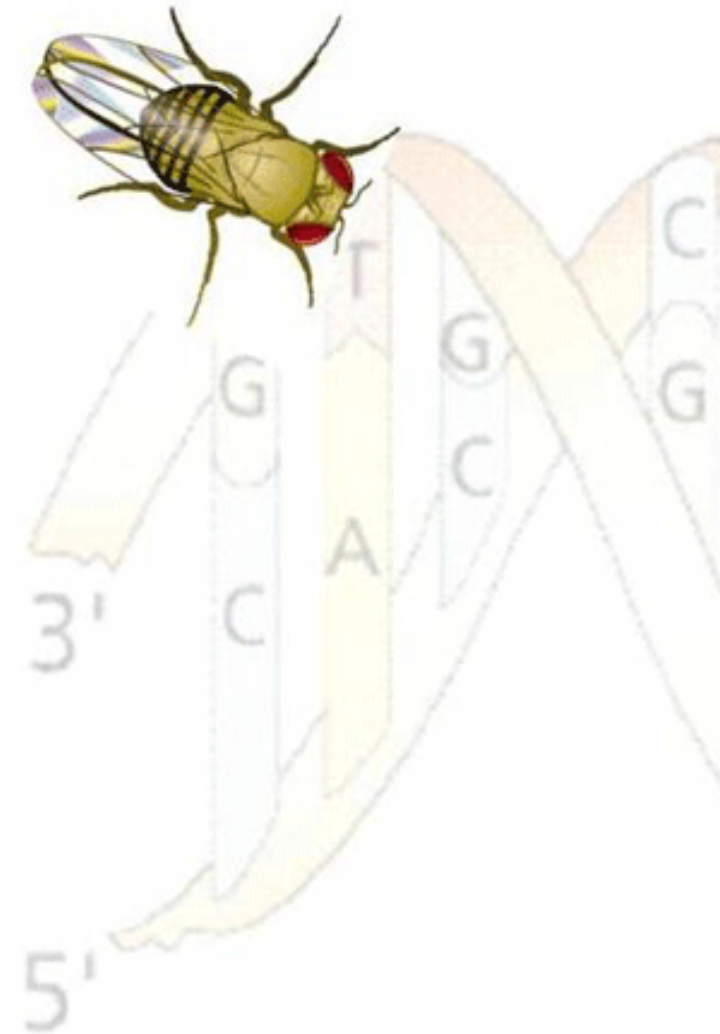


UABDIVULGA

BARCELONA RECERCA I INNOVACIÓ

07/2007

Intron content influences protein variation



Variation in DNA sequence is the basis of natural selection, and hence a study of the variation patterns in the gene sequences of a species is essential for the understanding of the mechanisms of evolution at molecular level. Introns are DNA sequences that form part of the genes

but do not codify proteins and previously were considered to be non-functional. The genetic study of the vinegar fly has shown that the degree of protein variation is inversely related to intron size.

Variation in the DNA sequence of a species is the "raw material" that natural selection works on, and therefore the study of the patterns of this variation in the gene sequences of a species is essential to our understanding of the mechanisms of evolution at a molecular level. Genes that are translated into proteins consist mainly of sequences that are translated into a protein (exons) and non-codifying sequences (introns). Introns were considered for many years as being non-functional; however at present the comparative analysis of complete genomes has shown that a great deal of the non-codifying part of the DNA of proteins is highly conserved between species, indicating some type of functionality of the same.

The Genome Biocomputer and Evolution research group in the Department of Genetics and Microbiology at the UAB, led by doctors Antonio Barbadilla and Alfredo Ruíz, have carried out an analysis using data on DNA variation in the model species used in genetic studies, *Drosophila melanogaster*, in which it has been shown that the degree of protein variation is inversely related to intron size. Long introns contain a greater amount of conserved sequence than short introns and the presence of these conserved sequences explains the relationship between intron size and protein variation.

There is a great deal of evidence that conserved non-codifying sequences function as regulators of a gene's expression. This has been confirmed by the set of data for *D. melanogaster*, given that the presence of these sequences in the genes is related to the complexity of the pattern of expression of a protein. A protein that has a complex pattern of expression is one that is expressed in various different tissues and stages of development, and consequently is more limited by selection, given that any variation in this protein could be prejudicial for the carrier organism. Consequently, genes that codify by means of proteins with complex patterns of expression have long conserved introns that contain the information necessary for the protein's expression in tissue and to determine the exact moment of development.

These results show that selection acting at the level of protein sequences and at the level of sequences that regulate protein expression is coupled; hence mutations in either of the two classes of sequences may be important both for the survival of the carrier organisms and for the evolution of the species.

Natalia Petit

Natalia.Petit@uab.cat

References

Petit, N; Casillas, S; Ruiz, A; Barbadilla, A. "Protein Polymorphism Is Negatively Correlated with Conservation of Intronic Sequences and Complexity of Expression Patterns in *Drosophila melanogaster*" *Journal of Molecular Evolution* Vol. 64, No. 5, pp 511-518.

[View low-bandwidth version](#)