Dark Matter of the Human Genome

Sarah Slavoff

Yale University, US
sarah.slavoff@yale.edu

Advanced methods in next-generation sequencing and proteogenomics have revealed thousands of previously invisible human genes, increasing the known size of the human genome by at least 10%. This previously unannotated "dark matter" of the human genome includes small open reading frames (smORFs) encoding polypeptides of fewer than 100 amino acids, and alternative open reading frames (alt-ORFs) encoding proteins 100 amino acids or larger. Sm/alt-ORFs previously escaped detection due to their short lengths, overlap with annotated protein coding sequences in different reading frames, and/or initiation with non-AUG start codons. Recent studies have shown that hundreds of smORFs are required for cell growth and survival, and some smORF-encoded polypeptides or "microproteins" bind to and regulate the activity of macromolecular complexes involved in critical cellular processes and disease. In this presentation, I will describe (1) proteomic discovery of alt-RPL36, which overlaps the coding sequence for human ribosomal protein L36 and binds to TMEM24 to regulate PI3K/AKT/mTOR pathway signaling, and (2) phosphoregulation of the intrinsically disordered NBDY microprotein that controls mRNA decay and membraneless organelle dynamics. Taken together, these findings demonstrate that microproteins represent a trove of new functional cellular molecules that function in disease-relevant processes.