GENETIC VARIATION

154

Potential Susceptibility Variants for Obesity and Type 2 Diabetes in the TALLYHO Mouse**James Denvir¹**, D. Primerano¹, G. Boskovic, J. Fan¹, J. H. Kim²¹*Department of Biochemistry and Microbiology, Joan C. Edwards School of Medicine, Marshall University,*²*Department of Pharmacology, Physiology and Toxicology, Joan C. Edwards School of Medicine, Marshall University*

The TALLYHO/Jng (TH) mouse is an inbred polygenic model for Type 2 diabetes characterized by obesity, insulin resistance, hyperlipidemia, and hyperglycemia. Genetic out-cross experiments with lean non-diabetic strain of C57BL/6 (B6) mice revealed major susceptibility loci for obesity on chromosome 6 (Tabw2) and for diabetes on chromosome 4 (Tanidd4). For both loci, the TH alleles are associated with disease susceptibilities. Currently, the Tabw2 locus has been fine mapped to 8-Mb interval and the Tanidd4 locus to 15.8-Mb interval.

In order to identify potential susceptibility variants for Tabw2 and Tanidd4 loci, we performed whole genome sequencing of the TH mouse genome in two 2 x 100 paired end read strategies using an Illumina HiSeq1000 next generation sequencer. Libraries were prepared and sequenced using standard Illumina protocols.

We generated ~60 Gb of TH sequence with an average coverage of ~25X. Seventy five percent of the readout aligned to the mouse reference B6 genome (Build 37). In our initial test, we compared the TH genome to the B6 genome using Illumina CASAVA software and identified 4,749,554 SNPs. Among the SNPs, 38,147 mapped to the Tabw2 interval and 83,122 to the Tanidd4 interval. In order to refine these sets, we compared the TH, which contains Swiss lineage, variant set to those variants in three Swiss-derived non-obese, non-diabetic strains (A/J, AKR/J and Balb/cJ). We identified 17,360 SNPs and 28,297 SNPs within the Tabw2 and Tanidd4 intervals, respectively, that were not found in any of the three non-diabetic strains.

In conclusion, we applied a whole-genome sequencing strategy in combination with mapping information to identify a set of candidate genes for the obese and diabetic phenotypes in TH mice. Further defined and systemic filtration of the variants (e.g. based on variant effect of function) will facilitate identification of the causal genes and variants.

155

Pathogenic Orphan Transduction Created by a Non-reference LINE-1 Retrotransposon**Szilvia Solyom**, Adam D. Ewing, Dustin C. Hancks, Yasuhiro Takeshima, Hiroyuki Awano, Masafumi Matsuo, Haig. H. Kazazian Jr.¹*Johns Hopkins University School of Medicine,* ²*Kobe University*

Long Interspersed Element-1 (LINE-1) retrotransposons comprise 17% of the human genome, and move by a potentially mutagenic "copy and paste" mechanism via an RNA intermediate. Recently, the retrotransposition-mediated in-

sertion of a new transcript was described as a novel cause of genetic disease, Duchenne muscular dystrophy, in a Japanese male. The inserted sequence was presumed to derive from a single-copy, non-coding RNA transcribed from chr. 11q22.3 that retrotransposed into the dystrophin gene on chr. X. Here, we demonstrate that a non-reference full-length LINE-1 (termed LRE4) is situated in the proband and maternal genome at chromosome 11q22.3, directly upstream of the sequence, whose copy was inserted into the dystrophin gene. LRE4 is highly active in a cell culture assay. LINE-1 insertions are often associated with 3' transduction of adjacent genomic sequences. Thus, the likely explanation for the mutagenic insertion is a LINE-1-mediated 3' transduction with severe 5' truncation. This is the first example of LINE-1-induced human disease caused by an "orphan" 3' transduction.

This work was carried out with the help of the Genetic Resources Core Facility (GRCF) at the Johns Hopkins University. LRE4 was found to be present in one Japanese individual out of 15 unrelated individuals in a LINE-1 targeted resequencing dataset. This dataset was generated by Illumina sequencing of the human-specific subfamily of LINE-1s (Ewing and Kazazian, 2010), using the service offered by the GRCF High Throughput Sequencing Center. Sequencing of the PCR products from the patient, as well as the cloned LRE4 sequence was done by the GRCF DNA Analysis Facility.

HISTIDINE TAGGED PROTEINS

156

Column Purification of Histidine-Tagged Proteins from Unclarified Samples**Lena Hornsten**, Marianne Carlsson, Jon Lundqvist, Therese Granér, Lars Andersson*GE Healthcare Bio-Sciences AB, Sweden*

The histidine-tag is today the most commonly used tag for recombinant proteins. It is widely used to facilitate purification of recombinant expressed proteins, but can also be useful for detection purposes. For immobilized metal-ion affinity chromatography (IMAC) the most used metal ion is nickel (Ni²⁺). However, also the cobalt ion (Co²⁺) has been shown useful in obtaining high purity with good yields of, especially low expressed, histidine-tagged target proteins in IMAC applications.

Pre-packed HiTrapTM IMAC crude column is a format which facilitates scalable (1-ml and 5-ml columns), fast and convenient protein purifications of histidine-tagged proteins. Un-clarified samples can be loaded directly to the column. The process time is decreased which helps to ensure high concentrated pure and intact target proteins with minimal degradation.

Several applications were performed to verify the performance of pre-packed HiTrapTM IMAC crude column. Two histidine-tagged proteins (produced in *E. coli* and *P. pastoris* expression systems) were purified ending up in 90% purity and yields of 90%. A fully automated (AKTApurifyTM) purification of a histidine-tagged kinase (1 mg/g *E. coli* cells) has been performed. The purification was done by IMAC and gel filtration. According to SDS-PAGE analysis the