PERSPECTIVE

Nagoya J. Med. Sci. **67**. 59 ~ 64, 2005

EVOLUTION OF HUMANS OUTSIDE THE GENOME*

IZUMI NAKASHIMA

Nagoya University, Nagoya;
International Institute for Advanced Studies, Kyoto;
and Chubu University, Kasugai

ABSTRACT

Complete sequencing of the whole genome of humans has revealed a surprisingly small difference between the genomes of humans and higher primates. I here propose that evolution has occurred in living organisms in two steps, first within the genome and then outside the genome. The first step of evolution is based on creation of new information in the genome (DNA), followed by selection after its vertical transmission to individuals of the next generation through reproduction. The second step includes accumulation of a huge amount of information translated into a "linguistic" code by memory and a natural process of computation that creates new patterns of information through thinking in the brain. The created patterns are selected positively or negatively by applying some criteria as to the usefulness for adaptation of humans to nature. When positively selected, the patterns are horizontally transmitted to other brains of a number of individuals, within and also beyond the generation recursively by the use of linguistic codes. This second step of evolution, in conjunction with the evolution of human language itself, has enabled the *homo sapiens* to attain an enormously high level of cognitive faculty for adaptation of the thought processes to needs in nature at an extraordinarily high speed.

Key Words: Evolution, Genome, Language, Human, *Homo sapiens*

INTRODUCTION

We now know that the number of genes in humans (up to 33,000, and recently reduced to 20,000~25.000) does not greatly differ from that in mice and is only a little over two times more than that in flies.[1,2] The difference between the genomes of humans and chimpanzees, of which sizes are approximately 300,000,000 base pairs (bp), has been reported to be only 1.2%[3] or even as little as 0.6%.[4] Thus, the presumed superiority of humans in terms of intelligence, in comparison with other primates, must be due to a very small difference in specific parts of the genomes[5] or to a significant number of insertions or deletions of DNA in particular species as have been identified by advanced studies.[6]

According to a recent report, *Homo sapiens* (humans) initially developed 160,000 yeas ago in Africa.[7] Certainly, some changes must have occurred in the genes of *Homo sapiens* since their first appearance on the earth. However, if an infant of modern man were brought up in ancient

Izumi Nakashima, M.D., Ph.D., Vice-President, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan
Present address: Executive Superintendent, Chubu University, 1200 Matsumoto-cho, Kasugai, Aichi 487-8501, Japan
Phone: +81.568.51.1111 Fax:+81.568.51.4001 email: inakashi@isc.chubu.ac.jp

*This paper is an English version of the article previously published in Japanese (Nakashima, I.: Gakushikai-Kaiho No. 843, 66–70, 2003).

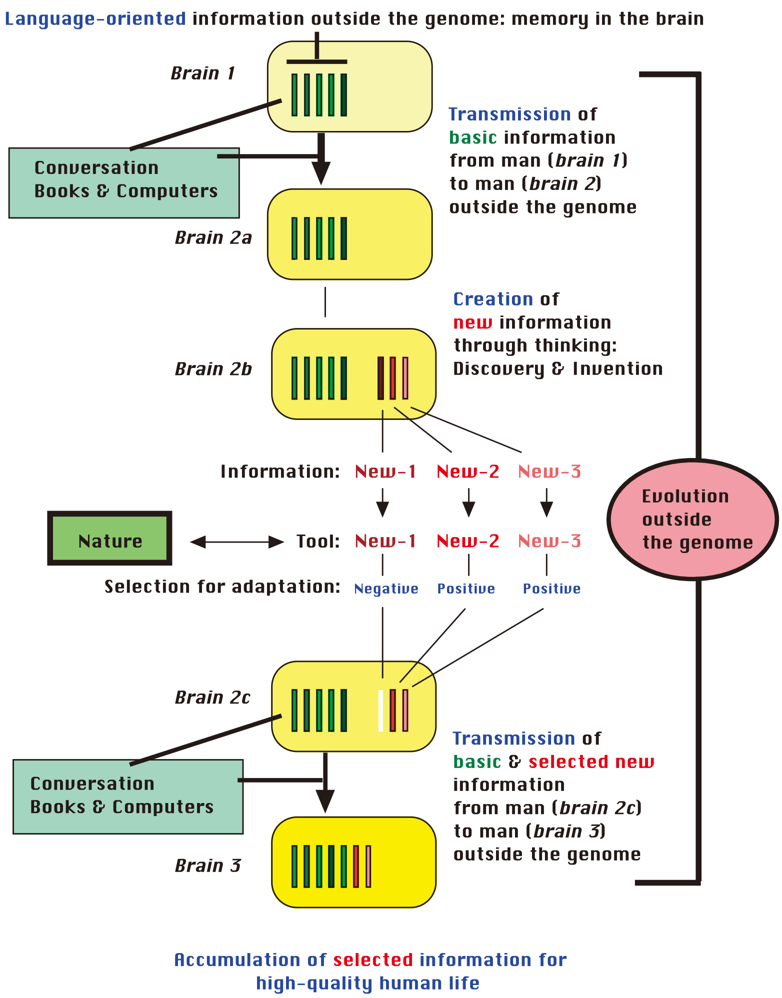

Fig. 1 **The concept of evolution outside the genome**. Two steps of evolution (only the concept of the second step is illustrated in the figure), one on the genome and the other outside the genome, have used distinct mediators of information of digital nature for selection: DNA in the first step and languages (vowels and consonants) and characters (letters in English, for example) in the second stage. Whereas DNA-mediated information is only vertically transmitted from one individual to another through reproduction of the organisms, language-mediated information can be horizontally transmitted from one brain (*brain 1* in Figure 1) to another (*brain 2*) among individuals of the same as well as different generations through conversation and writing/reading. New information arises in the first step through mutation or recombination of genes and is selected positively or negatively when it is transmitted to descendants, and the usefulness of the information-based new phenotype of the organism is tested for adaptation to nature. In the second step, new information, the amount of which is too large to be transmitted by DNA, develops when information that has accumulated in the brain (*brain 2a*) is re-organized through thinking (*brain 2b*). The usefulness of this newly developed information is tested for selection after it is translated for production of a new tool as an extension of the genome-oriented organism (*brain 2c*). The positively selected information for production of a new tool is transmitted to other brains (*brain 3*) and is accumulated for developing cultures in the society of humans where many individuals share new tools that are useful for high-quality human life. Because of its outstanding capacity to create, maintain and transmit a large amount of information, the second step of evolution has resulted in the rapid development of humans from 160,000 years ago, when *Homo sapiens* first appeared at the end of the first step of evolution that lasted for 4,000,000,000 years after the initial birth of living organisms on the earth.

human society, the individual would not become a well-qualified member of modern society. Therefore, in addition, there must have been some mechanisms that can account for the difference between ancient and modern humans in terms of their social and cultural potential.

The FOXP2 gene product was recently suggested to be the target of selection in the human lineage as a critical element for language development.[5,8] Combining these pieces of recently provided information I here propose that there are two steps of evolution for living organisms in general: The first step is classical Darwinism, in which the FOXP2 gene and other yet-unidentified genes that are human-specific evolved. Human-specific genes, which developed at the end of the first step of evolution, almost certainly prepared the unique faculty of the brain to develop the second step of evolution outside the genome, for storage of information in memory and for creation of new information by re-organizing the constantly accumulating information through thinking, namely discovery and invention. These human-specific abilities seem to be largely mediated by language.[5,8] The scheme for the concept of the second step of evolution is illustrated in Figure 1.

DEFINITION OF THE LANGUAGE-ORIETED EVOLUTION SYSTEM

An approximate picture of human natural language, which has been the standard linguistic description by tradition, is to represent speech signals abstractly as linearly concatenated strings of vowels and consonants and then by corresponding text using alphabetical letters. This conventional representation of human language by characters finds its parallel in how we understand a DNA can be represented.[9] Humans interpret phenomena observed in surrounding nature as abstract and discrete patterns of information by language.[10,11] These patterns are re-organized in the process of thinking for discovering principles that underlie observed phenomena and for inventing new tools based on discovered principles. Language-oriented information for creating new tools, which may replicate the genome-based information processing for producing new phenotypes, is subject to selection regarding the usefulness for adaptation of human life to surrounding nature, including their own society. The positively selected information is accumulated in the brain of the individual and his/her peers and not in his/her genome waiting for the manifestation in the next generation. It is also transmitted to other brains of humans of the same or next generation in adjacent or remote areas through communication and learning by use of speech and, in more advanced cultures, text and books/libraries and, most recently, computers. Accumulation of new information and tools results in the creation of new cultures.

TRANSMITTERS OF INFORMATION IN THE LANGUAGE-ORIETED EVOLUTION SYSTEM

There is a big difference between genome-mediated and language-oriented evolution in quality and quantity, and particularly in the speed of progression (Table 1).

The language system abstracts elements and organizational structure from the complex information in spatial images and temporally changeable amorphous patterns that are obtained by concrete experiences received via sensory organs. Then it finds corresponding linguistic expressions. For example, individual words may arise as a discrete pattern (in digital mode, as we might call it) through appropriate arrangements of vowels and consonants (in the case of English alphabetic text, for example, symbolized in 26 letters,) as the primary units of information. This representation in text may be compared to the arrangements of bases (in 4 types) of

Table 1 Comparison of genome-mediated and language-oriented evolution systems

		Genome-mediated evolution	**Language-oriented evolution**
Transmitters of Information (standard number of units carrying information)	1st	DNA: Bases (A, T, C, G) (4)	Vowels/Consonants & Letters (a, b, c, d,....) (26 in English alphabetic text)
	2nd	Amino acids (20)	Syllables (10,000)
	3rd	Proteins (33,000)	Words, Phrases & Sentences (indefinite)
Way of Creation of New Information		Point mutation & Recombination of Genes	Discovery & Invention through Thinking
Frequency of Selection of Information		Single Selection in One Sexual Reproduction	Multiple Selections in One Individual
Site of Accumulation of Information (amount of accumulated information)		Genome (limited)	Brain & Books/Computers (unlimited)
Way of Transmission of Information (frequency of transmission)		Sexual Reproduction (once per reproduction)	Conversation & Writing/Reading (frequent in individuals)

DNA. Linear arrangements of the alphabetic letters, if we consider these to be primary units, for the sake of current discussion, create syllables (about 10,000 types in English) as secondary units, which may be compared with the ordered string of three bases representing individual amino acids (20 types). Linear arrangements of syllables as meaningless phonological units, or correspondingly the codes representing a series of amino acids, may represent meaningful words, possibly corresponding to genes or proteins as biologically meaningful functional units (see Table 1). The formation of larger and more complex units such as phrases and sentences for organization of cognitive concepts in language does not seem to find known parallels in the current hereditary science but may be sought in future studies. Such a linguistic process crucially depends on applications of recursive processes dealing with hierarchically organized abstract structure[10,11] and the corresponding biological process may well be found to use a similar principle of organizing larger functional units as an inherent property of humans or similar advanced forms of biology.

WAY OF TRANSMISSTION OF INFORMATION IN THE LANGUAGE-ORIETED EVOLUTION SYSTEM

Unlike genome-mediated information, the huge amount of language-mediated information is quickly transmitted to distinct individuals of humans across generations as well as within each generation and is accumulated for creation of new information in the brain of each individual. The selection of information can be carried out instantaneously upon the individual's exposure to surrounding stimuli, verbal or nonverbal and from another individual or other natural environments. This information processing takes place continuously throughout the individual's life regardless of the biological reproduction cycle. The mutual transmission, *i.e.*, communication among individuals and groups of individuals by the use of language can be transmitted through linguistic codes, whether by speech or by text, written documents. Language-oriented cognitive processes may be found in other manifestations using other graphic patterns as symbolic codes, some directly representing words or constituent fragments of words (morphemes) in hierarchically organized complex structures.[12)]

WAY OF SELECTION OF INFORMATION IN THE LANGUAGE-ORIETED EVOLUTION SYSTEM

Beginning from primitive tools that are invented by one or many individuals in one generation, tools are tested for usefulness in the species' adaptation to nature. Since the evaluation and selection does not have to wait for physical elimination of relatively unfitted individuals by competition, but can be performed by limited experience of applications of the idea and principle for practical tools, combined with the cognitive interpretations of the test results or experiments, the second-stage evolution is fast and self-feeding. The size, quality, complexity, and the rate of the creation, selection and transmission of new information in the language-oriented system are thus incomparably larger than those in the genome-mediated system. This could be the reason why the new evolution system has allowed humans to evolve to an extraordinarily high level in a surprisingly short period of time. What endows the *homo sapiens* with this key element of explosive evolution? It is language. Is there a principle of potential linguistic systems that is incorporated in any of their hereditary codes? We do not know the answer yet.

ORIGIN OF THE SYSTEM OF EVOLUTION OUTSIDE THE GENOME

The amount of overall information for selection and accumulation has increased during evolution[13)] and probably has become too complex and too large to be handled in the one-dimensional DNA at the end of the first step of evolution for primates at least, probably resulting in progression to the second step.

Transmission of some abilities of individual living organisms to descendents outside the genome occurs in behaviors known as 'instincts' of birds and fishes, which, for example, make seasonal travels. This type of information was probably already too complex to be included in the genome for transmission and was thus likely transmitted from one brain to another brain. However, these 'instincts' do not seem to have changed greatly in time in the history of evolution. On the other hand, chimpanzees and, more recently, orangutans, have been found to create a new culture in one society that is transmitted to individuals in another society.[14)] This might be the prototype of the evolution outside the genome in the primate, which became exceptionally effective in humans, presumably because of the use of language as the means of communication. Even though a language spoken in a community differs considerably from one in another human community to the extent it is not mutually intelligible, recent studies in syntactic theory demonstrate a universal framework with system parameter settings for different languages.[11)] The commonly shared representation framework guarantees that languages are learnable given a rather limited exposure to linguistic examples, particularly for young enough individuals.

ACKNOWLEDGEMENTS

I thank O. Fujimura, International Institute for Advanced Studies (Emeritus Professor of Ohio University), for his critical reading of the manuscript and providing me with helpful comments

and suggestions from the linguistic point of view.

REFERENCES

1) Dennis, C., Gallangher, R. and Campbell, P. (Editors): The human genome. Nature, 409, 813–958 (2001) [International Human Genome Sequencing Consortium. Nature, 409, 860–921 (2001)].
2) Stein, L.D.: Human genome: end of the beginning. Nature, 431, 915–916 (2004) [International Human Genome Sequencing Consortium. Nature, 431, 931–945 (2004)].
3) Chen, F.-C. and Li, W.H.: Genomic differences between humans and other hominoids and the effective population size of the common ancestor of humans and chimpanzees. Am. J. Hum. Genet., 68, 444–456 (2001).
4) Wildman, D.E., Uddin, M., Liu, G., Grossman, L.I. and Goodman, M.: Implications of natural selection in shaping 99.4% nonsynonymous DNA identity between humans and chimpanzees: enlarging genus *Homo*. Proc. Natl. Acad. Sci. U.S.A. 100, 7181–7188 (2003).
5) Pennisi, E.: Genome comparisons hold clues to human evolution. Science, 302, 1876–1877 (2003).
6) Wessenbach, J.: Differences with the relatives. Nature, 429, 353–355 (2004).
7) White, T.D., Asfaw, B., DeGusta, D., Gilbert, H., Richards, G.D., Suwa, G. and Howell, F.C.: Pleistocene *Homo sapiens* from Middle Awash, Ethiopia. Nature. 423, 742–747 (2003).
8) Enard, W., Przeworski, M., Fisher, S.E., Lai, C.S., Wiebe, V., Kitano, T., Monaco, A.P. and Paabo, S.: Molecular evolution of FOXP2, a gene involved in speech and language. Nature, 418, 869–872 (2002).
9) Hood, L. and Galas, D.: The digital code of DNA. Nature, 421, 444–448 (2003).
10) Bever, T. and Montalbetti, M.: Noam's Ark. Science, 22, 1565–1566 (2002).
11) Hauser, D., Chomsky, N. and Fitch, T.: The faculty of language: What is it, who has it, and how did it evolve? Science, 298, 1569–1579 (2002).
12) Fujimura, O.: Graphic specification of Kanji. In K. Takami, A. Kamio and J. Whitman (eds). Syntactic and Functional Explorations: In Honor of Susumu Kuno. Kuroshio, Tokyo. pp. 3–22 (2000).
13) Britten, R.J. and Davidson, E.H.: Gene regulation for higher cells. Science, 165, 349–357 (1969).
14) van Schaik, C.P., Ancrenaz, M., Borgen, G., Galdikas, B., Knott, C.D., Singleton, I., Suzuki, A., Utami, S.S. and Merrill, M.: Orangutan cultures and the evolution of material culture. Science, 299, 102–105 (2003).