The Evolutionary Theory along the Phylogenetic Tree of Eukaryotes after the Acquirement of Mitochondria

ABSTRACT

According to the phylogeny of organisms reconstructed by the analyses on nucleotide base changes, the eukaryotes acquired the mitochondria first show the divergence of unicellular ones, then the divergence of fungi and sea algae occurs, and finally the animals and land plants appear from the unicellular eukaryotes other than fungi and sea algae. The main purpose of the present paper is to resolve the problems how the animal and land plant have acquired many kinds of genes necessary for multicellularity and cell differentiation and they have shown explosive divergence afterward. First, it is proposed theoretically that the conjugation of unicellular monoploids to exchange homologous chromosomes yields the monoploid variant receiving more new genes generated from gene duplication than the monoploid having solely experienced gene duplication in a stepwise manner. Then, the evolutionarty line from such a cell differentiated monoploid to multicellular diploids is illustrated from a series of present-day green plants. Secondarily, a theoretical method for explaining the explosive divergence of morphological characters is proposed on basis of the breeding style of diploid eukaryotes. This methodindicates thatdifferent combinational sets of new genes for cell differentiation arise from the successive hybridization of various silent variants with the latent variants yielded on the way to establish a new style diploid homologously when a new style diploid carrying a set of new genes heterogeneously appears under the high biological activity.Such hybridization process continues until the fraction of silent variants is decreased in the original population. In contrast, the fungi and sea algae have directly attempted the multicellularity and

cell differentiation in the diploid state but cannot accumulate the material and energy sufficient for further evolution of forming egg and seed, respectively. The landing of animals by metamophosis is also discussed from the aspect of genome size and biological activity.

Keywords:biological activity, cell differentiation, conjugation, explosive divergence, gene duplication, hybridization.

1. INTRODUCTION

The analyses of nucleotide base changes in ribosomal RNAs (rRNAs) reveal the phylogenetic relation of the present-day organisms in the DNA-RNA-protein world [1-5]. According to these results, the divergence of three kingdoms, archaebacteria, eubacteria and eukaryote, first occurred, but the divergence of present-day eukaryoteshas occurredmore recently after the ancestral eukaryote of them has acquired the mitochondria as the endosymbionts of O_2-respiratory *Proteobacteria* α. The genome sequencing finds that the archaebacteria mostly carry the genome of around $2x10^6$base pairs (bps) compactly encoding around $1.7x10^3$genes, the eubacteria carry slightly larger genome of $4\sim5x10^6$bps encoding$3\sim4x10^3$ genes, but the genome size is expanded to $1.7x10^7$bps encoding$6.3x10^3$genes in *Sacchaomyces cerevisia*, $1.2x10^8$ bps encoding$2.4x10^4$ genes in *Arabidopsis thaliana*, $1.4x10^8$ bps encoding$1.3x10^4$ genes in *Drosophila melanogaster* and $3x10^9$bps encoding $3x10^4$genes in *Homo sapiens* [6]. Among the genes increased in the above multicellular diploid eukaryotes, the outstanding feature is the expanded repertoire of genes of proteins associated with the intercellular and intracellular signal transmission represented by receptors and kinases. In accordance, many kinds of proteins including transcriptional regulators each carry serine-threonine repeats. The activity and specificity of such a protein are considered to be controlled by the phosphorylation [7] and/or glycosylation [8] of these special amino acid residues. With these characteristic proteins in mind, the cell differentiation is theoretically formulated in terms of the transition of cells from self-reproducing mode to differentiation mode, the long-range interaction between distinctive types of cells

and the short-range interaction between the same type of cells [9-11].These studies indicate that more than ten kinds of protein genes are needed even for the differentiation into two types of cells.

In the present paper, the evolutionary process fromthe unicellular monoploid eukaryote to the multicellular diploid eukaryote will be theoretically investigated along the divergence pattern of eukaryotes after the acquirement of mitochondria, succeedingly to the previous paper [12] of unicellular organisms. The molecular mechanism underlying the explosive divergence of multicullular diploid eukaryotes such as animals and land plants is also proposed theoretically.

2. THE DIVERGENCE OF EUKARYOTES AFTER THE ACQUIREMENT OF MITOCHONDRIA

The divergence pattern of eukaryotes after the acquirement of mitochondria is more clearly shown by the analysis on the base-pair changes in the stem regions of 5S rRNA [4] than that shown by the analysis on large and small ribosomal subunit RNA genes used in the previous review [12]. This result is shown in Fig. 1. One of the most remarkable points is that the eukaryotes having acquired the mitochondria have shown at least four stages of divergence. The first stage (*a*) is the divergence of unicellular monoploids, although some of *Ascomycota* 1 forms a thread-like connection of multicellular cytoplasm. The fungi and sea algae then diverged at the second stage (*b*) contain the lineages to evolve the alternation of unicellular monoploid generation with multicellular diploid generation. For example, many species of *Hymenocetes* in *Basidiomycota* repeat the life cycle of hypha (+ and – types), conjugation of hyphas to form fruit body (mushroom) and the production of hyphas in the basidium of the fruit body, although such evolution probably becomes possible after thegreen plants have landed. Many species of *Laminariales* in *Phaeophyta* repeat the life cycle of female and male types of zoospores, formation of female and male types of gametophytes to produce egg and sperm, respectively, fertilization of released egg to form a sporophyte in diploid state which produces both female and male types of zoospores. However, such evolved fungi and sea algae are not directly related with the higher animals (*Metazoa*) and higher plants (*Tracheophyta*) phylogenetically. As seen in Fig. 1, *Protista* 2 and 3 appear between thesecond branch (*b*) and the third branch (*c*) leading to the *Tracheophyta*.In *Protista* 3, the *Euglena* is assigned to the *Sarcomastigophora* by zoologists while it is assigned to the *Euglenophyta* by botanists, and the *Cyanophora* is assigned to the *Glaucophyta* by botanists and the *Teytrahymena* is to the *Ciliophora* by zoologists. This strongly suggests that the animals and green plants

are most closely related phylogenetically among the eukaryote and their ancestors were unicellular around $12x10^8$ years ago. However, the cell wall becomes rigid in the cell where photosynthesis takes place. This causes the division of animals and plants when they advance to multicellularity, just like the divergence of fungi and sea algae.In fact, the unicellular mobile *Dinophyta* and immature multicellular *Metazoa* (*Porifera*) first appear between the branch (*c*) and the branch (*d*) leading to *Metazoa* (animals).

The divergence pattern of different phyla inanimalsis resolved more clearly by the analsis on the base-pair changes in the stem regions of rRNAs in the mitochondria [14]. This result shown in Fig. 2 clearly reproduces the derivation of *Triploblastica* from *Diploblastica*, the divergence of *Protostomia* and *Deuterostomia*, and the divergence of many phyla in both *Protostomia* and *Deuterostomia*. Although the fossil records (ii) and (iii) indicated in Fig. 2 are biased to those of *Arthropoda* and shelly *Mollusca*, *Echinodermata* and *Chordata* had also begun the respective divergence at this time. Moreover, this result of reconstructed phylogenetic relationshows the characteristic feature that the divergence of morphological characters has been accelerated and explosive in both *Protostomia* and *Deuterostomia*. On this point, the theoretical consideration will be carried out in sections 5 and 6.

The divergence pattern of *Charophyta* and many phyla in land plants is shown in Fig. 3, which is obtained from the analysis on the nucleotide base changes at the third codon positions of *rbc* genes in chloroplast genome [19].As seen in this figure, the evolution from the unicellular monoploids to the multicellular diploids occurred more recently in the green plants than in the animals. Thus, the green plants are more suitable for inquiring into this evolution than the animals. In fact, the intermediate organisms on the way to this evolution can be found in green plants rather than in animals, as will be shown in the next two sections.

3. ACCUMULATION OF MANY KINDS OF NEW GENES GENERATED FROM GENE DUPLICATION THROUGH THE CONJUGATION OF MONOPLOIDS

The *Conjugatae* such as *Roya* and *Spirogyra*, which is treated as an independent plylum or incorporated into *Charophyta* by taxonomists, repeats the life cycle of unicellular monoploids, conjugation of two monoploids to form a zygote and return to two unicellular monoploids by meiosis. This eukaryote provides a hint for the problem how many kinds of genes necessary for multicellularity and cell differentiation have been accumulated.

As shown in the preceding review [12], the fraction $f(y_{dk})$of monoploid

variantsy_{dk}having experienced *k* kinds of gene duplication satisfies the following relation with the fraction $f(y_o)$ of dominant monoploids $y_o$in the population taking the common material and energy source *M*.

$$f(y_{dk}) = \prod_{m=1}^{k} \frac{q_{ydm,ydm-1} R(M,y_{dm-1})}{W(M;y_o)-W(M;y_{dm})} f(y_o) \qquad (1)$$

Here, $q_{ydm,ydm-1}$ is the probability that the *m*th step of gene duplication occurs. For investigating the effect of conjugation on the partition of duplicated genes, the self-reproducing rate $R(M;y_{dm-1})$ of monoploid $y_{dm-1}$is simply assumed to decrease by a reduction factor *r* in every step ofgene duplication, that is, $R(M, y_{dm-1})= R(M, y_o)\{1-(m-1)r\}$and the increase rate $W(M;y_{dm})$ defined by $R(M;y_m) –D(y_m)$ is also decreased by the factor *mr* in comparison with the increase rate $W(M; y_o)= R(M; y_o) – D(y_o)$ ofthe dominant monoploid,under the assumption that the death rate $D(y_m)$ is hardly influenced by gene duplication. Then, Eq.(1) is rewritten into the following way.

$$f(y_{dk}) = \frac{(1-r)(1-2r).....\{1-(k-1)r\}}{k!r^k} Q_k f(y_o) \qquad (2)$$

Here, $Q_k$is the probability of having experienced *k* kinds of gene duplication, i. e.,

$$Q_k \equiv \prod_{m=1}^{k} q_{ydm,ydm-1} \quad (3)$$

As seen from Eq. (2), the fraction $f(y_{dk})$ of variants y_{dk} having *k* kinds of gene duplication becomes smaller than the fraction $f(y_o)$ of dominant monoploids, as the number *k* becomes larger. However, the exchange of homologous chromosomes through conjugationyields the monoploid variant receiving the more kinds of duplicated genes with the higher probability than that expected from Eq. (2). This probability depends on the number of chromosomes carryingthese duplicated genes, and the following two extreme cases will be considered.

In the case when the monoploid variant carrying k_1 kinds of duplicated genes separately on k_1 kinds of chromosomes conjugates with another monoploid variant carrying $k_2$kinds of duplicated genes separately on otherk_2 kinds of chromosomes, the zygote produces the daughter monoploid receiving (k_1+k_2) kinds of duplicated genes with the following probability $P_{(k1+k2);(k1+k2)}$ by the random partition of homologous chromosomes in each pair.

$$P_{(k1+k2);(k1+k2)} = \left(\frac{1}{2}\right)^{k1+k2} \frac{(1-r)(1-2r).....\{(1-(k_1-)r\}}{k_1!r^{k1}} x \frac{(1-r)(1-2r).....\{1-(k_2-1)r\}}{k_2!r^{k2}} Q_{k1+k2} f^2(y_o) \qquad (4)$$

In another extreme case when the variant monoploid carrying k_1 kinds of duplicated genes in one kind of chromosome conjugates with other variant monoploid carrying k_2 kinds of duplicated genes in another kind of chromosome, the zygote produces

the monoploid receiving (k_1+k_2) kinds of duplicated genes with the probability $P_{2;(k1+k2)}$. In this probability $P_{2;(k1+k2)}$, the coefficient $(1/2)^{k1+k2}$ in Eq. (4) is replaced by $(1/2)^2$.

For the numerical comparison of $P_{(k1+k2);(k1+k2)}$with the fraction $f(y_{dk})$in Eq. (2), $f(y_o)$ is simply assumed to be nearly equal to one and the probabilities$P_{k;k}$, whose suffix k satisfies the relation $k/2 = k_1=k_2$ or $(k+1)/2 = k_1=k_2+1$, are selected.Then, the value of $P_{k;k}/Q_k$ becomes larger than the value of $f(y_{dk})/Q_k$,especially in the region of large k values,as shown already [29]. Moreover, it is clear that $P_{2;(k1+k2)}>P_{(k1+k2);(k1+k2)}$.

At the stage when the unicellular monoploid eukaryote began the partition of homologous chromosomes through conjugation, such a monoploid eukaryote would have carried only several kinds of chromosomes, and the probability of accumulating many kinds of new genes generated from the counterparts of duplicated genes would have been intermediate between the above two extreme cases. At any rate, the present result indicates that the partition of homologous chromosomes into daughter cells through conjugation is the competent strategy for accumulating many kinds of new genes generated from gene duplication avoiding much lowering of biological activity. Moreover, the specificities of receptor, transcriptional regulator and kinase, which have appeared in the ancestral eukaryote for its living style [12], are easily changed by slight changes of amino acid residues in the conterparts of duplicated genes. Thus, such counterparts would have been gotten together as a set of new genes for cell differentiation through conjugation between the branch points (b) and (c) denoted in Fig. 1.

4. EVOLUTION OF CELL DIFFERENTIATION FROM MONOPLOIDS TO DIPLOIDS

This evolutionary process can be also traced from the present-day green plants as listed below. Succeedingly to (α) the *Conjugatae* mentioned in the preceding section, (β) the *Chara* of *Charophta* develops the multicellularity and cell differentiation in the monoploid state. This adult form produces both eggs and sperms, whose fertilization yields the oospore. By the meiosis, the oospore produces the spores, each of which grows into the adult form. (γ) In the *Bryophyta*, the monoploid generation differentiates into female and types of gametophytes, and the fertilized egg on the female gametophyte grows to the sporangium which produces both female and male types of spores. (δ) The *Pterophyta* develops the cell differentiation in the diploid generation called the sporophyte to acquire the material and energy by itself and the monoploid generation called the prothallium is specialized to produce only eggs and sperms. In some species of *Pterophyta*, the

prothallium also differentiates into female and male types. (ε) The seed plants (*Coniferophyta* and *Anthophyta*), in which the monoploid generation (egg and sperm) is produced in the reproductive organ incorporated into the diploid body, take the material and energy from the outside by the cooperative action of differentiated cells in the diploid state and develop the albumen inthe seed for its germination.

The above series of example clearly shows that the multicellularity and cell differentiation first occur in the monoploid state by the acculmulation of sufficient kinds of new genes through conjugation. This means that the energy acquired by the cooperative action of differentiated cells in monoploid state becomes larger than the energy consumed for the development ofits cell differentiation. Then, such monoploid generation endows the egg with the material and energy necessary for the cell differentiation in the diploid state. After the alternation of such monoploid generation with the diploid generation, the cell differentiation in the diploid state gradually becomes predominant by the longer duration time against injurious nucleotide base changes, and the monoploid generation appears only as egg or sperm. This resolves the problem indicated in the previous paper [30]that the material and energy are needed for the development of cell differentiation until the differentiated cells begin to acquire the material and energy from the outside by their cooperative action. That is, the development of cell differentiation in the monoploid generation requires less material and energy than that in the diploid state and the material and energy acquired by the cooperative action of differentiated cells in the monoploid state have helped the differentiation of cells in the diploid state. These examples also indicate that the meiosis has first evolved and then the mitosis has evolved in the multicellular diploid eukaryotes.

Although the cell differentiation only in monoploid state is hardly found in the animals at the present time probably due to the severer struggle for existence, the *Cnidaria* alternates the monoploid generation differentiated into female and male types with the asexual diploid generation. Probably, the ancestral animal also began to form multicellularity and cell differentiation in the monoploid state to prey the increasing unicellular eukaryotes.

This is a contrast to the fungi and sea algae which have directly attempted the multicellularity and cell differentiation without the cell differentiation in the monoploid state.

5. THETHEORY OF THE EXPLOSIVEDIVERGENCE IN ANIMALS AND LAND PLANTS BY GENE DUPLICATION

The animals and land plants that have established the cell differentiation in

the diploid generation exhibit the explosive divergence of morphological and physiological characters due to the difference in cell differentiation. The resolution of this problem will be theoretically formulated in the following way.

As shown already [31], it is the same form as the case of monoploid organisms that the fraction $f(y_k,y_o)$ of new style diploid eukaryotescarrying k kinds of new genes heterogeneously first appears in the relation with the fraction $f(y_o,y_o)$ of the original style of dominant diploids (y_o,y_o). However, a remarkable difference in the evolutionary pattern between the diploid and monoploid eukaryotes comes from the process to fix the new genes arising from the counterparts of duplicated genes. Although a set of new genes suitable for expressing a new characteris immediately fixed in the monoploid eukaryote to form a new population apart from the original population, the breeding of new style of diploid eukaryotes each carrying a set of new genes heterogeneously still produces various latent variants. In such a diploid eukaryote, the number of homologous chromosomes may be increased to more than ten, and it is reasonable to consider the case when most of the new genes generated from duplicated genes are distributed separately on different kinds of chromosomes, as will be discussed in the last part of this section.

In this case, we first consider the ratios of variant children produced by the breeding of new stylediploid parents each carrying k kinds of new genes heterogeneously. When the homo, hetero, and vacant states concerning a new gene on the J th pair of homologous chromosomes are denoted by

$$a_J^2 = \binom{1}{1}_J, \quad a_J b_J = \binom{1}{0}_J = \binom{0}{1}_J, \quad b_J^2 = \binom{0}{0}_J \tag{5}$$

respectively, the ratios of different types of children born from suchparentsare expressed by those of the terms on the right hand side of the following expansion.

$$\prod_{J=1}^{k}\left(a_J^2 + 2a_J b_J + b_J^2\right) =$$

$$\prod_{J=1}^{k}\left(a_J^2 + 2a_J b_J\right) + \sum_{J=1}^{k} b_J^2 \prod_{G(\neq J)=1}^{k}\left(a_G^2 + 2a_G b_G\right) + \sum_{J\neq H=1}^{k} b_J^2 b_H^2 \prod_{I(\neq J,H)=1}^{k}\left(a_I^2 + 2a_I b_I\right) + \cdots + \prod_{J=1}^{k} b_J^2 \tag{6}$$

The first term on the right hand side of Eq. (6) corresponds to the children receiving a set of all new genes, and the number of such children amounts to 3^k when the total number of children on the left hand side is set to be 4^k. The second term corresponds to the children lacking one kind of new gene, and the number of such children is ${}_kC_1 3^{k-1}$. In the same way, the number of children lacking v kinds of new genes amounts to ${}_kC_v 3^{k-v}$and the last term corresponds to the children completely lacking new genes. These children lacking some of new genes return to the population of the original style diploid eukaryotes as the latent variants not

expressing a new cell differentiation, but serve to produce the second stage of divergenceby the hybridization with the silent variants carrying other kinds of new genes.

If any of new types of cell differentiation is assumed to be also expressed by k kinds of new genes for simplicity, the second stage of divergence occurs when the 3^{k-v}latent variantslacking v kinds of new genes produced at the first stage hybrize with the silentvariant carrying other v kinds of new genes. The fraction $f(y_v,y_0)$ of such silent variants must be larger than the fraction $f(y_k,y_0)$ of the first style of new diploids (y_k,y_0) which have carried out the first stage breeding. Thus, the hybridization of 3^{k-v} latent varients with the silent variants v produces secondarily new style children carrying new sets of genes as the combination of $(k-v)$ and v with the higher probability than the appearance of first new style diploids (y_k,y_0). This probability becomes higher as the number v is smaller, although the number of choice ${}_kC_v$is maximum for $v = k/2$ or $v = (k-1)/2$.

Moreover, the ratio of the children carrying a full set of new genes homologously is only$(1/4)^k$ and the ratio of children carrying a full set of new genes heterogeneously amounts to $(3/4)^k(2/3)^k = (1/2)^k$in the breeding between the parents each carrying new genes heterogeously, as seen in the first term on the right hand side of Eq. (6). This characteristic feature also holds in the new style of diploids carrying a full set of new genes heterogeneously which appear at later stages by the hybridization of latent variants with silent variants. Thesenew styles of eukaryotes carrying a full set new genes heterogeneously further produce the latent variants, which can hybridize with silent variants, on the way to establish the homologous set of new genes. Such successive divergence due to the partial replacement of new genes continues until the fraction $f(y_v,y_0)$ of silent variants (y_v,y_0) is decreased in the original population.

Although all these diploid eukaryotes receiving different sets of new genes do not necessarily survive, the survived ones are recorded as the divergence of different styles of diploid eukaryotes. When some of the survived new style diploid eukaryotes raises its biological activity to allow the existence of variants carrying furthernew genes generated from gene duplication, however, the divergence again occurs from this new style eukaryote. Thus, the explosive divergence mentioned above occurs with a punctuated mode.

Even in the case when the new genes generated from gene duplication are located on less number of chromosomes, the above formulation formally holds with less numbers of k and v. However, the increased difference between homologous

chromosomes makes it difficult to produce the children by the incompatibility of homologous chromosomes. On the other hand, the new genes generated from gene duplication tend to be scattered on different chromosomes, for example, by transposons. Thus, the explosive divergece of morphological and physiological characters tend to occur in the multicellular diploid eukaryote carrying alarge number of chromosomes.

The punctuated mode of evolution isfirst pointed out from thepaleontology of animals [32]. The paleontology and the phylogenetic relation of organisms reconstructed from the neutral nucleotide base changes are complementary to each other in investigating the divergence pattern of multicellular diploid eukaryotes. The former detects the fossils including the extinct organisms fragmentally while the latter traces back to the origin of present-day organisms continneously.

6. EXPLANATION OF DIVERGENCE PATTERNSIN ANIMALS AND LAND PLANTS

As seen in Figs. 2 and 3, the divergence pattern of different phyla in animals and land plants shows the tendency thata new hylumhas appeared at the relatively early stage of the preceding phylum. This supports the theory described in section 5, indicating that the new phylum has appeared through the hybridization of silent variants with latent variantsyielded during the process for the preceding phylum to become homogeneous,although the set of genes for the first cell differentiation into two types of cells seems to have been definite.

In the animals,the cell differentiation into ectoderm and endoderm first occurred in *Cnidaria* by the first set of new genes. When other new genes have been accumulated in the variants of ancestral*Cnidaria*, another set of new genes appeared heterogeneusly to yield ancestral *Platyhelminthes*, causing the differentiation into mesoderm and endoderm. During the process that the *Platyhelminthes* hasbecome homogeneous of this second set of new genes, other sets of new genes must have appeared, causing the divergence of ancestral *Brachiopoda* and the common ancestor of other *Protostomia* and *Deuterostomia*by the partial replacement of member genes in the second set of genes. In the appearance of *Deuterostomia*, the third and fourth sets of new genes must have been added to form the mouth instead of blastopore by the further cell differentiation in both ectoderm and endoderm. Although somewhat difference in the derivation of mesoderm between *Deuterstomia* and *Protostomia* has been only indicated from developmental biology [33], at least four classes of genes are recently suggested to be involved in the formaton of mesendoderm [34]. This difference between the *Protostomia* and the *Deuterostomia*has brought the decisive difference

in acceptance of new sets of genes for the further cell differentiationas well as inthe formation of mouth.In the *Deuterostomia,* the divergence of *Echinodermata* and *Chordata* has further occurred. The *Echinodermata* has accepted the set of genes to secrete the calcium carbide from the mesoderm to cover the body but the *Chordata* has accepted the set of genes to derive the chorda dorsalis from the mesoderm. The chordadorsalis thus generatedsupports the body from the inside and brings about the high biological activity enough to accept the other sets of new genes for the further cell differentiation. Thus, the successive divergence of *Hemichordata*, *Cephalochordata*, *Petromyzontiformes*, *Condrichthyes* and*Osteichthyes* has occurred,through the partial replacement of member genes in each set. Meanwhile, the set of genes for metamerism has been generated in the common ancestor of *Annelida* and *Arthropoda*and this style of body plan also raises the biological activity, leading to the successive divergence of *Chlicerata*, *Chilopoda*, *Branchiopoda,Malacostraca* and *Insecta* in *Arthropoda,* although the set of genes for metamerism has not been inherited to*Mollusca*. In this divergence of subphyla in *Arthropoda*, the partial replacement of member genes in each set may have also acceraleted the explosive divergence. The body of *Arthropoda* is supported by the layer of cutile that is formed by the secretion from epitheluim cells but most of *Mollusca* except for *Cephalopoda* are each covered by a shell mainly consisting of calcuim carbide secreted from mantle membrane. Such shells are also used in *Brachiopoda*, probably to protect themselves from foreign enemy. Although the ancestor of *Cephalopoda* was also covered by the shell of calcium carbide, it has taken the evolutionary route afterwards to move more freely by abondoning the shell.The carriers of O_2 molecule in animal's bodies are not systematically consistent with the above divergence pattern of cell differentiation. For example, they are hemoglobin in *Vertebrata* and chlorocuorin or eryththrocuorin in *Annelida*, all of which have been derived probably from the gene of hemoprotein participated in the electron transfer system of the mitochondria. On the other hand. the carrier of O_2 moleculeis hemocyanin in *Arthropoda* and most *Mollusca*, although its origin is unclear. This fact also indicates the diversity of body plans in animals.Among such diverged animals, *Insecta* and some of *Osteichthyes* have especially evolved for landing, as will be described in the next section.

The cutile layer is also used in land plants to protecttheir bodiesfrom dry. The diversity of such land plants has been increased after the third type of cells appeared to form vascular bundle. As seen in Fig. 3, this has appeared on the way for *Bryopsida* to diverge from *Marchantiidae*.Although the variant accepting theset

of genes for forming vascular bundlemay have first lowered its biological activity, the cells forming the vascular bundle can elongate both stem and root. The elongated root makes it possible to absorb much more water in the ground and extended branches make it possible for many leaves to accept light energy. Thus, the variantgradually raises its biological activity as the*Pterophyta* to allow the further generation of various new genes. Some of *Felicopsida* become a tree by the lignification of the stem with a high ability of regeneration and other grass-like ones can pursue the vegetative reproduction from the subterranean stem as well as the reproduction by forming spores. The such high ability of regeneration and vegetative reproduction arises from the faster proliferation rate of undifferentiated cells than the transition rate of these cells to the differentiation mode [11], while the cell differentiation in higher animals advances forming the so-called stem cells and the ability of regenerationbecomes lower [10]. These living styles of tree, grass and vegetative reproduction are also inherited to the seed plants that have appeared in the next stage of evolution. The set of genes for forming flowers and seeds has appeared from the ancestral *Pterophyta* before the divergence of *Fillicopsida* and *Ophioglosscopsida*. The seed plants have evolved the pistil and stamen, which are separated as unisexual flowers in *Coniferophyta* but are bunddled as a bisexual flower in *Anthophyta*, and supply the fertilized egg with the material and energy sufficient for its growth to the seed. Although most of land plants are monoecism and the divergence due to the hybridization is not outstanding in comparison with the animals, the evolution of flowers enhances the hybridization through the pollens cattered by wind and/or insects. The comparison of the set of genes responsible for performing the process from the reproductive organ of pistil and stamen to form the seed among *Gnetopsida*, *Podocarpaceae*, *Pinacea*, *Dicotyledonopsida* and *Monocotyledonopsida*may support the partial replacement of member genes by the hybridization of silent and latent variants mentioned in section 5.

7. FURTHER EVOLUTION OF ANIMALS; LANDING BY METAMORPHOSIS

As is well known from paleontology [35, 36], many lineages of animals began to land following the land plants, probably after the adaptation to fresh water like the green plants. In particular, the *Insecta* and some of *Osteichthyes* have succeeded in landing by metamorphosis,leaving the intermediates such as dragonfly and frog, respectively,although the metamorphosis is also seen in *Echinodermata*, *Arthropoda* and some of *Annelida* and *Mollusca*. The metamorphosis is the phenomenon within one generation that the cells differentiated with the expression of first sets of genes are replaced by the cells differentiated with the expression of second sets of

genes through the programmed death or apoptosis of the former cells [37].While the metamophosis is triggered by prothoracicotropic hormone and prothoracic gland hormonein most *Insecta*[38], the metamorphosis from tadpole to frog is triggered by thyroid hormone [39]. The thyroid hormone also plays an important role in maintaining various organs in the completely landed*Vertebrata* [40] where the trace of metamorphosis can be seen during the fetus.

The*Dipneustei* may be the survival of collateralfamily at the initial stage of evolution from the *Osteichthyes* on the way to *Amphibia*. Curiously, such collateral familyand intermediates upon the evolution for landing carry much expanded genomes. For example, the lungfish genome is $4.3x10^{10}$ base pair long, which is about 14 times larger than the human genome and *30* percent larger than the genome of axolotl [41], while the genome sizes of major fishes range from $4.4x10^{8}$ *to* $2.9x10^{9}$ bps, showing the maximum by salmon [42]. The genome sizes of *14* anuran species fall in the range of$1.1\sim6.8x10^{9}$ bps,and the abundance and density of simple repeat sequences positively correlate with the genome size [43]. Presumably, the biological activities of such collateral families and intermediates have been too much lowered by expanded genomes to advance the further evolution.

On the other hand, the *Amphibia*, which was the ancestor common to *Reptilia* and *Mammalia*, would have evolved fore- and hind-legs as well as lung by steadily generating new genes from gene duplication, in parallel to the abandonment of the genes for swimming in water, within the genome size of around 10^{9} bps. The biological activity of such ancestral *Amphibia*would have been further increased by the respiration of 0_2 molecules in the air, whose concentration has become higher than that in the water after the green plants have landed, and has allowed the gene duplication for the further evolution. Thus, the divergence of *Reptilia* and *Mammalia*occurred soon after the appearance of *Amphibia.* The *Reptilis* simply evolved the egg shell to protect the egg from dry and became first prosperous on the land, yielding the *dinosaurs*. Some of the *dinosaurs* has further evolved to the *Aves* which now constitute the groupof minimum genome size among the *Amniota*. During the prosperity of dinosaurs, the *Mammalia*were not outstanding in their size and number but they have evolved mammary gland and placenta probably by the mechanism described in section 5. In practice, the *Prototheria* and*Metatheria* may be the survival of collateral families on the way of evolution to the *Eutheria*. Such *Mammalia* were tolerant of the severer climate on the Earth due to the collision of a small planet which compelled the *dinosaurs* to extinction, and the *Eutheria* have

especially shown the explosive divergence, together with the *Aves*,after the climate has been recovered. They have become homeotherms by evolving the mechanism that a part of acquired energy is used to keep the temperature of a body constant regardless of the temperature change in the outside.

Among the genome sizes *1.6 ~ 9.8x10*^{8}bpsof *Insecta*[44, 45], the relatively large genomes are seen in the speciesthat show complete metamophosis, e. g., *6.6x10*^{8} bps in dragonfly and *9.8x10*^{8} bps in *Coleoptera*. The smaller genome size of *Insecta* than the genome size of *Vertebrata* may reflect the O_2 carrier ability of hemocyanine which is inferior to that of hemoglogin.

8. CONCLUSIONS AND DISCUSSION

The evolution of multicellular diploid eukaryotes is the active challenge to a new environment as well as the passive adaptation to a given environment.This has occurred by generating the sets of new genes from different origins of gene duplication under the raised biological activity due to the cooperative action of differented cells.As the cell differentiaton is advanced to the diploid generation, various combinationof new genes arises from the breeding style of diploid eukaryote that the latent variants carrying partial sets of new genes are generated on the way to establish a new style of multicellular diploids homologously and these latent variants successively hybridize with other silent variants to yield different sets of cell differentiations. Such divergence originates from the genome structure of eukaryote that consists of plural number of linear chromosomes and causes the exon shuffling due to the base change in introns, although the rate of base changes by the miss in proofreading is almost the same as that in the prokaryote. Moreover, the number of new genes from gene duplication increases as the the number of original genes is increased. However, the DNA genome size seems to approach a limit in land *Vertebrata* when we consider the base change rate of *10*$^{-9}$ per site per year and the prolonged life spans of some land *Vertebrata*. Thus,it also becomes an important strategy for retaining the high biological activity to abandon the genes and spacers which have become unuseful.

Under the above conclusion, several discussions will be given below.

Although the present study of multicellular diploid eukaryotes mainly focuses on the divergence of phyla and classes due to the difference in cell differentiations, the formulation described in section 5 is also applicable to the divergence of orders, families, genera or even species. Then, the category of new characters is lowered to the cell differentiation in the peripheral parts of a body, the ratio of differentiated cells, and the slight difference in the properties expressed by each type of

differentiated cells.Even in the divergence of species, multiple kinds of genes in variants are slightly different from those in the original species individually, and the combinational set of such genes produced by the hybridization between the variants can yield new species different from the original species with respect to habitat, breeding season, body smell etc.The fairly well investigatedexample is the divergence of human and chimpanzee, which is estimated to have occurred $7x10^6$ years ago [46]. It is now becoming clarified that many species of *Hominidae* first appeared upon this divergence [47~50]. After the successive partial replacement of new genes by the hybridization between the silent and latent variants probably due to the formulation described in section 5, *Homo sapience* has finally accepted the best combinational set of member genes forerect bipedalism, development of ligamentum vocale and brain, etc. for prosperity.

The Darwinian evolution, which is originally proposed from the observation of land vertebratesand plants [51], corresponds to the process for almost a definite set of genes to be elaborated by nucleotide base changes and the selection of organisms carrying changed bases. It depends on the living region what bases are selected, and the organisms living in an isolated region are selected to the region-specificspecies. On the contrary, the appearance of new body plan by the drastic change in cell differentiation tends to occur in a larger population of multicellular diploid eukaryotes with the higher biological activity. This is now becoming evidenced from the fossil record showing that the appearance of *Aves* from dinosaurs occurred in the wide area ranging from Europe to China [52-54].

The phylogenetic tree of many orders, families and species in the present-day *Aves* can be reconstructed from the comparison of base changes at the third codon positions between orthologous protein genes encoded on nuclear genome and mitochondrial genome. The base change rate in nuclear genome is almost the same as that of *rbc* genes in plant, and the analysis of these base changes is expected to reconstruct the phylogenetic tree with almost the same resolution as that for the land plants shown in Fig. 3, covering the range from the Jurassic period to $6x10^7$ years ago. On the other hand, the base change rate in mitochondrial genome of animals is more faster and a careful analysis on these base changes make it possible to reconstruct the phylogenetic tree during the period from $6x10^7$ to $1x10^7$ years ago [46]. This is also the case for the reconstruction of phylogenetic tree of many orders, families and species in *Insecta*. This is especially important in investigating the evolution from the phylogenetical divergence pattern, bcause the fossil records of *Insecta* are scare in comparison with

those of land *Vertebrata*.

The severe glacial age during *7.2 ~ 6.4x10*^{8} years ago is suggested to be rsponsible for the *Cambrian* explosion of animalsfrom the paleontology and geology [55, 56]. Certainly, the biological activity of an organism becomes higher under the lower temprature. Thus, this possibility cannot be excluded completely but it is also true that the concentration of O_2 molecules has entered into the stage 4 from *8.5x10*^{8} years ago to the present [57] probably by the prosperity of unicellular eukaryotes carrying chloroplasts and sea algae. Therefore, it is more plausible that the multicellular animals have appeared to prey them by activating their mitochondria under the high O_2 concentration. In fact, the divergence of *Protostomia* and *Deuterostomia* occurred before the glacial period, as seen in Fig. 2.

The author declare that no generative AI technologies have been used during the writing or eliciting the manuscript.

REFERENCES

1. Yang D, Oyaizu H, Olsen GJ, Woese CR. Mitochondrial origin. Proc. Natl. Acad. Sci. USA.1985;82: 443-447.
2. Woese CR. Bacterial evolution. Microbiol. Rev.1987; 51: 221-271.
3. Van den Eynde H, De Baere R, De Roeck E, Van de Peer Y, Van den Berghe A Willekens P, De Wachter R. The 5S ribosomal RNA sequences of a red algal rhodoplast and a gymnosperm chloroplast: Implication for the evolution of plastids and cyanobacteria. J. Mol. Evol.1988;27: 126-132.
4. Otsuka J, Nakano T, Terai G.A theoretical study on the nucleotide base changes under a definite functional constraint of forming stable base-pairs in the stem regions of ribosomal RNAs: Its application to the phylogeny of eukaryotes. J. Theor. Biol. 1997;184: 171-186.
5. Otsuka J, Terai G, Nakano T.Phylogeny of organisms investigated by the base-pair changes in the sytem regions of small and large ribosomal subunit RNAs. J. Mol. Evol. 1999; 48: 218-235.
6. Wheeler DL, Church DM, Edgar R, Federhen S, Helmberg W, Madden TL, Pontinus JU, Schular GD, Schriml LM, Sequeira E, Suzek TO, Tatusova TA, Wagner L. Nucl. Acid Res. 2004; 32: Darabase issue D35.
7. Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson JD.Mollecular Biology of the Cell. 1994;3rd edn. Garland Publishing Inc., New York.
8. Hart GW. Dynamic O-linked glycosylation of nuclear and cytoskeletal proteins. Annual Review of Biochemistry 1997;66: 315-335.

9. Otsuka J. A mathematical model of the cell differentiation in multicellular eukaryotes. Applied Mathematics 2020; 11(3): 157-171.
10. Otsuka J.A theoretical study on the cell differentiation forming stem cells in higher animals. Phys. Sci.& Biophys.2021; 5(2): 1-10.
11. Otsuka J. A theoretical study on the cell differentiation retaining meristems in higher plants. Phys. Sci.& Biophys.2022; 6(2): 1- 9.
12. Otsuka J. The evolutionary theory along the phylogenetic tree of unicellular organisms. Ann. Res. Rev. Biol. 2023; 38(9): 39-49.
13. Dickerson RE.The structure of cytochrome c and the rate of molecular evolution. J. Mol. Evol. 1971; 1:26-45.
14. Otsuka J, Sugaya N.Advanced formulation of base-pair changes in the stemregions of ribosomal RNAs; its application to mitochondrial rRNAs for resolving the phylogeny of animals. J. Theor. Biol.2003;222: 447-460.
15. Rasmussen B, Bengstom S, Fletcher IR, McNaughton NJ. Discoidal impressions and trace-like fossils more than 1200 milion years old. Science 2002; 296: 1112-1115.
16. Mathews SC, Missarzhevsky V. Small shelly fossils of late precambrian and early cambrian age: a review of recent work. Q. J. Geol. Soc. London 1975;131:289-304.
17. Rozanov AY, Zhuravlev AY. The Lower Cambrian fossil record of the Soviet Union. In: Lipps JH and Signor PW (eds.), Origin and Early Evolution of the Metazoa. Plenum Press, New York, London.1992;pp. 205-282.
18. Gould SJ. Wonderful Life. The Burgess Shale and the Nature of History. W. W. Norton & Company Inc., New York;1989
19.Kawai Y, Otsuka J. The deep phylogeny of land plants inferred from a full analysis of nucleotide base changes in terms of mutation and selection. J. Mol. Evol. 2004; 58: 479-489.
20.Garratt MJ, Rickard RB. Pridoli (Silurian) graptolites in association with *Baragwanathia* (Lycophytina). Bull. Geol. Soc. Denmark 1987; 35: 135-139.
21. Hueber FM. Thought on the early lycopsids and zosterophyls. Ann. Mo. Bot. Gard. 1992; 79: 474-499.
22. Galtier J, Scott AC. Diversification of early ferns. Proc. R. Soc. Edinb. [Biol.] 1985; B86: 289-301.
23. Fairon-Demaret M, Scheckler SE. Typification and rediscription of *Morasnetia zalesskyi* Stockmas, 1948; an early seed plant from the Upper Famennian of Belgium. Bull. Inst. Roy. Sci. Nat. Belg. Sci. Terre.1987; 57: 193-199.

24. Rowe NP. Winged late Devonian seeds. Nature 1992; 359: 682.
25. Rothwell GW, Scheckler SE, Gillespie WH.*Elkinsia gen*. nov., a late Devonian gymnosperm with cupulate ovules. Bot. Gazette 1989; 150: 170-189.
26. Kräusal R, Weyland H. Pflanzenreste aus dem Devon. X. Zwei Pflanzendunde im Oberdevon der Eifel. Senckenbergiana 1937; 19: 338-355.
27. Kenrick P, Crane RP. The origin and early diversification of land plants. Sumithsonian Institution Press, Washington DC; 1997:pp249-256.
28. Stewart NS, Rotwell GW. Paleobotany and the evolution of plants. Cambridge University Press, Cambridge; 1993; pp. 438-467.
29. Otsuka J. A mathematical formulation of evolution and innovation II. From unicellular monoploid eukaryotes to multicellulr diploid eukaryotes. Phy. Sci. Biophy. J. 2018; 2(2): 1-11.
30. Otsuka J. The concept of biological activity and its application to biological phenomena. J. Phys. Chem.& Biophys.2017; 7: 235-240.
31. Otsuka J. The large-scale evolution by generating new genes from gene duplication; similarity and difference between monoploid and diploid organisms. J. Theor. Biol. 2011; 278: 120-127.
32. Eldredge N, Goulds SJ. Punctuated equilibria; an alternative to phyletic gradualism. In: Schof TJM (edn.) Models in Paleobiology. Freeman and Cooper, San Francisco; 1972; pp. 82.
33. Salvini-Plawen L, Splechtna H.On the origin and evolution of lower metazoa. Z. f. Zool. Systematik Evolutionsforshurg 1979; 16:40-88.
34. Technau U, Scholz CB. Origin and evolution of endoderm and mesoderm. Inf. Dev. Biol. 2003; 47: 531-539.
35. Carroll RL. Pattern and Processes of Vertebrate Evolution. Cambridge University Press, New York; 1997.
36. Anderson JS, Sue, HD eds. Major Transitions in Vertebrate Evolution. Indiana University Press,Bloomington, Ind.; 2007.
37. Lockshin RA, Zakeri Z. Programmed cell death: early changes in metamorphosing cells. Biochem. Cell Biol. 1994; 72(11): 589-596.
38. Klowden JR. Physiological Systems in Insects. 2nd (Edn) Academic Press, USA; 2007.
39. Tata JR. Amphibian metamorphosis as a model for studying the developmental actions of thyroid hormone. Cell Research 1998; 8: 259-272.
40. Mullur R, Liu YY, Brent GA. Thyroid hormone regulation of metabolism.

Physiol. Rev. 2014; 94(2): 355-382.
41. Lu D. Australian lungfish has largest genome of any animals sequenced so far. New Scientist 2021; 18 January.
42. Lu G, Luo M. Genomes of major fishes in world fisheries and aquaculture: Status application and perspective. Aquaculture and Fisheries2020; 5(4): 163-173.
43. Zuo B, Micah L, Sun Y-B. Comparative genomics reveals into anuran genome size and evolution. BMC. Genomics 2023; 24: 379.
44. Lin H, Jiang F, Wang S. etc.Chromosome-level genome of globe skimmer dragonfly (*Pantale flavescens*). GigaScience vol. 11, giac009. http://doi.org/10.1093/gigascience/giga009.
45. He K, Lin K, Wang G, Li F. Genome sizes of nine species determined by flow ctometry and k-mer analysis. Frontiers in Physiology 2016; 7:565.
46. Otsuka J Kawai Y, Sugaya N. The influence of selection on the evolutionary distance estimated from the base changes observed between homologous nucleotide sequences. J. Ther. Biol. 2001; 213: 129-144.
47.Aiello L, Dean C. An Introduction to Human Evolutionary Anatomy. Elsevier Academic Press,London, Sandiego;1990.
48. Strait DS, Grin FE, Moniz MA.A reappraisal of early hominid phylogeny. Journal of Human Evolution 1997; 32(1): 17-82.
49. Grine FE, Fleagle JG. The First Humans: A Summary Perspective on the Origin and Early Evolution of the Genus Homo. Vertebrate Paleobiology and Paleoanthropology, Springer, Netherland; 2009: pp. 197-207.
50. Sayer K, Raghanti MA, Lovezoy CO. Human evolution and the chimpanzee referential doctrine. Annual Review of Anthropology 2012; 41: 119-138.
51. Darwin C. The Origin of Species. John Murry, London;1859.
52. Sero PC, Rao C. Early evolution of avian flight and perching: new evidence from the Lower Cretaceous of China. Science 1992; 255(5046): 845-848.
53. Mayer G, Pohl B, Peters, DS. A well preserved *Archaepteryx* specimen with theropod features. Science 2005; 310(5753): 1483-1486.
54. Lingham-Soliar T, Fuduccia A, Wang X. A new chinese specimen indicates that protofethers in the Early Cretaceous theropod dinosaurs *Sinosanropteryx* are degraded collagen fibers. Proceedings of the Royal Society B. 2007; 274 (1620): 1823-1829.
55. Van Andel TH. New views on an old planet: A History of Global Change (2nd edn.) Cambridge University Press, Cambridge UK;1994.

56. Rubin R et al. Climate cycles during a Neoproterozoic “snow ball” glacial epoch. Geology 2007; 35(4): 299-302.
57. Holland HD.The oxygenation of the atmospheres and oceans.
Philosophical Transactions of Royal Society B. Biological Science 2006; 361 (1470): 903-915.

Figure legends

Fig. 1. Divergence of eukaryotes after the acquirement of mitochondria. The divergence times in the abscissa are measured from the divergence time of green plant and animal, which is estimated to have been $1,2x10^9$ years ago [13]. At least fourbranch points *a*, *b, c* and *d* are recognized as the major divergence of eukaryotes; (*a*) the first divergence of unicellular eukaryotes, (*b*) the divergence of fungi and sea algae, (*c*) the appearance of ancestral green plants and (*d*) the appearance of ancestral animals, although the divergence of unicellular eukaryotes is still recognized between the branch points (*b*) and (*c*).It is also notable that the unicellular eukaryotesassigned to the same taxonomical category are divided into two or three lineages, as listed below. *Ascomycota* 1 (*Hemiascomycetes)*, *Ascomycota* 2 (*Archiascomycetes*), *Rhodophyta* 1 (*Florideophycidae*), *Rhodophyta* 2 (*Protoflorideophycidae*), *Chlorophyta* 1 (*Chlamydomonas*), *Chlorophyta* 2 (*Chlorella*, *Ulva*, *Spirogyra*), *Protista*1 (*Trypanozoma*), *Protista* 2 (*Acanthamoeba*),*Protista*3 (*Euglena*, *Cyanophora*, *Tetrahymena*).

Fig. 2. The phylogeny of main phyla in animals on the basis of the analysis on base-pair changes in mitochondrial rRNA. For reference, the dates of representative fossil records of animals in the range from $5x10^8$ to $12x10^8$ years ago are denoted by dotted vertical lines: (i) Stirling Range formation of south western Australia of $12.15x10^8$years ago containing discoidal impression and trace-like fossils intidal sand stones [15], (ii) *Ediacara* and *Avalon* faunas contaning small

shelly fossils of $5.7x10^8$ years ago [16, 17] and (iii) Cambrian Burgess Shale of $5.4x10^8$years ago recording various types of *Arthropoda*, some *Annelida* and an ancestral form of*Chordata* called *Pikaia* [18].

Fig. 3. Phylogenetic tree of *Charophyta*and main phyla and classes in land plants. The following fossil records of land plants are used for the dating: (1)the early macrofossils of free-sporing vascular plants from the Upper Silrian sediments of Australia [20, 21], (2) the early macrofossils of *Filicopsida*from the Upper Devonian throughto Lower Carboniferous [22], (3) the earliest seed plants from Upper Devonian sediment of Europe [23, 24] and of north America [25], (4) the earliest fossil assigned to *Lycopsida* from the Upper Devonian [26, 27], (5) the early fossil record of *Podocarpaceae* from the Triassic [28], (6) the early fossil record of *Pinaceae* from the Lower Cretaceous [28] and (7) the fossils of uniapertuate pollen and triaperturate pollen from the Lower Cretaceous [28].