Deep sequencing study reveals new insights into human transcriptome

July 8 2008

In a collaborative project scientists from the Max-Planck-Institute for Molecular Genetics in Berlin (MPI MolGen), Germany and Genomatix with a business in Munich, Germany and Ann Arbor, MI, USA, applied next generation sequencing and analysis methods to generate an unprecedented view at the human transcriptome.

Deep sequencing of transcripts from two human cell lines revealed so far unrecognized complexity and variability of the human transcriptome. They found that 34% of the polyadenylated transcriptome mapped to so far non-annotated genomic regions. Obviously a large number of novel gene candidates are active in the cell lines under study.

In addition, a global survey of mRNA splicing events identified 94,241 splice junctions, of which 4,096 are novel, and showed that exon skipping is the most prevalent form of alternative splicing.

Details are presented in the Science report of Sultan et al. "A Global View of Gene Activity and Alternative Splicing by Deep Sequencing of the Human Transcriptome", published online at *Science Express*.

Dr. Marie-Laure Yaspo, Group Leader at the MPI MolGen and head scientist of the study states: " Deep sequencing allows for the first time to explore directly the complexity and dynamics of the human transcriptome with a reasonable effort. This will lead to a new picture of the mammalian genome annotation far beyond the current state of the art. We provide here global features of alternative splicing events in

human cell lines. Such a comparison of within-cell and between-cell alternative splicing events, combined with the simultaneous analysis of gene expression has never been presented before. It becomes clear that the so far available methods only delivered a part of the transcriptional landscape of mammalian cells, especially if gene regulation analysis is considered"

Dr. Martin Seifert, Vice President Business Development and Consulting at Genomatix says:

"The main biological impact is the observation of a new dimension in complexity and variability. Based on the method we could find a significant number of new transcriptional units and splice variants. Our analyses clearly show that transcription is a highly dynamic and variable process. We learned a lot by having access to such high quality data and co-developed necessary new analysis strategies with the MPI MolGen. Especially users of our brand new Genomatix Genome Analyzer will benefit from our experiences along the project, since they have access to all developed strategies."

Annotation and data visualization is publicly available at www.genomatix.de/MPI.html .

Source: Genomatix Software