Name: Boris Tabakoff Email: boris.tabakoff@cuanschutz.edu

Revisiting the DNA "Dark Matter"

Spencer Mahaffey[1,2], Boris Tabakoff[1,3], Paula L. Hoffman[1], Laura M. Saba[1,2]

[1]Lohocla Research Corporation, Aurora CO; [2]Skaggs School of Pharmacy and Pharmaceutical Sciences, Anshutz Medical Campus, Aurora CO; [3]Institute for Behavioral Genetics, University of Colorado Boulder, Boulder CO

Early in this century an argument arose regarding the proportion of the genome that produces functional RNA (Kapranov & St. Laurent, Front. Genet. 2012). It has become obvious that much of the eukaryotic DNA is a blueprint for functional RNAs that do not code for proteins (e.g., miRNA, lncRNA, snoRNA, etc.) However, a significant portion of DNA remains that produces unannotated transcripts. To determine whether these unannotated transcripts are yet-unidentified functional elements, we reasoned that if the transcript levels are determined by genetic (heritable) factors they are not likely to reflect a random phenomenon. We quantified "total" RNA by deep sequencing (~ 70M reads/individual) from brain (74 strains), liver (62), kidney (28) and heart (21) of the HRDP rat panel and reconstructed the reads into transcripts. When the transcripts were aligned on the recently generated DNA sequence of the rat (RN7.2), the transcripts (including their excised introns) produced a substantial coverage of the DNA sequence of all autosomes. Brain coverage was 64-74% across the 20 autosomes. Without reconstructing transcripts, nearly all of the DNA bases were covered by RNA reads. Each identified transcript's broad sense heritability (h2) values were calculated. Given our available data, ~65% of the transcripts in brain had h2 values >0.2. The results indicate that that in rat over 50% of the DNA is utilized by various organs to produce "transcripts" and a substantial portion of these "transcripts" reach the threshold of heritability values indicating significant genetic control of their steady-state levels.

Supported by NIH/NIAAA AA013162 and NIH/NIDA P30 DA044223.