MEDWIN PUBLISHERS
Committed to create value for Researchers

Physical Science & Biophysics Journal

A Mathematical Formulation of Evolution and Innovation I. Unicellular Organisms

Otsuka J*

JO Institute of Biophysics, Japan

***Corresponding author:** Otsuka J, JO Institute of Biophysics, Nerima-ku, Tokyo 178-0063, Japan, Tel: +813 3921 1466; E-mail: jin.otsuka@kyj.biglobe.ne.jp

Review Article

Volume 1 Issue 1

Received Date: December 08, 2017

Published Date: December 20, 2017

Abstract

Most types of evolution indicated on unicellular organisms are shown to be derived mathematically from the time change equation of organisms that self-reproduce, occasionally mutate, and die. Darwinian evolution corresponds to the approximate solution with respect to the first order of mutation term, showing that the mutant with the higher self-reproducing rate and/or slower death rate increases its number. Such a mutant is mainly due to the nucleotide base substitutions. However, the scaffold, under which Darwinian evolution takes place, has been sometimes changed during the longer time. The well-known examples are the gene duplication to generate new genes and the gene transfer from the endosymbiont to the host genome of eukaryotes after the acceptance of endosymbionts. The mutants having experienced such drastic change in genome first decline to the minor members in the population but some of them recover the increase rate as a new style of organisms by Darwinian evolution under the reconstructed genome. The evolutionary process containing such innovation is formulated by successively solving the time change equations of satellite variants up to the higher order of mutation terms until the appearance of new style organisms. This mathematical formulation applied to the self-reproducing proto-cells in the RNA-protein world also provides a possibility to explain the origin of protein genes as well as of translation apparatus and the transition of the proto-cells to the DNA-RNA-protein world.

Keywords: Endosymbiosis; Gene duplication; Mutation; Organization; Selection; Self-reproduction

Introduction

In the famous book by Darwin [1], the gradual accumulation of selectively advantageous variants is proposed qualitatively for the generation of new species from the artificial selection of domestic animals and plants as well as from the observation of unique species in a geographically isolated region. The core of this proposal has become evident, after the re-discovery of Mendelian heredity, by the detection of hereditary variants, i. e., mutants, and extensive investigations have been carried out for the behavior of mutants especially in the Drosophila population [2-5]. In parallel, Darwinian evolution is mathematically formulated in population genetics to estimate the probability that a spontaneously arisen mutant is fixed in the population, depending on its selective parameter value [6,7]. This study also reveals that a selectively neutral mutant is fixed with the probability equal to the mutation rate, independently of population size.

The gene and genome sequencing since the latter half of the last century has brought new information about the evolution of organisms. First, the comparison of orthologous genes from different species of organisms finds the selectively neutral nucleotide base substitutions with the constant rate of 10^{-9} per site per year [8-10] and makes it possible to reconstruct the phylogeny of the present-day organisms, revealing the following striking evolutionary pattern. All free-living organisms are traced back to the ancient divergence of proteobacteria, eubacteria and eukaryotes [11-14]. The mitochondria in eukaryotes have been derived from the endosymbionts of O_2-respiratory eubacteria, and some of such eukaryotes have further acquired the photosynthetic plastids as the endosymbionts of cyanobacteria [15-18]. The multi-cellularity and cell differentiation have occurred in the eukaryotes after the acquirement of the mitochondria. Second, the amino acid sequence similarities of paralogous proteins strongly suggest that the repertoire of protein functions has been expanded by gene duplication and by the succeeding nucleotide base changes, partial deletion and/or insertion, and domain shuffling in the counterpart of duplicated genes [19-22]. Moreover, the studies of molecular biology begin to discuss the problems how the translation apparatus has been formed in the RNA-protein world and how the DNA-RNA-protein world has been realized [23]. These recent results indicate the necessity of the more general formulation of evolution than the population genetics.

In the present paper, the evolutionary types of unicellular organisms listed above are shown to be derived mathematically from the time change equation of the organisms that self-reproduce, occasionally mutate, and die. While Darwinian evolution is derived from the selective advantage of the first order of mutation term, the other types of evolution such as the gene duplication and the reception of endosymbionts must be formulated by considering the mutants first decreased in their increase rate and the succeeding steps of mutation to recover the increase rate. The change in the increase rate of such a mutant is measured by the thermodynamic quantity of biological activity proposed previously [24,25]. This formulation provides a possibility to explain the formation of translation apparatus in the self-reproducing proto-cells of the RNA-protein world and the transition of the proto-cells to the DNA-RNA-protein world.

Mathematical Formulation of Evolution and Innovation

In the population of unicellular organisms taking a common material and energy source M from the outside, the number $n_{xi}(t)$ of variants with the internal variable x_i of genetic information carriers and their systematization generally obeys the following time change equation;

$$\frac{d}{dt}n_{xi}(t)=\{Q_{xi}(t)R(M;x_i)-D(x_i)\}n_{xi}(t)+\sum_{j(\neq i)}q_{xi,xj}(t)R(M;x_j)n_{xj}(t) \quad (1)$$

where the self-reproducing rate and the death rate of the variant x_i are denoted by $R(M;x_i)$ and $D(x_i)$, respectively. The apparent decrease factor $Q_{xi}(t)$ for the self-reproducing rate of variant x_i means the mutation of variant x_i to other kinds of variants and is related with the mutation terms $q_{xj,xi}(t)$'s in the following way:

$$Q_{xi}(t)=1-\sum_{j(\neq i)}q_{xj,xi}(t) \quad (2)$$

As indicated by Eigen [26], the behavior of such population becomes transparent by transforming Eq.(1) into two types of equations; one concerning the total number $B(t)$ of all kinds of variants defined by

$$B(t)=\sum_i n_{xi}(t) \quad (3)$$

and another concerning the fraction $f_{xi}(t)$ of the variants x_i defined by $n_{xi}(t)/B(t)$. By simple calculation, these equations are obtained from Eq. (1) to the following forms, respectively, using the relation (2).

$$\frac{d}{dt}B(t)=W_{av}(M;t)B(t) \quad (4)$$

and

$$\frac{d}{dt}f_{xi}(t)=\{W(M;x_i)-W_{av}(M;t)\}f_{xi}(t)+\sum_j q_{xi,xj}(t)R(M;x_j)f_{xj}(t) \quad (5)$$

where $q_{xi,xi}(t)$ is defined by $Q_{xi}(t)$ *-1*. The increase rate $W(M;x_i)$ of the variant x_i and the average increase rate $W_{av}(M;t)$ of all organisms in the population are defined by

$$W(M;x_i)\equiv R(M;x_i)-D(x_i) \quad (6)$$

and

$$W_{av}(M;t) \equiv \sum_i W(M;x_i) f_{xi}(t) \qquad (7)$$

respectively.

Darwinian evolution corresponds to the following approximate evaluation of Eq. (5) by only focusing on the first order of mutation term. If the increase rate $W(M;x_i)$ of an occasionally arisen mutant x_i is larger than the average increase rate of the population, that is, $W(M;x_i) - W_{av}(M;t) > 0$, the fraction $f_{xi}(t)$ of the variants x_i increases with time, according to the first term on the right side of Eq. (5). This raises the average increase rate $W_{av}(M;t)$, resulting in the increase in the total number $B(t)$ of the organisms according to Eq. (4), although this increase is ultimately stopped by the decrease in available material and energy source M. On the other hand, the fraction $f_{xi}(t)$ decreases when $W(M;x_i) - W_{av}(M;t) < 0$. Thus, the organisms taking a common material and energy source M are elaborated by mutation and selection, and most of them reach the ones x_{opt} with the optimum increase rate. In such Darwinian evolution, the mutants are mostly due to the nucleotide base substitutions in an almost constant size of genome. It depends on the environment which variant is the most advantageous, and this evolution yields new species by the geographical isolation and/or climate change.

However, the organisms have sometimes experienced the more drastic change in genome as indicated in Introduction. By such drastic change, the mutants first decrease their increase rate, declining to the minor members in the population, but some of them recover the increase rate by Darwinian evolution under the reconstructed genome. For formulating mathematically such evolutionary process, the fraction of variants with the lower increase rate is considered more accurately after the organisms x_{opt} become dominant. For this purpose, Eq. (5) will be formally integrated with respect to time t, i. e.,

$$f_{xi}(t) = \exp[\int_0^t \{W(M;x_i) - W_{av}(M;\tau)\}d\tau][\int_0^t \sum_j q_{xi,xj}(\tau) R(M;x_j) f_{xj}(\tau) \exp[-\int_0^\tau \{W(M;x_i) - W_{av}(M;\tau')\}d\tau']d\tau + f_{xi}(0)] \qquad (8)$$

Among the fractions $f_{xi}(t)$'s in this expression, we focus on the fraction $f_{xi1}(t)$ of the variants x_{i1}, which have arisen from the dominant organisms x_{opt} by a drastic change in genome. Then, the average increase rate $W_{av}(M;\tau)$ and $W_{av}(M;\tau')$ are approximated to be $W(M;x_{opt})$ and the fractions of other variants except for x_{opt} are neglected on the right side of Eq. (8). The mutation terms $q_{xi1,xopt}(\tau)$ are averaged over a sufficiently long time:

$$q_{xi1,xopt} = \frac{1}{t}\int_0^t q_{xi1,xopt}(\tau)d\tau \qquad (9)$$

This quantity is used as the mutation rate instead of the mutation term. In this large time scale, the fraction $f(x_{i1})$ is present with the following relation to the fraction $f(x_{opt})$ of dominant organisms as a stationary state.

$$f(x_{i1}) = \frac{q_{xi1,xopt} R(M;x_{opt})}{W(M;x_{opt}) - W(M;x_{i1})} f(x_{opt}) \quad (10\text{-}1)$$

The fraction $f(x_{i2})$ of the second step of variants x_{i2}, which have experienced further genome change in the variants x_{i1}, is also obtained from Eq. (8). By focusing on the mutation term of $q_{xi2,xi1}(\tau)$ of $f_{xi1}(\tau)$ on the right side of this equation, the fraction of $f(x_{i2})$ of variants x_{i2} is also expressed as

$$f(x_{i2}) = \frac{q_{xi2,xi1} R(M;x_{i1})}{W(M;x_{opt}) - W(M;x_{i2})} f(x_{i1}) \quad (10\text{-}2)$$

By the repetition of the similar procedure, the fraction $f(x_{in})$ of variants x_{in}, which have experienced n steps of genome changes, is obtained as

$$f(x_{in}) = \frac{q_{xin,xin-1} R(M;x_{in-1})}{W(M;x_{opt}) - W(M;x_{in})} f(x_{in-1}) \quad (10\text{-n})$$

When a new style of organisms y appear from such minor members x_{in} with the probability $q_{y,xin}$ by reconstructing the genome, the probability $P_n(y<\!\!-x_o)$ of transition from the original style of organisms x_o to the new style of organisms y via n steps of genome changes is expressed as

$$P_n(y \leftarrow x_o) = q_{y,xin} \prod_{m=2}^{n} \frac{q_{xim,xim-1} R(M;x_{im-1})}{W(M;x_o) - W(M;x_{im})} \frac{q_{xi1,xo} R(M;x_o)}{W(M;x_o) - W(M;x_{i1})} \qquad (11)$$

Here, x_{opt} in Eqs. (10-1) ~ (10-n) is replaced by x_o.

If the new style organisms y are elaborated to utilize the material and energy source M more efficiently by Darwinian evolution, they compel the original style organism x_o to be extinct. If the new style organisms y

utilize a new material and energy source *L* other than *M* and/or move to a new area, on the contrary, they form a new population, where the fraction $f_{yk}(t)$ of variants y_k obeys the following equation, apart from the population of original style organisms x_i in Eq. (5).

$$\frac{d}{dt} f_{yk}(t) = \{W(L; y_k) - \overline{W}(t)\} f_{yk}(t) + \sum_l q_{yk,yl}(t) R(L; y_l) f_{yl}(t) \quad (12)$$

Here, the average increase rate of new and original styles of organisms is defined by

$$\overline{W} \equiv \sum_i W(M; x_i) f_{xi}(t) + \sum_k W(L; y_k) f_{yk}(t) \quad (13)$$

The total number *B(t)* of organisms is re-defined as

$$B(t) = \sum_i n_{xi}(t) + \sum_k n_{yk}(t) \quad (14)$$

and obeys the following equation.

$$\frac{d}{dt} B(t) = \overline{W}(t) B(t) \quad (15)$$

This divergence of new and original styles of organisms can occur without geographical isolation and/or climate change. The generation of the same style of new organisms *y* from the original style organisms *x* is ceased after the new style organisms *y* are established by Darwinian evolution. This is because the immature new style organisms *y* arising from the original style organisms *x* are compelled to be extinct by the struggle for existence with the new style organisms y_{opt} established already. In the present paper, the above type of evolution, which cannot be attained only by Darwinian evolution, will be called innovation.

Application to Four Examples of Innovation

In this application, the internal variables are specified before and after each innovation and the physiological problem resolved by this innovation is also pointed out.

(i) Evolution of Prokaryotes by Gene Duplication

In the unicellular monoploid organism such as prokaryote, the internal variable x_o corresponds to the DNA genome size N_o and the systematization S_{No} of its products [24,25]. The variants x_{in} having experienced *n* steps of gene duplication first decrease their increase rate by the load of carrying such an expanded genome $N_o + \Delta_1 N + \Delta_2 N + \ldots\ldots\ldots + \Delta_n N$, while the systematization remains S_{No}. However, some of them recover their increase rate as a new style of organisms *y* by the increase in acquired energy under the advanced systematization $S_{No+\Delta 1N+\Delta 2N+\ldots\ldots\ldots+\Delta nN}$ of new genes generated from the counterparts of duplicated genes. In the prokaryotes, the gene duplication severely lowers the increase rate, and the evolution by gene duplication must have been taken place by the recurrence of the evolutionary process from Eq. (5) to Eq. (12) with the transition probability $pn(y \leftarrow x_0)$ of relatively small number *n*. This is exemplified by the presence of eubacteria in the intermediate stages of evolution to the O_2-respiration and O_2-releasing photosynthesis, which have much influenced upon the recent evolution of eukaryotes.

According to the similarity search of amino acid sequences, some ubiquitous permeases show considerable similarities to the Sdh *C* and Sdh *D* chains of succinate oxidoreductase (complex II), Frd *C* and Frd *D* chains of quinol-fumarate reductase, cytochrome *b*'s in cytochrome bc_1 complex (complex III) and cytochrome b_6f complex, the subunit of cytochrome *c* oxidase and the core proteins in the photoreaction centre, while the DNA associated proteins Rec *A* and Rho show much similarity to the mutually similar subunits forming the catalytic domain of ATP synthase [27,28].

From the degree of sequence similarity between these proteins and the reconstructed phylogeny of eubacteria carrying these proteins [18], the evolutionary steps of eubacteria to the O_2-respiration can be traced; first, ATP synthesis utilizing the proton concentration difference across the cytoplasmic membrane, which is caused by the pumping out of protons coupled with the electron transport from NADH dehydrogenase (complex I) to complex II, thereby requiring the resolution of reactive oxygen species, as seen in anaerobic eubacteria, and then the appearance of bc_1 complex and terminal oxidase to produce H_2O molecule from protons, transported electrons and O_2 molecules without the generation of reactive oxygen.

The photosynthetic system also seems to have started from the ATP synthesis. In the purple bacteria, the simple core protein consisting of mutually similar subunits *L* and

M functions as the proton pumping using the photon energy. The proton concentration difference thus produced across the membrane is used to synthesize ATP molecules by the ATP synthase very similar to the ATP synthase in the respiratory system. The electron released by the photolysis return to the core protein through quinones, cytochrome *bf* complex and cytochrome *c*. The cytochromes *b* and *f* are very similar to the cytochrome *b* in the complex III and cytochrome c_1, respectively. The principal component analysis of amino acid residues suggests that the *L* and *M* subunit genes have evolved to the genes of *D1* and *D2* subunits of photosystem II and to the genes of *A1* and *A2* subunits of photosystem I in cyanobacteria by additional insertion of new domains [29].

However, the excellent abilities of O_2-respiration and O_2-releasing photosynthesis cannot be fully displayed in the simple cell structure of eubacteria. In the O_2-respiration, the ATP synthesis is carried out by using the lower proton concentration in the cytoplasm. This is incompatible with the syntheses of other biomolecules using protons. Although the O_2-releasing photosynthesis taken place in the thylakoid of cyanobacteria produces protons to synthesize the sugars by decomposing H_2O molecules, the cyanobacteria also carry the O_2-respiratory system in their cytoplasmic membrane. Thus, the genome of eubacteria with these abilities is also limited to the order of $10^6 bp$ compactly encoding 3000 ~ 4000 genes, like other prokaryotes [30].

The evolution of chemical syntheses in other prokaryotes may also be traced along the line of gene duplication.

(ii) Innovation of Eukaryotes by the Acquirement of Endosymbionts

When the ancestral eukaryote received O_2-respiratory eubacteria as endosymbionts, the host cell would have its nutrients more or less stolen, and its increase rate must have been first decreased. The internal variable of host-endosymbionts at this stage corresponds to x_{i1} in the formulation of the preceding section, while the internal variable of the ancestral eukaryote corresponds to x_o. However, such a host cell is hopeful of utilizing the ATPs synthesized in the endosymbionts and the protons released from them. To realize this utilization and decrease the stolen nutrients, the host-endosymbionts must further experience at least following steps of variation; (*a*) the generation of new genes of proteins to transfer the ATP molecules from the endosymbiont to the cytoplasm of the host cell and the ADP molecules to the endosymbiont, (*b*) the transfer of most protein genes except for those participating in O_2-respiration to the host genome, (*c*) deletion of transferred genes overlapping the host genes, and (*d*) putting the endosymbionts under the regulation and control of the host cell by accepting some of the protein genes for the transcription and replication of endosymbiont genome. The internal variables of host-endosymbiont in these steps of variation correspond to x_{i2}, x_{i3},.....,x_{in}, and the new style eukaryote having transformed the endosymbionts into the mitochondria appears with the probability of Eq. (11). The mitochondrion of the present-day eukaryote only carries 67 kinds of genes at most [31]. By this innovation, the new style eukaryotes *y* become the recipients of favours of the O_2-respiration performed by mitochondria, resolving the incompatibility of ATP synthesis with the synthesis of other biomolecules using protons, and compel the original style eukaryote *x* to be extinct.

When some of such new style eukaryotes further receives the cyanobacteria as endosymbionts, the increase rate of the host cell is first decreased, but is gradually recovered as the host genome acquires the protein genes for utilizing the synthesized sugars and most genes except for those participating in the photosynthesis are deleted in the endosymbiont genome. In particular, the electron transport system for O_2-respiration in cyanobacteria is completely deleted in the photosynthetic plastids. Thus, the photosynthetic plastid genome only encodes 209 proteins at most [31], although the gene transfer between mitochondria and photosynthetic plastids and that from the host to mitochondria are slightly detected [32]. This process of innovation is also expressed by the formulation in the preceding section with the transition probability of Eq. (11), if *x* and *y* are assigned to the internal variable of a eukaryote carrying the mitochondria and that of a eukaryote carrying both mitochondria and photosynthetic plastids, respectively. By this innovation, the eukaryotes *y* evolve as autotrophs, while the eukaryotes *x* continue to evolve as heterotrophs.

(iii) Formation of translation apparatus and RNA genes of proteins

Although the RNA replicase has been proposed as the start of life [33], various organic compounds including nucleotides and amino acids would have been synthesized non-biologically when the RNA replicases lived. Some satellite variants of RNA replicases would have catalyzed the polymerization of amino acids as primitive ribosomal RNA (rRNA) and transfer RNA (tRNA) [23,34]. Even if several amino acids are only polymerized at this stage, the

random sequences of 20 kinds of amino acid residues amount to enormous kinds of polypeptide chains. When five amino acids are randomly polymerized, for example, 20^5 kinds of polypeptide chains are yielded and they sufficiently cover the active centres of enzyme proteins in a free-living organism at the present time. By the catalytic actions of these polypeptides, therefore, it is plausible that primitive cells are formed and they self-reproduce by the spontaneous cell division due to the antagonistic balance of lipid synthetase and cell wall construction, i. e., matured cell wall elements are automatically transferred across the lipid bilayer to construct the cell wall covering the lipid vesicle [35-38]. The ancient origin of cell wall before the appearance of translation apparatus is suggested from the prokaryote where proteins catalyse the polymerization of L- and D-types of amino acids to form cell wall elements as a molecular fossil [39,40].

The formation of self-reproducing cells is essential in producing the special arrangement of nucleotide bases for the translation apparatus and RNA genes of proteins [25], when the whole world is pointing toward disorder under the second law of thermodynamics. This is because the systematic force of selection acts on such self-reproducing cells. To illustrate a possible route of this innovation, the first formed proto-cells will be specified by x_i's on their RNA contents. These cell contents would have changed upon their replication by the primitive RNA polymerase as well as RNA replicase activity, and most of them have been first elaborated to x_{opt} by Darwinian evolution, especially with respect to the primitive tRNAs and rRNAs to catalyze the polymerization of amino acids carried by tRNAs.

However, this evolution also increases the concentration of primitive RNA polymerases and thus increases the concentration of non-functional RNAs as well as the primitive tRNAs and rRNAs. This yields the variant cells x_{i1}, in which increased non-functional RNAs interfere with the primitive rRNAs and tRNAs. These variants cells decline to the minor members in the population and further experience at least following steps of variation until the appearance of translation apparatus. (α) Generation of anticodons. A part of tRNA originally embracing the side chain of amino acid residue bounded by primitive aminoacyl-tRNA synthetase alternatively attaches to the non-functional RNAs, which become later the RNA genes of proteins. (β) Contact of such ancestral RNA genes with the second type of RNAs, which become later 30S initiation complex. (γ) The interaction between the ancestral RNA genes carrying charged tRNAs and the primitive rRNA under the assistance of primitive initiation complex. (δ) The enlargement of the rRNA to form the A and P sites for the successive polymerization of amino acid residues carried by tRNAs aligned on ancestral RNA genes of proteins.

After the above series of variant cells x_{i1}, x_{i2},.....,x_{in}, a new style self-reproducing cells y, in which L-type of amino acids are polymerized dependently on the template of ancestral RNA gene, appear with the transition probability of Eq. (11). This mechanism of translation not only raises the rate of polymerizing amino acid residues but also directs the substituted bases in ancestral RNA genes so as to encode the proteins for raising the increase rate of cells by Darwinian evolution. This conversion of non-functional RNAs to the RNA genes of proteins would have taken place relatively faster, because the replication of RNAs by RNA polymerase is accompanied by more frequent nucleotide base substitutions than the replication of DNAs evolved afterwards. Thus, the new style cells y_{opt} compel the proto-cells x_o to be extinct.

(iv) Innovation from the RNA-protein world to DNA-RNA-protein world

The deoxidization of RNAs would have then occurred in some of such self-reproducing cells that began the glycolysis but still lacked the electron transport system to pumping out the protons. In fact, the amino acid sequences showing the similarities to membrane synthetic proteins are concentrically found in the enzyme proteins acting in and around the glycolytic pathway [41]. The DNAs thus generated would have been first rubbish and decreased the increase rate of such variant cells. If the optimum RNA content of the self-reproducing cell with translation apparatus is rewritten into x_{opt}, the internal variable of the variant cell suffering deoxidization corresponds to x_{i1}.

For utilizing the DNAs as genes, the variant cells x_{i1} must have experienced the succeeding steps of variation x_{i2}, x_{i3},....,x_{in} to induce the genes of proteins for the transcription and replication of DNAs as well as of auxiliary proteins for the unwinding of double-stranded DNAs from the RNA genes. Such induction is suggested from the fact that DNA-dependent DNA polymerases and RNA polymerases form a superfamily together with the RNA-dependent RNA polymerases in RNA viruses [42] and the DNA helicases show similarities to RNA helicases in super family II [43,44]. After the parallel usage of RNA genes and DNA genes, the decisive turning point for the variant cells x_{in} to enter the DNA-RNA-protein world would have been the direct or indirect attachment of replicating DNAs to the cytoplasmic membrane. This mechanism not only makes the cell division mechanical

but also makes it possible for the replicated DNAs to be equi-partitioned into daughter cells.

In the case of direct attachment to the membrane, the DNA genes are gradually fused to a single circular molecule, leading to the appearance of ancestral prokaryote. On the other hand, the DNA genes in the ancestral eukaryote get together to the plural number of chromosomes and two sets of them arising from replication are equi-partitioned into daughter cells by the attachment of each chromosome to the membrane through contractive microtubules. In this innovation, the probability $P_n(y <\!\!-\!\!- x_o)$ is divided into $P_n(y_p <\!\!-\!\!- x_o)$ and $P_n(y_e <\!\!-\!\!- x_o)$ where y_p and y_e denote the prokaryoytic style of DNA genome and the eukaryotic style of DNA genome, respectively. Accordingly, Eq. (12) of the fraction $f_{yk}(t)$ is also divided into the equation of the fraction $f_{ypk}(t)$ and that of the fraction $f_{yek}(t)$.

On the basis of the stable genetic information carrier of DNA and under the regulation and control of deoxidization of nucleotides, the prokaryotes further evolve the O_2-respiration and photosynthesis as well as various chemical syntheses as described in subsection (i), and the eukaryotes evolve the nuclear membrane, endocytosis and exocytosis, which become useful for the acceptance of mitochondria and photosynthetic plastids in subsections (ii). Meanwhile, the organic compounds synthesized non-biologically have been decreased, and the self-reproducing cells x_o in the RNA-protein world would have turned their strategy to survive as RNA phages and viruses, under the codon usage established in subsection (iii). In this connection, it should be noted that the biological activity is degenerate, that is, almost the same strength of activity is attained either by the larger amount of acquired energy, larger stored energy and higher systematization or by smaller amount of acquired energy, smaller stored energy and lower systematization [25].

Although the proteobacteria are proposed to be nearer to the eukaryotes from the protein clock [12,13,45], the more careful analyses on rRNAs indicate that the prokaryote and the eukaryote have first diverged and then the divergence of archaebacteria and eubacteria has occurred in the lineage of prokaryotes [14,46].

Conclusions and Discussion

Darwinian evolution is the fundamental process of evolution to raise the increase rate of an organism by selecting changed nucleotide bases, and also resolves the paradox of Maxwell's demon discussed in physics [25]. Moreover, the scaffold, under which Darwinian evolution takes place, has been sometimes changed during the long time of evolution. The organisms experiencing such a drastic change in the scaffold are first lowered in their increase rate and declined to the minor members in the population. As long as such variants retain the ability of self-reproduction, however, some of them recover the increase rate as a new style of organisms by Darwinian evolution under the new scaffold. Throughout such innovation, organisms have advanced their organization. Although this process of innovation is expressed by the transition probability in the present paper, the time course of this process can be also described as the solution of simultaneous equations such as Eq. (5) concerning the internal variables x_{i1}, x_{i2},.....,x_{in}. Such description may become useful especially for the case (iii) of the preceding section, in the future when the self-reproducing proto-cells are artificially reconstructed and/or the process of forming the translation apparatus is investigated by computer simulation.

The further innovation has occurred in the eukaryotes after the acquirement of the mitochondria. In particular, the transient conjugation of monoploid eukaryotes to exchange homologous chromosomes can accumulate multiple kinds of new genes generated from different origins of gene duplication [47]. This makes it possible for some of such eukaryotes to take the multicellularity and cell differentiation by the set of new genes for ligands, receptors, intracellular signal transduction and transcriptional regulators as well as the respective characters of differentiated cells, although other eukaryotes evolve intracellular organs in the unicellular form. Some of multicellular eukaryotes further innovate to take the diploid state, which elongates the duration time of enlarged genome of 10^8 *bp* or more against nucleotide base substitutions, through the intermediate stages of eukaryotes alternating the monoploid generation with the diploid one [48]. In the diploid eukaryotes, multiple kinds of new genes can be accumulated through the hybridization of variants having experienced different origins of gene duplication, and various combinational sets of new genes are yielded in the succeeding process to establish these genes homologously [49]. This explains the punctuated mode of explosive divergence of body plans in diploid eukaryotes, especially in animals. The mathematical formulation of such innovation from unicellular monoploid to multicellular diploid eukaryotes will be reported in the near future.

References

1. Darwin C (1859) The origin of species. John Murry, London.

2. Dobzhansky T (1941) Genetics and the origin of species. 2nd (edn.), Columbia University Press, New York.

3. Mayer E (1942) Systematics and the origin of species; a correlation of the evidence and points of view of systematics with those of other biological discipline, particularly genetics ecology. Columbia University Press, New York.

4. Huxley J (1943) Evolution: the modern synthesis. The grandson of TH Huxley explores mendelism, evolutionary trends, and genetic systems. Harper and Row, New York.

5. Simpson GG (1944) Tempo and mode in evolution: a synthesis of paleontology and genetics. Columbia University Press, New York.

6. Fisher RA (1930) The general theory of natural selection. Oxford University Press, London, New York.

7. Wright S (1949) Adaptation and selection. In: Jepson GL, Simpson GG and Mayer E (eds.), Genetics, paelontology and evolution. Princeton University Press, Princeton, New Jersey, pp: 365-389.

8. Kimura M (1968) Evolutionary rate at the molecular level. Nature 217: 624-628.

9. Kimura M (1980) A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J Mol Evol 16(2): 111-120.

10. King JL, Jukes TH (1969) Non-Darwinian evolution. Science 164(3881): 788-789.

11. Woese CR (1987) Bacterial evolution. Microbiol Rev 51(2): 221-271.

12. Iwabe N, Kuma K, Hasegawa M, Osawa S, Miyata T (1989) Evolutionary relationship of archaebacteria, eubacteria, and eukaryotes inferred from phylogenetic trees of duplicated genes. Proc Natl Acad Sci USA 86(23): 9355-9359.

13. Brown JR, Doolittle WF (1995) Root of the universal tree of life based on ancient aminoacyl-tRNA synthetase gene duplication. Proc. Natl Acad Sci USA 92: 2441-2445.

14. Sugaya N, Otsuka J (2002) The lineage-specific base-pair contents in the stem regions of ribosomal RNAs and their influence on the estimation of evolutionary distances. J Mol Evol 55(5): 584-594.

15. Margulis L (1981) Symbiosis in cell evolution: life and its environment on the early earth. WH Freeman, San Francisco.

16. Yang D, Oyaizu Y, Oyaizu H, Olsen GJ, Woese CR (1985) Mitochondrial origin. Proc Natl Acad Sci USA 82(13): 4443-4447.

17. Van den Eynde H, De Baere R, De Roeck E, van de Peer Y, Vandenberhe A, et al. (1988) The 5S ribosomal RNA sequences of a red algal rhodoplast and gymnosperm chloroplast: implication for the evolution of plants and cyanobacteria. J Mol Evol 27(2): 126-132.

18. Otsuka J, Terai G, Nakano T (1999a) Phylogeny of organisms investigated by the base-pair changes in the stem regions of small and large ribosomal subunit RNAs. J Mol Evol 48(2): 218-235.

19. Ingram VM (1963) The haemoglobin in genetics and evolution. Columbia Press, New York.

20. Ohno S (1970) Evolution by gene duplication. Spring-Verlag, Berlin.

21. Gilbert W (1978) Why genes in pieces? Nature 271: 501.

22. Ferris SD, Whitt GS (1979) Evolution of the differential regulation of duplicated gene after polyploidization. J Mol Evol 12(4): 267-317.

23. Watson JD, Hopkins NH, Roberts JW, Steitz JA, Weiner AM (1987) Molecular biology of the gene. 4th (edn.), volume II. The Benjamin/Cumming Publishing Company Inc, California. pp: 1098-1161.

24. Otsuka J (2008) A theoretical approach to the large-scale evolution of multicellularity and cell differentiation. J Theor Biol 255(1): 129-136.

25. Otsuka J (2017) The concept of biological activity and its application to biological phenomena. J Phys Chem & Biophys 7: 235-241.

26. Eigen E (1971) Selforganization of matter and evolution of biological macromolecules. Die Naturwissenshaften 58: 465-523.

27. Otsuka J (2002) An inquiry into the evolutionary history of photosynthetic and respiratory systems from the similarity relationships of member proteins. Recent Res. Devel. Proteins 1: 229-256.

28. Otsuka J, Kawai Y (2006) Phylogenetical relationships among permeases and the membrane proteins in photosynthetic and respiratory systems. Trends in Photochemistry & Photobiology 11: 1-21.

29. Otsuka J, Miyachi H, Horimoto K (1992) Structure model of core proteins in photosystem I inferred from the comparison with those in photosystem II and bacteria; an application of principal component analysis to detect the similar regions between distantly related families of proteins. Biochim Biophys Acta 1118(2): 194-210.

30. Wheeler DL, Church DM, Edgar R, Federhen S, Helmberg W, et al. (2004) National center of biotechnology information. Nucleic Acids Res 32(Database issue): D35-D40.

31. Keeling PJ, Palmer JD (2008) Horizontal gene transfer in eukaryotic evolution. Nat Rev Genet 9(8): 605-618.

32. Kleine T, Maier UG, Leister D (2009) DNA transfer from organelles to the nucleus: the idiosyneratic genetics of endosymbiosis. Annu Rev Plant Biol 60: 115-138.

33. Cech TR (1986) A model for the RNA-catalyzed replication of RNA. Proc Natl Acad Sci USA 83(12): 4360-4363.

34. Weiner AM, Maizels N (1987) tRNA structures tag the 3' ends of genomic RNA molecules for replication: implications for the origin of protein synthesis. Proc Natl Acad Sci USA 84(21): 7383-7387.

35. Botta GA, Park JT (1981) Evidence for involvement of penicillin-binding protein 3 in murein synthesis during septation but not during cell elongation. J Bacteriol 145(1): 333-340.

36. Begg KJ, Donachie WD (1985) Cell shape and division in Escherichia coli experiments with shape and division mutants. J Bacteriol 163(2): 615-622.

37. Cooper S (1988) Rate and topology of cell wall synthesis during the division cycle of Salmonella typhmurium. J Bact 170: 422-430.

38. Otsuka J (2001) Cell division potentiality arising from the antagonistic balance of lipid synthesis and cell wall construction. Research Communications in Biochemistry, Cell and Molecular Biology 5: 49-73.

39. Izaki K, Matsuhashi M, Strominger JL (1968) Biosynthesis of the peptidoglycan of bacterial cell wall. XIII. Peptidoglycan transpeptidase and D-alanine carboxypeptidase: penicillin-sensitive enzymatic reaction in strains of Esherichia coli. J Biol Chem 243(11): 3180-3192.

40. Van Heijenoort Y, Gemez M, Derien M, Ayala J, van Heijenoort J (1992) Membrane intermediates in the peptidoglycan metabolism of Esherichia coli: possible roles of PBP1b and PBP3. J Bacteriol 174(11): 3549-3557.

41. Fukuchi S, Otsuka J (1996) Evolution of the self-reproducing system to the biosynthesis of the membrane: an approach from the amino acid sequence similarity in proteins. J Theor Biol 182(2): 117-136.

42. Otsuka J, Kikuchi N, Kojima S (1999b) Similarity relations of DNA and RNA polymerases investigated by the principal component analysis of amino acid sequences. Biochim Biophys Acta 1434(2): 221-247.

43. Bork P, Koonin EV (1993) An expanding family of helicases within DEAD/H superfamily. Nucleic Acid Res 21(3): 751-752.

44. Pyle AM (2008) Translation and unwinding mechanisms of RNA and DNA helicases. Annu Rev Biophys 37: 317-336.

45. Doolittle WF, Feng DF, Tsang S, Cho G, Little E (1996) Determining divergence times of the major kingdoms of living organisms with a protein clock. Science 271(5248): 470-471.

46. Saccone C, Gissi C, Lanave C, Pesole G (1995) Molecular classification of living organisms. J Mol Evol 40(3): 273-279.

47. Otsuka J (2011a) A theoretical scheme of the large-scale evolution by generating new genes from gene duplication. In: Friedberg F (ed.), Gene duplication, InTech, Janeza Trdine, Croatia, pp: 3-26.

48. Otsuka J (2011b) The large-scale evolution by generating new genes from gene duplication; similarity and difference between monoploid and diploid organisms. J Theor Biol 278(1): 120-126.

49. Otsuka J (2014) The large-scale evolution by generating new genes from gene duplication and the divergence pattern of organisms. In: Urbano KV (ed.), Advances in genetics research 11, Nova Science Publisher, Inc., NY USA, pp: 85-116.