Scientists discover new genetic sub-code

April 16 2010

In a multidisciplinary approach, Professor Yves Barral, from the Biology Department at ETH Zurich and the computer scientists Dr. Gina Cannarozzi and Professor Gaston Gonnet, from the Computer Science Department of ETH Zurich and the SIB Swiss Institute of Bioinformatics, joined forces to chase possible sub-codes in genomic information. The study, which will be published in today's issue of the journal *Cell*, led to the identification of novel sequence biases and their role in the control of genomic expression.

Each cell of an organism contains a copy of its genome, which is a sequence of deoxyribo nucleotides, also called DNA. The cell is able to translate some of the coding sequences into different proteins, which are necessary for an organism's growth, the repair of some tissues and the provision of energy. For this translation work, the cell follows a decoding procedure provided by the "genetic code", which tells what protein is made from a given sequence. The genetic code has been known since the early 1960's.

The researchers from ETH and SIB now identified a new sub-code that determines at which rate given products must be made by the cell. This information has several interesting implications. First, it provides novel insights into how the decoding machinery works. Secondly, and more pragmatically, it makes possible to read information about gene expression rates directly from genomic sequences, whereas up to now, this information could only be obtained through laborious and expensive experimental approaches, such as microarrays.

"A cell must respond very quickly to injuries such as DNA damage and to potent poisons such as arsenic. The new sub-code enables us to know which genes are turned-on quickly after these insults and which are best expressed slowly. One benefit of this study is that we now can get this information using only analysis of the coding sequence", said Dr. Gina Cannarozzi.

Additionally, the new sub-code provides insight into cellular processes at the molecular level. In every living cell, the translation allowing the production of proteins takes place at specialised factories, the ribosomes. The discovery of this novel sub-code will therefore also provide more information about the functioning of these ribosomes. Indeed, all the data gathered up to now indicate that these factories recycle their own components, the tRNAs, to optimize the speed of protein synthesis. This discovery of a new way to regulate translation could potentially be exploited to more efficiently produce therapeutic agents and research reagents. For example, many therapeutic agents, such as insulin, are produced by expressing a protein in a foreign host such as *E. coli* or *S. cerevisiae*. The new sub-code can be now used to rewrite the information such as to optimize in a much more rational manner the amount of product delivered by the foreign host.

Provided by Swiss Institute of Bioinoformatics