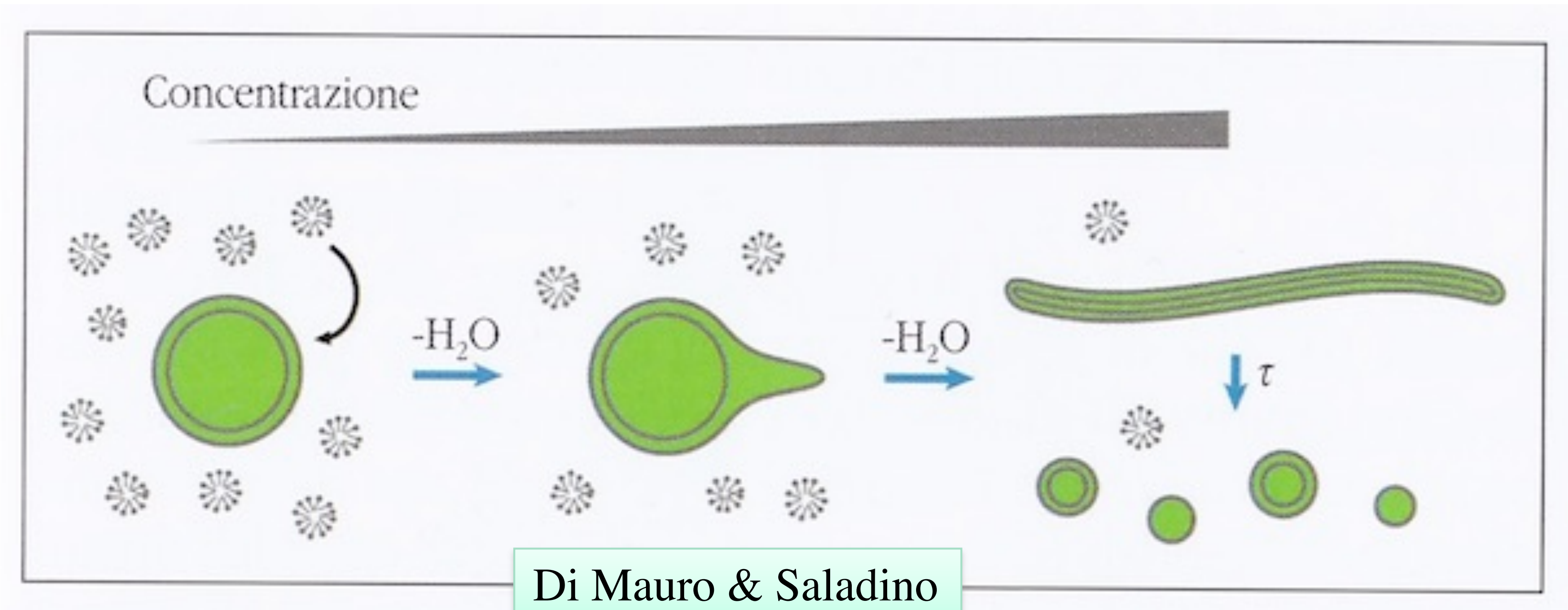
Figure: Di Mauro & Saladino

Protocellular vesicles

Laboratory experiments demonstrate that simple amphiphilic molecules, resulting from prebiotic processes, can give rise to vesicle structures

Variations of the *concentration* of amphiphilic molecules and of *ambient conditions* drive the formation and destruction of vesicle structures that can grow and duplicate

Jack Szostak demonstrated that protocellular vesicles better replicate if they contain RNA and, at the same time, RNA better replicates if it is enclosed in lipidic vesicles



Di Mauro & Saladino

Casting light on the first living cells

- "Top-down" approach
  - From the study of present-day living organisms, we try to characterize the properties of the first terrestrial organisms proceeding backwards in evolution

- One of the methods being employed is the comparison of genetic sequences of present-day living organisms
  - Thanks to this comparative analysis, we can trace backwards the evolution <u>at the molecular level</u>
  - The results are visualized in the "phylogenetic tree", where the distances between different species are proportional to the differences found in the genetic sequences

Genetic sequences and classification of organisms

The techniques of molecular biology allow us to classify organisms on the basis of their genetic sequences, rather than on their morphology or phenotype (composite of observable traits and behaviour of organisms)

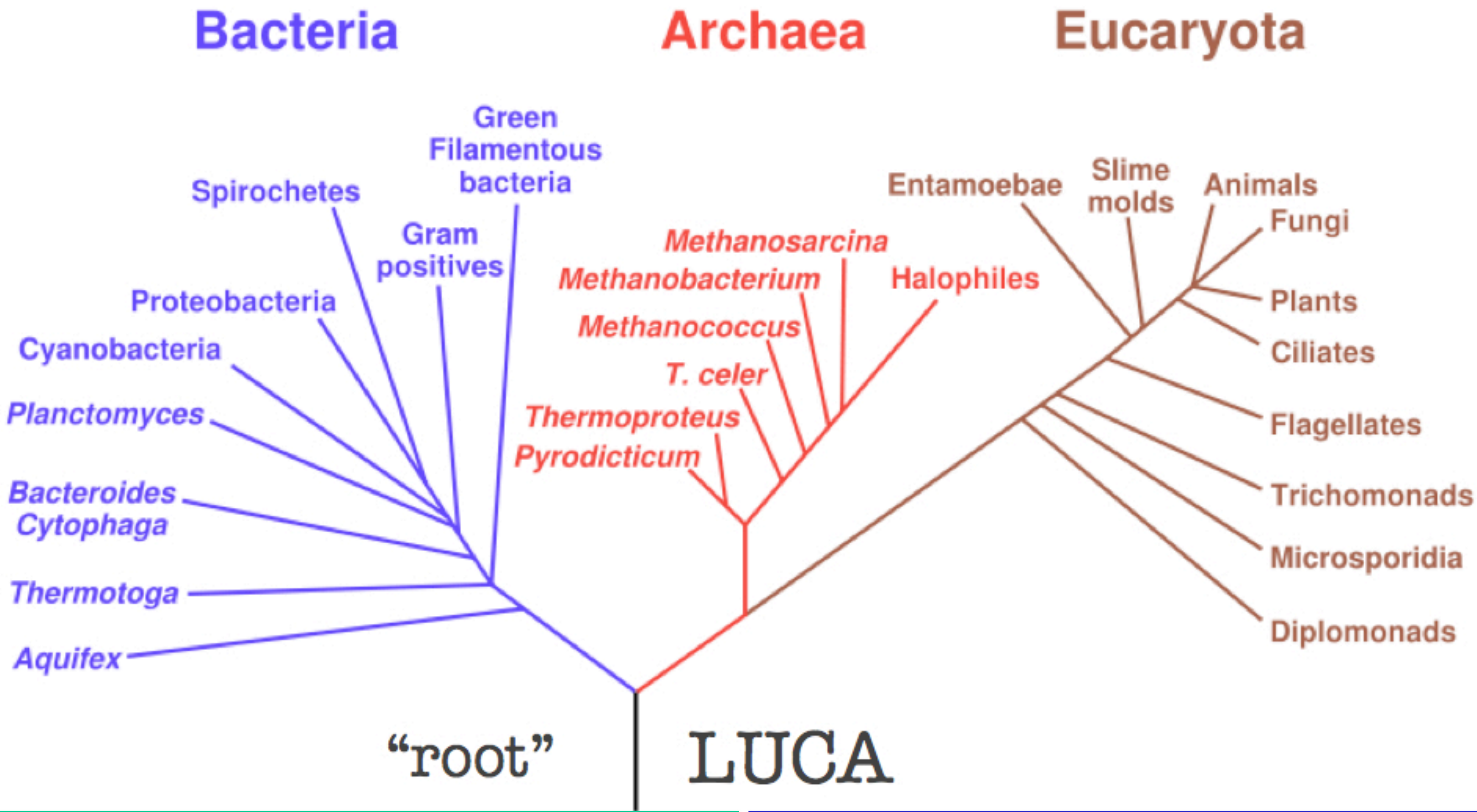
The classification based on genetic sequences has revolutionized our understanding of unicellular organisms

The classification based on genetic sequences has lead to distinguish three different types of unicellular organisms:

archaea, eubacteria and eukaryotes

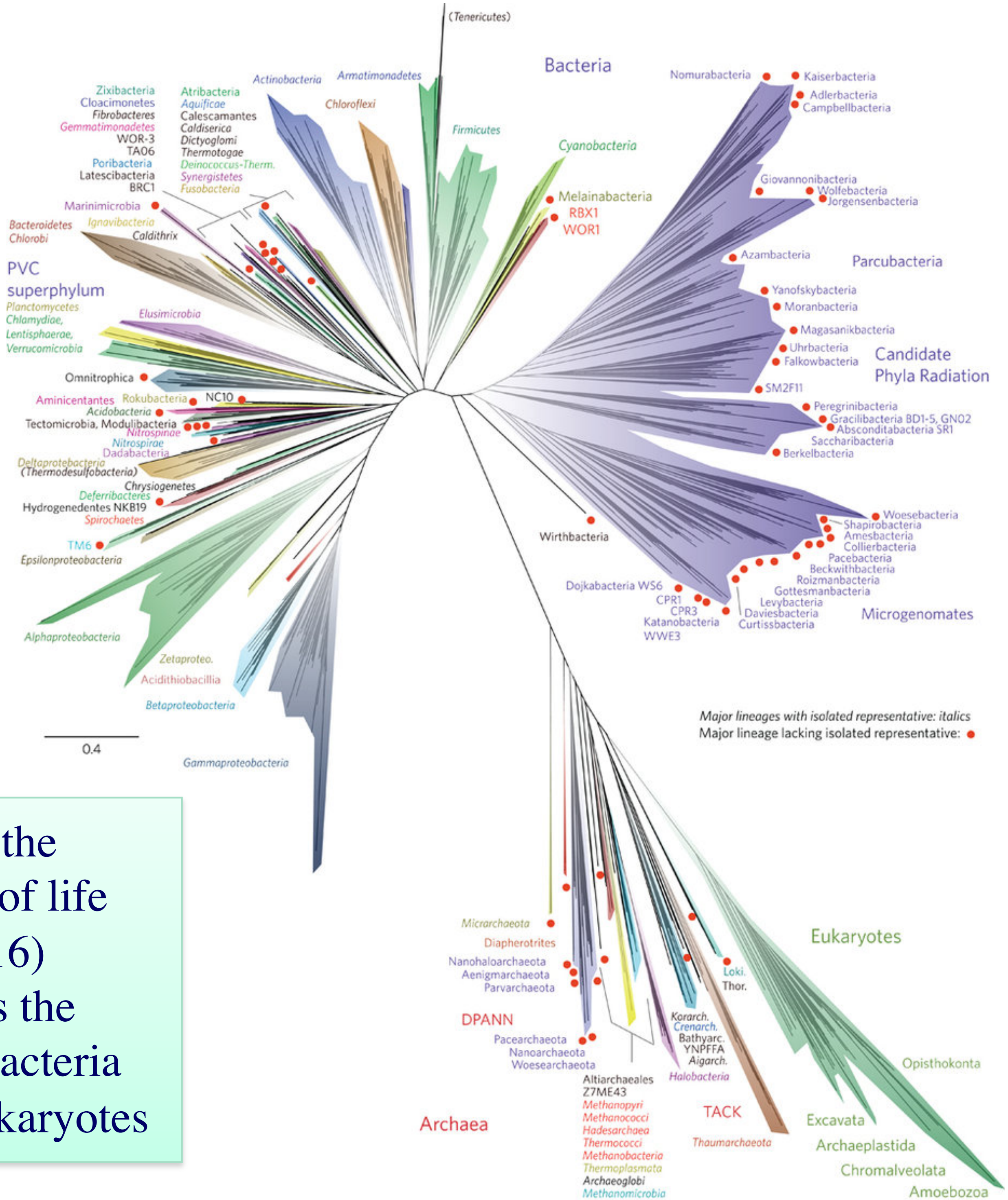
Archaea have been discovered through genetic classification

The phylogenetic tree of life



LUCA = Last Universal Common Ancestor of present-day living organisms, also called Cenancestor

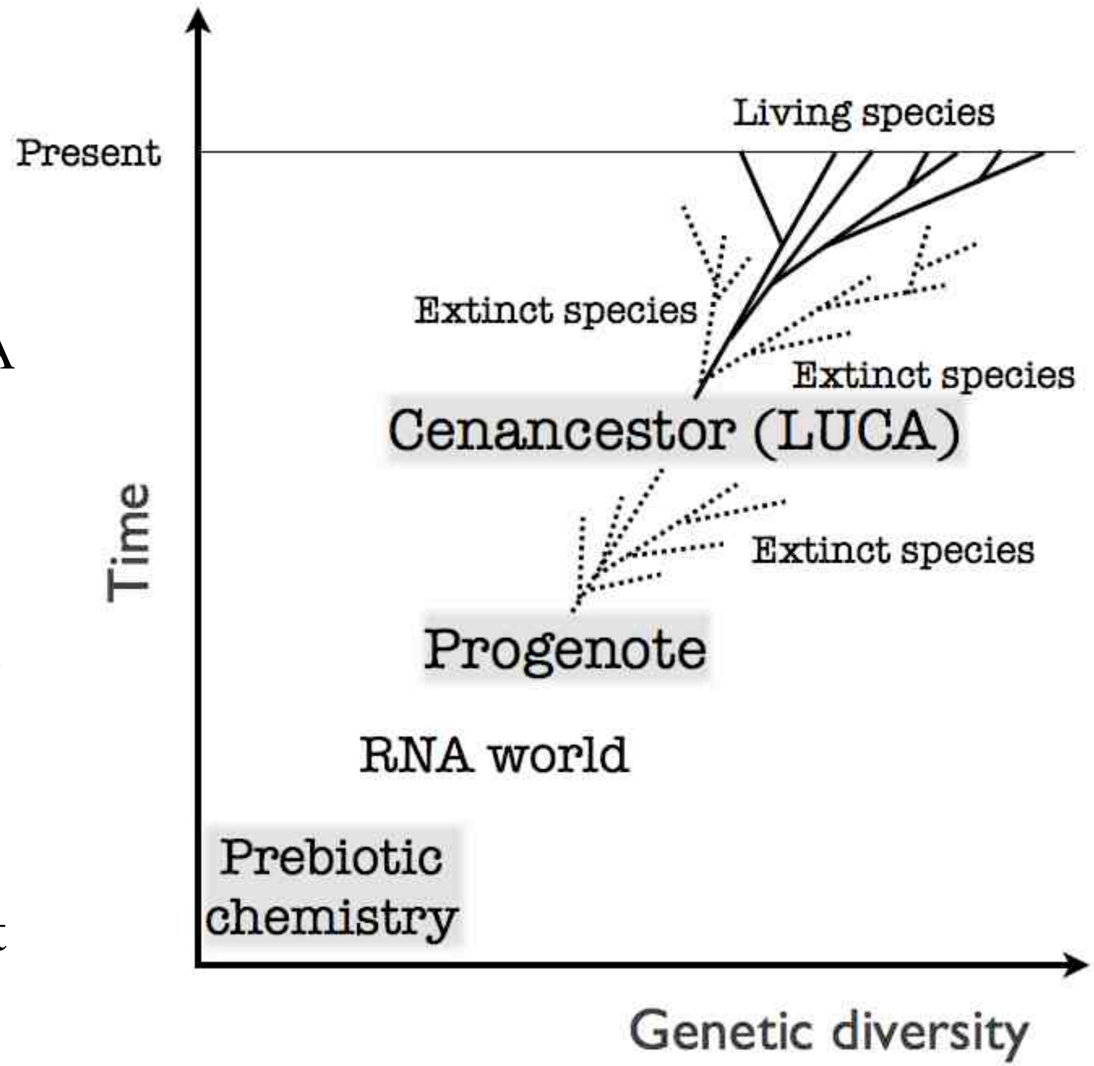
Close to the "root" of the tree, we find thermophilic Archaea and Bacteria



A new view of the phylogenetic tree of life (Hug et al. 2016) which highlights the predominance of bacteria over archaea and eukaryotes

The gap between the RNA world and the LUCA

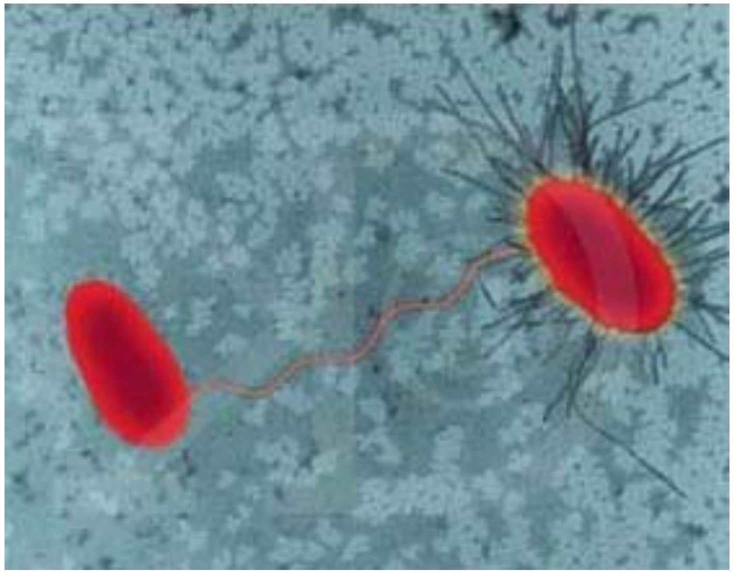
- The root of the phylogenetic tree is not representative of the oldest living cell
  - Other forms of life, extinct in the course of the evolution, must have preceded the LUCA
  - This early form of life is sometimes called the “progenote”
  - The early life could have been a collection of somewhat different cells, rather than a single type of cell
  - Detailed analysis suggests that *early life was mesophilic, rather than thermophilic*

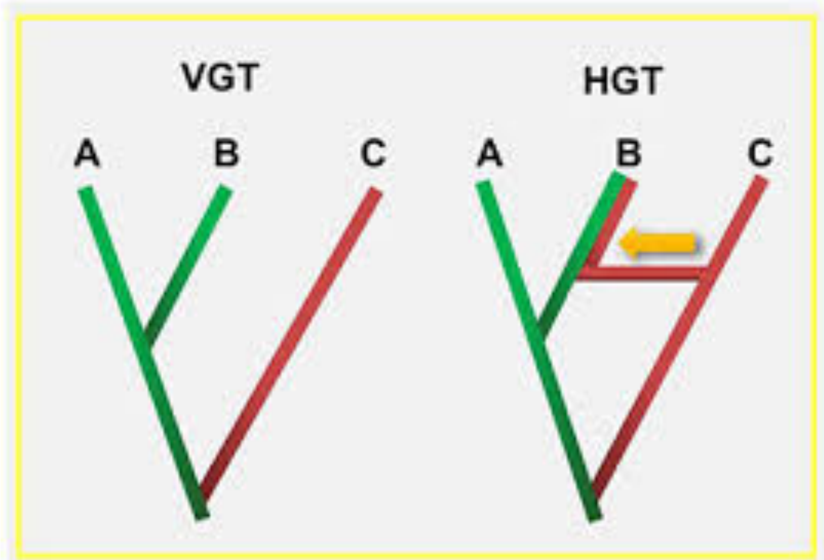


Horizontal gene transfer

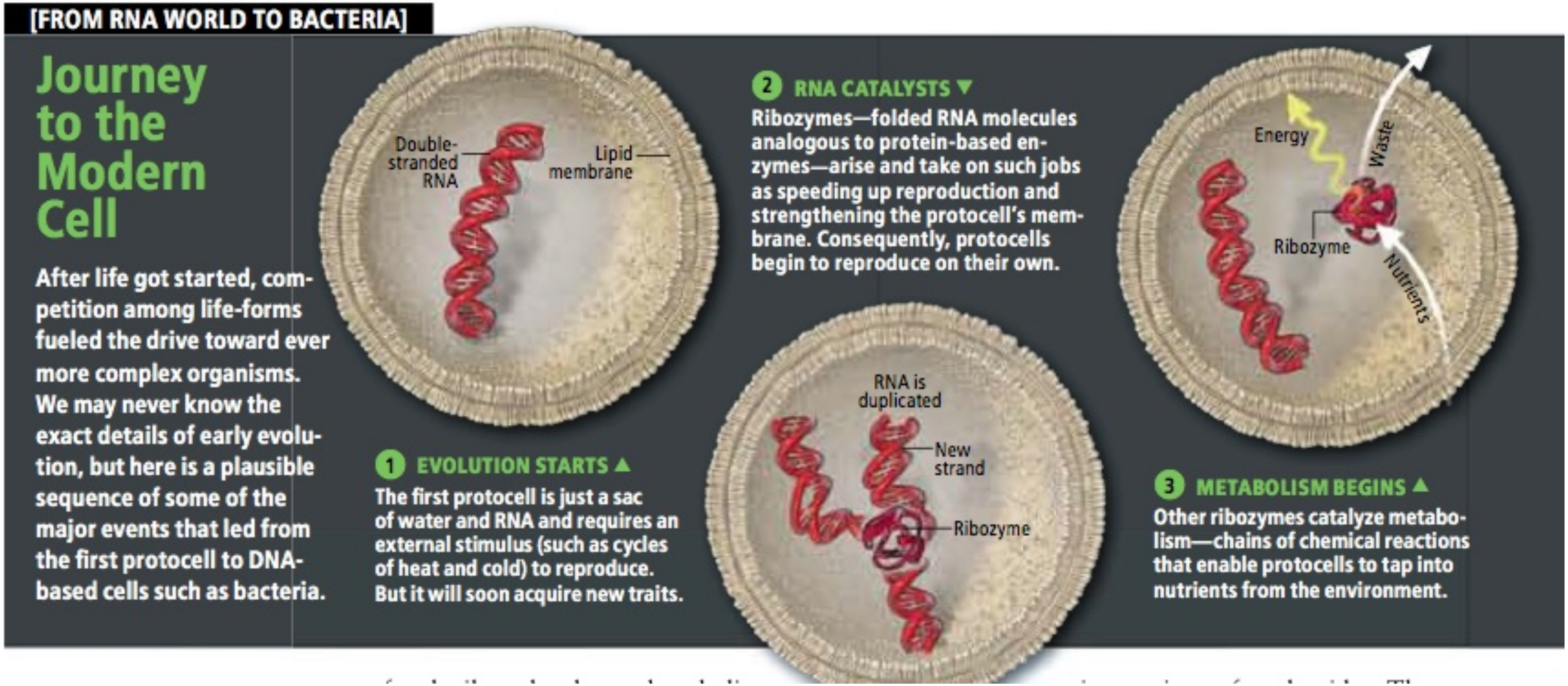
(also called Lateral Gene Transfer)

- Bacteria can exchange genetic material not only during their reproduction ("vertical gene transfer", VGT) but also via direct exchange from one cell to another ("horizontal gene transfer", HGT)
- The existence of HGT complicates the reconstruction of the philogenetic tree, which is based on the VGT scenario
- HGT must have played an essential role in the early stages of life, providing a simple mechanism to exchange genetic material before more complex mechanisms of "vertical" transmission were set in place



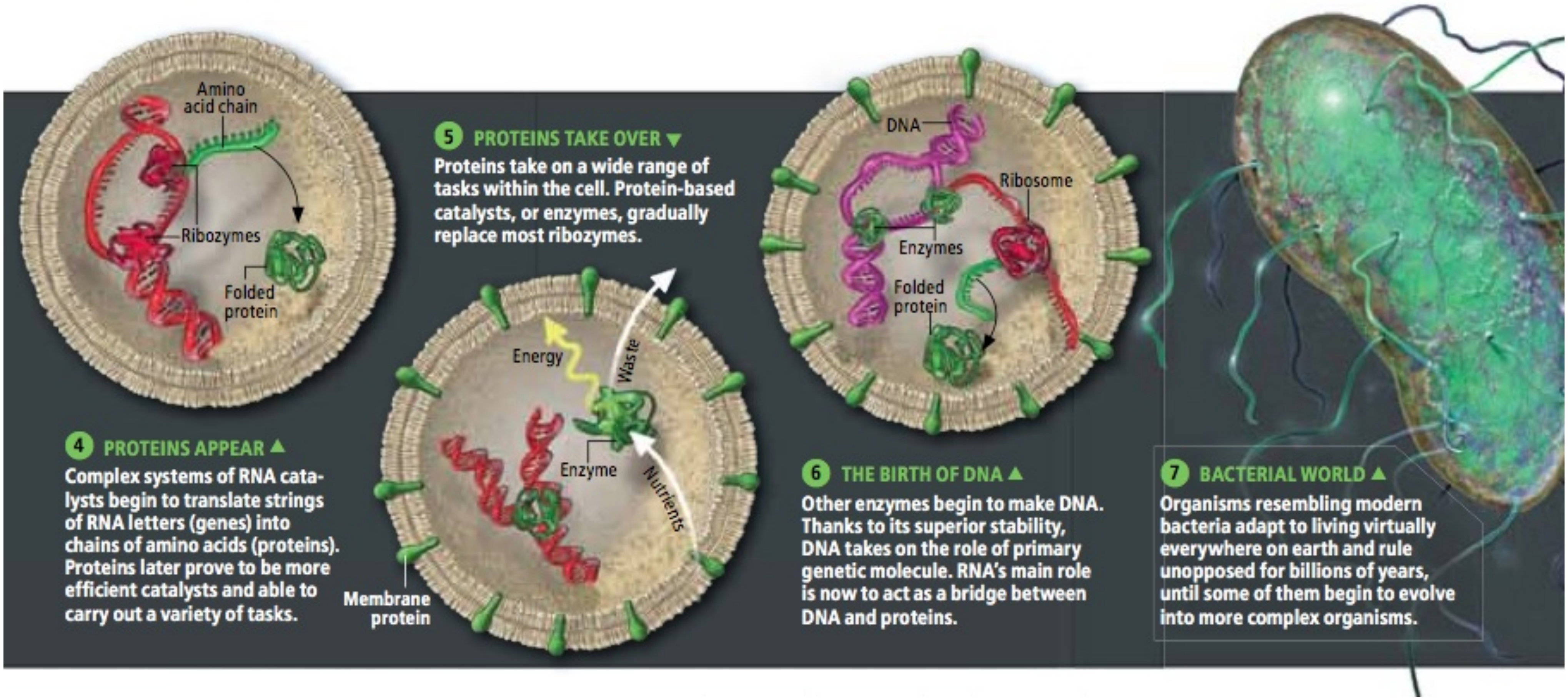


Filling the gap from the RNA world to protocells



Ricardo & Szostak (2009)

Filling the gas from protocells to prokaryotes



Ricardo & Szostak (2009)

Heterotrophic versus autotrophic origin of life

Heterotrophic hypothesis

The first organisms harvest organic material and energy from prebiotic molecules that are already present in the environment

The molecular ingredients could have been delivered on Earth from space or could also have been synthesized on the primitive Earth

This hypothesis does not require a specific enviromental niche

Autotrophic hypothesis

The first organisms extract energy and synthesise organic material from the abiotic world

The early life forms would have emerged in the proximity of redox or pH gradients, using the harvested energy to feed biosynthesis reactions.

These processes require extremely reactive chemical environments.

This scenario can only take place in specific thermodynamical niches

The cradles of life

- Deep sea hydrothermal vents (autotrophic origin)
  - Hydrothermal vents at the bottom of the oceans could provide inorganic compartments, versatile catalysis and sources of organic matter
  - An origin in the oceans, often considered in the past, poses the problems of the containment of the reactions in an open water environment
  - The presence of Na salts, typical of the oceans, would hinder the formation of biological membranes (Natochin 2010)
- Anoxic geothermal fields (heterotrophic origin)
  - Supported by geochemical data and phylogenomic analysis (Mulkidjanian et al. 2012)
  - Geothermal fields are conducive to condensation reactions and enable the involvement of solar light as an energy source and as a selector factor of stable nucleotides
  - Geothermal vapour is enriched in phosphorus compounds that could be essential for the emergence of the first RNA-like oligomers