Evolution of DNA Structure: Direction, Mechanism, Rate

ALEXANDER L. MAZIN

A.N. Belozersky Laboratory of Bioorganic Chemistry and Molecular Biology, Bldg. "A", Lomonosov State University, Moscow 117234, USSR

Received May 2, 1975; revised May 29, 1976

Summary. On the basis of the results of an analysis of frequencies of pyrimidine oligonucleotides, the degree of pyrimidine clustering of DNA in species from different taxa has been determined. A tendency for an increase in the index of clustering of DNA was revealed in the sequence: invertebrates, fishes, amphibians, reptiles, birds, mammals. A mechanism is postulated, according to which the increase in the degree of clustering of DNA during the evolution may be associated with the accumulation of mutations, Purine => Pyrimidine transversions, resulting in a selective enrichment of one of the chains of DNA with pyrimidines and the other - with purines, i.e. in an increase in the degree of purine-pyrimidine imbalance (asymmetry) of DNA complementary chains. This mechanism of DNA evolution is supported by the presence of positive correlation between the degree of clustering and the degree of the chain asymmetry of natural DNAs, as well as the character of the amino acid substitutions in cytochromes c in different species. The progressive evolution of different groups of organisms on the whole may have been accompanied by an acceleration of the rates of evolution of the DNA structure.

On the basis of the amino acid sequence of cytochromes c in different species the degree of clustering and the degree of the chain asymmetry of the corresponding structural genes of DNA was found to have a general tendency towards an increase in the following order: invertebrates, fishes, amphibians, reptiles, birds, mammals. Thus, evolution of cytochrome c cistron is a vector process based on a selection of mutations which, on the one hand, are neutral to protein, and, on the other hand, result in the sense chain of DNA being enriched with pyrimidines and the nonsense one (and the corresponding mRNA) - with purines. Hence, it is the polynucleotide template rather than protein, that must have been the "object of selection". The frequency of substitutions in cytochromes c cistron

Abbreviations Used: Py - pyrimidine; Pu - purine; H - "heavy", i.e. the pyrimidine rich strand of DNA; L - "light", i.e. the purine rich strand of DNA.
for vertebrates is $1.56 \times 10^{-9}$ per nucleotide per year. It is believed that the evolutionary modification of the DNA structure may be associated with an increase in the interference resistance of the translation, i.e. with selection for codons of highest readout stability.

Key words: Pyrimidine Oligonucleotide Frequencies/Asymmetry of DNA Chains/Evolution of Cytochrome c Cistron

INTRODUCTION

Protein molecules were the first and major area of investigation in molecular evolution. Analysis of divergence of protein sequences reinforced the concept that evolution at the molecular level did not involve natural selection and occurred due to random drift and fixation of neutral mutations (Kimura, 1968; King & Jukes, 1969; Kimura & Ohta, 1971). On the other hand, there is just as indisputable proof of the existence of a selective process during the evolution of proteins. Among other things, abundance of data supports the non-randomness in evolution of protein molecules (Fitch, 1967; Zuckerkandl et al., 1971; Dickerson, 1971 and others). At present there is a tendency for the two points of view to converge. An attempt to reconcile "neutralism" and "selectionism" was recently made by Homquist who put forward the "selective-stochastic" hypothesis (Homquist, 1975).

In my view, neutral amino acid substitutions in proteins which is usually believed to be a random process, may be a result of natural selection operating at the DNA, rather than at the polypeptide level. The absence of data on the nucleotide sequence of DNA together with the persistence of the idea about its conservative role as the preserver and the carrier of genetic information has for a long time inhibited evolutionary studies of these biopolymers. Undoubtedly, it is only a comprehensive approach to the evolution of the genome as a whole, including the major steps of transfer and realization of genetic information in the cell that may throw light on the principles and mechanisms underlying molecular evolution. The present paper is an attempt at a coordinated analysis of the evolution of total DNA and of a protein, cytochrome c.

I. EVOLUTION OF DNA STRUCTURE

At present the regularities of DNA evolution may only be inferred from indirect data, such as the degree of reiterated base sequences and also DNA-DNA hybridization. Some insight into the character of nucleotide sequences of DNA may be gained from an analysis of different oligonucleotide frequencies ob-