The influence of intron length on the intron phase distribution

Tatiana Astakhova[1], Ivan Tcitovich[2], Mikhail Roytberg[1]

[1]*Institute of Mathematical Problems of Biology, 4, Institutskaja str., 142290, Pushchino, Moscow Region, Russia, mroytberg@impb.psn.ru*
[2]*Institute of Information Transmission Problems, RAS, Moscow, Russia*

We have studied the dependence of intron phase distribution on intron length (intron phase is the number of nucleotides in an incomplete codon preceding the intron) in various species.

Species	Pos	40-200		5000+		T	p
		N1	F1	N2	F2		
Apis mellifera	1	1060	0,34	249	0,42	2,53	5,8E-03
	2	1363	0,32	147	0,41	2,36	9,1E-03
	3	1241	0,32	94	0,40	1,66	4,8E-02
	4	1025	0,30	68	0,44	2,46	7,0E-03
Drosophila. Melano-gaster	1	3222	0,33	761	0,45	6,17	3,9E-10
	2	2726	0,29	216	0,40	3,18	7,3E-04
	3	2143	0,29	145	0,39	2,48	6,7E-03
	4	1561	0,29	142	0,49	5,15	1,5E-07
Anopheles gambiae	1	785	0,31	201	0,39	2,09	1,9E-02
	2	530	0,25	117	0,36	2,30	1,1E-02
	3	388	0,27	67	0,30	0,53	3,0E-01
	4	268	0,28	60	0,37	1,31	9,6E-02
Nasonia vitripennis	1	1514	0,36	268	0,44	2,38	8,8E-03
	2	1630	0,31	131	0,36	1,11	1,3E-01
	3	1393	0,29	91	0,35	1,21	1,1E-01
	4	1253	0,31	63	0,35	0,62	2,7E-01
Tribolium castaneum	1	1978	0,35	314	0,46	3,67	1,2E-04
	2	1536	0,31	147	0,36	1,37	8,5E-02
	3	1225	0,31	104	0,40	2,03	2,1E-02
	4	946	0,30	59	0,37	1,19	1,2E-01

Table 1. Legend: Pos - position of an intron in the gene, N - number of introns of the specified length in the 1 st phase; F1 - the frequency of introns in the 1 st phase among all introns of the specified length; T - value of Student's t test, p – confidential probability.

It was shown that for insects the frequency of introns in phase 1 increases for "abnormally long" introns (length ≥ 5000, most introns are about 80 bp), see Tab.1. In [1] it was shown that for different taxa the frequency of introns in phase 1 depends of the position of the intron in the gene. Therefore we separately considered introns in different positions. The effect holds for all 4 positions of introns holds, but most strong is for the first intron.

This effect is most significant for D. melanogaster, however, it is observed in all considered insects. The same effect can be observed also for “medium length” introns (200-1000 and 1000-5000 bp), see Fig.1.

For other taxa situation differs from the insects' one. The effect was not observed (*Ciona_intestinalis)* or was very weak (*Cearnohabditis elegans* and *Xenopus Silurana tropicalis).* In genome *Hydra magnipapillata* one can observe the opposite effect (the frequency of introns in phase 1 decreases with icreasing of introns length), in genomes, *M_musculus, Macaca mulatta, Homo sapiens* and plants (*Populus trichocarpa, Arabidopsis thaliana , Oryza_sativa, Vitis vinifera)* this opposite effect can be noted only for the first introns. Detailed information can be found at http://lpm.org.ru/~mroytberg/Introns_phases.

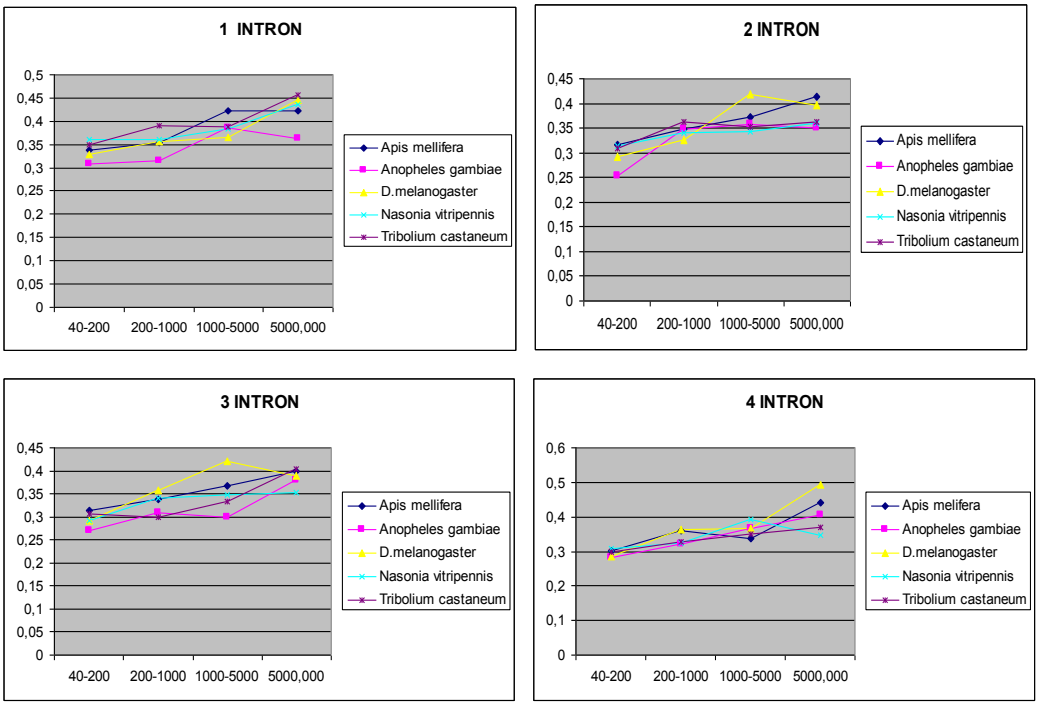

Fig.1. Frequency of phase 1 introns for various species, intron lengths and intron positions

1. A. Ruvinsky, W. Ward A Gradient in the Distribution of Introns in Eukaryotic Genes. (2006) J Mol Evol 63:136–141.