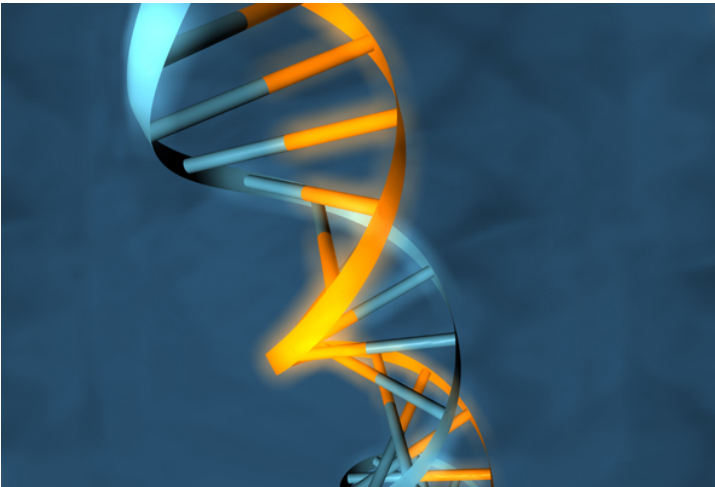


Study on efficient way to crunch sequencing data

30 May 2012 | News | By BioSpectrum Bureau

Study focuses on technique to crunch genome sequencing data



Singapore: Hitachi Asia and Data Storage Institute (DSI), a research institute of the Agency for Science, Technology and Research (A*STAR), are devising a data compression technique to tackle the increasing volume of genome sequencing data generated by the healthcare and biomedical industry. As the volume of such data has been forecasted to double annually, the collaborators aim to develop a more efficient data storage technology that will compress genome sequencing data more effectively than existing methods.

This is an extension of an earlier partnership, where Hitachi and DSI researchers discovered the pattern of typical genome data transactions that would enable current storage systems to function optimally.

Genome sequencing is a data intensive process and high-powered machines are required to decipher the order of deoxyribonucleic acid (commonly known as DNA) nucleotide bases - Adenine (A), Cytosine (C), Guanine (G), and Thymine (T) - that consist within a DNA molecule. A human genome of an individual contains over three billion of these genetic letters and occupies up to 725 MB of uncompressed data.

Scientists and medical practitioners rely on genome sequence to decode the string of letters and gain a clearer understanding of the human anatomy, how genes interact and affect growth and development of an organism.

In a bid to address the current computational and scalability limitations, DSI researchers were commissioned to study how genome sequencing data is optimised by researchers from Genome Institute of Singapore (GIS), another A*STAR research institute.

"By raising compression capacity, we can envision smaller genome sequencing facilities to handle petabytes of data in a year compared to current terabytes levels which are mostly restricted to large genome sequencing centres due to storage limitations. DSI will continue to play a pivotal role in enabling new storage technologies for the biomedical research and healthcare industry to accelerate research findings and discoveries," said Dr Pantelis Alexopoulos, DSI's executive director.