Genes and Environment, Vol. 36, No. 3 pp. 99–102 (2014)

Review

Homo sapiens under Neutral Evolution

Naruya Saitou[1]

Division of Population Genetics, National Institute of Genetics, Mishima, Japan

Received March 3, 2014; Revised April 26, 2014; Accepted April 28, 2014
J-STAGE Advance published date: May 3, 2014

People often would like to think themselves as special existence. In studies on organismal evolution, unexpectedly many researchers believe in contribution of positive selection for evolution of characters which made us human. However, most of evolution at genomic level is neutral process, and this is also true for evolution at protein level. There is no exception in evolution toward human lineage. I would like to show various examples on this point, including studies of my group, and would like to confirm natural phenomenon that most of evolution which produced *Homo sapiens* was neutral. It should also be noted that majority of this paper is from ''Introduction to Evolutionary Genomics'' written by Saitou (Saitou N. Berlin: Springer; 2014).

Key words: evolutionary genomics, natural selection, human evolution

Neutral Evolution as Default Process of the Genome Changes

It is now established that the majority of mutations fixed during evolution are selectively neutral, as amply demonstrated by Kimura (1) and by Nei (2). Reports of many genome sequencing projects routinely mention neutral evolution in the 21st Century. I thus discuss neutral evolution as one of the basic processes of genome evolution in this paper.

Neutral evolution is characterized by the egalitarian nature of the propagation of selectively neutral mutants. Mutation is the ultimate source of diversity of organisms. If a mutation occurring in some gene modifies gene function, there is a possibility of heterogeneity in terms of number of offsprings. This is the start of natural selection. However, some mutations may not change gene function, and although they are somewhat different from parental type DNA sequences, mutants and parental or wild types are equal in terms of offspring propagation. We meet the egalitarian characteristic of the selectively neutral mutants. If all members of evolutionary units, such as DNA molecules, cells, individuals, or populations, are all equivallent, the frequency change of these types are dominated by random events. It is therefore logical that randomness is the most important factor in organismal evolution.

Our World is Finite

Randomness also comes in when abiotic phenomena are involved in organismal evolution. Earthquakes, volcanic eruptions, continental drifts, meteorite hits, and many other geological and astronomical events are not the outcome of biotic evolution, and they can be considered to be stochastic from organismal point of view.

Before proposal of the neutral theory of evolution by Kimura (3), randomness was not considered as the basic process of evolution. Systematic pressure, particularly natural selection, was believed to play the major role in evolution. This view is applicable if the population size, or the number of individuals in one population, is effectively infinite. However, the earth is finite, and the number of individuals is always finite. Even this whole universe is finite. This finiteness is the basis of the random nature of organismal evolution.

Basic Concept of Natural Selection

The fundamental source of evolution is mutation, or any change of genome sequences. Therefore, natural selection is tightly connected with the effect of mutations. If mutations are highly deleterious, they will soon disappear though natural selection. Before Darwin (4) proposed the possibility of natural selection as a creative power of evolution, natural selection in modern term has been considered as the mechanism to keep status quo, as initially created by the divine power. Elimination of deleterious mutations is now called 'negative' selection after Kimura (1). When the conservative nature of this process is stressed, it is called 'purifying' selection (5). To keep the current genetic entity thorough elimination of deleterious mutations is the core of the ''struggle for existence''.

Most of mutations become extinct simply by chance, and only a small fraction exists for a long evolutionary

[1]Correspondence to: Naruya Saitou, Division of Population Genetics, National Institute of Genetics, 1111 Yata, Mishima 411-8540, Japan. Tel: +81-55-981-6790, Fax: +81-55-981-6789, E-mail: saitounr@nig.ac.jp
doi: org/10.3123/jemsge.2014.010

time. If we focus on these mutations that contribute to evolution for a certain period, say 1000 generations, the majority is selectively neutral, and only a small fraction is kept through natural selection that favors advantageous mutations. Charles Darwin, together with Alfred Russel Wallace, proposed this type of natural selection as the mechanism of evolution in 1858. This process is therefore sometimes called 'Darwinian' selection, but nowadays 'positive' selection, after Kimura (1), is commonly used. Negative and positive selection sandwich the selective neutrality, the dominant situation in evolution.

Natural selection is tightly connected with two concepts: adaptation and fitness. Although 'adaptation' is widely used in many books on evolution, it is often not quantitatively defined, and we should be careful for using this word in the context of natural selection unless clearly defined. This word is also used to describe various characters of many organisms without testing whether those characters were products of positive selection. For example, 'adaptive radiation' has often been used on the evolution of many mammalian lineages as if adaptation caused by positive selection is the main factor for radiation. Recent studies, however, suggest that the geographical isolation caused by the continental drift seems to be the main cause of mammalian radiation. Any evolutionary biologist should have a discreet attitude on the usage of 'adaptation'.

In contrast, 'fitness' is quantitatively defined in population genetics theory. There are two kinds of fitness; absolute and relative. Because the basis of evolution is the change of genetic constituents of organisms, natural selection should be discussed in terms of differential rates of reproduction among genotypes. The absolute fitness is the mean number of offsprings for one particular genotype. If this is larger than 1, the genotype is expected to increase its offspring if the effect of random genetic drift is negligible. Consideration of the absolute fitness is important, for we can directly discuss the temporal change of population size. It should be noted that the absolute fitness may change without mutations, if the environmental condition changes.

We are often interested in the relative success of one genotype to other ones, and consider the relative fitness. The relative fitness of one particular genotype is usually set to be 1, and that of other genotypes are expressed using selection coefficients. Genotypes are identical with alleles in haploid organisms, and the relative fitness for allele i may be written as $1+s_i$, where s_i is selection coefficient for allele i. Let us assume that the relative fitness of allele 0 is 1 ($s_0=0$). If s_i ($i>0$) is positive, allele i is advantageous compared to allele 0 and is positively selected, while allele i is deleterious compared to allele 0 and is negatively selected when s_i is negative. Allele i is selectively neutral with allele 0 when $s_i=0$.

Positive Selection for Ape and Human Genes

We now move to discuss positive selection. Because the author is familiar with studies on primates, lest us first consider positive selection in primates. Hughes and Nei (6,7) showed that both MHC class I and class II genes experienced positive selection, probably of the overdominant type, through the comparison of synonymous and non-synonymous nucleotide substitutions (Dn/Ds test). This test is quite powerful to detect positive selection based on the neutral theory, and many studies were conducted to detect such selection. However, one drawback of this test is that it can detect only a limited type of genes under positive selection, and to obtain a statistically significant result, we need to have many substitutions. Therefore, a protein gene must have many variable amino acids, a pre-condition that may be satisfied in protein-coding genes that are involved in the immune system, such as MHC. If a single amino acid change was responsible in enhancing fitness, however, it would be difficult to detect it though the Dn/Ds test.

The majority of protein-coding genes have lower nonsynonymous substitutions than synonymous ones, and only a small proportion is under positive selection. Although most of the genes were found to have experienced positive selection thorough the Dn/Ds test, the FOXP2 gene studied by Enard *et al.* (8) and Zhang *et al.* (9) did not show significantly higher Dn than Ds.

The FOXP2 gene was initially found in mouse (10), and its highly homologous human ortholog was shown to be responsible for hereditary orofacial dyspraxia associated with dysphasia (11). Interestingly, two nonsynonymous nucleotide substitutions at codon 304 (ACC:Thr⇒AAC:Asn) and at codon 326 (AAT:Asn⇒AGT:Ser) occurred after the human and chimpanzee divergence, and the DNA polymorphism study of modern human populations suggested that the FOXP2 gene region experienced a selective sweep after mutations occurred within modern humans (8,9). However, one of the two amino acid substitutions occurred in the human lineage after the divergence of the human-chimpanzee common ancestor also occurred in the carnivores (9). It is therefore not clear if these amino acid changes were truly responsible for the emergence of language. There was a report that FOXP2 of Neandelthals also had these two amino acid substitutions (12). If this finding is true, two nonsynonymous substitutions occurred in the common ancestor of modern humans and Neandelthals. Coop *et al.* (13) suggested two alternative scenarios for this; low rates of gene flow between modern humans and Neanderthals or contamination of modern human DNA in the putative Neanderthal genome. Mouse FOXP2 gene knockin experiment by Fujita *et al.* (14) and Enard *et al.* (15) both showed some phenotypic differences in vocalization. Konopka *et al.*

(16) reported that the FoxP2 gene is responsible for the human-specific transcriptional regulation of the central nervous system.

Pollard *et al.* (17,18) found a series of short DNA sequences which are highly conserved in vertebrates but show accelerated evolution only in human, and named them HARs (human accelerated regions). Prabhakar *et al.* (19) and Bird *et al.* (20) also conducted similar genome-wide studies. Prabhakar *et al.* (21) found one such sequence from the human genome, HACNS1 which includes 119-bp HAR2, and showed it to act as a limb bud enhancer with enhanced limb enhancer activity specifically in human. This change was caused by 13 human-specific substitutions within that region, and they interpreted that accumulation of these positively selected substitutions created multiple novel transcription factor binding sites (gain-of-function) and that the deposition of those facilitated the human-specific enhanced activity (21).

However, a GC-biased gene conversion may be an alternative explanation for fixation of such mutations causing loss-of-function in a repressor element within HACNS1 (22–24). GC-biased gene conversion is a consequence of DNA double strand break repair between homologous chromosomal regions, and the alleles from one chromosome are converted to the other with a bias of A or T to G or C (25). Neutral or even deleterious alleles could be fixed by GC-biased gene conversions (26). It is possible that the 13 human-specific substitutions were caused by a GC-biased gene conversion and resulted in a disruption of repressor function of the 81 bp region (loss-of-function), which may eventually enhance the activity of human HACNS1.

To evaluate the function of HACNS1, Sumiyama and Saitou (27) performed transgenic mouse assay by using the HACNS1 construct lacking the 13 human-specific substitutions. The deleted construct showed similar enhancer activity to the intact human HACNS1. This result suggests that the function of HACNS1 is not an activating enhancer, but rather a disrupted repressor. If so, loss-of-function in the HACNS1, possibly via a GC-biased gene conversion, not via positive selection, played an important role in human-specific evolution.

Kryukov *et al.* (28) reported that the selective pressure affecting the evolution of regulatory elements in the hominid lineage was significantly relaxed compared with that of the rodent lineage. Keightley *et al.* (29) suggested that regulatory elements in hominids may be diverging at a neutral evolutionary rate. All these studies discussed in this section revealed the difficulty in detecting evidence of positive selection in one lineage. We therefore should be careful for any study which jumped to a conclusion of accelerated evolution without carefully examining an alternative neutral evolution scenario.

Detection of Positive Selection through Genome-wide Searches

Let us now discuss on positive selection on modern humans. Adaptation to a high altitude environment has long been an interesting subject in human genetics. Recently, three independent genome-wide studies (30–32) found that the EPAS1 gene showed the highest differentiation between Han Chinese and Tibetans who are believed to adapted to high altitude. This gene codes a transcription factor HIF-2α involved in the induction of genes regulated by oxygen. The EGLN1 gene, which encodes HIF-prolyl hydroxylase 2, was the second-best differentiated gene. EGLN1 catalyzes the post-translational formation of 4-hydroxyproline in HIF-α proteins.

Moreno-Estrada *et al.* (33) compared $\sim$11,000 human genes with their orthologs in chimpanzee, mouse, rat, and dog, and found 11 genes as showing the signatures of positive selection on the human lineage through a branch-site likelihood method (34). These genes were then analyzed for signatures of recent positive selection using SNP data in modern humans. One SNP every 5–10 kb inside each candidate gene and up to around 30 kb in both upstream and downstream flanking regions, plus additional SNPs around 200 kb in both flanking regions were selected, and a total of 223 SNPs were typed for 39 worldwide populations from the HGDP-CEPH diversity panel (35). They also analyzed 4,814 SNP data distributed along 2 Mb centered on each gene from the HGDP-CEPH panel. Through examination of allele frequency spectrum, population differentiation, and the maintenance of long unbroken haplotypes, they found signals of recent adaptive phenomena in only one gene region. The signal of recent positive selection came from a neighboring gene CD5, which codes a transmembrane receptor expressed in the T-cell surface (33). This careful study suggests that most of positively selected genes in modern humans are involved in the immune system. It is not surprising, for our ancestors have suffered many sorts of infectious diseases, and the human immune system-related genes confronted the battle against bacteria, virus, or parasitic eukaryotes. When we discuss positive selection on modern humans, we should consider natural selection on the interaction with other organisms, before attempting to apply other types of natural selection such as sexual selection.

References

1 Kimura M. The neutral theory of molecular evolution. Cambridge: Cambridge University Press; 1983.
2 Nei M. Molecular evolutionary genetics. New York: Columbia University Press; 1987.
3 Kimura M. Evolutionary rate at the molecular level. Nature. 1968; 217: 624–6.
4 Darwin C. On the origin of species. London: John Mur-

ray; 1859.
5 Lewontin R. The genetic basis of evolutionary change. New York: Columbia University Press; 1974.
6 Hughes A, Nei M. Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. Nature. 1988; 335: 167–70.
7 Hughes A, Nei M. Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection. Proc Natl Acad Sci USA. 1989; 86: 958–62.
8 Enard W, Przeworski M, Fisher SE, Lai CS, Wiebe V, Kitano T, Monaco AP, Paabo S. Molecular evolution of FOXP2, a gene involved in speech and language. Nature. 2002; 418: 869–72.
9 Zhang J, Webb DM, Podlaha O. Accelerated protein evolution and origins of human- specific features. Foxp2 as an example. Genetics. 2002; 162: 1825–35.
10 Shu W, Yang H, Zhang L, Lu MM, Morrisey EE. Characterization of a new subfamily of winged-helix/forkhead (Fox) genes that are expressed in the lung and act as transcriptional repressors. J Biol Chem 2001; 276: 27488–97.
11 Lai CSL, Fisher SE, Hurst JA,Vargha-Khadem F, and Monaco AP. A forkhead-domain gene is mutated in a severe speech and language disorder. Nature. 2001; 413: 519–23.
12 Krause J, Lalueza-Fox C, Orlando L, Enard W, Green RE, et al. The derived FOXP2 variant of modern humans was shared with Neandertals. Curr Biol. 2007; 17: 1908–12.
13 Coop G, Bullaughey K, Luca F, Przeworski M. The timing of selection at the human FOXP2 gene. Mol Biol Evol. 2008; 25: 1257–9.
14 Fujita E, Tanabe Y, Shiota A, Ueda M, Suwa K, et al. Ultrasonic vocalization impairment of Foxp2 (R552H) knockin mice related to speech-language disorder and abnormality of Purkinje cells. Proc Natl Acad Sci USA. 2008; 105: 3117–22.
15 Enard W, Gehre S, Hammerschmidt K, Holter SM, Blass T, et al. A humanized version of Foxp2 affects cortico-basal ganglia circuits in mice. Cell. 2009; 137: 961–71.
16 Konopka G Bomar JM, Winden K, Coppola G, Jonsson ZO, et al. Human-specific transcriptional regulation of CNS development genes by FOXP2. Nature. 2009; 462: 213–7.
17 Pollard KS, Salama SR, Lambert N, Lambot MA, Coppens S, et al. An RNA gene expressed during cortical development evolved rapidly in humans. Nature. 2006; 433: 167–72.
18 Pollard KS, Salama SR, King B, Kern AD, Dreszer T, et al. Forces shaping the fastest evolving regions in the human genome. PLoS Genetics. 2006; 2: e168.
19 Prabhakar A, Noonan JP, Pääbo S, Rubin EM. Accelerated evolution of conserved noncoding sequences in humans. Science. 2006; 314: 786.
20 Bird CP, Stranger BE, Liu M, Thomas DJ, Ingle CE, et al. Fast-evolving noncoding sequences in the human genome. Genome Biol. 2007; 8: R118.
21 Prabhakar S, Visel A, Akiyama JA, Shoukry M, Lewis KD, et al. Human-specific gain of function in a developmental enhancer. Science. 2008; 321: 1346–50.
22 Galtier N, Duret L. Adaptation or biased gene conversion? Extending the null hypothesis of molecular evolution. Trends Genet. 2007; 23: 273–7.
23 Galtier N, Duret L, Glemin S, Ranwez V. GC-biased gene conversion promotes the fixation of deleterious amino acid changes in primates. Trends Genet. 2009; 25: 1–5.
24 Katzman S, Kern AD, Pollard KS, Salama SR, Haussler D. GC-biased evolution near human accelerated regions. PLoS Genet. 2010; 6: e1000960.
25 Strathern JN, Shafer BK, McGill CB. DNA synthesis errors associated with double-strand-break repair. Genetics. 1995; 140: 965- 72.
26 Galtier N, Piganeau G, Mouchiroud D, Duret L. GC-content evolution in mammalian genomes: the biased gene conversion hypothesis. Genetics. 2001; 159: 907–11.
27 Sumiyama K, Saitou N. Loss-of-function mutation in a repressor module of human-specifically activated enhancer HACNS1. Mol Biol Evol. 2011; 28: 3005–7.
28 Kryukov GV, Schmidt S, Sunyaev S. Small fitness effect of mutations in highly conserved non-coding regions. Hum Mol Genet. 2005; 14: 2221–9.
29 Keightley PD, Lercher MJ, Eyre-Walker A. Evidence for widespread degradation of gene control regions in hominid genomes. PLoS Biol. 2005; 3: e42.
30 Beall CM, Cavalleri GL, Deng L, Elston RC, Gao Y, et al. Natural selection 11 on EPAS1 (HIF2alpha) associated with low hemoglobin concentration in Tibetan 12 highlanders. Proc Natl Acad Sci USA. 2011; 107: 11459–64.
31 Yi X, Liang Y, Huerta-Sanchez E, Jin X, Cuo ZX, et al. Sequencing of 50 human exomes reveals adaptation to high altitude. Science. 2010; 329: 75–8.
32 Xu S, Li S, Yang Y, Tan J, Lou H, et al. A Genome-wide search for signals of high altitude adaptation in Tibetans. Mol Biol Evol. 2011; 28: 1003–11.
33 Moreno-Estrada A, Tang K, Sikora M, Marquès-Bonet T, Casals F, et al. Interrogating 11 fast-evolving genes for signatures of recent positive selection in worldwide human populations. Mol Biol Evol. 2009; 26: 2285–97.
34 Zhang J, Nielsen R, Yang Z. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. Mol Biol Evol. 2005; 22: 2472–9.
35 Li JZ, Absher DM, Tang H, Southwick AM, Casto AM, et al. Worldwide human relationships inferred from genome-wide patterns of variation. Science. 2008; 319: 1100–4.