

## THE DARK PROTEOME

In humans, the number of annotated canonical protein-coding genes is estimated to range between 20,000 and 25,000. However, alternative splicing dramatically increases the complexity of the proteome, with estimates surpassing 100,000 unique isoforms. Recent research has begun uncovering a “dark proteome,” composed of small proteins, as short as a dozen or fewer amino acids, derived from non-canonical open reading frames (ncORFs). A preprint published on BioRxiv (1) reports that at least 25% of the 7,264 identified ncORFs generate translational products, resulting in over 3,000 peptides.

The importance of these findings has been emphasized in a commentary featured in *Science* (2), underlining the potential implications for understanding human biology and identifying new therapeutic targets.

1. <https://www.biorxiv.org/content/10.1101/2024.09.09.612016v1>
2. <https://www.science.org/content/article/dark-proteome-survey-reveals-thousands-new-human-genes>