

**Assignment #2**

**Physics 498: Statistical Physics of Biological Complexity and Information**

**Molecular Phylogenetic Analysis**

Date of Submission: 31<sup>th</sup> October, 2001

**Rahul Roy**  
**Center for Biophysics and Computational Biology**  
**University of Illinois at Urbana-Champaign**  
email: rroy1@students.uiuc.edu

## **Evolve by Borrowing ?**

*Rahul Roy*

*Center for Biophysics and Computational Biology*

*University of Illinois at Urbana Champaign*

### **Introduction**

Molecular evolution over a period of four billion years has resulted in the development of present myriad of species from a hot soup of molecules. It has resulted in the evolution of complex multicellular organisms (Eukaryotes) along with single-celled organisms like prokaryotes at the same time. The question that has baffled scientists for long is: how did this incredible transition from a soup to this complex state take place? The discussion, that still continues, as how life started from the primordial soup is not the emphasis of the present debate, though Stanley Miller showed in, as back as 1950s, that organic chemicals can be synthesized from completely inorganic inputs. It has been proposed based on structural and functional complexity and fossil evidence that prokaryotes must have predated the eukaryotes by at least 1.0-2.0 billion years [Knoll, 1992]. This has led to the notion that eukaryotic cells have evolved from more simple prokaryotic organisms with which they share numerous common (or related) molecules [Margulis, 1970; Zillig, 1991]. Indeed, studies have shown that a number of eukaryotic organelles (namely mitochondria and chloroplasts) bear a close evolutionary relationship to specific groups of prokaryotes (namely  $\alpha$ -proteobacteria and cyanobacteria, respectively) [Gray, 1992]. So, how did the complex multicellular eukaryotes evolve from prokaryotes?

To understand the origin of the eukaryotic cells a number of theories have been proposed. Several of these theories are conjectural in nature, whereas others are based on limited experimental evidence. However, with the availability of huge quantities of sequence data and computational tools for phylogenetic analysis, the reconstruction of the evolutionary history of eukaryotic cells seems more plausible. In this submission, we discuss the proposed theories of the eukaryotic cell evolution and test them with phylogenetic analysis tools.

### **Evolution by borrowing**

The most crucial phenomena that led to the evolution of the eukaryotes is proposed to be endo-symbiotic contact of two species which contribute some of their cellular material and characteristics to the newly evolved product [Margulis, 1970]. The ancestors of eukaryotic cells formed symbiotic contacts with other organisms of other species and derived their cellular machinery along the way. By the proficient and widespread use of molecular systematics, Woese and colleagues found that prokaryotes formed two distinct phylogenetic groups, originally termed archaeobacteria and eubacteria [Woese, 1987].

How did we acquire the nucleus?

The origin of the eukaryotic nucleus, which is the hallmark distinction between eukaryotes and prokaryotes has been the center for debate for long. The recent widely accepted view represents eukaryotic nucleus is a chimera formed by the primary fusion of

an archaeobacteria and a bacteria. This chimeric model proposes that the ancestral eukaryotic cell arose as a result of a unique fusion event between a Gram-negative eubacteria and an 'eocyte' archaeobacterium [Iwabe et al, 1989; Gogarten et al, 1989; Gupta and Golding, 1996]. So, we borrowed our nucleus from this two organisms. Another recent hypothesis proposes that the eukaryotic nucleus evolved from a complex DNA virus [Bell, 2001]. According to this theory the virus established a persistent presence in the cytoplasm of a methanogenic mycoplasma and evolved into the eukaryotic nucleus by acquiring a set of essential genes from the host genome and eventually usurping its role. Again on the basis of  $\alpha$ -DNA polymerase phylogeny, Takemura proposed that the eukaryotic nucleus, derived its  $\alpha$ -polymerase from symbiotic contact of an orthopoxvirus ancestor with an archae-bacterium, whose genome already had a  $\delta$ -like polymerase [Takemura, 2001].

#### How did we acquire cellular organelle ?

We will discuss mainly two organelles here, namely Mitochondria and the Chloroplasts. Mitochondria are the eukaryotic organelles which carry out oxidative respiration, the final step in cellular respiration. Oxidative respiration breaks down the pyruvate formed from glycolysis to form carbon dioxide and produces the majority of the cell's ATP. The presence of its own genetic material, considerable sequence homology and similar morphology to bacteria led researchers to believe that mitochondria were the remnants of bacteria which had an endo-symbiotic contact with the eukaryotic ancestor [Margulis, 1970]. On a similar note, the chloroplast have been proposed to have evolved from prokaryotic ancestors of cyanobacteria through endo-symbiotic capture by eukaryotic hosts lacking such structures. Other organelle like cytoskeleton might also have been derived from prokaryotes.

To test the hypothesis of eukaryotic nucleus origin, which generated phylogenetic trees for DNA Polymerase (both alpha and delta forms) from human, drosophila, yeast, archaeobacteria (*Pyrococcus furiosus* and *Sulfolobus solfataricus*), Gram negative bacteria (*Escherichia coli*) and Poxvirus (*Vaccinia*).

#### **Materials and Methods**

The nucleotide and protein sequences for the phylogenetic analysis were obtained from the GenBank. (Eukarya) *Homo sapiens* alpha- and delta-polymerases, X06745 and M80397; *Drosophila melanogaster* alpha- and delta-polymerases, D90310 and X88928; *Saccharomyces cerevisiae* alpha- and delta-polymerases, J03268 and X15477; (Archaea) *Pyrococcus furiosus*, D12983; *Sulfolobus solfataricus* P2 B1, U92875; (Eubacteria) *Escherichia coli* polymerase II, X54847; and (Viruses) vaccinia virus, M13213. The phylogenetic trees were constructed using Phylip program [Thompson et. al., 1994; Felsenstein, 1989] from the Biology Workbench hosted by San Diego Supercomputer Center.

## Results

To test the chimera eukaryotic nuclear hypothesis, we generated a phylogenetic tree comprising of the  $\delta$ -DNA polymerases of the eukaryotes with DNA polymerases from archaea, eubacteria and poxvirus. The output trees (Figure 2.2.1.(a) unrooted and (b)rooted) show that the  $\delta$ -DNA polymerases are phylogenetically linked more closer to eubacteria and archaea than  $\alpha$ -DNA polymerases and poxvirus. This though not an sufficient proof validates the chimeric nuclear model. But it would be preliminary to predict or ascertain the ancestor of the fusion incase we assume the hypothesis to hold.

Viral Eukaryogenesis was tested by generating phylogenetic tree comprising of the  $\alpha$ -DNA polymerases of the eukaryotes with DNA polymerases from archaea, eubacteria and poxvirus. The output trees (Figure 2.2.2.(a) unrooted and (b)rooted) show that the  $\alpha$ -DNA polymerases are phylogenetically linked to poxvirus than eubacteria and archaea. This along with presence of linear chromosome and mRNA capping enzymes does validify the viral eukaryogenesis theory.

## Discussion

The seemingly difficult job of explaining the case when both the theories of nucleus evolution hold ground simplifies when seen from the perspective of the time of event. Both the theories will hold without contradicting the experimental results if we assume the nuclear fusion was followed by the symbiotic contact with complex DNA virus (like poxvirus). This will explain the presence of a DNA polymerase (delta form) derived from the fusion product of eubacteria and archaea and another alpha form which was acquired after a endo-symbiotic existence with a poxvirus ancestor. The evolutionary process described above generates an organism with many eukaryotic characteristics. In particular, it provides an origin for the nuclear structure, linear chromosomes, membrane fusion processes, mRNA capping, endosymbionts, and cellular compartmentalization. It also provides a rationale for the distribution of archaeal information-processing genes and bacterial metabolic genes found within the eukaryotic genome. The model predicts that many unique eukaryotic features will be related to those of viral lineages, rather than cellular lineages. So, we find that the cellular machinery has been borrowing a thing or two from the organisms it comes in contact for some critical time. The question is: are we still in the phase of borrowing and evolving?

## References

- Bell, P J L (2001) *J Mol Evol* 53: 251-256
- Gupta, R S and Golding, G B (1996) *TIBS* 21: 166-171
- Takemura, M (2001) *J Mol Evol* 52: 419-425
- Felsenstein, J (1989) *Cladistics* 5: 164-166
- Thompson, J D, Higgins, D G, and Gibson, T J (1994) *Nucleic Acids Res* 22: 4673-4680
- Margulis, L (1970) *Origin of Eukaryotic Cells*, Yale University Press
- Knoll, A H (1992) *Science* 256:622-627
- Zillig, W (1991) *Curr Opin Genet Dev* 1: 544-551
- Gray, M W (1992) *Int Rev Cytol* 141: 233-357
- Woese, C R (1987) *Microbiol Rev* 51: 221-271
- Iwabe, N et al (1989) *PNAS (USA)* 86: 9355-9359
- Gogarten, J P et al (1989) *PNAS (USA)* 86: 6661-6665