



DATE: Wednesday, July 20, 2006
CONTACT: Yivsam Azgad, Tel: 972-8-934-3856 / 52
EMAIL: Yivsam.azgad@weizmann.ac.il / news@weizmann.ac.il

SCIENTISTS DISCOVER A GENETIC CODE FOR ORGANIZING DNA WITHIN THE NUCLEUS

DNA – the long, thin molecule that carries our hereditary material – is compressed around protein scaffolding in the cell nucleus into tiny spheres called nucleosomes. The bead-like nucleosomes are strung along the entire chromosome, which is itself folded and packaged to fit into the nucleus. What determines how, when and where a nucleosome will be positioned along the DNA sequence? Dr. Eran Segal and research student Yair Field of the Computer Science and Applied Mathematics Department at the Weizmann Institute of Science have succeeded, together with colleagues from Northwestern University in Chicago, to crack the genetic code that sets the rules for where on the DNA strand the nucleosomes will be situated. Their findings appeared today in *Nature*.

The precise location of the nucleosomes along the DNA is known to play an important role in the cell's day to day function, since access to DNA wrapped in a nucleosome is blocked for many proteins, including those responsible for some of life's most basic processes. Among these barred proteins are factors that initiate DNA replication, transcription (the transfer of genetic information from DNA to RNA) and DNA repair. Thus, the positioning of nucleosomes defines the segments in which these processes can and can't take place. These limitations are considerable: Most of the DNA is packaged into nucleosomes. A single nucleosome contains about 150 genetic bases (the "letters" that make up a genetic sequence) while the free area between neighboring nucleosomes is only about 20 bases long. It is in these nucleosome-free regions that processes such as transcription can be initiated.

For many years, scientists have been unable to agree whether the placement of nucleosomes in live cells is controlled by the genetic sequence itself. Segal and his colleagues managed to prove that the DNA sequence indeed encodes "zoning" information on where to place nucleosomes. They also characterized this code and then, using the DNA sequence alone, were able to accurately predict a large number of nucleosome positions in yeast cells.

Segal and his colleagues accomplished this by examining around 200 different nucleosome sites on the DNA and asking whether their sequences have something in common. Mathematical analysis revealed similarities between the nucleosome-bound sequences and eventually uncovered a specific "code word." This "code word" consists of a periodic signal that appears every 10 bases on the sequence. The regular repetition of this signal helps the DNA segment to bend sharply into the spherical shape required to form a nucleosome. To identify this nucleosome positioning code, the research team used probabilistic models to characterize the sequences bound by nucleosomes, and they then developed a computer algorithm to predict the encoded organization of nucleosomes along an entire chromosome.

The team's findings provided insight into another mystery that has long been puzzling molecular biologists: How do cells direct transcription factors to their intended sites on the DNA, as opposed to the many similar, but functionally irrelevant sites along the genomic sequence? The short binding sites themselves do not contain enough information for the transcription factors to discern between them. The scientists showed that basic information on the functional relevance of a binding site is at least partially encoded in the nucleosome positioning code: The intended sites are found in nucleosome-free segments, thereby allowing them to be accessed by the various transcription factors. In contrast, spurious binding sites with identical structures that could potentially sidetrack transcription factors are conveniently situated in segments that form nucleosomes, and are thus mostly inaccessible.

Since the proteins that form the core of the nucleosome are among the most evolutionarily conserved in nature, the scientists believe the genetic code they identified should also be conserved in many organisms, including humans. Several diseases, such as cancer, are typically accompanied or caused by mutations in the DNA and the way it organizes into chromosomes. Such mutational processes may be influenced by the relative accessibility of the DNA to various proteins and by the organization of the DNA in the cell nucleus. Therefore, the scientists believe that the nucleosome positioning code they discovered may aid scientists in the future in understanding the mechanisms underlying many diseases.

Dr. Eran Segal's research is supported by the Willner Family Leadership Institute for the Weizmann Institute of Science, the Arie and Ida Crown Memorial Charitable Fund and the Estelle Funk Foundation.

The Weizmann Institute of Science in Rehovot, Israel, is one of the world's top-ranking multidisciplinary research institutions. Noted for its wide-ranging exploration of the natural and exact sciences, the Institute is home to 2,500 scientists, students, technicians and supporting staff. Institute research efforts include the search for new ways of fighting disease and hunger, examining leading questions in mathematics and computer science, probing the physics of matter and the universe, creating novel materials and developing new strategies for protecting the environment.

Weizmann Institute news releases are posted on the World Wide Web at <http://wis-wander.weizmann.ac.il>, and are also available at <http://www.eurekalert.org>.