SPECIAL ARTICLE
UDC: 576.311.347:611.018.13

Arch Oncol 2006;14(1-2):11-4.
DOI: 10.2298/AOO0602011S

Eukaryotes arose after genetic recombination

Milanko R. Stupar

ABSTRACT

Division of ancestral prokaryotic pragenome into two circular double-stranded DNA molecules by genetic recombination, is a base for future separate evolution of nuclear and mitochondrial gene compartment. This suggests monophyletic origin of both, mitochondrion and nucleus. Presumed organism which genome undergoes genetic recombination has to be searched among an aerobic, oxygen non-producing, archaeon with no rigid cell wall, but a plasma membrane. Plastid evolves from an aerobic, oxygen producing protoeukaryote, after mitoplastid genome duplication and subsequent functional segregation.

KEY WORDS: *Genome; Recombination, Genetic; Phylogeny; Gene Expression Regulation, Developmental; Plastids; Replicon; Evolution, Molecular; DNA, Mitochondrial*

Institute for Nuclear Science Vinča, Laboratory for Molecular Biology RMG, Belgrade, Serbia; Address correspondence to: Stupar Milanko, 30, rue Sainte Marguerite, 67000 Strasbourg, France;
E-mail: jasna.stupar@laposte.net;
The manuscript was received: 14.05.2006, Provisionally accepted: 18.05.2006, Accepted for publication: 22.05.2006

INTRODUCTION

Pragenome evolves by acquiring new sequences, by mutation and by recombination of existing ones. Genome evolves in the way that genes participated in the same pathway are clustered, with compact DNA sequences. At the beginning of the evolution, 15-30 nucleotides were sufficient to allow functional activity of corresponding peptide, like YPITP motif of oxoacid: ferredoxin oxidoreductase (1), including another 3319 proteins with functional single domain (2). After almost 1.2 billion years of acellular (metabolism fixed to the solid iron-sulfide surface) and cellular (genetic) evolution, the pragenome was composed of hundreds of genes.

In the universal phylogenetic tree deepest branches consist of hypertermofiles. This suggests that last common ancestor of all life on Earth my have been an Archaea. Archaeal cell my display a mixture of features from Eubacteria and Eukaryotes. Archaea possess multiple types of energy productions. Genes for the gluconeogenic (glycolytic precursor) pathway (Embden-Meyerhof-Parnas) (3), reductive tricarboxylic acid cycle (2), reductive pentose phosphate cycle (4), Entner-Doudoroff pathway (5), Fe-hydrogenese catalytic H2 production (6), the sulphate assimilation (7), nitrogen fixation (8), are widely distributed among archaea and are often considered to be central to the origin of metabolism. Like TCA cycle, reverse TCA cycle may work together with phosphoenol pyruvate carboxylase and enzymes of glico neogenesis to fix inorganic carbon into sugars (2). Partial reactions of Entner-Doudoroff pathway may interact with the Embden-Meyerhof-Parnas pathway to replace steps of the standard glycolytic scheme in the way that alfa-protobacterial assistance is not necessary for mitochondrial function or existence. Gene for cytochrome oxidase, which presence indicated that aerobic metabolism is possible in an environment with a low level of oxygen, was present in common ancestor of archaea and eubacteria (9). That means that aerobic respiration was a monophyletic and ancient enzymatic system before oxygenic photosynthesis. Ribulose bisphosphate carboxylase, central enzyme in carbon fixation, was present in archaeal genomes (4). The genes of the electron transport chain (complex I-V), which are responsible for oxidative phosphorylation, are examined in archaeon *Natronomonas pharaonic* (10). Genes for NADH-dehydrogenase type II, succinate dehydrogenase, terminal oxidase, ATP synthase and equivalent for cytochrome-C reductase have been identified. Experimental studies provide the existence of a functional respiratory chain in these archaeon.

Genetic apparatus for all of this metabolic activity existed within the broad populations of Archaea before the last common ancestor.

To consolidate evolution by genetic recombination, ancient pragenome was developed genes for the family of site specific nickase-ligase enzymes and corresponding recognition sequence inverted and direct repeats (11,12). Helicase (13), topoisomerase I and II (14), gyrase, reverse gyrase (15), resolvase (16), integrase (12) have been find out in archaea. Following role of compact DNA sequence, reverse gyrase exhibits helicase and topoisomerase domain (17). *Sulfolobus solfaticus*, as others archaea which lives in extreme conditions has 1000 times higher risk of DNA damage than *E. coli*. To diminish this risk, expression of recombinase genes is permanent in ancient pragenome, so that they became housekeeping genes. In the putative replicon of *Sulfolobus* there is perfect direct repeats of 11bp:TCTATACCCCC, similar types of motifs can provide possible recombination site (18). One of the typical examples is a reciprocal genetic recombination between two copies of direct repeats which allow "figure eight" form of DNA. During the performance of its functions, prokaryotic DNA is typically attached to plasma membrane. It is possible that plasma membrane with attached DNA, where "figure eight" was resolved by recombination, could have invaginated and formed two-layered envelope surrounding two DNA molecules. Archaeal ftsZ , eubacterial ftsZ and eukaryal tubuline genes are homologue (19), which lead to conclusion that cell division apparatus are similar and originated from archaea. The dynamin has no mitochondrial import sequence and regulate membrane squeezing (20) and

peroxisomal (organelle surrounded by one membrane) fission (21). An ancestral polypeptide encoded by nuclear loci which can exhibits function of FtsZ and Dnm 1 domains, may operate in dual division mechanism of invagination, sucking down and constructing, to envelope DNA molecules. This envelop is presumed to roll on producing two DNA compartments, one with "luxury" (nuclear) and another with "working" (mitochondrial) genome. DNA sequence in prokaryotic genome tends to change relatively rapidly in the course of evolution. Three billion years after this recombination event, many kinds of modern prokaryots do not resemble organisms evolutionary proximal to the last common ancestor. By now 89 prokaryotic genomes have been sequenced (only 16 archaeals) and extrapolation of the results from comparative genome analysis way back in time is ungrateful.

RESULTS

Plastids appear to have retained more bacterial features than the mitochondria, particularly when considering transcription initiation and translation apparatus. The proteins identified in the mRNA processing/stabilizing complex in plastids are likewise mostly clearly recognizable homologs of bacterial factors then mitochondrial. That means that plastid genome was evolved after mitochondrial.

Mitochondrial origin

Monophyletic origin of mitochondrion and possibility that this organelle originated at the same time as the nuclear compartment, suggest that both arose after single event in prokaryotic cell. Suppose that pragenome of an aerobic, non-oxygen producing, archaeon, with no rigid cell wall but a plasma membrane, was a maxi circle with two replicons (Fig.1).

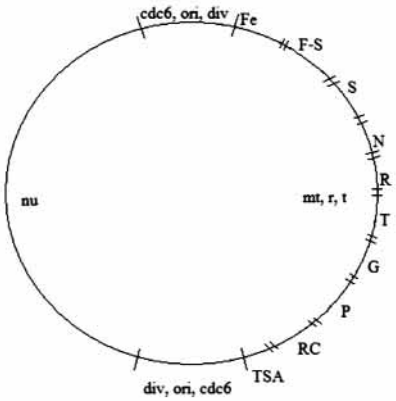

Figure 1. Possible genome organization of the last common ancestor, an aerobic non-oxygen producing archaeon, prior to the division by genetic recombination. Fe = Fe-hydrogenase operon, F-S=Fe-S protein assembly cluster, S=sulphate assimilation cluster, N=nitrogen fixation cluster, R=rRNA genes, T=tRNA genes, G=glycolytic pathway, P=pentose phosphate pathway, RC=respiratory chain, TCA=tricarboxylic acide cycle, nu=†nulear gene content, mt, r, t=mitochondrial gene content, div=dividing sequences, ori=origin of replication, cdc6=cell division gene. With ef-rpo-hsp operon in R cluster

One which activates the origin recognition complex for nuclear and another for mitochondrial gene compartment. At or in the vicinity of each replicon there are two "hotspots", two direct repeat sequences. The initiation of replication occurs through loading of the minichromosome maintenance protein (Mcm), which can activate cell division cycle protein Cdc6. In the most archaea cdc6 gene is adjacent to the origin of replication. Frequent exposure to DNA damage in ancient archaea result in frequent recombination and genomic rearrangement. It can be that initiation of replication process trigger these genomic rearrangements. If the direct repeat sequence are near or at the origin of replication, this can lead to the functional segregation of the maxi circle. During recombination DNA molecule is relaxed, unwound.

Ones reaction was terminated, DNA molecule tends to supercoil and, still attached to the membrane and coupling with activity of cdc6 gene and replication, cause its fission and enveloping. Once maximal capacity has been reached, additional sequences tends to interrupt such DNA molecule. To prevent integrity, pragenome has been divided, allowing stability of growing nuclear genome, as well as mitochondrial specialization.

In this transition phase, immediately after genome separation, when intensive transfer of information between mt and nu has not yet established, mitochondria needs its own elementary replication-transcription-translation machinery and protein degradation-secretion. This is the reason why these pathways exist in some of mt genome (22). Each product of reciprocal genetic recombination has one copy of the origin of replication, i.e. cdc6-oridiv cluster. Archaeal primase may act as a bifunctional replicase to incorporated both NTPs and dNTPs to bringing together replication and transcription (23). If the gene for primase is at origin of replication, the impact is obvious. Genes involved in protein synthesis, including tuf, fus, rpoA, rpoB, rpoC and some chaperons are encoded within or proximal to the large ribosomal protein cluster in bacteria but not found in archaea. It is reasonable to predict presence of : -tuf (ef-1 alpha)-rpo-hsp gene operon in the ribosomal cluster of the last common ancestor "mitochondrial genome".

Pragenome organization are in a tightly correlation with order of genes expression. Genes located at the start of the cluster are activated earlier than genes further along. To avoid progressive decrease in their efficiency of expression, as one moves from the start to the end of cluster, clusters are separated by nontranscribed spacers (Figure 1).

Plastid origin

All extant photosynthetic cells descend from a single common ancestor that possessed a primeval photosynthetic mechanism. Essential components of photosynthetic apparatus are present in both archaea and eubacteria, including the nonphotosynthetic eubacteria, suggesting that photosynthesis could be a primitive property of both groups.

Photosynthesis was developed in archaea more then 3 billion of years ago. Bacteriorhodopsin, observed in archaea, act as a light-activated proton pump but do not produce storable energy (24). Nitrogenase iron protein (Nif H) gene from *M. jannischii* has remarkable similarity with gene bch L involved in bacteriochlorophyll biosynthesis (from 24). In the archaeon *Haloarcula marismorti* (2), nine plastocyanin precursor-like proteins were identified, as well as phycocyanobilins, which are major components of macromolecular complex termed phycobilisomes that are involved with capture of light energy. *H. marismorti* has at least 29 unique proteins containing a light-responsive domain motif (GAF) found in plant and cyanobacterial phytochromes, and invertebrate cGMP-stimulated phosphodiesterase, of which several are likely to be phytochrome or phytochrome-like genes. Set of opsins facilitate phototaxis in this archaeon which are capable to regulate metabolism in response to day/night cycle grace a circadian clock regulator-like gene (kai C). Phylogenetic analysis of the core antenna domain (25) support a single phylogeny in which PS II core antenna proteins (Psb C, Psb B) arose from duplications of reaction centre (RC) 1 (=PS I), associated core antenna and accessory antenna proteins (Isi A, Pcb A, PcbC) arose from duplication of Psb B. RC 1 are parts of a single polypeptide already present in an ancestor of cyanobacteria. So that chlorophyll based photosynthesis evolved from homodimeric protein encoded by ancient archaea (bacteriochlorophyll biosynthesis evolved before chlorophyll biosynthesis) (Figure 2).

Mg-tetrapyrrole-based photosynthesis starts to develop in a single, continuous evolutionary archaeal line of cells and was a final invention of mitoplastid genome, prior to the chloroplast/cyanobacteria lineage. Acquisition of complete photosynthetic pathway was a signal for mitoplastid genome duplication and subsequent functional fission.

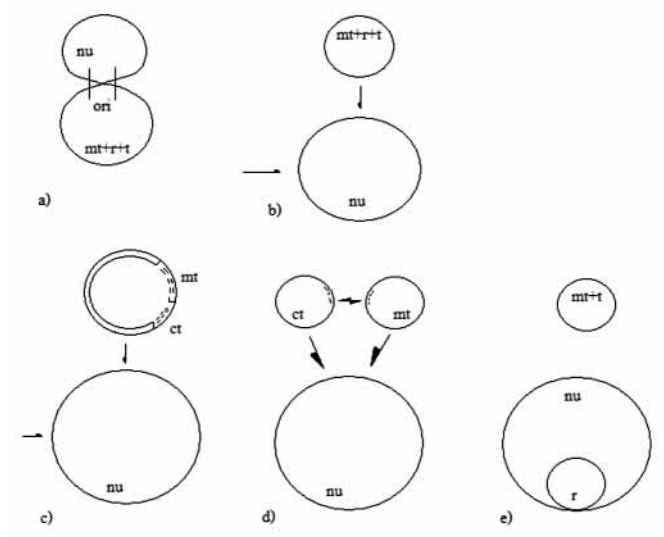

Figure 2. Evolution of mitochondrial, plastide and ribosomal compartments. a) Bipartite genome with nu (nuclear) and mt+r+t (mitochondrial) replicons. b) Division and enveloping of both compartments, c) Mitoplastide genome (ct=photocluster +mt=mitochondria) duplication, d) Division of mitoplastide genome with partial sisters genome deletion and subsequent organelle fission. e) Complete ribosomal genes transfer to the nucleus and ribosomale compartmenet evolution (r). Arrows indicate gene transfer and acquisition of a new genes. Photocluster include: hli (high light inducible gene),set of opsins genes for phototaxis, kaiC (circadian clock regulator), photoreceptor operon (pgb-protoglobin, plastocyanin, ubiquinone, phycocyanobilins, phycobilisome homodimeric reaction center), nif genes, cytb-c,rubisco, ferredoxin, wox (water oxidation-like gene), psa, psb...

DISCUSSION

This hypothesis proposed existence of a single evolutionary line of cells, with basic characteristic which passed, as a backbone trough all living organisms, and that the direct eukaryotic ancestor should look for within archaea. The major source of motivation for this hypothesis is "figure eight" forming (precisely mathematical symbol for infinity) by double circles of DNA molecule. An explanation for this enigma has been search for almost 30 years.

Later and recent discoveries in molecular biology was lead to the understanding and possible solution, by genetic recombination, of "figure eight" foggy. Fundamental similarities in the metabolic pathways in cellular organisms provide a starting evidence for evolutionary conservatism operating over four billion years. To give proof of the genetic recombination engagement in the origin of eukaryotes, several facts has to be considered. If mitochondrial genome arises after separation from nuclear, than archaea is the oldest group of cellular form of life and worldwide ancestor for all living organisms. Which mechanism decide the way for new gene acquisition? Good example is transition from aerobic respiration to the oxygenic photosynthesis. Selective pressure favored acquisition of a new - water oxidation gene (wox), but under the formerly prepared background ("when gene is present, that means that coding pathway is in advance stage"), i.e. photocluster (gens for day/night cycle, O_2 detection and transport, mobile electron carriers, including duplication of already existing quinine, ferredoxine, NADP reductase). The recent discovery of photosynthetic core genes in viruses (26,27) present the proof for lateral gene transfer (LGT). Many viruses that infect cyanobacteria carry photosynthesis genes. So that similarity between chloroplasts and modern cyanobacteria is due to LGT from mitoplastid photocluster to cyanobacterial precursor. It can be the same for the similarity between mitochondria and alpha-proteobacteria. Fe-S proteins are crucial in respiration, nitrogen fixation and photosynthesis. Two groups of 6-7 genes each have been identified that are required for Fe-S cluster assembly, isc and suf. isc (iron-sulfur cluster) genes are present in yeast and animals, and the biosynthetic pathway is located in mitochondria. In plants, both the isc and suf genes have been found, some of the isc gene products have been localized to the mitochondria whereas all 6 suf gene products are likely to reside in the plastids. On the other hand, isoforms of a protein NFU occur in both organelles. It is not known why Fe-S cluster assembly in each type of organelle is carried out by different sets of proteins. If genes participated in Fe-S cluster biosynthesis possess a common ancestor, this fact can contribute to the mitoplastid origin and its genome duplication with subsequent fission giving birth to the mitochondrial and plastid gene compartment in plants.

Phylogenetic evidence presumed that primitively anaerobic amitochondriated eukaryotes containing the nucleus, cytoskeleton and endomembrane system may have never existed (28). The results presented in this article placed origin of mitochondrion in prokaryotic world, this is what step-by-step hypothesis did as well.

Usually the replicons arose after segregation contain lower G+C content than major archaeal chromosome. G+C content in mitochondria is less than 50% and the G+C contents in archaealís major chromosomes is above 60%. Origin of plastid take place in an aerobic, 02-producing protoeukaryote by duplication of mitoplastid (mt genome + photosynthetic mini replicon) genome and subsequent organelle fission with sisters genome partial deletion (Figure 2d). Traces of this event is visible in plan mt genome where plastid-like sequences exist, non-functional pieces of psa, ndh, rbc, rpo, psb D... genes (29), extensive sequences exchange between mt and pt genome. In plant mitochondrial genome there are plastid-like sequences ranging from 1%-4% of the total genome, most of them are non-functional. In the *Arabidopsis* and rapeseed there is the same sets of genes for rRNAs and tRNAs (29), probably vestige of ancient partial duplication of "mitoplastid" genome (Figure 2).Wheat mitochondrial genome has 55 sequences homologous (mostly with 80% or higher homology on the nucleotide basis) to the corresponding sequences of the wheat chloroplast genome, the total size, 26264 bp, corresponds to 25% of the mt coding region (30). Nitrogenase operon and respiratory chain (reverse photosynthesis) are mitochondrial encoded and photosystems I and II proteins are plastid encoded. Homology and evolutionary connection between these three groups of genes indicate their ancestral and common origin which occurred in a single, continuous line of cells, encoded by mitoplastid genome. In modern organisms glycolysis is catalyzed in cytosol and photosynthesis initiated in chloroplast (instead in macromolecular complex - phycobillisome in archaea) and continued in cytosol, because coupling hexose-pentose pathway is archaeal vestige. Two of the most important pieces of biochemical innovation that occurred in early biosphere - the development of photosynthesis and nitrogen fixation may be related to each other because some of their key genes [nitrogenase gene and bacteriochlorophylle gene (both with circadian rhythm)] appear to have evolved from a common ancestor that may be part of a third, significantly different, biochemical process, encoded by the unique genome. Direct repeat (div-dividing) sequences at the replication origin, which can represent recombination site has to be searched in nuclear/mitochondrial genomes. The "Fe-only" hydrogenase gene from *Nyctotherus ovalis* (31) hybridizes to genomic fragment which terminates in G3T4G3(T4G4)5 repeat that is very similar to the telomere sequences. In the *N. ovalis*, hydrogenosome genome starts with Fe-hydrogenase gene. To speculate, telomere-like sequence, which can be one of the div-sequence candidate, starts the mt replicon and precede Fe-hydrogenase gene in the last common ancestor.

In the human mt genome there are two rRNA genes, 12S rRNA and 16S rRNA. Transferred to the nucleus, they can cause new organelle origin - ribosomal compartment, infrastructurally formerly defined by nucleolar organizers and nucleolus. Just as it is the case with

transition from aerobic respiration to the oxygenic photosinthesis, this transition will create new outlook of the human evolution.

Communication between mitochondria and nucleus is important for carbohydrate and nitrogen metabolism, cell cycle and proliferation. Mitochondria-to-nucleus stress activates the genes implicated in carcinogenesis. As a result of retrograde regulation, the tumor cell transforms to a unicellular lifestyle and descents on one's back in evolution till the last common ancestor, just prior to the separation of mitochondrial from nuclear genomes.

Acknowledgement

To the Prof. Dr. Vladimir Leskovac and family Stupar.

REFERENCES

1. Fukuda E, Kino H, Matsuzava H, Wakagi T. Role of a highly conserved YPITP motif in 2-oxoacid-ferredoxin oxidoreductase. Eur J Biochem 2001;268:5639-46.
2. Baliga N, Bonneau R, Facciotti MT, Pan M, Glusman G, Deutsch EW, et al. Genome sequence of *Haloarcula marismortui*: A halophilic archaeon from the Dead Sea. Genome Res 2005;14:2221-34.
3. Ronimus RS, Morgan HW. Distribution and phylogenesis of enzymez of the EmbdenMeyerhoff-Parnas pathway from archaea and hyperthermophilic bacteria support a gluconeogenic origin of metabolism. Archaea 2003;1:199-221.
4. Soderberg T. Biosynthesis of ribulose-5-phosphate and erythose-4-phosphate in Archaea: a phylogenetic analysis of archaeal genomes. Archaea 2005;1:347-52.
5. Canback B, Andersson SGE, Kurtland CG. The global phylogeny of glycolytic enzymes. Proc Natl Acad Sci 2002;99:6097-102.
6. Kanai T, Ito S, Imanaka TJ. Characterization of a cytosolic NiFe-hydrogenase from the hyperthermophilic Archaeon *Termococcus kadakarensis* KOD 1. J Bacteriol 2003;185:1705-11.
7. Teske A, Dhillon A, Sogin ML. Genomic markers of ancient anaerobic microbial pathways: sulphate reduction, methanogenesis and metane oxidation. Biol Bull 2003;204:186-91.
8. Chien YT, Zinder SH. Cloning, functional organization, transcript studies and phylogenetic analysis of the complete nitrogenase structural genes (nif HDK2) and associated genes in the archaeon *Methanosarcina bakeri* 227. J Bacteriology 1996;178:143-8.
9. Castresana J, Lubben M, Saraste M, Higgins DG. Evolution of cytochrome oxidase, an enzyme older than atmospheric oxygene. EMBO J 1994;13:2516-26.
10. Falb M, Pfeiffer F, Palm P, Rodewald K, Hickmann V, Tittor J, et al. Living with two extremes: Conclusions from the genome sequence of *Natronomonas pharaonis*. Genome Res 2005;15:1336-43.
11. Greve B, Susane J, Brugger K, Zillig W, Garret RA. Genomic comparison of archaeal conjugate plasmids from *Sulfolobus*. Archaea 2004;1:231-9.
12. She Q, Singh RK, Confaloneri F, Zivanovic Y, Allard G, Awayez MJ, et al. The complete genome of the cranarchaeon *Sulfolobus solfaticus* P2. Proc Natl Acad Sci 2001;98:7835-40.
13. Constantinesco F, Forterre P, Koonin EV, Aravind L, Elie C. A bipolar DNA helicase gene, herA, clusters with rad50, mre11 and nur A genes in thermophile Archaea. Nucleuc Acid Res 2004; 32:1439-47.
14. Forterre P, Bergerat A, Lopez-Garcia P. The unique DNA topology and DNA topoisomerase of hyperthermophilic archaea. FEMS Microbial Rev 1996;18:237-48.
15. Guipalid O, Marguet E, Noll KM, de la Tour, Forterre P. Both DNA gyrase and reverse gyrase are present in the hyperthermo-philic bacterium *Thermotoga maritime*. Proc Natl Acad Sci 1997;94:10606-11.
16. Komori K, Sakae S, Shinagava H, Morikawa K, Ishio Y. A Hollyday junction resolvase from Pyrococcus furiosus: Functional similarity to Escherichia coli RuvC provides evidence for conserved mechanism of homologous recombination in bacteria, eukarya and arhaea. Proc Natl Acad Sci 1999;96:8873 -8.
17. de la Tour CB, Portemer C, Kaltoum H, Duguet M. Reverse gyrase from the hyperthermophilic bacterium *Thermotoga maritima*: properties and gene structure. J Bacteriol 1998;180:274-81..
18. Berquist B, DasSarma. An archaeal chromosomal autonomously replicating sequences element from an extreme halophile, *Halobacterium sp.* strain NRC-1. J Bacteriol 2003;20:5959-66.
19. Bauman P, Jackson SP. An archaebacterial homologue of the essential eubacterial cell division protein Fts Z. Proc Natl Acad Sci 1996;93:6726-30.
20. Erickson HP. Dynamin and Fts Z: missing links in mitochondrial and bacterial division. J Cell Biol 2000;148:1103-5.
21. Fujimoto M, Arimura S, Tsutsumi N. Two *Arabidopsis* dynamin-related proteins 3, DRP3A and DRP3B are involved in mitochondrial and peroxisomal fission. ICPMB; 2005 May 28-June 2, Obernai, France; 2005.
22. Lang FB, Burger G, O'Kelly CJ, Cedergren R, Golding GB, Lemieux C, et al. An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. Nature 1997;387:493-7.
23. Bocquier AA, Liu L, Cann IK, Komori K, Kohda D, Ishino Y. Archaeal primase: briging the gap between RNA and DNA polymerase. Curr Biol 2001;11:452-6.
24. Xiong J, Kazuhito I, Bauer CE. Tracking molecular evolution of photosynthesis gene cluster from *Heliobacillus mobilis*. Proc Natl Acad Sci 1998;95:14851-6.
25. Mix LJ, Haig D, Cavanaugh CM. Phylogenetic analysis of the core antenna domain: investigating the origin of photosystem I. J Mol Evol 2005;60:153-63.
26. Mann NH, Cook A, Millard A, Bailey S, Clokie M. Bacterial photosynthesis genes in a virus. Nature 2003;424:741.
27. Chio Charles G. Virus may aid photosynthesis. News from The Scientist 2005;6:20051013-01 page.
28. Emelyanov VV. Mitochondrial connection to the origin of the eukaryotic cell. Eur J Biochem 2003;270:1599-618.
29. Handa H. The complete nucleotide sequence and RNA editing content of the mitochondrial genome of rapeseed (*Brasica napus* L): comparative analysis of the mitochondrial genomes of rapeseed and Arabidopsis thaliana. Nucleic Acid Res 2003;31(20):5907-16.
30. Ogihara Y, Yamazaki Y, Murai K, Kanno A, Terachi T, Shiina T, et al. Structural dinamics of cereal mitochondrial genomes as reveald by complete nucleotide sequencing of the wheat mitochondrial genome. Nucleic Acids Res 2005;33:6235-50.
31. Akhamova A, Vonckeen F, vanAlen T. A hydrogenosome with genome. Nature 1998;396:527-8.