Computational model of translation initiation leaky scanning and its application to ribo-seq data.

Audrey M. Michel, Pavel V. Baranov

Biochemistry Department, University College Cork, Ireland, brave.oval.pan@gmail.com

Ribosome profiling (ribo-seq) is a recently developed technique that provides Genome Wide Information on Protein Synthesis (GWIPS) in vivo. It is based on the deep sequencing of ribosome protected mRNA fragments which allows the ribosome density along all mRNA transcripts present in the cell to be quantified. Recent studies have adapted the ribosome profiling technique to specifically capture initiating ribosomes [1,2,3]. As well as revealing surprisingly widespread initiation at non-AUG codons, ribosome profiling of initiating ribosomes allowed the identification of multiple translation initiation sites (TISs) for many mRNA transcripts in human and mouse. We have utilised initiating ribosome density data from these studies [1,2] to develop a computational model of translation initiation based on the process of leaky scanning [4]. The model allows to calculate the probability of initiation at a given start codon. Application of our model demonstrated that despite the high frequency of non-AUG initiation, the probability of initiation at non-AUG codons is considerably lower than at AUG codons. This suggests that initiation at most non-AUG codons contributes to the noise of gene expression rather than to the generation of protein products with functions that are distinct from AUG initiated products. Our computational model of leaky scanning provides a simple and efficient method for evaluating the strength of individual initiation codons based on ribosome profiling data.

1. A.Uthor, O.N.E.More (2003) Title of the paper, *Journal*, **12:**13–16.

2. A.U.Thor et al. (2003) Title of the paper, *Journal*, **12:**13–16.

3. A.Uthor (2003) Title of the chapter, In: *The Book*, E.D.Itor et al. (Eds.), 13–16 (Publisher).

1. Ingolia NT, Lareau LF, Weissman JS. (2011) Ribosome profiling of mouse embryonic

stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell*. **147:**789-802.

2. Lee S, Liu B, Huang SX, Shen B, Qian SB. (2012) Global mapping of translation initiation sites in mammalian cells at single-nucleotide resolution. *Proc Natl Acad Sci U S A*, **109**:E2424-2432.

3. Fritsch C, Herrmann A, Nothnagel M, Szafranski K, Huse K, Schumann F, Schreiber S, Platzer M, Krawczak M, Hampe J, Brosch M. (2012) Genome-wide search for novel human uORFs and N-terminal protein extensions using ribosomal footprinting. *Genome Res*, **22**:2208-2218.

4. Kozak M. (1999) Initiation of translation in prokaryotes and eukaryotes. *Gene,* **234**: 187-208.