

## **Deep insights into the history of hexokinases evolution based on mutational pressure theory**

S.V. Lelevich

*Grodno State Medical University, Belarus, Grodno, Gorkogo, 80, slelevich@yandex.by*

V.V. Khrustalev, E.V. Barkovsky

*Belarussian State Medical University, Belarus, Minsk, Dzerzinskogo, 83, vvkhrustalev@mail.ru, barkovsky@hotmail.ru*

There are four genes coding for four types of hexokinases containing two homologous halves in genomes of vertebrates: hexokinase I (HKI); hexokinase II (HKII), hexokinase III (HKIII) and hexokinase domain containing protein I (HKDCI) [1]. It was shown that N-halves of HKI and HKIII are not catalytically active, unlike their C-halves [1]. N-half of HKDCI also possesses mutations which should lead to the loss of its catalytic function [1]. In contrast, both halves of HKII are able to catalyze phosphorylation of hexoses [1].

Nucleotide sequences of coding regions have been downloaded from Ensembl database and from GenBank. N-halves have been separated from C-halves with the help of REPRO algorithm (<http://www.ibi.vu.nl/programs/reprowww/>). Then they have been aligned by Muscle algorithm included in MEGA 5 program [2]. JTT amino acid evolutionary distances have been calculated between all the N-halves and compared by two-tailed paired t-test with evolutionary distances between all the C-halves for each enzyme.

N-half of HKDCI evolves faster than C-half (average ratio between evolutionary distances is equal to 1.147;  $P = 2.459 \cdot 10^{-15}$ ; number of pairwise distances is equal to 171). N-half of HKIII evolves 1.853 times faster than its C-half ( $P = 4.241 \cdot 10^{-48}$ ;  $N = 120$ ). Although both halves of hexokinase II are catalytically active, evolutionary distances between N-halves of this enzyme are significantly higher ( $P = 6.437 \cdot 10^{-13}$ ;  $N = 231$ ) than evolutionary distances between its C-halves (average ratio is equal to 1.107). Surprisingly, evolutionary distances between catalytically inactive N-halves of hexokinase I are significantly lower ( $P = 2.035 \cdot 10^{-10}$ ;  $N = 190$ ) than evolutionary distances between its catalytically active C-halves

(average ratio is equal to 0.886). This fact is in conflict with widely accepted hypothesis stating that N-halves of all hexokinases evolve faster than C-halves [1].

The most of the sequences encoding hexokinases are under the influence of symmetric mutational GC-pressure [3], since GC-content in their third codon positions (3GC) is higher than 50%. A few sequences with 3GC < 50% were excluded from phylogenetic analyses.

We calculated average GC-content in first and second codon positions in homologous parts of N- and C-halves for each sequence coding for HKI, HKDCI, HKII and HKIII. Average “(1GC + 2GC) / 2” levels for regions coding for N-halves of HKI, HKDCI and HKIII are higher than those for regions coding for their C-halves (P-values are:  $5.752 \cdot 10^{-6}$ ;  $3.766 \cdot 10^{-3}$  and  $5.066 \cdot 10^{-11}$ , respectively). The level of “(1GC + 2GC) / 2” is lower in regions coding for N-halves of HKII than in regions coding for C-halves of HKII (P-value is equal to 0.02118). Since average GC-content in first and second codon positions is a kind of a retrospective index, we may suggest that there was a period of time when negative selection was weaker for C-half of HKII than for N-half. Then amino acid sequence of HKII C-half has been stabilized by negative selection and the rates of amino acid substitutions fixation became lower. There had been a period of the “drift” towards higher GC-content in first and second codon positions in the evolutionary history of a region encoding N-half of HKI. Then amino acid sequence of HKI N-half has been stabilized by negative selection, probably, after the acquirement of new functions.

1. D.M. Irwin, H. Tan, (2008) Molecular evolution of the vertebrate hexokinase gene family: Identification of a conserved fifth vertebrate hexokinase gene, *Comp. Biochem. Physiol., Part D* **3**: 96-107.
2. K. Tamura, D. Peterson, N. Peterson, G. Stecher, M. Nei, S. Kumar, (2011) MEGA5: Molecular Evolutionary Genetics Analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods, *Mol. Biol. Evol.*, **28**: 2731-2739.
3. N. Sueoka (1998) Directional mutation pressure and neutral molecular evolution, *Proc. Natl. Acad. Sci. USA*, **85**: 2653–2657.