Gene 534 (2014) 444–448

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Short Communication

Extreme nearly neutral evolution in mitochondrial genomes of laboratory mouse strains

Takahiro Yonezawa *, Masami Hasegawa

School of Life Sciences, Fudan Univeristy, HanDan Rd. 220, Shanghai, 200433, China
Institute of the Statistical Mathematics, Midori-cho 10-3, Tachikawa, Tokyo, 190-8562, Japan

ARTICLE INFO

Article history:
Accepted 8 August 2013
Available online 16 August 2013

Keywords:
Mus musculus domesticus
Slightly deleterious mutation
Branch model

ABSTRACT

Relaxation of the selective constraint during the domestication process is known. In this study, we report unexpected closeness to neutral evolution of mitochondrial genomes of laboratory mouse strains: estimated nonsynonymous/synonymous rate ratio being very close to 1 ($\widehat{\omega} = 1.32$). Probably it is due to the extreme inbreeding extending over 100 years as well as to their recent origin (middle of the last millennium). There is no rate difference observed among three codon positions as well as ribosomal RNA and control regions. However, the amino acid substitutions occurred not randomly, and substitutions were more frequent between physico-chemically similar amino acids than between dissimilar ones. Probably this is inevitable consequence caused by the codon table itself, but not by selections. This implies that a large portion of the new mutations are conservative, and most of them are slightly deleterious and not lethal. It seems that, even though the selection pressures do not hold normally, the function of genes may not be impaired in most cases.

1. Introduction

As Charles Darwin got the inspiration of the evolution from the variation of the domesticated animals, the domestication process is ruled by a very similar principal as the evolution. The domesticated animals are defined as the animals whose reproduction is under human control (Spurway, 1955).

Hence, for the better understanding of the breeds of the domestic animals, the most important thing is to understand the human societies surrounding them. Aim and motivation of human mind determine the direction of the evolution of the breeds. In this meaning, the laboratory mouse strains (*Mus musculus domesticus*) are very interesting. Compared with the other domesticated animals, the aim of the mouse breeding is totally different. In this case, aim and direction of breeding were to generate and to maintain the strains that are genetically identical and their genomes are homozygous at all loci (Beck et al., 2000). The inbred laboratory mouse strains were defined as "a strain shall be regarded as inbred when it has been mated brother × sister for twenty or more consecutive generations," and theoretically in average at least 98.6% of the loci in each mouse are homozygous at the 20th generations (Beck et al., 2000). In the case of the maternally inherited mitochondrial genes, theoretically they are fixed at the F1 generation by brother × sister inbred. Since William Castle began studying inheritance in mice, about 100 years have passed. The inbred laboratory mouse strain has played an important role in genetics, immunology, etiology, developmental biology, and other disciplines. Now they are essential for the life sciences.

The evolutional studies of them were mainly focused on their phylogeny (Goios et al., 2007; Nunome et al., 2009; Suzuki et al., 2004; Yu et al., 2009).

However, it is interesting in the evolutionary aspect as well as in the thremmatology aspect to consider about what was raised by the extreme inbreeding. We carried out a detailed analysis of the selection pressure operating on their mitochondrial genomes during the evolution, and here we report their unexpectedly nearly neutral evolution.

2. Materials and methods

2.1. Sequence data and alignment

Seventy-seven mitochondrial genomes of *Mus musculus* (69 for *domesticus*, 3 for *musculus*, 3 for *molossinus*, and 2 for *castaneus*), one mitochondrial genome of *Mus terricolor* and one mitochondrial genome of *Rattus norvegicus* were retrieved from NCBI. Accession numbers were summarized in the supplemental Table S1. Twelve protein genes on the H strand were manually aligned and concatenated. The initiation and termination codons as well as the overlapping regions between genes were excluded from the alignment. In this data set, 32 haplotypes were identified in *M. musculus*.

Abbreviations: ML, maximum likelihood; CLS, common laboratory strain; Ma, mega annum; Ka, kilo annum.

* Corresponding author at: School of Life Sciences, Fudan University, 220, Han Dan Road, Shanghai, 200433, China. Tel./fax: +86 21 5566 5264.
E-mail address: cyclotis@gmail.com (T. Yonezawa).

0378-1119/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.gene.2013.08.022

2.2. Data analyses

After removing all identical sequences, ML tree was inferred by RAxML ver. 7.0.3 (Stamatakis et al., 2008) with GTR+I+Γ model. The branch model (Yang, 1998) that allows the different level of heterogeneity in the ω (non-synonymous/synonymous rate ratio) in lineages was applied using the CODEML program of PAML ver. 4.4 (Yang, 2007). The coalescent times were estimated under the strict clock method by the BASEML and CODEML programs of PAML. The median joining (MJ) network (Bandelt et al., 1999) was estimated by using NETWORK 4.6.0.0 program (http://www.fluxus-engineering.com/sharenet.htm) based on the complete mitochondrial protein coding genes of the common laboratory mouse strains (68 individuals with 23 haplotypes).

3. Results

3.1. Phylogenetic relationships and estimation of selection pressures

ML tree and the ω ratios of each lineage were shown in the Fig. 1 (panel A). The median joining network based on the CLS (common laboratory strain) was shown in the supplemental Fig. S1. Phylogenetic relationships among subspecies of *M. musculus* as well as relationships among *M. m. domesticus* are basically concordant with previous studies (Goios et al., 2007; Yu et al., 2009).

The ω ratios of each lineage were estimated as follows: $\hat{\omega}_{DEEP}$ (ω ratio of the deep branches among subspecies) is 0.040, $\hat{\omega}_{WILD}$ (ω ratio of wild subspecies) is 0.069, $\hat{\omega}_{NZB}$ (ω ratio of the lineage of strain NZB/JN, the oldest known lineage of *M. m. domesticus*) is 0.172, $\hat{\omega}_{ANC}$ (ω ratio of the ancestral lineage of the CLS) is 0.013, and $\hat{\omega}_{CLS}$ (ω ratio of the lineage of the CLS) is 1.321 (Fig. 1, panel B). The high ω ratio of the CLS is harmonious with the previous studies (Goios et al., 2007; Yu et al., 2009). To evaluate the difference of these ω ratios, the likelihood ratio test (LRT) was applied (Table 1). In most cases, the difference among $\hat{\omega}_{DEEP}$, $\hat{\omega}_{WILD}$, and $\hat{\omega}_{ANC}$ were not significant, but the differences between these estimates and $\hat{\omega}_{CLS}$ were highly significant. The best model in terms of AIC was "3ω model 2" ($\hat{\omega}_{DEEP} = \hat{\omega}_{WILD} = \hat{\omega}_{ANC} \neq \hat{\omega}_{NZB} \neq \hat{\omega}_{CLS}$). Subsequently, we evaluated whether $\hat{\omega}_{CLS}$ is significantly larger than 1 or not. Based on the 23 sequences of the CLS, two different models were compared: (1) optimizing the ω ratio and (2) fixing the ω ratio as 1. AIC preferred to the latter model.

3.2. Coalescent times estimations

Ho et al. (2005) demonstrated the time dependent discrepancy of the molecular evolutionary rate, high evolutionary rate (≈mutation rate) in the short term and low evolutionary rate (≈substitution rate) in the long term. Hence, if such time dependent discrepancy is not taken into account, the estimates are not reliable. For this reason, coalescent (or divergence) time estimations should be separately carried out for the different time scales using appropriate method and data. By this procedure, the effect of the time dependent discrepancy of evolutionary rates on time estimation might be minimized.

In this study, time estimations were separately carried out in the following three different time scales: [1] inter-species level, [2] intra-species level (except for among common laboratory strains), and [3] among common laboratory strains.

[1] At the inter-species level, time dependent discrepancy of the evolutionary rates can be ignored due to long time scale (>1 Ma). In this study, the divergence time between *M. musculus* and *M. terricolor* was estimated using *R. norvegicus* as outgroup. The ω ratios of these three species were very low (0.015–

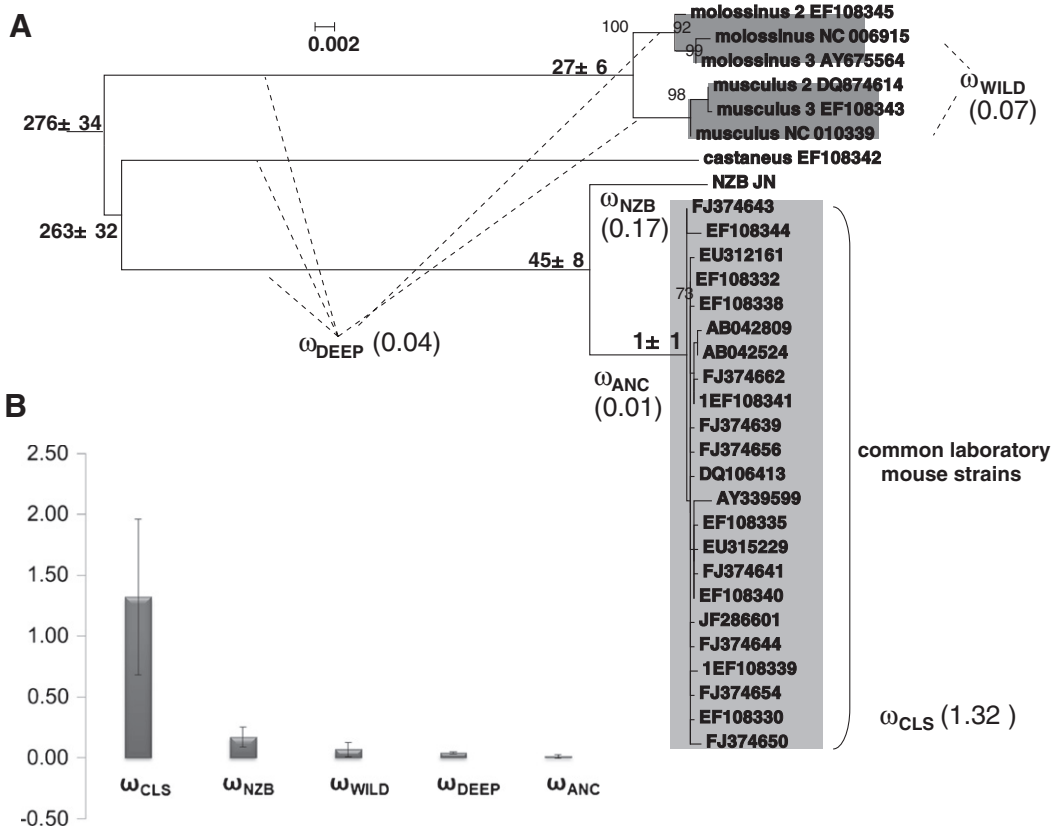

Fig. 1. The ML tree of *Mus musculus* based on the complete mitochondrial protein genes. Branch lengths were proportional to numbers of codon substitutions. The ω ratios in each lineage were shown along branches or near the clades. The nodal numbers were estimated coalescent times (in Ka: kilo annum) with standard deviations after ± (panel A). Maximum likelihood estimation of ω ratios of CLS, NZB, ancestral branch of laboratory strains, wild subspecies, deep branches of *Mus musculus* (thick bars) and their standard deviations (thin bars). Standard deviations were estimated by the bootstrap method (100 replications) (panel B).