

## TECHNICAL BRIEF

# GWIPS-viz as a tool for exploring ribosome profiling evidence supporting the synthesis of alternative proteoforms

Audrey M. Michel, Anna M. Ahern, Claire A. Donohue and Pavel V. Baranov

School of Biochemistry and Cell Biology, University College Cork, Cork, Ireland

The boundaries of protein coding sequences are more difficult to define at the 5' end than at the 3' end due to potential multiple translation initiation sites (TISs). Even in the presence of phylogenetic data, the use of sequence information only may not be sufficient for the accurate identification of TISs. Traditional proteomics approaches may also fail because the N-termini of newly synthesized proteins are often processed. Thus ribosome profiling (ribo-seq), producing a snapshot of the ribosome distribution across the entire transcriptome, is an attractive experimental technique for the purpose of TIS location exploration. The GWIPS-viz (Genome Wide Information on Protein Synthesis visualized) browser (<http://gwips.ucc.ie>) provides free access to the genomic alignments of ribo-seq data and corresponding mRNA-seq data along with relevant annotation tracks. In this brief, we illustrate how GWIPS-viz can be used to explore the ribosome occupancy at the 5' ends of protein coding genes to assess the activity of AUG and non-AUG TISs responsible for the synthesis of proteoforms with alternative or heterogeneous N-termini. The presence of ribo-seq tracks for various organisms allows for cross-species comparison of orthologous genes and the availability of datasets from multiple laboratories permits the assessment of the technical reproducibility of the ribosome densities.

Received: December 16, 2014

Revised: February 3, 2015

Accepted: February 26, 2015

**Keywords:**

Bioinformatics / Non-AUG initiation / ribo-seq / Ribosome profiling / Translatomics / uORF

Usually only a single translation initiation site (TIS) is annotated for each protein coding ORF in the mRNAs of most eukaryotic organisms. However, due to leaky scanning [1], downstream AUG codons are also frequently used as TISs, see [2] for a review. Proteins produced as a result of such translation are truncated at the N-terminal ends in comparison with annotated protein sequences. Alternative TISs may also be found upstream of annotated TISs, for example, when translation initiation takes place on upstream in-frame AUG codons. This situation, however, is very rare since the first

in-frame AUG codon is usually annotated as the TIS. More frequently, proteoforms with N-terminal extensions are not detected because non-AUG codons are used for initiation such as CUG [3], or when non-AUG initiation is modulated by RNA structures formed by nucleotide repeats [4–6], or when they involve tRNAs other than initiator tRNA-Met [7], or if initiation takes place in the ribosome A site as in some viral mRNAs [8, 9]. The functionality of many N-terminal extensions produced by such initiation events is often strongly supported by the signatures of protein evolution provided with phylogenetic analysis [10]. A remarkable example occurs in the tumor suppressor *PTEN*, where an N-terminal extension allows an alternative *PTEN* proteoform to be secreted [11]. The diversity and complexity of translation initiation mechanisms complicates the detection of TISs when based purely on sequence analysis even in the presence of phylogenetic data.

Ribosome profiling (ribo-seq) [12] is a promising experimental technique that helps to identify TISs that are operative under given conditions. The technique is based on

**Correspondence:** Dr. Pavel Baranov, School of Biochemistry and Cell Biology, Western Gateway, University College Cork, Cork, Ireland

**E-mail:** P.Baranov@ucc.ie

**Fax:** +353 21 4205462

**Abbreviations:** CDS, coding sequence; ribo-seq, ribosome profiling; TIS, translation initiation site; uORF, upstream ORF; 3' UTR, 3' untranslated terminal region

using translation elongation inhibitors, RNA digestion, and massively parallel sequencing of ribosome protected mRNA fragments followed by their alignment to reference sequences (see [13, 14] for reviews). Attempts to enrich ribosomes at TIS locations have been done using specific drugs that block ribosomes before the formation of the first peptide bond [15, 16] or by a combination of drugs [17, 18]. While useful, these techniques suffer from high levels of signal noise, for example, elongating ribosomes pausing at particular codons [19, 20]. Therefore, the use of ribo-seq data is particularly powerful in combination with other techniques, for example, phylogenetic approaches [10] or MS [21–23].

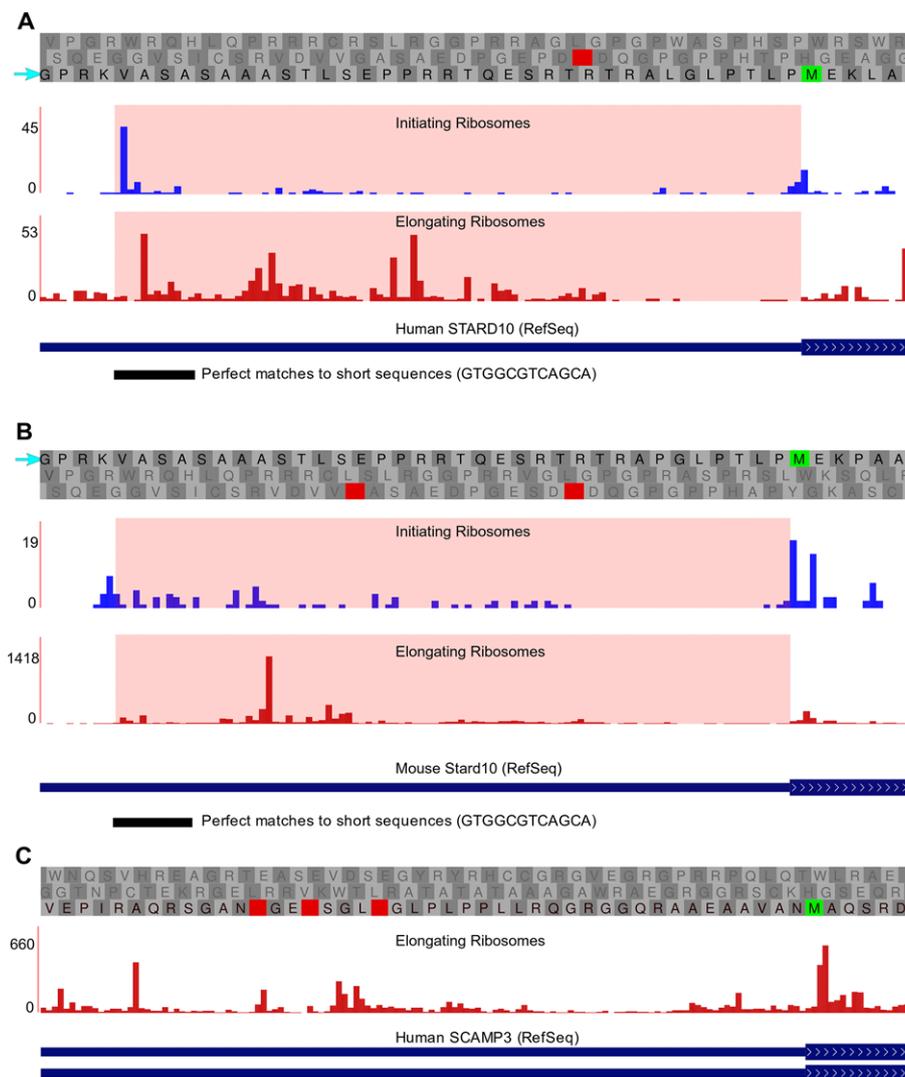
The GWIPS-viz (Genome Wide Information on Protein Synthesis visualized) browser [24] is a specialized ribo-seq browser available at <http://gwips.ucc.ie>. At present it provides access to the genomic alignments of public ribo-seq reads in conjunction with mRNA-seq reads along with relevant annotation tracks. Thus, GWIPS-viz is a powerful tool for researchers seeking supporting ribo-seq evidence for alternative proteoforms inferred from phylogenetic analysis or detected with proteomics or other experimental techniques. The GWIPS-viz genome browser is based on the University of California Santa Cruz (UCSC) Genome Browser (<http://genome.ucsc.edu/>) [25], but it is specifically tailored for the visualization of ribo-seq data. The alignments are visualized in two modes. One mode is a coverage plot, where the number of sequence reads aligning to each coordinate is displayed as a bar. Coverage plots are provided for both ribo-seq data and mRNA-seq data. The other mode (profile) is used to display the inferred positions of the decoding ribosomes. For the eukaryotic datasets, the coordinates of the A-site (for elongating ribosomes) or the P-site (for initiating ribosomes) are inferred for each sequence read by adding a specific offset to the coordinate of the most 5' nucleotide of the read. For the prokaryotic datasets, a center-weighted approach [26] is used to indicate the most probable positions of the ribosome A-site. It needs to be noted that this approach predicts the most likely location of the A-site, rather than its real location, because the exact length of the footprints as well as the symmetry of the regions flanking the A-site location are influenced by sequence dependent interactions between ribosomal RNA and mRNA [27]. In the current color scheme in GWIPS-viz, elongating ribosomes data are shown as red, initiating ribosomes as blue, and mRNA-seq data as green.

Separate tracks for each ribo-seq study enable cross-study and even cross-species comparisons. In addition, to improve the overall signal, we have aggregated ribo-seq data from several studies into *Global aggregate* tracks. These cumulative global tracks are currently provided for human, mouse, yeast, zebrafish and nematode and new datasets are integrated when available. We would recommend using the global tracks for initial exploration of translated regions, while the individual study tracks can be consulted for more detailed examination.

In the following section, we explain in detail how to use GWIPS-viz to manually examine available ribo-seq data

for supporting evidence for the synthesis of proteoforms extended at the N-terminus and how to prepare publication quality screenshots. For this purpose, we use the human *STARD10* gene for which phylogenetic data have been predicted to encode a GUG-initiated proteoform with an N-terminal extension [10]:

- (1) Under the *Genomes* page of GWIPS-viz, select your organism of interest in the genome drop-down menu (human in this case) and in the gene search box enter the gene symbol (*STARD10* in this example).
- (2) Scroll down to the following track sections *ribosome profiles—initiating ribosomes*, *ribosome profiles—elongating ribosomes* and *mRNA-seq coverage plots* and set the corresponding *global aggregate* track to full. In addition, set the *base position* and *Refseq genes* tracks from the *annotation tracks* group to full.
- (3) There is one RefSeq [28] transcript variant for the human *STARD10* gene (see Fig. 1A). The transcript representation is inherited from the UCSC Genome Browser, that is, exons are shown as blocks connected by introns represented as lines with arrowheads displaying the direction of transcription/translation. The annotated coding sequences (CDSs) are shown as thicker blocks than the rest of the transcript exons. The annotated start codon for *STARD10* is located in the second exon.
- (4) To zoom into the region of interest (the second exon in this case), click and hold the left mouse button on one edge of the second exon in the *base position* track and drag the mouse to the other edge of the second exon and then release the mouse button.
- (5) To display the location of the predicted TIS, click on the *short match* track link in the *annotation tracks* group and in the *short (2–30 base) sequence* search box on the *perfect match to short sequence* page, enter “GTG” followed by a string of nucleotides, which is long enough to be unique within the displayed locus (in our case GTGGCGTCAGCA—the location of the predicted TIS is annotated in the supplementary material of [10]). Note that since GWIPS-viz uses DNA sequence as a reference, sequences of DNA rather than RNA nucleotides should be used, for example, GTG for GUG. Also note that the *short match* feature only works for nucleotide sequences. After entering the sequence, set the display mode to full and click submit. The location of this sequence should be displayed as a black block in the corresponding track (see Fig. 1A). An alternative approach is to configure the *base position* track from the *annotation tracks* group to display the sequence of interest. Click on the *base position* link and enter the sequence in the *Motifs to highlight* search box in the configuration window of the track. Note that in order to see the sequence highlighted, the zoom should be at the base level.
- (6) To explore ribo-seq data from the individual studies that contribute to the *Global Track*, additional tracks can be displayed.



**Figure 1.** (A) Initiating (blue) and elongating (red) ribosomes profiles for the second exon of the human *STARD10* gene. The region of the predicted GUG-initiated CDS extension [10] is highlighted in pale pink. In the reading frame organization, AUG codons are represented by green boxes and stop codons are represented by red boxes. The blue arrow points to the reading frame of the CDS extension. Also displayed is the *short match* detected location of the GUG start codon and part of its 3' flanking nucleotide sequence. The initiating ribosome profile was generated using the *Global aggregate* track in GWIPS-viz (<http://gwips.ucc.ie>) using data from four studies [15, 17, 18, 29]. The elongating ribosomes profile was generated using data aggregated from 13 studies [15, 17, 18, 29, 38]. (B) Initiating ribosomes profile and elongating ribosomes profile for the second exon of the mouse *Stard10* gene. The description is the same as for panel A except that the initiating ribosomes profile was generated from two studies [15, 18] and the elongating ribosomes profile from five studies [15, 16, 37, 39, 40]. (C) Elongating ribosomes profile for the first exon of the human *SCAMP3* gene. The ribosome footprints in the region upstream of the annotated AUG start codon (highlighted in pale pink) cannot originate from the translation of an extended CDS as there are three upstream stop codons in the same reading frame. The profile was generated from the same studies as panel A.

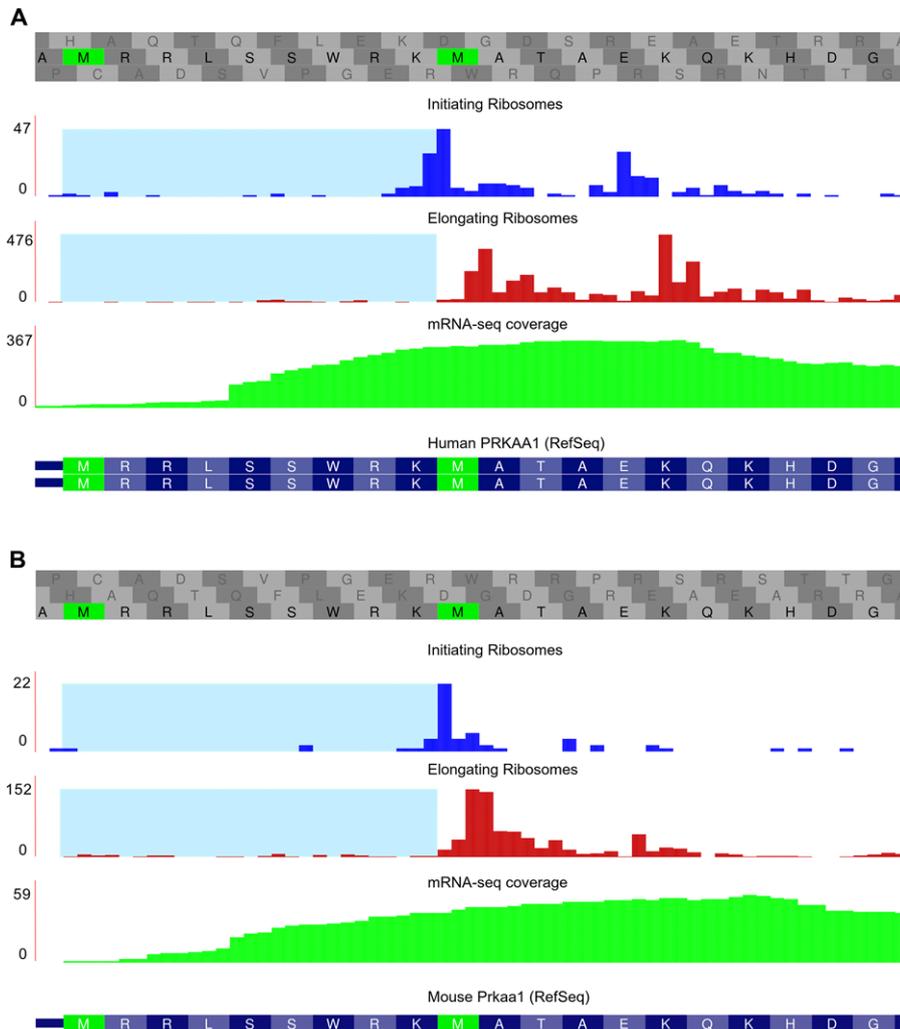
As can be seen in Fig. 1A, which displays the second exon of the human *STARD10* locus, there is an enrichment of initiating ribosomes at the location of the predicted GUG TIS (blue profile), while elongating ribosomes (red profile) span the entire downstream region up to, and past, the annotated AUG start codon (highlighted in pale pink in Fig. 1A).

The availability of data for several organisms allows us to explore whether the corresponding region is translated in orthologs of the human *STARD10* gene. The display in Fig. 1B is generated by a similar procedure to above, but for the mouse *Stard10* gene. Thus, Fig. 1A and B suggests that the synthesis of the N-terminal extended products of *STARD10* and *Stard10* is a conserved feature.

How to determine if the footprints aligning to a 5' leader region originate from ribosomes translating the same ORF as the one to which the annotated CDS belongs or from ribosomes translating a different ORF? The answer to this question may be determined with subsequent analysis of the triplet periodicity within the corresponding region [41, 42].

However, even in the absence of periodicity analysis, a simple examination of the ORF organization could rule out the possibility of an extended CDS. In both the human *STARD10* gene and the mouse *Stard10* gene, there is an absence of stop codons in the reading frame upstream of the annotated start codon (Fig. 1A and B). Hence, it is possible that footprints are generated from the same ORF as that of the annotated CDS. Contrast this with the sequence of the first exon of the human *SCAMP3* gene for which ribo-seq data are displayed in Fig. 1C. The stop codons in the same reading frame upstream of the annotated start codon rule out the possibility of an extended CDS and so the footprints do not belong to the same ORF as the annotated CDS. Instead the ribosome footprints likely originate from an overlapping upstream ORF (uORF) as reported earlier, see Supporting Information of [41].

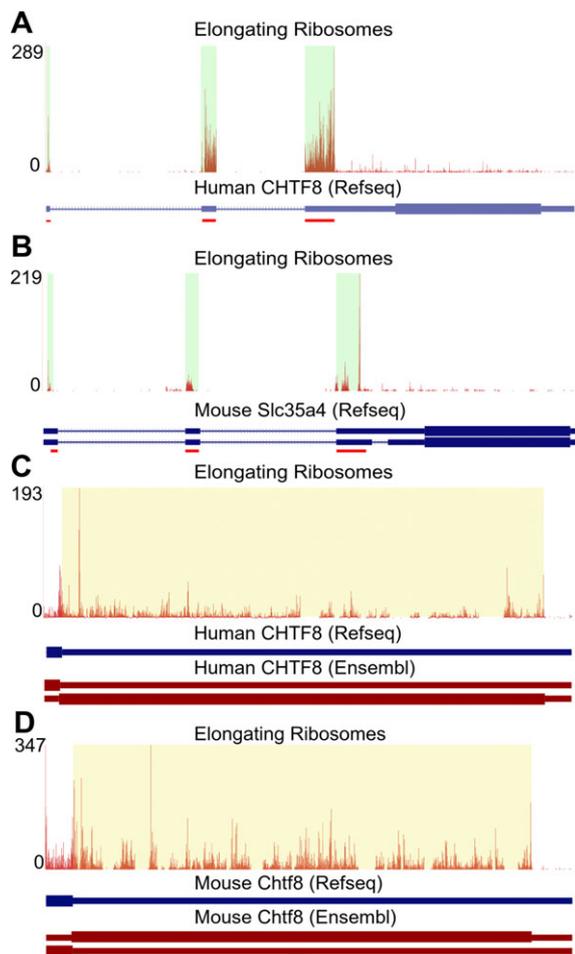
GWIPS-viz can also be used for the identification of TISs responsible for the synthesis of proteoforms truncated at the N-terminus, particularly when initiation at the downstream TIS is substantially more efficient. This is the case with the



**Figure 2.** (A) Ribosome profiles and mRNA-seq coverage plot for the first exon of the human *PRKAA1* gene. The density of elongating ribosomes (red profile) in the predicted CDS truncated region is very low (highlighted in pale blue), while there is an increase in density downstream of the second AUG codon. In addition, there is an enrichment of initiating ribosomes (blue profile) at the second AUG. The mRNA-seq coverage plot (green) suggests an alternative transcript isoform truncated at the 5' end. The ribosomes profiles were generated from the same studies as in Fig. 1A and the mRNA-seq coverage plot was generated using data from seven studies [29, 32–37]. (B) Ribosome profiles and mRNA-seq coverage plot for the first exon of the mouse *Prkaa1* gene. The description is the same as for panel A except that the ribosome profiles were generated from the same studies as in Fig. 1B and the mRNA-seq coverage plot was generated using data from four studies [16, 37, 39, 40].

human *PRKAA1* gene for which the synthesis of a truncated proteoform has been validated previously at the protein level [15, 22]. Figure 2A and B shows ribo-seq data and mRNA-seq data aligned to the first exon of the human *PRKAA1* gene and the mouse *Prkaa1* gene, respectively. In both cases, the density of elongating ribosomes (red) in the region between the first and the second AUG (highlighted in pale blue) is much lower than downstream of the second AUG. This suggests that the second AUG corresponds to a more efficient TIS. In addition, there is an enrichment of initiating ribosomes at the second AUG (blue plot). Interestingly, the mRNA-seq density is distributed in a similar manner and therefore it is likely that the reason is not leakiness of the first AUG, but due either to a misannotation of the transcription start site and the translation start site in the current databases or due to an alternative transcript isoform truncated at the 5' end in comparison with the RefSeq transcript either as a result of alternative transcription initiation or post-transcriptional processing of the 5'-end.

In addition to facilitating the exploration and authentication of truncated or extended proteoforms, GWIPS-viz can be used to identify or corroborate the translation of ORFs alternative to the one corresponding to the CDS. Figure 3A and B shows ribo-seq data aligned to the locus of *SLC35A4* in human and *Slc35a4* in mouse, respectively. Most of the ribo-seq density occurs in a uORF of the corresponding transcripts rather than its annotated CDS. The product of this uORF (MADDKDSLPLKDLAFLKNQLESQRVEDEVNSGVGQDGSLLSPFLKGLAGYVAVKLRAVAVLGFVGTCTGIYAAQAYAVPNVEKTLRDYLLRKGPD for *SLC35A4* in human and MADDKDSLPLKDLTFLKNQLERLQQRVEDEVNSGVGQDGSLLSPFFKGLAGYVAVKLRAVAVLGFVGTCTGIYAAQAYAVPNVEKALKNYFRSLRKGPD for *Slc35a4* in mouse) was detected earlier with MS [43, 44]. Phylogenetic evidence also strongly supports the functional importance of this uORF product in addition to the uORF's regulatory role in providing stress resistance to the translation of the annotated CDS [45]. The detection of a product



**Figure 3.** (A) Elongating ribosomes profile for the human *SLC35A4* gene. The density of ribosome footprints in the 5' leader region (highlighted in pale green) is clearly much higher than in the predicted CDS (the lighter blue color of the RefSeq mRNA transcript indicates that this transcript annotation is “provisional”). The red lines under the RefSeq transcript indicate the location of the predicted uORF. The profile was generated from the same studies as in Fig. 1A. (B) Elongating ribosomes profile for the mouse *Slc35a4* gene. The profile was generated from the same studies as in Fig. 1B. (C) Elongating ribosomes profile for the last exon of the human *CHTF8* gene. There is clearly a high ribosome occupancy (highlighted in pale yellow) in the 3'UTR region of this gene. An alternative Ensembl transcript for this locus contains a CDS that corresponds exactly to the location of the ribosome footprints in the highlighted region. The profile was generated based on the data from the same studies as in Fig. 1A. (D) Elongating ribosomes profile for the last exon in the mouse *Chtf8* gene. The description is the same as for panel C except that the profile was generated from the same studies as in Fig. 1B.

translated from an ORF in the 3'UTR of the human *CHTF8* gene has also been reported [43, 44]. Indeed in GWIPS-viz, a high ribosome occupancy can be seen in the 3'UTR region of the human *CHTF8* transcripts (Fig. 3C) as well as for the mouse orthologs (Fig. 3D). These ribosomes likely originated

from the translation of this downstream ORF that is located in an alternative reading frame to the canonical ORF of this gene. This, however, does not necessarily indicate that *CHTF8* mRNA is bicistronic as it is possible that the annotated CDS and downstream ORF are translated in different transcripts originated at the *CHTF8* locus. Indeed a number of alternative transcripts exist among Ensembl [46] genes (including transcript ENST00000306585 shown for human in Fig. 3C and transcript ENSMUST00000169312 shown for mouse in Fig. 3D) in addition to the Refseq [28] transcripts.

Newly published ribo-seq datasets are continuously added to GWIPS-viz. Researchers can also explore their own ribo-seq data by using the *add custom tracks* functionality. The custom track is only visible to the owner and it allows the researcher to explore their own data in the context of other published datasets. Snapshot figures of the ribosome profiles can be generated using the *PS/PDF* link in the top menu bar and the alignment count data for all tracks are available to download from the *Tables* link for defined regions or for the entire genome.

It is clear that the ribo-seq technique has a number of limitations in relation to the prediction of alternative proteoforms. The technique does not allow for the discrimination between reads produced from alternative transcript isoforms if they share the same sequences. Often it can be difficult to detect precisely which ORF is being translated due to the convolution of the triplet periodicity signal that occurs when several short overlapping ORFs are translated simultaneously. Nonetheless, as we demonstrated, GWIPS-viz can be used as a powerful support tool for predictions based on other approaches as well as for generating hypotheses that can be tested using methods other than ribo-seq.

*This work is supported by Science Foundation Ireland (12/IA/1335) and the Wellcome Trust (094423) to PVB. The authors also thank everyone who contributed to the ribosome profiling tracks in the GWIPS-viz browser.*

*The authors have declared no conflict of interest.*

## References

- [1] Kozak, M., Initiation of translation in prokaryotes and eukaryotes. *Gene* 1999, **234**, 187–208.
- [2] Kochetov, A. V., Alternative translation start sites and hidden coding potential of eukaryotic mRNAs. *Bioessays* 2008, **30**, 683–691.
- [3] Kozak, M., Context effects and inefficient initiation at non-AUG codons in eucaryotic cell-free translation systems. *Mol. Cell. Biol.* 1989, **9**, 5073–5080.
- [4] Kearse, M. G., Todd, P. K., Repeat-associated non-AUG translation and its impact in neurodegenerative disease. *Neurotherapeutics* 2014, **11**, 721–731.
- [5] Cleary, J. D., Ranum, L. P. W., Repeat-associated non-ATG (RAN) translation in neurological disease. *Hum. Mol. Genet.* 2013, **22**, R45–R51.

- [6] Wojciechowska, M., Olejniczak, M., Galka-Marciniak, P., Jazurek, M., Krzyzosiak, W. J., RAN translation and frameshifting as translational challenges at simple repeats of human neurodegenerative disorders. *Nucleic Acids Res.* 2014, *42*, 11849–11864.
- [7] Starck, S. R., Jiang, V., Pavon-Eternod, M., Prasad, S. et al., Leucine-tRNA initiates at CUG start codons for protein synthesis and presentation by MHC class I. *Science* 2012, *336*, 1719–1723.
- [8] Wilson, J. E., Pestova, T. V., Hellen, C. U., Sarnow, P., Initiation of protein synthesis from the A site of the ribosome. *Cell* 2000, *102*, 511–520.
- [9] Spahn, C. M. T., Jan, E., Mulder, A., Grassucci, R. A. et al., Cryo-EM visualization of a viral internal ribosome entry site bound to human ribosomes: the IRES functions as an RNA-based translation factor. *Cell* 2004, *118*, 465–475.
- [10] Ivanov, I. P., Firth, A. E., Michel, A. M., Atkins, J. F., Baranov, P. V., Identification of evolutionarily conserved non-AUG-initiated N-terminal extensions in human coding sequences. *Nucleic Acids Res.* 2011, *39*, 4220–4234.
- [11] Hopkins, B. D., Fine, B., Steinbach, N., Dendy, M. et al., A secreted PTEN phosphatase that enters cells to alter signaling and survival. *Science* 2013, *341*, 399–402.
- [12] Ingolia, N. T., Ghaemmaghami, S., Newman, J. R. S., Weissman, J. S., Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* 2009, *324*, 218–223.
- [13] Michel, A. M., Baranov, P. V., Ribosome profiling: a Hi-Def monitor for protein synthesis at the genome-wide scale. *Wiley Interdiscip. Rev. RNA* 2013, *4*, 473–490.
- [14] Ingolia, N. T., Ribosome profiling: new views of translation, from single codons to genome scale. *Nat. Rev. Genet.* 2014, *15*, 205–213.
- [15] Lee, S., Liu, B., Lee, S., Huang, S.-X. et al., Global mapping of translation initiation sites in mammalian cells at single-nucleotide resolution. *Proc. Natl. Acad. Sci. USA* 2012, *109*, E2424–E2432.
- [16] Ingolia, N. T., Lareau, L. F., Weissman, J. S., Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* 2011, *147*, 789–802.
- [17] Fritsch, C., Herrmann, A., Nothnagel, M., Szafranski, K. et al., Genome-wide search for novel human uORFs and N-terminal protein extensions using ribosomal footprinting. *Genome Res.* 2012, *22*, 2208–2218.
- [18] Gao, X., Wan, J., Liu, B., Ma, M. et al., Quantitative profiling of initiating ribosomes in vivo. *Nat. Methods* 2014, *12*(2), 147–153.
- [19] Dmitriev, S. E., Akulich, K. A., Andreev, D. E., Terenin, I. M., Shastky, I. N., The peculiar mode of translation elongation inhibition by antitumor drug harringtonin. *FEBS J.* 2013, *280*, 51–51.
- [20] Michel, A. M., Andreev, D. E., Baranov, P. V., Computational approach for calculating the probability of eukaryotic translation initiation from ribo-seq data that takes into account leaky scanning. *BMC Bioinform.* 2014, *15*, 380–389.
- [21] Menschaert, G., Van Criekeing, W., Notelaers, T., Koch, A. et al., Deep proteome coverage based on ribosome profiling aids mass spectrometry-based protein and peptide discovery and provides evidence of alternative translation products and near-cognate translation initiation events. *Mol. Cell. Proteomics* 2013, *12*, 1780–1790.
- [22] Van Damme, P., Gawron, D., Van Criekeing, W., Menschaert, G., N-terminal proteomics and ribosome profiling provide a comprehensive view of the alternative translation initiation landscape in mice and men. *Mol. Cell. Proteomics* 2014, *13*, 1245–1261.
- [23] Koch, A., Gawron, D., Steyaert, S., Ndah, E., Crappé, J., A proteogenomics approach integrating proteomics and ribosome profiling increases the efficiency of protein identification and enables the discovery of alternative translation start sites. *Proteomics* 2014, *14*, 2688–2698.
- [24] Michel, A. M., Fox, G., M Kiran, A., De Bo, C. et al., GWIPS-viz: development of a ribo-seq genome browser. *Nucleic Acids Res.* 2014, *42*, D859–D864.
- [25] Karolchik, D., Barber, G. P., Casper, J., Clawson, H. et al., The UCSC genome browser database: 2014 update. *Nucleic Acids Res.* 2014, *42*, D764–D770.
- [26] Oh, E., Becker, A. H., Sandikci, A., Huber, D. et al., Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor in vivo. *Cell* 2011, *147*, 1295–1308.
- [27] O'Connor, P. B. F., Li, G.-W., Weissman, J. S., Atkins, J. F., Baranov, P. V., rRNA: mRNA pairing alters the length and the symmetry of mRNA-protected fragments in ribosome profiling experiments. *Bioinformatics* 2013, *29*, 1488–1491.
- [28] Pruitt, K. D., Brown, G. R., Hiatt, S. M., Thibaud-Nissen, F. et al., RefSeq: an update on mammalian reference sequences. *Nucleic Acids Res.* 2014, *42*, D756–D763.
- [29] Stern-Ginossar, N., Weisburd, B., Michalski, A., Le, V. T. K. et al., Decoding human cytomegalovirus. *Science* 2012, *338*, 1088–1093.
- [30] Loayza-Puch, F., Drost, J., Rooijers, K., Lopes, R. et al., p53 induces transcriptional and translational programs to suppress cell proliferation and growth. *Genome Biol.* 2013, *14*, R32–R43.
- [31] Gonzalez, C., Sims, J. S., Hornstein, N., Mela, A. et al., Ribosome profiling reveals a cell-type-specific translational landscape in brain tumors. *J. Neurosci.* 2014, *34*, 10924–10936.
- [32] Wein, N., Vulin, A., Falzarano, M. S., Szegarty, C. A.-K. et al., Translation from a DMD exon 5 IRES results in a functional dystrophin isoform that attenuates dystrophinopathy in humans and mice. *Nat. Med.* 2014, *20*, 992–1000.
- [33] Stumpf, C. R., Moreno, M. V., Olshen, A. B., Taylor, B. S., Ruggero, D., The translational landscape of the mammalian cell cycle. *Mol. Cell* 2013, *52*, 574–582.
- [34] Stadler, M., Fire, A., Wobble base-pairing slows in vivo translation elongation in metazoans. *RNA* 2011, *17*, 2063–2073.
- [35] Reid, D. W., Nicchitta, C. V., Primary role for endoplasmic reticulum-bound ribosomes in cellular translation identified by ribosome profiling. *J. Biol. Chem.* 2012, *287*, 5518–5527.

- [36] Hsieh, A. C., Liu, Y., Edlind, M. P., Ingolia, N. T. et al., The translational landscape of mTOR signalling steers cancer initiation and metastasis. *Nature* 2012, *485*, 55–61.
- [37] Guo, H., Ingolia, N. T., Weissman, J. S., Bartel, D. P., Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature* 2010, *466*, 835–840.
- [38] Liu, B., Han, Y., Qian, S.-B., Cotranslational response to proteotoxic stress by elongation pausing of ribosomes. *Mol. Cell* 2013, *49*(3), 453–463.
- [39] Thoreen, C. C., Chantranupong, L., Keys, H. R., Wang, T. et al., A unifying model for mTORC1-mediated regulation of mRNA translation. *Nature* 2012, *485*, 109–113.
- [40] Howard, M. T., Carlson, B. A., Anderson, C. B., Hatfield, D. L., Translational redefinition of UGA codons is regulated by selenium availability. *J. Biol. Chem.* 2013, *288*, 19401–19413.
- [41] Michel, A. M., Choudhury, K. R., Firth, A. E., Ingolia, N. T. et al., Observation of dually decoded regions of the human genome using ribosome profiling data. *Genome Res.* 2012, *22*, 2219–2229.
- [42] Bazzini, A. A., Johnstone, T. G., Christiano, R., Mackowiak, S. D. et al., Identification of small ORFs in vertebrates using ribosome footprinting and evolutionary conservation. *EMBO J.* 2014, *33*, 981–993.
- [43] Vanderperre, B., Lucier, J.-F., Bissonnette, C., Motard, J. et al., Direct detection of alternative open reading frames translation products in human significantly expands the proteome. *PLoS One* 2013, *8*, e70698.
- [44] Kim, M.-S., Pinto, S. M., Getnet, D., Nirujogi, R. S. et al., A draft map of the human proteome. *Nature* 2014, *509*, 575–581.
- [45] Andreev, D. E., O'Connor, P. B., Fahey, C., Kenny, E. M. et al., Translation of 5' leaders is pervasive in genes resistant to eIF2 repression. *Elife* 2015, *4*:e03971.
- [46] Flicek, P., Amode, M. R., Barrell, D., Beal, K. et al., Ensembl 2014. *Nucleic Acids Res.* 2014, *42*, D749–D755.