INTRON LENGTH DEPENDS ON PHASES OF SURROUNDING INTRONS

Roytberg M.A.*[1], Tsitovich I.I.[2], Astakhova T.V. [1]
Institute of Mathematical Problems of Biology, RAS, Moscow Region, Russia;
Institute of Information Transmission Problems, RAS, Moscow, Russia
e-mail: mroytberg@impb.psn.ru
** Corresponding author*

Key words: *exon, intron, intron phase*

Motivation and Aim: Phase of intron is a remainder of the total length of preceding exons divided by three. Type of intron is a triple xyz, where y is a phase of the intron under consideration, x is a phase of the preceding intron (0 for the first intron),, z is a phase of the subsequent intron (intron 0 for the last intron). The relation between lengths and phases of introns was studied in [1]. The aim of the work was to study relation between the length of the intron of its type.

Methods and Algorithms: We analyzed a set of insect and vertebrate introns, 17 organisms and 2036516 introns in total [ftp.ncbi.nih.gov/genomes]. The range of possible lengths of introns were divided into five intervals, the boundaries of the intervals were different for vertebrates and insects, because of difference in the avearage lengths of the introns. Given sample of introns and partition of all possible intron lengths, an R-value is a ratio of number of introns with lengths from the maximal length interval to the number of introns with lengths from the minimal length interval. For each organism examined the R-value was calculated for two samples of introns: (1) all introns (2) start introns

Results: For each species the R-values for different types were ordered by decrement, the results are given in the table. For all mammals and birds the best values are achieved for same intron types; the "R-value" column shows average values for mammals ad birds resp.

	First MAX		Second MAX		Third MAX		Fourth MAX	
Organizms	Type	R-value	Type	R-value	Type	R-value	Type	R-value
Mammals	0 1 1	34.41%	0 2 1	31.29%	0 0 1	29.56%	2 1 1	25.03%
Birds	0 1 1	17.46%	0 2 1	13.45%	0 0 1	13.03%	2 1 1	11.26%
Xenopus	0 1 1	21.91%	0 2 1	15.07%	0 0 1	14.17%	2 1 1	11.41%
Danio rerio	0 1 1	14.91%	2 1 1	11.69%	0 2 1	11.68%	0 0 1	9.85%
Lizard	0 1 1	52.45%	0 2 1	42.17%	0 0 1	37.57%	2 1 1	30.56%
Insects	0 1 1	12.48%	0 0 1	8.26%	0 2 1	7.74%	1 1 1	7.57%

For the start introns in all species except mammals and xenopus the maxomal R-value corresponds to the type 0 1 1; for mammals and xenopus it corresponds to 0 2 1.

Conclusion: Average intron length depends not only its phase but also of phases of intron neighbors. The maxima of R-values was observed for the introns of type X1Y and XY1.

Availability: http://lpm.org.ru/~mroytberg/intron_phase

References:

1. T. Astakhova, I. Tsitovich, M. Roytberg. Proceedings of the International Moscow conference on computational molecular biology MCCMB'11 Moscow, Russia, July 21-24, 2011. c.321-322.