Backtracking on DNA

June 23 2009

(PhysOrg.com) -- Accuracy is essential for life, so in converting the information stored in DNA into a form in which it can be used, a high level of precision is required. Dr Tanniemola Liverpool from the Department of Mathematics, working with colleagues from the University of Leeds, has developed a mathematical model for how the required accuracy is achieved.

A gene is the basic physical and functional unit of heredity. Genes are made up of DNA, which is the hereditary material in humans and almost all other organisms. Nearly every cell in a person's body has the same DNA, and the information in DNA is stored as a code made up of four chemical bases: adenine (A), guanine (G), cytosine (C), and thymine (T). Human DNA consists of about three billion bases, and more than 99 percent of those bases are the same in all people. The order, or sequence, of these bases determines the information available for building and

maintaining an organism, similar to the way in which letters of the alphabet appear in a certain order to form words and sentences.

DNA bases pair up with each other, A with T and C with G, to form units called base pairs. Each base is also attached to a sugar molecule and a phosphate molecule. Together, a base, sugar, and phosphate are called a nucleotide. Nucleotides are arranged in two long strands that form a spiral called a double helix. The structure of the double helix is like a ladder, with the base pairs forming the ladder's rungs and the sugar and phosphate molecules forming the vertical sidepieces of the ladder.

An important property of DNA is that it can replicate, or make copies of itself. Each strand of DNA in the double helix can serve as a blueprint for duplicating the sequence of bases. This is critical when cells divide because each new cell needs to have an exact copy of the DNA present in the old cell. The journey from gene to protein is complex and consists of two major steps: transcription and translation, which together are known as gene expression.

During the process of transcription, the information stored in a gene's DNA is transferred to a similar molecule called RNA in the cell nucleus. During transcription errors occasionally occur, and these can lead to defects in the protein being manufactured. In fact, an error of only 1 in 100,000 nucleotides is enough to give rise to proteins that do not function, which would lead ultimately to cell death.

The process of transcription is carried out by specialized enzymes known as RNA polymerases (RNAP) that move along the DNA, base by base. However, due to thermal fluctuations within the cell, the chemical reactions involved in transcription don't always follow the most likely (minimum energy) path and consequently there is a probability of around 1 in a 1,000 that a base-pair is incorrectly transcribed. Therefore, in order for cells to maintain the high level of accuracy required for life,

they must have a mechanism for dealing with errors. Liverpool and colleagues have developed a theoretical model for how the required accuracy is achieved.

In the early 70s it was pointed out that in order to keep track of the information required to proof read (ie, check for and correct errors), the cell had to have processes that dissipate (waste) energy, otherwise they would contradict the second law of thermodynamics which states that unless energy is supplied, disorder (ie, lack of information) tends to increase. Experiments have suggested a number of possible mechanisms by which this might occur, but it is difficult to figure out which, if any, is the primary one, so exactly how these processes occur remains a mystery.

Liverpool's paper, published today in Physical Review Letters, develops a mathematical model for proofreading in the transcription process, based on the fact that the RNAP does not move only in one direction along the DNA, but often makes random backward excursions as it transcribes the gene. The model shows how these 'backtracks' can improve the accuracy of transcription, and predicts the dependence of the probability of finding errors on the backtracking dynamics.

The results of the model suggest future experiments which can be used to discriminate between the different possible mechanisms. They should also shed light on error correction in other biological processes such as the translation of RNA to protein.

More information: Backtracking and Proofreading in DNA Transcription. Margaritis Voliotis, Netta Cohen, Carmen Molina-París, and Tanniemola B. Liverpool. *Physical Review Letters*, 26 June 2009, Volume 102, Number 2. Published online ahead of print 22 June 2009.

Provided by University of Bristol (news : web)