

INTRONS ARE NOT GARBAGE

LOS INTRONES NO SON BASURA

Jordano-Barea, D.

Senior Professor of Applied Biology and Systems Analyst. University of Cordoba. Spain.
E-mail: djordano@lucano.uco.es

ADDITIONAL KEYWORDS

Genetics. Genome. Informatics. Bioinformatics.
Oriented-to-Objects Programming.

PALABRAS CLAVE ADICIONALES

Genética. Genoma. Informática. Bioinformática.
Programación orientada a objetos.

Many scientists believe that more than 95 percent of the pairs of nucleotides contained in genomes are garbage. My long experience in Bioinformatics and Cell Biology refuses the idea that a system so highly complex that need millions of sentences to simulate in real time the simplified multiprocesses of replication, auto-poiesis, growth, homeostasis, adaptation to the environment, enzymatic metabolism, etc., can be accomplished with 5 percent of information, dragging a 95 percent of informatics' burden.

This idea contradicts the general principle of parsimony usually accepted in modern Biology. Furthermore, some viroids are a tape of a few hundred of nucleotides, without genes, and nevertheless contains an efficient project to invade and enslave a host cell and a specific virus. A statistical analysis of correlation on a sample of 8 human proteins (data from Internet, NCBI) gives the results shown in **table I**.

The high correlation between size of gene, number of introns, and

Table I. Correlations gen size-number of introns, gen size-mRNA size and mRNA size-number of introns. (Correlaciones entre el tamaño de los genes y el número de intrones, el tamaño de los genes y el de mRNA y el tamaño de mRNA y el número de intrones).

Variables	Coefficient	Z-Value	P-Value	Lower p.100	Upper p.100
Gen size-size of mRNA	0.899	3.278	< 0.0010	0.530	0.982
Gen size-number of introns	0.941	3.913	< 0.0001	0.703	0.990
mRNA size-number of introns	0.929	3.698	< 0.0002	0.651	0.987

complexity of protein is clearly seen in **figure 1**.

The increments of the variables *size of genes* and *number of introns* are exponential functions that have highly significant correlations. This means

that the more complex a protein is the more long is their gene and the greater the number of introns needed for their biosynthesis. A genome is an informatics project, and a budget too. Everybody knows that the personal

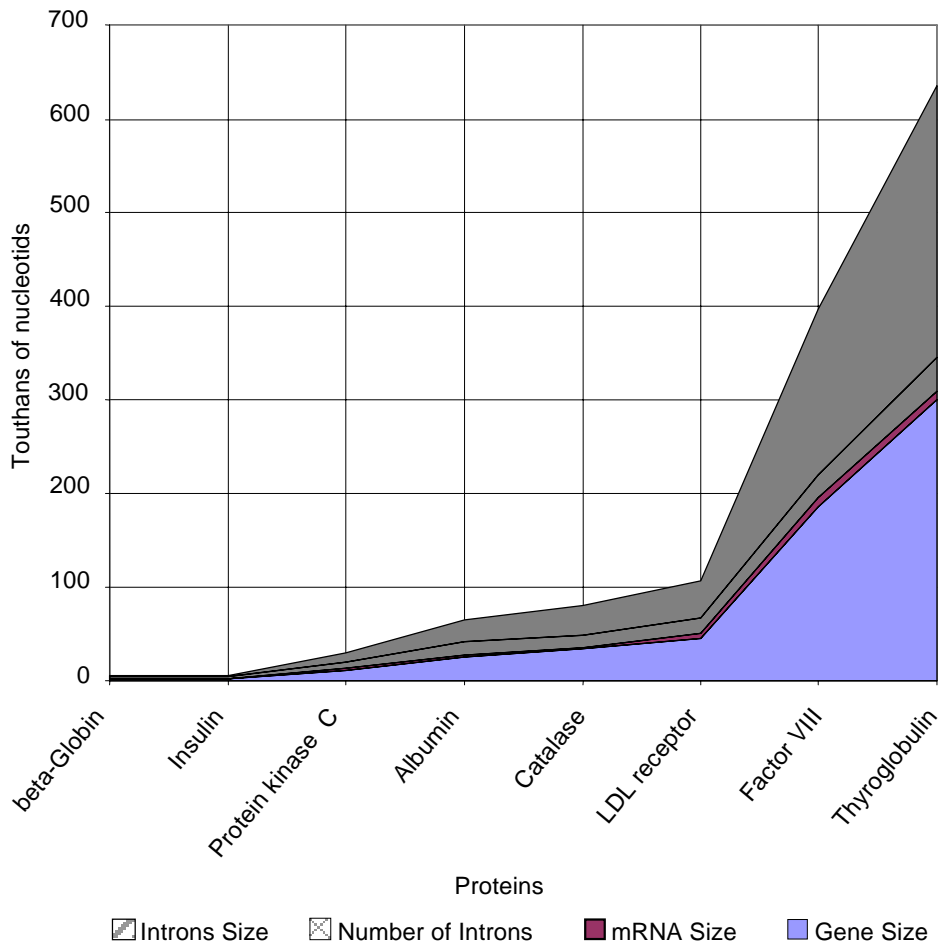


Figure 1. The number of introns (middle) grows exponentially as the size of genes (bottom) goes up. Length of mRNAs (upper) is slightly shorter than size of genes (bottom). (El número de introns (en medio) crece exponencialmente cuando el tamaño de los genes (abajo) aumenta. La longitud de los mRNA (arriba) es ligeramente más corto que el tamaño de los genes (abajo).

INTRONS ARE NOT GARBAGE

budget of a family can be written in a few pages whilst the State budget needs many big books. The *garbage hypothesis* doesn't explain this obvious fact; Bioinformatics and common sense do.

That's why Bioinformatics and the Oriented-to-Object Programming are very promising tools whereas the garbage hypothesis is an appalling mistake.

BIBLIOGRAFÍA

- Adleman, L.M. On constructing a molecular computer. Draft. L.M. Adleman. Univ. Southern California. Los Angeles. CA 90089. USA.
- Anónimo. 2000. Genetic and Evolutionary Computation Conference (GECCO-2000/GP-2000). July 8-12, Las Vegas.
- Framiñán Torres, J.M. 1997. Java al día en una hora. Ed. Anaya Multimedia. Madrid.
Internet : //www.ncbi.nlm.nih.gov/
- Ji, S. 1999. The Linguistics of DNA: Words, Sentences, Grammar, Phonetics, and Semantics. *Apud Molecular Strategies in Biological Evolution*, 870: 411-417.
- Jordano-Barea, D. 1966. Ordenadores de tamaño microscópico. Reto de la Biología molecular a la Electrónica. *Anal. Univ. Hispalense*, 26: 99-110.
- Jordano-Barea, D. 1967. Aportaciones a la teoría del kybernón: las bandas cromosómicas alternantes, posibles zonas declarativas y operativas del programa de ordenación electrónica del trabajo celular. IV Jornadas de Genética Luso-Españolas. Oeiras (Portugal). *Portugaliae Acta Biológica*, 1 Serie A; 10: 162-167.
- Jordano-Barea, D. 1979. Genetics and Bioinformatics. Proceedings of the Symposium on Biology and Ethics; páginas 351-357. UNESCO-CSIC. Madrid (10-14 oct.).

Recibido: 27-3-01. Aceptado: 6-4-01.

Archivos de zootecnia vol. 50, núm. 191, p. 409.