Biotechnology Frontier
March 2014, Volume 3, Issue 1, PP.1-7

Natural Selection on Human Mitochondrial DNA

Siqi Huang, Chuanchao Wang, Hui Li [#]

MOE Key Laboratory of Contemporary Anthropology, School of Life Sciences, Fudan University, Shanghai 200433, China

[#]Email: lihui.fudan@gmail.com

Abstract

As an important tool for population genetics and molecular anthropology, mitochondrial DNA (mtDNA) has long been used for the studies of human origin and evolution. However, previous studies were mostly based on simple models of evolution, ignoring the possible effects of natural selection upon mtDNA. In recent years, human mitochondrial genome was proved having undergone strong effect of natural selection by an increasing number of evidences. Climate-induced adaptive selection has contributed to shaping the current continent-specific distribution of mtDNA lineages. Strong purifying selection has also been suggested by statistics of nonsynonymous versus synonymous mutations on mtDNA. Therefore, possible effects induced by natural selection should be taken into account when we use mtDNA in human population genetic studies.

Keywords: Mitochondrial DNA; Purifying Selection; Adaptive Selection; Synonymous Mutation; Climate; Evolution

1 INTRODUCTION

Maternal inherited mtDNA has long been used to investigate the human evolution and origin because of its high mutation rate and the lack of recombination. Early studies neglected the natural selection because of their simple models of evolution including the assumption that the divergence of human mtDNA is due to a molecular clock [1-3]. However, recent experiments find that the current distribution of mtDNA has been shaped by external factors, such as climate [4]. More and more population geneticists and anthropologists have paid attention to natural selection on mtDNA and begun studying the effect of both adaptive and purifying selection [5-11]. With the development of biological technology in recent years, mtDNA data have been gradually accumulated and enable us to address this issue step by step.

2 FEATURES OF MTDNA

Human mtDNA is a small (16.6-kb) circular genome coding for 13 subunits of the mitochondrial oxidative phosphorylation system (OXPHOS: ATP6, ATP8, CO1, CO2, CO3, CYB, ND1, ND2, ND3, ND4, ND4L, ND5, and ND6), 2 rRNAs (RNR1, and RNR2), and 22 tRNAs. Most unique features of mtDNA are its maternal inheritance and lack of recombination [12]: the complexities imposed by recombination of paternal and maternal genomes can be excluded in reconstructing evolutionary histories. Besides, groups of related haplogroups and haplogroup clusters have defined branches of the phylogenetic tree for mtDNA, and many previous studies have indicated that the geographical distribution of haplogroups in aboriginal populations is continent-specific [13].

Continent-specific distribution is one of the distinguished features of human mtDNA regardless of genetic admixture in certain populations [14]. The seven most ancient haplogroups (designated by L0-L6) are observed specifically in Africa [15-19]. All non-African clades can be subdivided into two L3-derived haplogroups, M and N [20]. Afterwards, the haplogroup clusters HV, JT, KU, and IWX [1] were derived primarily from N and populated Europe, while M and N dispensed equally to the radiation of mtDNA into Asian-specific haplogroups A, C, D, G, Z, and Y. The American continent was covered from north-eastern Asia by individuals with haplogroups A, B, C, and D [14,21].

3 CLIMATE-INDUCED ADAPTIVE SELECTION

MtDNA population studies [19] have shown that human migrations and genetic drift were responsible for the current distribution of maternal lineages across the world as well as the regional variation of human mtDNA's. However, it cannot explain the enrichment of lineages A, C, D, and G in arctic region and the fact that only two mtDNA lineages (M and N) left Africa to colonize Eurasia. That contradiction thus gave rise to an alternative hypothesis: natural selection.

In 2003, Mishmar et al. [21] analyzed 104 complete mtDNA sequences from all global regions and lineages (56 mtDNA sequences from the literature [22-26] and 48 additional individuals from African, Asian, European, Siberian and Native American populations). It turned out to be a notable deviation from the standard neutral mode among European, Asian, Siberian and Native American mtDNA variations, but not in African variation. They then hypothesized that natural selection may play a role in shaping the regional differences between mtDNA lineages. Principally speaking, natural selection would act through amino acid variants in the mtDNA OXPHOS polypeptides. To detect how and to what extent selection influences, Mishmar et al. used DNASP (DNA Sequence Polymorphism) software to do the *Ka/Ks* analysis (the ratio of the number of substitutions causing amino acid replacements (nsyn sites) per total possible nsyn sites (*Ka*) was divided by the number of silent substitutions (syn sites) per possible syn sites (*Ks*)) [27,28]. The result showed that the value of *Ka/Ks* wasn't 1, reflecting the effect of natural selection on a particular mtDNA protein gene. What's more, the study also found a dramatic correlation between increased amino acid substitutions in particular genes of mitochondrial DNA and climate zone. By comparing the distribution of the *Ka/Ks* ratios of the 13 mtDNA genes for mtDNA haplogroup lineages from three different geographic regions: the tropical and subtropical zones (L0–L3); the temperate zone (H, V, U, J, T, I, X, N1b, W); the subarctic and arctic zones (A, C, D, G, Z, Y, X), Mishmar et al. discovered dramatic differences in the amino acid sequences of particular mtDNA genes. Increased amino acid sequence variation was found in ATP6, ATP8, CO3 and ND6 among arctic mtDNA lineages. However, Africans were labelled high amino acid variation for ATP6, CYB, CO1, CO2, ND1, ND2 and ND5.

While this experimental result pointed a new direction in analyzing human evolution and origin (afterwards many other anthropologists casted light on natural selection [29,30]), it raised some novel questions. What role does climate play in human mitochondrial DNA evolution? Whether the current studies and technology can support the selection hypothesis substantially? Which plays a predominant role in mtDNA, adaptive selection or purifying selection?

Previous studies have testified that selection has played a role in the differentiation of human mtDNA and climate was probably the main driving factor [18,21,29,30]. However, C. Sun et al. refuted that these equivocal findings were not supported by the subsequent studies [11,23,28]. One argument was the lack of a systematic investigation on South Asia and Oceania—two important (sub) continents with distinct climates and matrilineal components [31-33].

To clarify this problem, Sun et al. collected 237 complete mtDNA sequences of indigenous lineages from South Asia, Oceania and East Asia for the nonsynonymous (N) and synonymous (S) substitutions analysis [34]. They identified an abnormal excess of nonsynonymous substitutions in ATP6 from East Asia haplogroup N (13/2 in N/S ratio) and a relative surplus of nonsynonymous substitutions in ND5 from South Asia N, which might be interpreted as climate adaptation. However, its counterpart, haplogroup M, did not show a similar pattern in the analysis, suggesting the failure in detecting the imprint of climate adaptation. The result of this experiment did not support the selection hypothesis. Sun et al. attributed this discrepancy to some limitations in Mishmar et al.[21] and Ruiz-Pesini et al.[30], including the overweight sites of Nei-Gojobori, simplistic climate assumption, and overlooking of other effects[6]. Other demerits also stood out: (a) E. Ruiz-Pesini, et al. might incorrectly take the possibility of observed data as the Fisher's exact test value, causing the inaccuracy in comparing RFi and RF (replacement mutation frequency in internal and terminal groups respectively) [30]. A revaluation of their data indicated that almost no meaningful difference was detected, which agreed with Sun's test results. (b) The pairwise comparison data sets from Mishmar et al. evidently did not meet the prerequisite of the Wilcoxon ranked-sum test that the data taken into consideration should be generated or obtained independently. Consequently, there would be a biased or void result. (c) The lineages belonging to haplogroup X, which were mainly sampled from Europe (according to the temperate zone as

suggested by Ruiz-Pesini et al.) were, however, inconsistently arctic-specific and grouped with the East Asian haplogroups. As a result, the role of climate in human mtDNA wasn't substantiated evidently.

Sun et al. also found evidence in time axis that the increase in N/S ratio was not so striking in their study (0.46 vs. 0.42) and Ruiz-Pesini et al. (0.53 vs. 0.41). They attributed this discrepancy to the underestimation of the N/S ration, in that Elson et al. counted each homoplasy mutation only once. This result was definitely against the selective hypothesis, since the deleterious mutations in terminal branches had been exposed to selection pressure for a relatively shorter time and thus might not have been cleaned by selection. It was possible that terminal branches contain more deleterious or nonsynonymous mutations than the internal branches [35,36]. So Sun et al. speculated that the increase in N/S ratio in terminal branches might be slight.

According to Sun et al., climate was unlikely to be the driving force that shapes modern human mtDNA due to the lack of the notable difference in N/S ratio in both spatial and time dimensions.

Along with Sun et al., many other anthropologists [37-39] also suggested that methods used to detect the selection had deficiency and the results were not supported by the reanalysis of larger data sets [40]. However, in 2009, by using a geographical framework, Balloux carried out the first direct test of the role climate and past demography have played in shaping the current spatial distribution of mtDNA sequences worldwide [4].

To test the extent to which demography and climate have shaped the current distribution of mitochondrial sequence diversity, Balloux used physical distance as a proxy and then analyzed the linear relationship between genetic differentiation and geographical distance along landmasses as well as geographical distance from sub-Saharan Africa. It turned out to well predict the genetic diversity of individual populations all over the world. In dealing with the data (over 5000 hyper-variable segment I (HVS-I) sequences downloaded from the HVRBase++ online resource and 628 whole mtDNA genomes), Balloux used the MITOMAP database [41] to identify synonymous and non-synonymous SNPs (Single Nucleotide Polymorphism). He applied a linear model with the distribution of worldwide mitochondrial sequence diversity as the response variable and with distance and temperature as predictors to plot the data. According to the results, distance from Africa along landmasses explained 18.3% of the variance in within-population genetic diversity for the first hypervariable segment of mtDNA (HVS-I) ($F_{1,10}$= 25.2, $p<0.001$). Besides, HVS-I genetic diversity was found to be strongly related to minimum temperature as well, which explained 7.8% of variance ($F_{1,107}$=10.2, p=0.002). Balloux also pointed out that distance from Africa and minimum temperature for the sampled populations are not correlated (R^2 =0.010, $F_{1,107}$ =1.3, p=0.254), in that when they accounted for the effect of distance from Africa, the relationship between HVS-I diversity and minimum temperature remained significant (R^2=0.055, $F_{1,106}$=8.7; p=0.004). Considering that, Balloux explained that both distance from Africa (bottleneck or founder effect) and low temperatures have reduced within-population mtDNA diversity at HVS-I.

4 PURIFYING SELECTION PLAYS THE PREDOMINANT ROLE

After testifying the effect of natural selection, scientists began to weigh which selection contributed more to the shaping of human mtDNA distribution pattern. In an early study, Bazin et al. [42] compiled published animal mitochondrial DNA sequences and observed that mtDNA polymorphism was independent from population size. Thus they suggested that frequent episodes of natural selection are the cause of this discrepancy. When testing the signature of positive selection in the mtDNA, Bazin et al. considered the ratio of *Ka/Ks* in whole mitochondrial genomes of taxa with small versus large population sizes. It turned out that all *Ka/Ks* ratios were well below one, indicating that purifying selection is the predominant force in mtDNA evolution. We further testified this result by analyzing the ratio of *Ka/Ks* using the 16808 complete human mtDNA sequences downloaded from Phylotree (http://www.phylotree.org/). The ratios were calculated in the way: *Ka/* (*Ka* +*Ks*) [Fig1].

With many studies concluded that mtDNA is subject to strong purifying selection, Stewart et al. discussed this issue from the perspective of mtDNA transmission in the mammalian female germ line [43]. By using the mtDNA mutator mouse to studying the fate of random mtDNA mutations in the mouse female germ line [44], they found that the frequency of mutations which would change amino acids in mtDNA‐encoded proteins was lower than expected under a neutral model. This implies that mutations which are likely to affect an mtDNA‐encoded protein are being

strongly selected against. In trans-mitochondrial mice created by Fan et al., a point mutation that severely impaired the function of a respiratory chain subunit was found to be eliminated rapidly from the maternal germ line [45]. That severe mtDNA mutations are not transmitted to the offspring was also discovered in other laboratorys [46]. These findings provide a strong evidence for the existence of purifying selection on mtDNA in the mouse germ line.

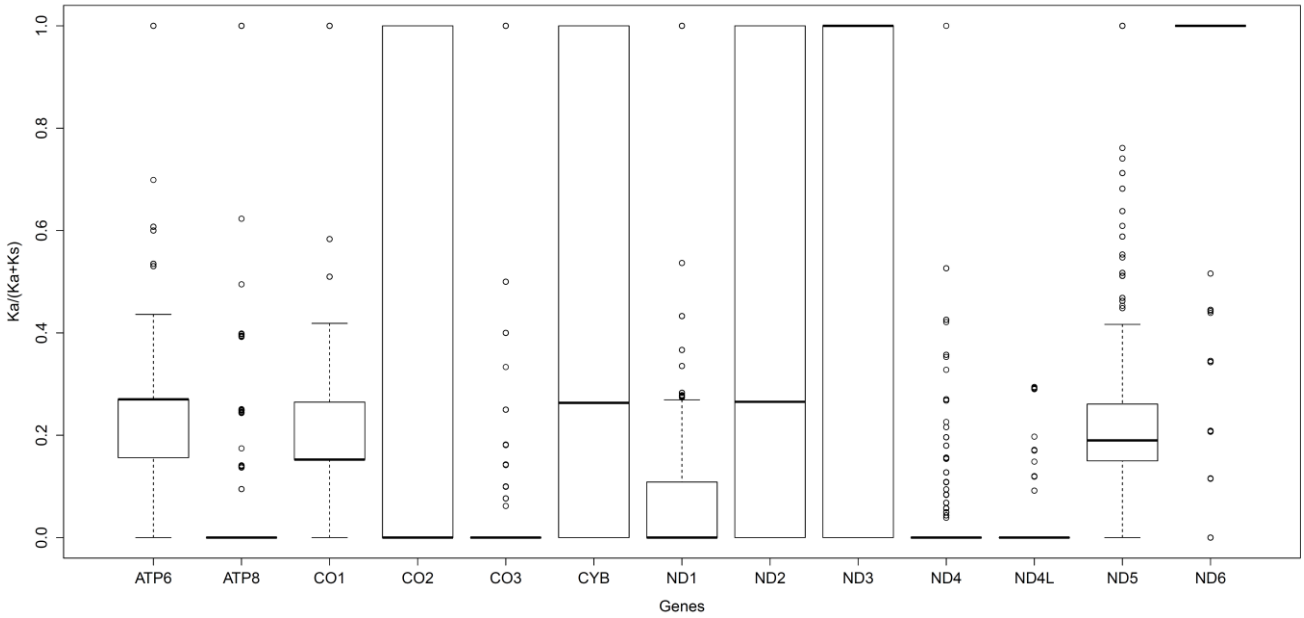

Fig1. Distribution of the relative selective constraints [*Ka/ (Ka +Ks)*] of the 13 human mtDNA genes calculated from the 16808 complete human mtDNA sequences. If *Ka/ Ks*>1, the positive selection would be predominant, while *Ka/Ks*=1 suggests a neutral effect. However, the ratios of most genes in the figure are below 0.5, indicating the effect of purifying selection. The bottom and top of the box are the first and third quartiles, and the thick lines inside the box are the medians. The ends of the whiskers represent the minimum and maximum of all the data. Outliers are plotted as individual points.

Correspondingly, purifying selection may also has an important role in shaping human mtDNA sequence variation, since there are pronounced similarities in the pattern of neutral versus non-neutral substitutions in protein coding genes in mtDNA sequences from humans. For example, Elson et al. [6] assembled a data set of 560 coding region sequences and calculated *Ka/Ks* ratios and the neutrality index (NI) values to consider the action of selection on individual mitochondrial genes. In this experiment, the NI for the African superclade sequences is 4.75, while the value for the European mtDNAs is 1.39. The numbers of substitutions in the Asian mtDNAs are much smaller. Thus, the relatively lower proportion of nonsynonymous substitutions in the evolutionarily older African mtDNAs indicates that purifying selection has acted on the human mitochondrial genome during evolution. Evidence for purifying selection was also obtained from their phylogenetically stratified contingency analyses of synonymous and nonsynonymous substitutions. There was a highly marked “loss” of nonsynonymous substitutions among relatively old site changes. The preponderance of evidence indicated the operation of purifying selection on the human mitochondrial genome during evolution.

According to those studies, we could conclude that while both purifying and adaptive selection play an important role in shaping human mtDNA, and the former one is the main force. However, as what tested in Bazin et al.’s study, neither negative nor positive selection should be ignored during the research.

5 CONCLUSIONS

Supported by many solid evidences, natural selection has been confirmed to shape the evolution of the human mitochondrial genome and purifying selection is the predominant force. Further researches should take great care when using mtDNA data to date phylogenetic events, since the previous assumption of neutral theory may not hold any longer. And as we keep digging the role of selection, this tiny piece of genetic information is likely to play a more important role in the study of human evolution and population dispersal.

ACKNOWLEDGMENT

This work was supported by the National Excellent Youth Science Foundation of China (31222030), National Natural Science Foundation of China (31071098, 91131002), Shanghai Rising-Star Program (12QA1400300), Shanghai Commission of Education Research Innovation Key Project (11zz04), and Shanghai Professional Development Funding (2010001).

REFERENCES

[1] Finnil äA., Lehtonen M S., Majamaa K. “Phylogenetic network for European mtDNA.” Am J Hum Genet, 2001, 68: 1475-1484

[2] Herrnstadt C., Elson J L., Fahy E. et al. “Reduced-median-network analysis of complete mitochondrial DNA coding-region sequences for the major African, Asian, and European haplogroups.” Am J Hum Genet, 2002, 70: 1152-1171

[3] Torroni A., Wallace D C. “Mitochondrial DNA variation in human populations and implications for detection of mitochondrial DNA mutations of pathological significance.” J Bioenerg Biomembr, 1994, 26: 261-271

[4] Balloux F, Handley L J L, Jombart T, et al. “Climate shaped the worldwide distribution of human mitochondrial DNA sequence variation.” Proc R Soc B, 2009, 276: 3447-3455

[5] Blier P U, Dufresne F, Burton R S. “Natural selection and the evolution of mtDNA-encoded peptides: evidence for intergenomic co-adaptation.” Trends Genet, 2001,.17: 400-406

[6] Elson J L., Turnbull D M., Howell N. “Comparative genomics and the evolution of human mitochondrial DNA: assessing the effects of selection.” Am J Hum Genet, 2004, 74: 229-238

[7] Gu M, Dong X, Shi L, et al. “Differences in mtDNA whole sequence between Tibetan and Han populations suggesting adaptive selection to high altitude.” Gene,2012, 496: 37-44

[8] Mishmar D, Ruiz-Pesini E, Mondragon-Palomino M, et al. “Adaptive selection of mitochondrial complex I subunits during primate radiation.” Gene, 2006, 378: 11-18

[9] Pereira L, Soares P, Radivojac P, et al. “Comparing Phylogeny and the Predicted Pathogenicity of Protein Variations Reveals Equal Purifying Selection across the Global Human mtDNA Diversity.” Am J Hum Genet,2011, 88: 433-439

[10] Ruiz-Pesini E, Wallace D C. “Evidence for Adaptive Selection Acting on the tRNA and rRNA Genes of Human Mitochondrial DNA .” Hum Mut,2006, 27: 1072-1081

[11] Soares P, Ermini L, Thomson N, et al. “Correcting for Purifying Selection: An Improved Human Mitochondrial Molecular Clock.” Am J Hum Genet,2009, 84: 740-759

[12] Elson J L., Andrews R M., Chinnery P F. et al. “Analysis of European mtDNAs for recombination.” Am J Hum Genet, 2001, 68: 145-153

[13] Torroni A., Huoponen K., Francalacci P. et al. “Classification of European mtDNAs from an analysis of three European populations.” Genetics,1996, 144: 1835-1850

[14] Wallace D C., Brown M D., Lott M T. “Mitochondrial DNA variation in human evolution and disease.” Gene, 1999, 238:211-230Howell N., Elson JL., Turnbull DM. et al. “African haplogroup L mtDNA sequences show violations of clock-like evolution.” Mol Biol Evol, 2004, 21: 1843-1854

[15] Kivisild T, Reidla M, Metspalu E, et al. “Ethiopian mitochondrial DNA heritage: tracking gene flow across and around the gate of tears.” Am J Hum Genet, 2004, 75: 752-770

[16] Salas A, Richards M, De la Fe T, et al. “The making of the African mtDNA landscape.” Am J Hum Genet, 2002, 71: 1082-1111

[17] Torroni A, Rengo C, Guida V, et al. “Do the four clades of the mtDNA haplogroup L2 evolve at different rates?” Am J Hum Genet, 2001, 69: 1348-1356

[18] Torroni A., Achilli A., Macaulay V. et al. “Harvesting the fruit of the human mtDNA tree.” Trends Genet, 2006, 22: 339-345

[19] Quintana-Murci L, Semino O, Bandelt H J, et al. “Genetic evidence of an early exit of Homo sapiens sapiens from Africa through eastern Africa.” Nat Genet, 1994, 23: 437-441

[20] Mishmar D, Ruiz-Pesini E, Golik P, et al. “Natural selection shaped regional mtDNA variation in humans.” Proc Natl Acad Sci USA, 2003, 100: 171-176

[21] Arnason U, Xu X, Gullberg A. “Comparison between the complete mitochondrial DNA sequences of Homo and the common chimpanzee based on nonchimeric sequences.” Journal of Molecular Evolution, 1996, 42(2): 145-152.

[22] Andrews RM., Kubacka I., Chinnery PF. et al. “Reannalysis and revision of the Cambridge reference sequence for human mitochondrial DNA” Genet, 1999, 23: 147

[23] Anderson S, Bankier A T, Barrell B G, et al. “Sequence and organization of the human mitochondrial genome.” Nature, 1981, 290: 457-465

[24] Horai S, Hayasaka K, Kondo R, et al. “Recent African origin of modern humans revealed by complete sequences of hominoid mitochondrial DNAs.” Proceedings of the National Academy of Sciences, 1995, 92(2): 532-536.

[25] Ingman M., Kaessmann H., Paabo S. et al. “Mitochondrial genome variation and the origin of modern humans.” Nature, 2000, 408(6813): 708-713.

[26] Nei M., Gojobori T. “Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions.” Mol Biol Evol, 1986, 3: 418-426

[27] Rozas J., Rozas R. “DNASP version 3: an integrated program for molecular population genetics and molecular evolution analysis.” Bioinformatics, 1999, 15: 174-175

[28] Ballard J W O., Whitlock M C. “The incomplete natural history of mitochondria.” Mol Ecol, 2004, 13: 729-744

[29] Ruiz-Pesini E., Mishmar D., Brandon M. et al. “Effects of purifying and adaptive selection on regional variation in human mtDNA.” Science, 2004, 303: 223-226

[30] Sun C, Agrawal S, Bandelt H J, et al. “Phylogeny of mitochondrial DNA macrohaplogroup N in India, based on complete sequencing: implications for the peopling of South Asia.” Am J Hum Genet, 2004, 75: 966-978

[31] Macaulay V, Hill C, Achilli A, et al. “Single, rapid coastal settlement of Asia revealed by analysis of complete mitochondrial genomes.” Science, 2005, 308: 1034-1036

[32] Sun C. et al. “The dazzling array of basal branches in the mtDNA macrohaplogroup M from India as inferred from complete genomes.” Mol Biol Evol, 2006, 23: 683-690

[33] Sun C, Kong Q P, Zhang Y P. “The role of climate in human mitochondrial DNA evolution: A reappraisal.” Genomics, 2007, 89: 338-342

[34] Moilanen J S., Majamaa K. “Phylogenetic network and physicochemical properties of nonsynonymous mutations in the proteincoding genes of human mitochondrial DNA.” Mol Biol Evol, 2003, 20: 1195-1210

[35] Torroni A, Campos Y, Rengo C, et al. “Mitochondrial DNA haplogroups do not play a role in the variable phenotypic presentation of the A3243G mutation.” Am J Hum Genet, 2003, 72: 1005-1012

[36] Amo T., Brand M D. “Were inefficient mitochondrial haplogroups selected during migrations of modern humans? A test using modular kinetic analysis of coupling in mitochondria from cybrid cell lines.” Biochem J, 2007, 404: 345-351

[37] Ingman M., Gyllensten U. “Rate variation between mitochondrial domains and adaptive evolution in humans.” Hum. Mol. Genet, 2007, 16: 2281-2287

[38] Kivisild T, Shen P, Wall D P, et al. “The role of selection in the evolution of human mitochondrial genomes.” Genetics, 2006, 172: 373-387

[39] Kryazhimskiy S., Plotkin J B. “The population genetics of dN/dS.” PLoS Genet, 2008, 4: e1000304

[40] Ruiz-Pesini E, Lott M T, Procaccio V, et al. “An enhanced MITOMAP with a global mtDNA mutational phylogeny.” Nucleic Acids Res, 2007, 35: D823-D828

[41] Bazin E, Glémin S, Galtier N. “Population size does not influence mitochondrial genetic diversity in animals.” Science, 2006, 312: 570-572

[42] Stewart J B, Freyer C, Elson J L, et al. “Strong purifying selection in transmission of mammalian mitochondrial DNA.” PLoS Biol, 2008, 6: e10

[43] Trifunovic A, Wredenberg A, Falkenberg M, et al. “Premature ageing in mice expressing defective mitochondrial DNA polymerase.” Nature, 2004, 429: 417-423

[44] Fan W, Waymire K G, Narula N, et al. “A mouse model of mitochondrial disease reveals germline selection against severe mtDNA mutations.” Science, 2008, 319: 958-962

[45] Kasahara A, Ishikawa K, Yamaoka M, et al. “Generation of trans-mitochondrial mice carrying homoplasmic mtDNAs with a missense mutation in a structural gene using ES cells.” Hum Mol Genet, 2006, 15: 871-881

[46] Balloux F, Handley L J L, Jombart T, et al. “Climate shaped the worldwide distribution of human mitochondrial DNA sequence variation.” Proc R Soc B, 2009, 276: 3447-3455

AUTHORS

[1]***Siqi Huang*** (Shanghai, 1994-), female, China, undergraduate of Life Sciences, School of Life Sciences, Fudan University, Shanghai, China.

Email: leonardosiqi@sina.com

[2] ***Chuanchao Wang*** (Shandong, 1987-), male, China, Ph.D. Student of Human Biology, School of Life Sciences, Fudan University, Shanghai, China.

Email: CChao.Wang@gmail.com

[3]***Hui Li*** (Shanghai, 1978-), male, China, Professor of Human Biology and Anthropology, School of Life Sciences, Fudan University, Shanghai, China.

Email: LiHui.Fudan@Gmail.com